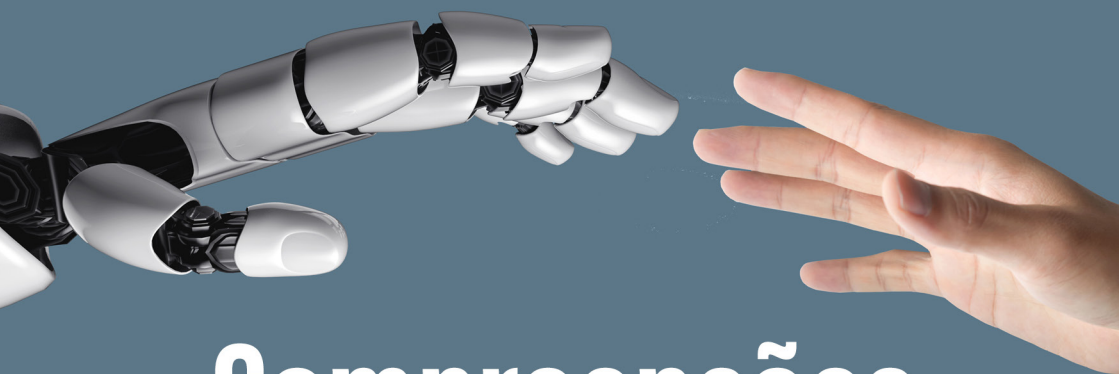


Angela Grossi, Pollyana Ferrari e Xabier Martínez-Rolán (Coords.)



Compreensões sobre a IA Generativa

Compreensões sobre a IA Generativa

Angela Grossi
Pollyana Ferrari
Xabier Martínez-Rolán
(Coordenação)

RIA
Editorial

Ria Editorial - Conselho Editorial

PhD Abel Suing (UTPL, Equador)

PhD Andrea Versuti (UnB, Brasil)

PhD Anton Szomolányi (Pan-European University, Eslováquia)

PhD Carlos Arcila (Universidade de Salamanca, Espanha)

PhD Catalina Mier (UTPL, Equador)

PhD Denis Porto Renó (Universidade Estadual Paulista - UNESP, Brasil)

PhD Fátima Lopes Cardoso (ESCS/IPL, Portugal)

PhD Fernando Gutierrez (ITESM, México)

Ms. Fernando Irigaray (Universidade Nacional de Rosario, Argentina)

PhD Florian Andrei Vlad (Ovidius University of Constanta, Romania)

PhD Gerson Martins (Universidade Federal de Mato Grosso do Sul - UFMS, Brasil)

PhD Jerónimo Rivera (Universidade La Sabana, Colombia)

PhD Jesús Flores Vivar (Universidade Complutense de Madrid, Espanha)

PhD João Canavilhas (Universidade da Beira Interior, Portugal)

PhD John Pavlik (Rutgers University, Estados Unidos)

PhD Joseph Straubhaar (Universidade do Texas - Austin, Estados Unidos)

PhD Juliana Colussi (Universidade Rey Juan Carlos, Espanha)

PhD Koldo Mesos (Universidade do País Vasco, Espanha)

PhD Lionel Brossi (Universidade do Chile, Chile)

PhD Lorenzo Vilches (Universidade Autônoma de Barcelona, Espanha)

PhD Manuela Penafria (Universidade da Beira Interior, Portugal)

PhD Marcelo Martínez (Universidade de Santiago de Compostela, Espanha)

PhD Marcos Pereira dos Santos (Univ. Tec. Federal do Paraná - UTFPR e
Fac. Rachel de Queiroz, Brasil)

PhD Maria Cristina Gobbi (Universidade Estadual Paulista - UNESP, Brasil)

PhD Oksana Tymoshchuk (Universidade de Aveiro, Portugal)

PhD Paul Levinson (Fordham University, Estados Unidos)

PhD Pedro Nunes (Universidade Federal da Paraíba - UFPB, Brasil)

PhD Piero Dominici (Universidade de Perugia, Italia)

PhD Raquel Longhi (Universidade Federal de Santa Catarina - UFSC, Brasil)

PhD Ricardo Alexino Ferreira (Universidade de São Paulo - USP, Brasil)

PhD Sergio Gadini (Universidade Estadual de Ponta Grossa - UEPG, Brasil)

PhD Thom Gencarelli (Manhattan College, Estados Unidos)

PhD Vicente Gosciola (Universidade Anhembi Morumbi, Brasil)

FICHA TÉCNICA

Copyright 2025 ©Autoras e autores. Todos os direitos reservados

Foto de capa: ©InfiniteFlow - stock.adobe.com (arquivo nº 330431165)

Design da capa: ©Denis Renó

Diagramação: Luciana Renó

ISBN 978-989-9220-50-8

Título: Compreensões sobre a IA Generativa

Coordenadores: Angela Grossi, Pollyana Ferrari e Xabier Martínez-Rolán

1.ª edição, 2025.



Esta obra tem licença Creative Commons **Attribution-NonCommercial-NoDerivatives**. Você tem o direito de compartilhar, copiar e redistribuir o material em qualquer suporte ou formato sempre que seja feito o reconhecimento de seus autores, não utilizá-la para fins comerciais e não modificar a obra de nenhuma forma.

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

©Ria Editorial

RIA
Editorial

riaeditora@gmail.com

<http://www.riaeditorial.com>

ESSA OBRA FOI AVALIADA POR PARECERISTAS

Todos os textos foram avaliados e selecionados pelos organizadores da obra. Os comentários dos organizadores foram enviados aos autores, que, mediante a aprovação, receberam tempo hábil para eventuais correções.

O livro foi posteriormente avaliado e aprovado pelo avaliador externo Dr. Denis Renó, que informou parecer positivo à publicação da seguinte forma:

O livro aborda temas atuais relacionados à Inteligência Artificial Generativa, perpassando por questões éticas, sociais e tecnológicas. Trata-se de uma obra que atualiza os estudos sobre o tema, algo necessário, dada a velocidade do seu desenvolvimento. De igual maneira, oferece leituras por olhares diversos, pois resulta de uma organização de obra. Recomendo a sua publicação, com a certeza de que será de grande valor científico.

O parecer foi enviado previamente ao lançamento.

Autoras e autores

Abel Suing

Alan César Belo Angeluci

Alexandre Farbiarz

Aloísio Corrêa de Araújo

Amanda Cristina Trentin

Angela Maria Grossi

Carolina Gois Falandes

Cláudia Albuquerque Thomé

Claviano Nascimento de Sousa

Dick Ronald Cáceres Navarro

Diogo de Almeida Camelo

Eli Borges Junior

Guilherme Gonçalves Mello

Jesús Miguel Flores-Vivar

João Pedro de Toledo Piza

Laryssa da Costa Gabellini

Leandro Marlon Barbosa Assis

Lucas Lima Coaracy

Luis Alberto Vásquez Ampudia

Margareth Boarini

Pablo Daniel Gutiérrez Gutiérrez

Pollyana Ferrari Teixeira

Rafael de Araújo Mélo

Régis Martins

Renato Sobhie Zambonato

Rodrigo Daniel Levoti Portari
Sabrina Marcandali
Vânia Cristina Pires Nogueira Valente
Vicente Gosciola
Victor de Souza Crispim
Vincenzo Miracula
Xabier Martínez Rolán

SUMÁRIO

Introdução a uma vida desacelerada.....	13
<i>Denis Renó</i>	

PARTE 1 - OLHARES CRÍTICOS

Aceleração predatória e agentes inteligentes como “amigos”.....	16
<i>Pollyana Ferrari Teixeira</i>	

Modulação algorítmica e novos fluxos informacionais: um estudo sobre a influência do algoritmo no ecossistema da desinformação.....	34
<i>Angela Maria Grossi</i>	
<i>João Pedro de Toledo Piza</i>	
<i>Amanda Cristina Trentin</i>	

Os efeitos da inteligência artificial generativa na era da desinformação.....	58
<i>Régis Martins</i>	

Alfabetización de la inteligencia artificial, clave para contrarrestar la desinformación. Caso España e Italia.....	79
<i>Jesús Miguel Flores-Vivar</i>	
<i>Vincenzo Miracula</i>	
Análisis del uso de la inteligencia artificial generativa para producir fake news en la elección presidencial de Donald Trump.....	107
<i>Dick Ronald Cáceres Navarro</i>	
<i>Luis Alberto Vásquez Ampudia</i>	
Percepciones de regulación de la desinformación creada con GenAI en elecciones, caso de Ecuador.....	130
<i>Abel Suing</i>	
Entre a normatividade e o conseqüencialismo: perspectivas contemporáneas da ética em inteligência artificial.....	148
<i>Guilherme Gonçalves Mello</i>	
<i>Eli Borges Junior</i>	
Satiko e o imaginário feminino em plataformas digitais: entre subjetividade e mercantilização.....	174
<i>Laryssa da Costa Gabellini</i>	
<i>Cláudia Albuquerque Thomé</i>	
Mídia sintética: entre a inovação e a manipulação.....	203
<i>Margareth Boarini</i>	

PARTE 2 - PROPOSIÇÕES

Mil docentes, mil horas: una cartografía situada de la inteligencia artificial generativa en educación.....	223
<i>Xabier Martínez Rolán</i>	

Educação na era da IA: remodelando o aprendizado e a cidadania na ecologia dos meios.....	247
<i>Sabrina Marcandali</i>	
<i>Vânia Cristina Pires Nogueira Valente</i>	
Letramento midiático crítico e formação docente: experiências com TDIC e inteligência artificial em contextos escolares.....	267
<i>Leandro Marlon Barbosa Assis</i>	
<i>Lucas Lima Coaracy</i>	
<i>Alexandre Farbiarz</i>	
Métodos digitais aplicados às imagens em 360 graus: análise de dados visuais e interativos.....	298
<i>Carolina Gois Falandes</i>	
<i>Alan César Belo Angeluci</i>	
Da teoria à prática: o uso de inteligência artificial generativa de textos na implementação da linguagem simples em documentos jurídicos.....	337
<i>Renato Sobhie Zambonato</i>	
O uso da Inteligência Artificial para melhoria da redação jornalística: desenvolvimento de uma ferramenta de revisão automatizada.....	372
<i>Rodrigo Daniel Levoti Portari</i>	
Comunicación, colaboración y ChatGPT en el proceso de diseño: integración de herramientas digitales.....	391
<i>Pablo Daniel Gutiérrez Gutiérrez</i>	

Autoria inacabada(?) Reflexões e apontamentos acerca da atração no filme “Cartas telepáticas” de Edgar Pêra.....	431
<i>Aloísio Corrêa de Araújo</i>	
<i>Vicente Gosciola</i>	
<i>Victor de Souza Crispim</i>	
 A estetização algorítmica do audiovisual: ultrarrepresentação e modos de presença em vídeos gerados por IA.....	 450
<i>Rafael de Araújo Mélo</i>	
<i>Claviano Nascimento de Sousa</i>	
<i>Diogo de Almeida Camelo</i>	
 Índice Remissivo.....	 478

Compreensões sobre a IA Generativa

INTRODUÇÃO A UMA VIDA DESACELERADA

A sociedade midiaticizada urge movimentos denominados por Zygmunt Bauman (2017), em sua obra póstuma, como retrotópicos. Para o autor, em sua obra póstuma, é necessária uma revisão do mundo contemporâneo, observando o que abandonamos com o objetivo de voltar a caminhar adiante. Ainda, segundo o autor, não se trata de uma nostalgia sem sentido, mas a oportunidade de recuperar algumas características essenciais que sustentam a vida, dentre elas o tempo. Tal ideia está conectada diretamente a algumas das propostas do português António Fidalgo (2013), que defende a necessidade de uma ética mínima para a sociedade contemporânea, especialmente no tratamento à saúde e a necessidade de desacelerar, num movimento *Slow Living* que já está presente em outros setores da sociedade contemporânea.

Diante disso, o tema que mais podemos absorver das ideias do português quando pensamos na ecologia dos meios está na ética do descanso. O descanso é parte da vida, mas parece ter se transformado em privilégio. E a falta do descanso está diretamente conectado à obra *Retrotopia*, que oferece ao leitor a última grande reflexão do polonês Zygmunt Bauman. O livro, publicado pela primeira vez em fevereiro de 2017, um mês após o falecimento do pensador, traz à tona reflexões sobre a humanidade, com olhares (topia) para o passado (retro) em busca dos pontos pelos quais ela se perdeu. A introdução da obra e o primeiro capítulo dedicam-se a debater sobre o que é nostalgia para, em seguida, apontar os caminhos aparentemente equivocados. Neste espectro de erros direcionais, Bauman nos apresenta exatamente a velocidade

pela qual a sociedade passa pela vida, apesar da crescente expectativa de vida. Uma busca pela rapidez, que não é exclusivo da fotografia, e que, de fato, precisa de uma revisão. Para o autor, “A nostalgia - como sugere Svetlana Boym, professora de literatura eslava comparada em Harvard – ‘é um sentimento de perda e de deslocamento, mas também é um romance da pessoa com sua própria fantasia’” (Bauman, 2017, p. 8). Neste cenário, encontramos um ecossistema midiático contemporâneo que resulta de transformações tecnológicas marcantes, somadas às alterações comportamentais no uso destas novas tecnologias. De fato, essa não é uma característica somente da contemporaneidade, mas da própria ecologia dos meios (Postman, 2015). Porém, parece-nos ser obrigatório o reconhecimento de que a tecnologia digital nos levou a alterações mais rápidas e expressivas (Renó & Flores, 2018). Uma sociedade acelerada, portátil, móvel, multiplataformizada, cada vez mais dependente das tecnologias e com a instantaneidade como característica básica.

No panorama desta aceleração, propomos o debate desta obra, que reúne textos que se relacionam ao olhar retrotópico sobre a ecologia dos meios e os temas do livro. A partir de uma seleção rigorosa e realizada às cegas pelas coordenações, chegou-se à composição capitular. Espero, com este conteúdo, que pensamentos sobre o ecossistema em que vivemos sejam alavancados e soluções para um cotidiano eticamente saudável sejam encontradas. Boa leitura.

Denis Renó
Diretor Geral

PARTE 1 - OLHARES CRÍTICOS

ACELERAÇÃO PREDATÓRIA E AGENTES INTELIGENTES COMO “AMIGOS”

Pollyana Ferrari Teixeira¹

Mark Zuckerberg, CEO da Meta Platforms, acredita que as amizades por Inteligência Artificial (IA) deverão superar as humanas e que as experiências com *chatbots* poderão substituir terapeutas, agências de publicidade, professores, programadores etc., fazendo com que a IA assuma papel central na experiência humana (Basilio, 2025). A banda larga, aliada aos dispositivos móveis, transformou o ‘estar conectado’ em bolhas individuais cada vez mais fechadas e homogêneas. Esse fenômeno, que começou a se disseminar a partir do ano 2000, chegou ao seu auge em 2010, com a chegada dos *smartphones*, alcançando o atual cenário de hiperconexão.

1. Livre Docente em Comunicação e Educação; Doutora em Comunicação Social. Professora da Pontifícia Universidade Católica de São Paulo (PUC-SP). pollyana.ferrari@gmail.com

Nos espaços mediados por telas, percebe-se o aumento da participação de agentes inteligentes, seja na mediação, na interação e no ‘aconselhamento’, bem como o aumento dos discursos de ódio, polarização e cultura do cancelamento. Segundo o sociólogo Harmut Rosa, um dos sintomas mais visíveis destas décadas do século XXI é a aceleração que, conforme a metáfora por ele apresentada, nos faz correr por uma escada rolante invertida. O autor nos questiona sobre quanta velocidade podemos suportar (IHU, 2023).

Estamos indo muito mais rápido do que pensavam que seria possível no século XVIII. Quando a ferrovia foi inventada, os médicos diziam que era impossível que o corpo humano tolerasse aquelas velocidades, que iria destruir nossos cérebros. E é verdade que se não estamos acostumados a caminhar em uma velocidade tão alta, ficamos tontos. Mas os seres humanos se adaptam a tudo. (IHU, 2023, par. 12)

O psicólogo norte-americano Jonathan Haidt, autor do best-seller *A geração ansiosa*, na abertura do *Ciclo Fronteiras do Pensamento 2025*, em São Paulo, asseverou que “ficamos mais estúpidos justo quando as máquinas ficam inteligentes” (Azevedo, 2025). Esse evidente paradoxo é retomado por Prioste (2025, p. 20), ao afirmar que “a Inteligência Artificial Generativa (IAG) pode, simultaneamente, democratizar o acesso à criação visual e distorcer narrativas históricas e culturais, perpetuando preconceitos e desigualdades para atender interesses escusos de controle e persuasão”.

Segundo pesquisa do Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação (Cetic.br & NIC.br, 2025), com base em dados da TIC Domicílios, entre 2009 e 2024 a proporção

de usuários de Internet no país passou de 39% para 84%, enquanto a de domicílios com acesso à rede aumentou de 24% para 83% (Nic.br, 2009, 2024a). Os resultados também evidenciam mudanças nas formas de acesso: em 2009, apenas 21% dos usuários utilizavam o celular para se conectar, percentual que chegou a 88% em 2024. Além disso, praticamente a totalidade dos usuários brasileiros passou a acessar a rede diariamente: se, em 2009, 58% se conectavam todos os dias, em 2024 esse número alcançou 96%.

Os desafios sociais e midiáticos são imensos, sobretudo em uma realidade na qual 96% da população acessa a rede diariamente, muitas vezes mergulhada em bolhas digitais e em contextos de proliferação da desinformação, nos quais algoritmos controlam grande parte do conteúdo consumido — ainda sem a devida regulação. Questões como privacidade e práticas predatórias que envolvem tecnologia e poder podem prejudicar a cognição, como apontam cientistas do Laboratório de Mídia do Instituto de Tecnologia de Massachusetts (MIT) (Kosmyna et al., 2025). Segundo o estudo, o uso frequente do ChatGPT, ferramenta de inteligência artificial da OpenAI, afeta o desempenho cerebral e desestimula o pensamento crítico. Participaram 54 norte-americanos, com idades entre 18 e 39 anos. No início do experimento, eles foram divididos em três grupos e convidados a escrever, em 20 minutos, uma redação no estilo do vestibular tradicional dos Estados Unidos.

O primeiro grupo deveria utilizar o ChatGPT; o segundo, recorrer ao Google; e o terceiro realizar a tarefa sem qualquer auxílio. Por meio de encefalogramas, os pesquisadores mediram o engajamento e carga cognitivos durante as sessões em cada grupo, chegando à conclusão

de que o grupo que utilizou o ChatGPT teve reduções significativas na conectividade cerebral.

Os magnatas da tecnologia [...] desconfiam dos especialistas e das elites, de todos que representam o velho mundo e que poderiam impedi-los de perseguir seus sonhos. Como eles, gostam de ação e estão convencidos de que podem modelar a realidade de acordo com seus desejos; a viralidade prevalece sobre a verdade, e a velocidade está a serviço do mais forte. (Empoli, 2025, p. 91)

A delimitação do corpus foi direcionada para a problematização de como o processo de aceleração, moldado de forma predatória pelas *big techs*, influencia a forma como nós, usuários, recebemos e interpretamos as informações. A revolução digital, analisada por pensadores como Pierre Lévy (1999) e Manuel Castells (1999), descreveu a sociedade hiperconectada como um ambiente com infinitas opções de informação, especialmente de entretenimento, em que múltiplas vozes podiam interagir sem passar pela mediação da grande mídia (impressa, televisiva ou radiofônica). No entanto, duas décadas depois, vemos que, em vez de promover maior interação e empatia, essa hiperconexão tem nos tornados mais solitários, egoístas e emocionalmente dessensibilizados, guiados muitas vezes pelo culto às celebridades, refletindo exatamente a lógica apontada por Empoli: a velocidade e a viralidade se sobrepõem à verdade e moldam comportamentos conforme os interesses das plataformas digitais.

O filme *Eles Vivem* (1988), dirigido por John Carpenter, oferece uma crítica contundente ao sistema capitalista por meio de uma narrativa de ficção científica (Carpenter, 1988). Na trama, seres extraterrestres

se disfarçam como humanos, ocupam posições de poder e influenciam a sociedade, incentivando consumo e obediência sem questionamento. Combinando elementos cômicos e trágicos, o filme critica os Estados Unidos sob a administração de Ronald Reagan, mas sua mensagem permanece atual. Hoje, podemos traçar um paralelo com a influência das redes sociais e das *big techs*: assim como os alienígenas do filme moldam a percepção das pessoas, plataformas digitais utilizam algoritmos e recursos transmidiáticos para direcionar comportamentos, priorizando viralidade e atenção em detrimento da verdade e da reflexão crítica.

Narrativas complexas, vídeos de múltiplas camadas, trilhas sonoras e outras distrações digitais funcionam como mecanismos de persuasão que, muitas vezes, reforçam padrões de consumo, egoísmo e isolamento social — fenômenos já apontados por Empoli (2025) e que se conectam às previsões de Lévy (1999) e Castells (1999) sobre a sociedade hiperconectada.

A equação é simples: quanto mais isolado e dependente de interações com agentes inteligentes — seja para agendar um exame, namorar, fazer terapia, pedir comida, elaborar relatórios ou almoçar com companhia — mais distante do contato humano real a pessoa se torna, pois, agentes inteligentes não substituem pessoas. Gradualmente, perde-se também a habilidade de interagir em comunidade, reforçando padrões de solidão e individualismo associados à lógica das *big techs*, em que a velocidade e a viralidade prevalecem sobre a verdade e a reflexão crítica.

Na Coreia do Sul, a popularização do “mukbang” — prática em que se assiste outra pessoa comer enquanto se alimenta diante da tela — ilustra como interações mediadas por tecnologias podem substituir

relações humanas reais; muitos “companheiros” de refeição são IAs. De maneira semelhante, vídeos de bebês *reborn*, que já alcançaram mais de 4,2 milhões de visualizações no TikTok, incentivam comportamentos de imersão em experiências simuladas e distantes do mundo social imediato. Frente a isso, o padre Chrystian Shankar, da Diocese de Divinópolis (MG), divulgou em 16 de maio de 2025 uma nota recusando-se a batizar bebês *reborn* ou realizar outros ritos católicos em bonecos realistas, sugerindo encaminhamento dos “pais” a psicólogos ou à psiquiatria (Poder 360, 2025). Esses fenômenos exemplificam como a aceleração digital e a persuasão transmidiática das plataformas podem remodelar práticas sociais e afetivas, substituindo a presença humana por interações mediadas por tecnologia.

Para o casal de influenciadores que discute judicialmente para saber quem ficará com a monetização gerada pelo “filho” *reborn*, está evidente que as *trends* geram engajamento e dinheiro no banco. No entanto, e para as brasileiras que contraem dívidas, por exemplo, para conseguir visitar a loja temática Minha Infância — Maternidade de Bebês Reborns, em Belo Horizonte (MG)? É preciso lembrar que, fora das redes, há seres humanos reais que, influenciados por um ambiente virtual altamente manipulador, podem ser prejudicados de diversas maneiras. Quantas pessoas gastam dinheiro em cursos ou mentorias com *coaches* dos mais variados ramos, impulsionadas pelo *marketing* digital e movidas pela expectativa de alcançar aquela vida ficcional moldada pelos influenciadores digitais?

Han (2017, p. 32) alerta que, “hoje, a existência está marcada pela aceleração, pelo excesso e pelo cansaço. O tempo não se desenrola, ele colapsa”. Esse cenário contribui para uma condição de vulnerabilidade

diante da desinformação; resta pouco espaço e tempo para a reflexão crítica. Nesse sentido, Ferrari (2023, p. 13) elucida que a

desinformação não acontece só no exercício diário da política. Táticas usadas para reescrever a História são tão perigosas quanto as narrativas que hoje colocam em risco as democracias ao redor do planeta. Uma delas, entre tantas — das que causam mais polarização e sofrimento — é a história que aprendemos na escola e que constrói um belo caminho de melhoria das condições do negro no século XIX no Brasil rumo à abolição. Graças aos novos estudos sobre escravidão e racismo desenvolvidos desde meados do século passado, sabemos que o discurso do apaziguamento racial pré-abolição foi uma falácia planejada pela elite e pela mídia da época.

Corporações que dominam os mercados de tecnologia da informação, conhecidas como *big techs*, são empresas fundadas a partir da década de 1970, como a Microsoft (1975) e a Apple (1976), seguidas pela Amazon (1994), Google (1998) e Meta (2004, inicialmente como Facebook). No entanto, passaram a ser denominadas *big techs* a partir da consolidação de estratégias de abertura de capital, fusões e aquisições, que se tornaram mais evidentes na década de 2010. Para Faustino e Lippold (2023), essas empresas reforçam uma lógica que produz colonialidade, desigualdade e desinformação — uma aceleração que molda novas formas de colonialidade sem que a maioria dos usuários perceba.

Segundo Abidin (2017, p. 294), “às vezes os influenciadores representam a realidade, às vezes eles moldam a realidade, às vezes eles refletem a realidade e há muitos tipos de influenciadores. É difícil dizer qual será o mais popular”. E os algoritmos das *big techs* lucram com esta realidade moldada.

Simular um parto não dói, o bebê não chora, não há perrengues em casa nem noites mal dormidas, e tudo se transforma em cenário instagramável — visualmente atraente e projetado para ser fotografado e compartilhado nas redes sociais, seja em posts de roupinhas, do berço, dos utensílios ou de vídeos de mamadas. “Mesmo que a relação com bonecas seja antiga, a prática com os bebês *reborn* inaugura um tipo de brincadeira que ultrapassa o lúdico, a ficção e transborda um hiper-realismo que responde à demanda de um tipo de performance de maternidade celebrada nas redes sociais” (Zurita, 2025, par. 7). Essa lógica de performance não se limita à maternidade digital. A prática da esposa perfeita, conhecida como *tradwife* — neologismo em inglês para “esposa tradicional” —, consiste em mostrar a rotina doméstica para seguidores, gerando engajamento e fãs.

Uma busca pela hashtag *tradwife* no Instagram retorna mais de 80 mil resultados. Ou seja, não se trata apenas de fazer o bolo de chocolate para o café, mas de mostrar o bolo em um cenário instagramável para os seguidores. O que ambas as práticas revelam é como a aceleração digital, impulsionada por algoritmos de engajamento das *big techs*, transforma experiências cotidianas em performances públicas, priorizando visibilidade e viralidade em detrimento da intimidade e da autenticidade.

O conceito de “instagramável” surgiu com a popularização do Instagram e a busca por conteúdos esteticamente agradáveis, sempre apresentando uma performance positiva criada para viralizar. Vilém Flusser (1920-1991), filósofo tcheco que se mudou para o Brasil em 1941 para escapar do nazismo e atuou aqui como professor de filosofia, jornalista e escritor, argumenta que “jogar luz na caixa preta” significa escancarar

não apenas a mediação técnica, mas também a mediação ideológica presente nos aparelhos (Flusser, 2024).

É nesse contexto que o instagramável captura o usuário, mantendo-o imerso por horas em um passado ou realidade cuidadosamente construída para seu deleite — cenário ideal para o avanço das “alucinações” da inteligência artificial. Nos *prompts* (textos de entrada que orientam um modelo de linguagem a produzir uma resposta), a IA gera como resultado textos, vídeos, relatórios em 3D, vozes, novos seres e outras formas de conteúdo, reforçando essa imersão e a experiência de hiper-realidade.

Por algum tempo, acreditou-se que, para iluminar a caixa preta, seria necessário aprender a programar. Em *O mundo codificado* (Flusser, 2024), fica evidente que iluminar a caixa preta vai além do conhecimento e da manipulação do código. Acredita-se que uma das formas de lançar luz sobre os algoritmos e seus bilionários é por meio da desaceleração, pois, com foco e atenção plena, conseguimos perceber a encruzilhada entre a materialidade temporal e a imaterialidade eternizada à qual nossa cultura parece estar se dirigindo, conforme previsto por Flusser.

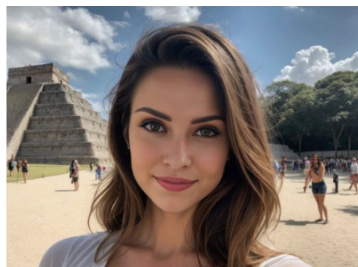
Sobre o processo de desacelerar, Ferrari (2025, p. 34) explica que

a nossa pressa não nos permite perceber sequer o gosto da comida que chega, rápida, pelo entregador do aplicativo. Há quanto tempo você não cozinha seu alimento? Descasca sua batata ou faz uma sopa que não seja pronta? Por que precisamos ser tão velozes? Estamos economizando tempo para quê? Questiona-se. O processo de desacelerar é também um ato de autocompaixão, uma abertura da sensibilidade ao sofrimento — o nosso e o dos outros —, acompanhada do desejo genuíno de aliviá-lo, assim como às suas causas. A sociedade do consumo, do empreendedorismo e do sucesso excluiu o sofrimento, tão necessário para nos lembrar que somos humanos.

A partir do momento em que se acredita nesse cenário instagrável; interage-se com agentes inteligentes, como a marroquina Kenza Layli (Figura 1) e a californiana Emily Pellegrini (Figura 2), como se fossem amigos; utiliza-se filtros que emagrecem, deixam as bochechas rosadas; pede-se à Alexa para organizar a lista de supermercado e administrar as aulas de idiomas; ou ainda se produz, via *prompt*, imagens de uma praia com o mar azul do Caribe que nunca se visitou, para compartilhar. Começa a fazer cada vez mais sentido a ideia de ter um bebê *reborn* e acreditar que ele merece ser atendido no pronto-socorro.

Figuras 1 e 2

Instagram



. A marroquina Kenza Layli; Emily Pellegrini.

Ou seja, o ato ou efeito de descolar da realidade, de arrancar ou separar algo que estava “colado”, tem conquistado cada vez mais adeptos. Trata-se de descolamentos que vão desde a crença na Terra plana, passando por supostas vacinas contendo chips comercializados pela Microsoft durante a pandemia de Covid-19, até o consumo de desinformações sobre autismo, que alcançaram 4 milhões de usuários na América Latina e no Caribe entre 2015 e 2025, totalizando 99.318.993 visualizações (Silva et al., 2022), para citar apenas alguns exemplos.

A neutralidade tecnológica é um mito, “precisamos entender o modo pelos quais o racismo se imbrica nas tecnologias digitais através de processos invisíveis nos recursos automatizados e/ou definidos pelas plataformas, tais como recomendação de conteúdo, moderação, reconhecimento facial e processamento de imagens” (Silva, 2022, p. 26)

A desumanização em curso

A perspectiva é que, ao longo dos próximos cinco anos, estaremos cada vez mais sob a influência das grandes empresas de tecnologia. Essas corporações já são transnacionais e competem com estados-nação, enquanto a IAG continua a catalisar seu poder. “Como inventamos uma sociedade de consumo, consumista, e a economia tem que crescer porque, se não crescer, é uma tragédia, inventamos uma montanha de consumo supérfluo. E tem que viver comprando e jogando fora, e o que estamos gastando é tempo de vida”, alertava o ex-presidente do Uruguai, Pepe Mujica, em trecho do documentário *Os Sonhos de Pepe* (Trobo, 2024). Uma pesquisa da Unesco (O’Hagan, 2025), publicada em janeiro de 2025, realizada com 500 influenciadores de 45 países, revelou que 62% deles não realizam uma verificação rigorosa e sistemática das informações antes de compartilhar conteúdos. No entanto, 73% expressaram interesse em receber treinamento para isso. Em resposta, a Unesco lançou o primeiro curso global para abordar essa questão, já com mais de 9.000 participantes de 160 países inscritos.

Para Audrey Azoulay, Diretora-Geral da Unesco,

os criadores de conteúdo digital adquiriram um lugar importante no ecossistema de informações, envolvendo milhões de pessoas com notícias culturais, sociais ou políticas. Mas muitos estão lutando contra a desinformação e o discurso de ódio online e

pedindo mais treinamento. Como parte de seu mandato para a alfabetização midiática e informacional, a Unesco os apoiará por meio do primeiro curso de treinamento global. (O’Hagan, 2025, par. 2)

Na Figura 3, observa-se uma reportagem do *El País* sobre desinformação, denominada ‘doutrina Banon’ (referência a Steve Banon), demonstrando a complexidade da era atual e os atores, inclusive políticos, responsáveis pela impossibilidade de banir *fake news* do nosso dia a dia.

Figura 3
Reportagem La era de la desinformación



Salas (2024).

Os bonecos *reborn* funcionam como sintomas de uma desumanização. Nas duas últimas décadas, com a presença e o alcance das redes sociais em nossas vidas, forjamos uma civilização que prefere fabricar realidades instagramáveis a cultivar um olhar afetivo pelo ser humano — aquele que abraça os refugiados; os fora dos padrões de beleza, expostos

pela indústria do consumo; os tristes, pois ser feliz e jovem o tempo todo só é possível em cenários instagramáveis. Precisamos desacelerar para conseguir nos humanizar novamente, ou nos tornaremos cada vez mais dependentes de realidades criadas artificialmente.

Ferrari (2025, p. 88) pensa em

uma sobrevivência que não dependa de telas, da inteligência artificial e do armazenamento [da vida] na nuvem. Em 2030, o padrão poderá ser os carros autômatos, mas é bom saber dirigir um carro manual, se precisar. Do mesmo modo que precisamos de músculos fortes para dar sustentação ao corpo na velhice, precisamos de autonomia mental e de decisão para não sermos manipulados por desinformações.

Vimos que normatizar ações para extrair informações sobre elas é um princípio epistemológico que caracteriza os modos de relação mediados pela dataficação. Nas palavras de Latour (2019, p. 55), “uma coisa é percorrer com coragem uma grande trilha bem-sinalizada, mas outra é decidir, no início da jornada, diante de placas de sinalização cujas indicações não se entendem muito bem, qual caminho tomar”. O usuário, leitor ou telespectador da segunda década do século XXI não sabe qual caminho seguir e se depara constantemente com placas cujas sinalizações são propositalmente confusas.

Dessa forma, acreditamos que estamos em processo de consolidação de um presentismo social onde só a aceleração e os assuntos que viram *trends* — que nos chegam pelas ferramentas de busca turbinadas por IA — interessam aos consumidores brasileiros que, segundo pesquisa da OpenAI, está em terceiro lugar no uso de agentes inteligentes no mundo. Brasileiros enviam cerca de 140 milhões de mensagens

por dia na plataforma de IA, segundo o primeiro estudo, em 2025, da norte-americana OpenAI sobre o país.

O dado faz parte do primeiro estudo dedicado ao país. “A pesquisa aponta que os brasileiros que usam o ChatGPT tendem a ser mais jovens: 27% têm entre 18 e 24 anos, e 33% entre 25 e 34 anos. Segundo a empresa, isso indica que muitos estudantes e profissionais em início de carreira estão se tornando o que se chama de ‘nativos da IA’, segundo noticiou reportagem do G1 Tecnologia (Helder, 2025).

Moldados por algoritmos, que cada vez sabem mais sobre nós mesmos, promovendo um apagamento histórico de acordo com interesses de algumas *big techs*, a geração Z (pessoas nascidas aproximadamente entre 1997 e 2012), deve ser a primeira a sofrer os danos desta redução cognitiva, pois quanto mais usam agentes inteligentes, menos capacidades pessoais e pensamento crítico desenvolvem.

Considerações finais

Neste cenário observado, que objetiva um pequeno panorama sobre a aceleração fabricada pelo avanço da IA e o crescente número de agentes inteligentes com funções de “amigos e conselheiros” percebemos que a capacidade cognitiva e retenção de informações tem caído vertiginosamente e com ela a busca por notícias e fatos. Parafraseando um trecho do livro *Nós: tecnoconsequências sobre o humano* (Ed. Fi, 2020), acreditamos que mapear a distopia vigente vai te ajudar a achar nas calçadas das cidades, na fila do banco, nos governos, exemplos de humanos como Joaquin Phoenix, que ganhou o Oscar (2020) de Melhor Ator por “Coringa” e num dos trechos de seu discurso diz: “quando usamos amor e compaixão como nossos guias principais, nós criamos

sistemas de mudança”. Desacelerar, escolher o que clicar e/ou assistir são sistemas de resistência e preservação do humano.

Milhares de anos antes que o mundo moderno revelasse a sua eletrizante e sedutora face digital e algorítmica, e muito antes que sua pegada dominadora e escravizadora pudesse ser sentida por bilhões de cérebros humanos indefesos, as sementes deste futuro já haviam sido semeadas [...] Dos ventres prenhes destas cidades, muitas hoje quase totalmente esquecidas, as primeiras Brainets, verdadeiras redes de cérebros altamente sincronizados e capazes de fazer história, emergiram, floresceram e passaram a dominar o processo civilizatório humano. (Nicolelis, 2025, p. 12)

Não caminhamos para um futuro sem aquecimento global, sem agentes inteligentes e sem o avanço vertiginoso da IA. Isso, infelizmente, está posto. Conseguir perceber a manipulação cognitiva por meio da aceleração e conteúdo que nos distraem e tiram o foco em memes, vídeos e *trends* oferecidas 24 horas, 7 dias por semana com conteúdos dispostos em telas que trazem “amigos” virtuais embalados a priori para nos entender, pode nos ajudar a prolongar nosso pensamento crítico, tão esgarçado nesta primeira metade do século XXI.

Referências

- Abidin, C. (2017). #familygoals: Family influencers, calibrated amateurism, and justifying young digital labor. *Social Media + Society*, 3(2). <https://doi.org/10.1177/2056305117707191>
- Azevedo, C. (2025). Ficamos mais estúpidos justamente quando as máquinas ficam inteligentes, diz Jonathan Haidt em SP. *Folha de S. Paulo*. <https://www1.folha.uol.com.br/cotidiano/2025/05/ficamos-mais-estupidos-justo-quando-as-maquinas-ficam-inteligentes-diz-jonathan-haidt-em-sp.shtml>

Basilio, P. (2025). Para Mark Zuckerberg, a maior parte de seus amigos no futuro será IA; saiba por quê. *Época Negócios*. *Globo*. <https://epocanegocios.globo.com/inteligencia-artificial/noticia/2025/05/para-mark-zuckerberg-a-maior-parte-de-seus-amigos-no-futuro-sera-ia-saiba-por-que.ghtml>

Carpenter, J. (Diretor). (1988). *Eles Vivem* [Filme]. Universal Pictures.

Castells, M. (1999). *A sociedade em rede*. Paz e Terra.

Cetic.br & Nic.br. (2025). Infraestrutura digital: avanços e desafios para a universalização da conectividade. Recuperado de <https://www.cetic.br/media/docs/publicacoes/6/20250512105226/ano-xvii-n-1-infraestrutura-digital-avancos-desafios-universalizacao-conectividade.pdf>

Empoli, G. (2025). *A hora dos predadores*: como autocratas e magnatas digitais estão levando o mundo à beira de um colapso orquestrado. Vestígio.

Faustino, D.; & Lippold, W. (2023). *Colonialismo digital*. Boitempo.

Ferrari, P. (2023). *Descolonizar pelo afeto*. Veríssima.

Ferrari, P. (2025). *Desacelerar*. RIA.

Flusser, V. (2024). *O mundo codificado: Por uma filosofia do design e da comunicação*. Ubu.

Haidt, J. (2024). *Geração ansiosa: como a infância hiperconectada está causando uma epidemia de transtornos mentais*. Cia das Letras.

Han, B. (2017). *Sociedade do cansaço*. Vozes.

Helder, D. (2025). Brasil está entre os 3 países que mais usam o ChatGPT, diz OpenAI. *G1*. <https://g1.globo.com/tecnologia/noticia/2025/08/12/brasil-esta-entre-os-3-paises-que-mais-usam-o-chatgpt-diz-openai.ghtml>

IHU - Instituto Humanitas Unisinos. (2023). “*Parar esse crescimento que nos levou a níveis absurdos*”: entrevista com Hartmut Rosa. IHU On-Line. Universidade do Vale do Rio Dos Sinos. <https://www.ihu.unisinos.br/categorias/623461-parar-esse-crescimento-nos-levou-a-niveis-absurdos-entrevista-com-hartmut-rosa>

Latour, B. (2019). O tecido social: Rastros digitais e métodos qualitativos. In J. J. Omena (Org.), *Métodos digitais: Teoria-prática-crítica* (pp. 37-46). Icnova.

Lévy, P. (1999) *Cibercultura*. Ed. 34.

O’Hagan, C. (2024). *2/3 of digital content creators do not check their facts before sharing, but want to learn how to do so (UNESCO survey)*. UNESCO. <https://www.unesco.org/en/articles/2/3-digital-content-creators-do-not-check-their-facts-sharing-want-learn-how-do-so-unesco-survey>

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025). *Your Brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task* [Preprint]. arXiv. <https://arxiv.org/abs/2506.08872>

Nicolelis, M. (2025). *Nada mais será como antes*. Planeta Minotauro.

Poder360. (2025). Padre se recusa a batizar bebês reborn e diz ser “caso de psiquiatra”. Recuperado de <https://www.poder360.com.br/poder-gente/padre-se-recusa-a-batizar-bebes-reborn-e-diz-ser-caso-de-psiquiatra/>

Prioste, M. (2025). *A crise da imagem*. Fi.

Rosa, H. (2019). *Aceleração: A transformação das estruturas temporais na modernidade*. UNESP.

Salas, J. (2024). De Bannon a Musk: la década que convirtió la desinformación en la nueva normalidad. *El País*. <https://elpais.com/tecnologia/2024-11-24/de-bannon-a-musk-la-decada-que-convirtio-la-desinformacion-en-la-nueva-normalidad.html>

Silva, E. C. M., Garcia, A. A. F., Almeida, G. de, & Ricard, J. (2025). *Desinformação sobre Autismo na América Latina e no Caribe: mapeando 150 falsas causas e 150 falsas curas do TEA em comunidades conspiratórias no Telegram*. Fundação Getulio Vargas; Autistas Brasil. <https://portal.fgv.br/sites/default/files/uploads/desinformacao-sobre-autismo-na-america-latina-e-no-caribe.pdf>

Silva, T. (2022). *Racismo Algorítmico: inteligência artificial e discriminação nas redes sociais*. Edições Sesc São Paulo.

Trobo, P. (Diretor). (2024). *Os sonhos de Pepe* [Documentário]. Prime Video.

Zurita, V. (2025). *De onde vêm os bebês reborn?* Outras Palavras. <https://outraspalavras.net/alemdamercadoria/de-onde-vem-os-bebes-reborns/>

MODULAÇÃO ALGORÍTMICA E NOVOS FLUXOS INFORMACIONAIS: UM ESTUDO SOBRE A INFLUÊNCIA DO ALGORITMO NO ECOSSISTEMA DA DESINFORMAÇÃO

Angela Maria Grossi¹
João Pedro de Toledo Piza²
Amanda Cristina Trentin³

A transformação digital, impulsionada pelas Tecnologias de Informação e Comunicação (TIC), remodelou profundamente as dinâmicas de interação social e o acesso à informação. O ecossistema comunicacional contemporâneo é marcado pela intermediação algorítmica, que coleta dados, classifica e distribui conteúdos a partir da promessa de

-
1. Docente do Departamento de Jornalismo e do Programa de Pós-graduação em Mídia e Tecnologia (PPGMiT) da Universidade Estadual Paulista (UNESP).
angela.grossi@unesp.br
 2. Jornalista pela Universidade Estadual Paulista (UNESP).
jp.piza@unesp.br
 3. Graduanda em Jornalismo pela Universidade Estadual Paulista (UNESP).
Bolsista de Iniciação Científica na FAPESP.
ac.trentin@unesp.br

personalização da experiência on-line. Essa nova forma de mediação desloca a lógica da comunicação do tradicional modelo de massa para fluxos organizados por plataformas digitais, cujos algoritmos privilegiam conteúdos que maximizam engajamento, em detrimento de critérios como relevância social ou factualidade (Jensen & Helles, 2017; Silveira, 2019; Souza et al., 2021; Zuboff, 2019).

Esse cenário potencializa um dos desafios mais centrais das democracias contemporâneas: a disseminação da desinformação. Embora práticas de manipulação informacional não sejam inéditas na história (Burke, 2021), a escala e a velocidade com que conteúdos desinformativos circulam nas redes digitais são fenômenos recentes, impulsionados pela arquitetura algorítmica das plataformas, que favorece a viralização de mensagens polarizadoras e sensacionalistas (Vossoughi et al., 2018; Wardle & Derakhshan, 2017). Além de fragilizar os parâmetros de verificação da verdade factual, esse processo compromete a confiança pública e corrói a coesão social e democrática (Habermas, 2023; Post, 2023).

Diante desse contexto, o presente artigo tem como objetivo geral compreender como a modulação algorítmica atua como vetor estruturante da disseminação da desinformação, reconfigurando fluxos informacionais e impondo novos riscos epistêmicos e políticos às democracias contemporâneas. Especificamente, busca-se: discutir a modulação algorítmica enquanto lógica de poder e mediação na circulação digital da informação; observar as condições estruturais que favorecem a viralização de conteúdos desinformativos nas mídias sociais; avaliar os impactos dessa dinâmica para a esfera pública e para os parâmetros de confiança pública.

A metodologia adotada é qualitativa e exploratória, baseada em revisão de literatura interdisciplinar, articulando autores da economia política da comunicação, da teoria social e dos estudos sobre desinformação; o *corpus* reúne contribuições que discutem os fluxos comunicacionais contemporâneos, a centralidade das plataformas e a influência da modulação algorítmica na formação das crenças e nas práticas políticas.

O artigo está organizado em três seções principais: a primeira apresenta a sistemática algorítmica como um novo regime de mediação informacional, destacando como as plataformas operam a partir de lógicas de engajamento e vigilância; a segunda discute a reconfiguração dos fluxos comunicacionais na era das mídias sociais, ressaltando a coexistência dos modelos *many-to-many* e *many-to-one* e as assimetrias de poder que estruturam essa dinâmica; a terceira examina a desinformação no contexto da modulação algorítmica, evidenciando seus efeitos sobre a confiança epistêmica, a coesão social e a vitalidade democrática; por fim, as considerações finais sintetizam os achados e discutem os desafios que emergem para a elaboração de políticas públicas e para as estratégias jornalísticas diante do ambiente informacional contemporâneo.

A sistemática algorítmica como meio precursor da desinformação

Por meio das plataformas digitais, a esfera pública transposta ao ambiente digital agrupa pessoas que têm o intuito de “pertencer às redes de amizade, negócios, afetos, entretenimento” (Silveira, 2021, p. 39). A distribuição de conteúdos publicados no ecossistema digital é organizada e definida pelos gestores das plataformas para obedecer a um fluxo de informação sustentado pela captação e centralização da

atenção (Castells, 2009). Esse processo é sistematizado pela modulação algorítmica, a qual controla o monitoramento dos discursos, imagens e sons para cristalizar as subjetividades dos usuários (Lazzarato, 2006) coletadas pelos dados. Assim, personaliza-se um espaço virtual pervasivo e limitado a caminhos de interação contíguos ao padrão de preferências dos utilizadores das plataformas.

A ubiquidade da mediação algorítmica, que elabora o método preditivo pelo qual as redes digitais operam, constrói a uniformização dos usuários agrupados pelas semelhanças de consumo (Seaver, 2017). Nesse contexto, a informação pessoal atua como matéria-prima que lubrifica a engrenagem algorítmica e a alimenta com os dados coletados, a fim de antecipar potencialidades similares e orientar uma regularidade de perfis, por intermédio da introdução de comportamentos arquitetada pela recomendação algorítmica. Dessa maneira, os sistemas algorítmicos são programados para detectar, filtrar e classificar mensagens e sentimentos, a fim de organizar um enquadramento mental à distância (Deleuze, 1992) para o reconhecimento dos sujeitos moduláveis.

As plataformas digitais atuam como “um novo tipo de empresa” (Srnicek, 2018, p. 49). Isto é, são organizações capitalistas privadas caracterizadas pela ferocidade no investimento em tecnologia associada à operacionalização de dados, já que o valor atribuído a elas é construído na coleta de informações dos internautas que interagem nesse meio (Carelli et al., 2020). Essa forma de interação compete ao estágio de monopólio do capital que funciona por domínios privados, os quais segregam grupos e culturas em bolhas configuradas por padronizações de pensamentos e comportamentos (Antunes, 2023).

Sérgio Amadeu da Silveira (2021, pp. 40-41) explica que o processo de modulação dos usuários possui uma sequência de eventos bastante particular:

Começa por identificar e conhecer precisamente o agente modulável. O segundo passo é a informação do seu perfil e o terceiro é construir dispositivos e processos de acompanhamento cotidiano constantes, se possível, pervasivos. O quarto é atuar sobre o agente modulável para conduzir seu comportamento ou opinião.

Essa prática de personalização da experiência cibernética do usuário apoia-se em tendências que implicam na seleção de entrega dos conteúdos. Santaella (2018) frisa que os filtros em que os utilizadores das plataformas estão inseridos definem o acesso às informações na medida em que os conduz a discussões estreitas e consonantes às próprias crenças.

Há, portanto, uma exposição informacional seletiva alimentada por algoritmos preditivos que castram o contato de indivíduos com debates discordantes. Essa movimentação “intensifica as tendências homofílicas, ou seja, aquelas de só se buscar concordâncias e fugir das discordâncias, tendências, de resto, que já fazem parte do funcionamento do psiquismo humano” (Kaufman, 2020, p. 8). Sendo assim, a personalização começa na pegada digital de cada utilizador e no manuseio do acervo de dados que é definido e programado previamente.

Nessa instância, é importante compreender que os algoritmos são a vitalidade das plataformas, visto que eles determinam a circulação dos conteúdos, ao estipular a relevância do grau de visibilidade e a área de rede em que determinados tópicos podem circular (Gillespie, 2024;

Santos, 2019). Caribé (2019) evidencia que o fenômeno da modulação algorítmica é auxiliado por mediações de naturezas algorítmicas, por pares, técnica e midiática.

A primeira refere-se a um procedimento de mediação tecnológica e não humana da informação, que absorve um volume exacerbado de dados pessoais para o encaminhamento sob medida de um conteúdo. A segunda, por pares, relaciona-se ao valor simbólico atribuído ao relacionamento entre sujeitos, ao notabilizar o convívio social. Nessa mediação, aproveita-se da troca de informação consumada entre os usuários para mapear o direcionamento dos conteúdos, ademais, essa orientação pode promover a manutenção de um estado de ignorância, por intermédio do empoderamento de sensacionalismos e desinformação, por exemplo. A terceira, técnica, preocupa-se com atividades editoriais e usabilidade da informação, em que desenvolve-se a diagramação e disposição das informações na tela. Nessa mediação, captura-se dados para elaborar métodos eficazes de retenção da atenção por sensores inteligentes que aferem o tempo do usuário na página, os cliques, o histórico de busca. Por fim, a midiática sustenta a modulação algorítmica ao efetivar a curadoria e filtragem invisível da informação. Nesse sentido, ela arquiteta um rol de informações e conteúdos embasados no acesso que o sujeito tem ou deixa de ter em ambiente e contexto determinados.

Percebe-se as mediações supracitadas na implementação de sistemas inovadores no funcionamento da maquinaria jornalística. A exemplo disso, Yang (2024) evidencia que o jornal estadunidense *New York Times*, que iniciou no mercado jornalístico no formato impresso e adaptou-se para o digital, investe em inovações tecnológicas para assegurar uma experiência relevante e atualizada ao leitor, tornar a curadoria editorial

mais eficiente e maximizar o alcance dos conteúdos divulgados. A programação de módulos algorítmicos do *New York Times* é definida para polir a atuação da redação e dos editores, por meio de métodos de pesquisa de informações do usuário, como o *bandido textual* — acesso ao histórico de leitura —, ou, o *bandido geolocalizador* — acesso a localização geográfica. Dessa forma, possibilita-se ao editor o exercício de *fixação*, no qual ele organiza o manuseio da diagramação na página inicial, ao passo em que prioriza-se uma informação em detrimento de outra. Assim, personaliza-se a experiência de navegação do leitor por meio da categorização e processamento das subjetividades coletadas.

Compreende-se, então, que a performance da máquina algorítmica afasta-se veementemente da neutralidade, em virtude da programação por um profissional humano que estabelece uma funcionalidade àquela, para a projeção do sucesso almejado (Amaral, 2025). Em convergência, O’Neil (2017) salienta que algoritmos são opiniões incorporadas em códigos, uma vez que a dinâmica de tradução e transferência do código computacional é indissociável dos posicionamentos, ideologias e idiosincrasias da origem humana que programa esse sistema.

À luz desse referencial, Santana e Neves (2022, p. 57) afirmam que “o algoritmo consegue probabilizar o conjunto de informações ou conteúdos a que o usuário potencialmente têm interesse, ao mesmo tempo que inviabiliza ou dificulta o acesso ao conjunto que, na sua concepção, é irrelevante para o indivíduo em questão”. Nesse viés, observa-se que o maquinário algorítmico projeta a prioridade do encaminhamento de conteúdos de alta demanda de interação, como acontece com os conteúdos com alto potencial de sensacionalismo.

Essa lógica pode, por conseguinte, favorecer a viralização de conteúdos falsos ou manipulados disseminados com o intuito de enganar e amplificar as narrativas simplistas, em função do alcance de uma ampla audiência de maneira rápida (Wardle & Derakhshan, 2017). Esse procedimento fortalece-se na frequente projeção da inteligência algorítmica, com o fito de maximizar o envolvimento dos utilizadores, ao passo em que escolhe-se o enquadramento dos conteúdos sensacionalistas e polarizadores que são privilegiados e geram interações rápidas (Zeynep, 2020).

Uma nova organização dos fluxos comunicacionais a partir das mídias sociais

Os fluxos comunicacionais e, conseqüentemente, as maneiras pelas quais os indivíduos consomem e trocam informações vêm se transformando ao longo da história humana. Briggs e Burke (2006) apontam, por meio de uma leitura sócio-histórica da mídia, que diferentes fatores influenciam a transformação desses fluxos ao longo do tempo. Alguns deles possuem maior relevância para o escopo da discussão presente neste texto: estruturas políticas, condições econômicas, mudanças culturais e tecnologias em questão.

Com base nesse entendimento, as formas de comunicação e o trânsito das informações em uma sociedade estão diretamente relacionados às suas estruturas de poder político, econômico e simbólico (Innis, 2011). Desenvolvendo essa percepção, McLuhan (1979) enfatiza o caráter formador das mediações, argumentando que os meios moldam tanto o conteúdo quanto a percepção das mensagens. Assim sendo, a hegemonia do modelo *broadcast* durante quase todo o século XX exemplifica

a consolidação de uma estrutura de poder específica, centralizada no papel dos meios de comunicação de massa, que se organizavam a partir da concentração da posse desses canais em torno de uma elite e conformavam uma audiência consumidora bastante singular por meio do rádio, da televisão e dos jornais impressos.

Com a ascensão das tecnologias de informação e comunicação (TIC), em especial da comunicação digital e, posteriormente, da popularização das mídias sociais na segunda década do século XXI, as transformações nos fluxos de comunicação também se delinearam. As características desse estágio englobam variados aspectos da cultura digital, os quais se manifestam na configuração e operação dos próprios fluxos e sua relação com os públicos on-line (Staab & Thiel, 2022). É importante mencionar que a manifestação dominante da comunicação digital pelas mídias sociais não exclui os meios de comunicação de massa desse cenário, sendo possível notar uma reconstrução deles pelas próprias mídias sociais a partir de um modelo híbrido (Chadwick, 2013).

Nesse contexto, as mídias digitais apresentam uma lógica de circulação da informação bastante particular. A literatura contemporânea identifica dois modelos que orientam esses novos fluxos, distintos, mas interdependentes. Primeiramente, com a emergência da Web 2.0 e, em seguida, nas primeiras redes sociais, foi possível observar a estruturação do modelo *many-to-many*. Nele, usuários e grupos podem interagir, produzir e compartilhar conteúdos em rede, sem depender de emissores ou mediadores fixos. Em contraste com a lógica do *one-to-many* típica do *broadcast*, o *many-to-many* se estrutura pela interatividade, pela descentralização e pela capacidade de articulação coletiva em torno de fluxos horizontais de informação. Nesse ponto, é importante destacar

que Bennett e Manheim (2006) revisitam o modelo clássico do *two-step flow* e identificam, no ambiente digital, o predomínio de um fluxo em um único estágio (*one-step flow*), em que mensagens circulam diretamente entre indivíduos. Posteriormente, Bennett e Segerberg (2012) aprofundaram essa discussão ao formular a noção de “ação conectiva”, ressaltando como as redes digitais on-line favorecem a organização descentralizada da comunicação coletiva.

A princípio, esse modelo foi interpretado a partir das possibilidades de democratização da produção e circulação da informação, permitindo a formação de arenas discursivas em que cidadãos podem trocar informações, articular argumentos e participar de debates com menor dependência dos meios tradicionais e de suas configurações editoriais (Maia, 2002; Marques et al., 2011; Papacharissi, 2002). Essa configuração possibilita que grupos historicamente marginalizados conquistem voz própria, exponham suas pautas e ampliem sua participação pública, ao mesmo tempo em que novos veículos digitais, engajados e especializados, reforçam esse processo ao oferecer cobertura e visibilidade às demandas desses atores, ampliando e incrementando exercícios de democratização (Bicalho, 2018; Felix, 2023; Lourenço & Carvalho, 2013; Sousa & Freitas, 2020).

Contudo, esse novo fluxo comunicacional, centrado na troca e interação direta entre usuários, sem a mediação clara de atores responsáveis por tal exercício, apresenta uma lacuna significativa. Segundo Miguel (2022) e Cesarino (2021), a estruturação desse fluxo proporciona uma desordem informacional orientada pela desintermediação e limitação do papel de sistemas peritos no ambiente digital. Tais sistemas correspondem a atores e instâncias especializados na mediação — como

o jornalismo — que selecionam, verificam e organizam informações, oferecendo aos cidadãos parâmetros de confiança e reduzindo a incerteza sobre o mundo social. Na ausência ou fragmentação dessas instâncias, a circulação digital favorece a projeção e a viralidade de desinformações e discursos de ódio, justamente porque não há mediações capazes de freá-los.

Junto ao modelo *many-to-many*, as mídias digitais também se orientam por uma comunicação baseada no *many-to-one*. Esse modelo configura-se como uma lógica de circulação da informação em que conteúdos produzidos por múltiplos usuários ou fontes convergem para um ponto central - por exemplo, uma plataforma ou algoritmo - que, com base nos dados e excedentes comportamentais resultantes da interação dos usuários com o ambiente virtual (como cliques, compartilhamentos e curtidas), seleciona, organiza e direciona informações a determinados destinatários (Jensen & Helles, 2017).

Essa abordagem tem ganhado vigor quando se investiga criticamente o papel das plataformas tanto pela economia política da comunicação (Martins & Bolaño, 2025), quanto pela influência desse modelo no comportamento (Morozov, 2018; Zuboff, 2019). Zuboff (2019) entende que essa mecanização integra o que classifica como “capitalismo de vigilância”. Nessa sua crítica da formatação contemporânea do capital, atenta-se à forma com que ele está inserido no cenário econômico das tecnologias digitais, modelando-se não apenas sobre a posse de determinado meio técnico, mas ao controle dos meios de modulação do comportamento humano. Portanto, muitas vezes cedidos de forma passiva, os dados pessoais dos usuários são a matéria-prima desse modelo.

Embora se estruture sobre um discurso que prioriza a personalização da experiência virtual dos usuários, emergem perspectivas críticas sobre como essa relação entre plataformas e usuários pode, na realidade, traduzir uma assimetria de poder. Trata-se de uma forma de dominação do comportamento, em que os dados pessoais, muitas vezes cedidos de forma passiva, funcionam como matéria-prima do modelo. Essa centralização do algoritmo e a coleta massiva de dados, elementos centrais do capitalismo de vigilância, serão abordadas em detalhe na seção seguinte, dedicada à modulação algorítmica e seus impactos na circulação da informação.

Desinformação no cenário da modulação algorítmica

Diante do entendimento acerca do conceito de modulação algorítmica como algo intimamente relacionado à infraestrutura informacional das mídias sociais, as quais possuem centralidade e protagonismo no consumo de informações (Newman et al., 2025), nota-se a possibilidade de tratá-lo em conjunto com a desinformação. Embora práticas semelhantes e o próprio conceito já tenham sido tratados (Burke, 2021; Serva, 2005), com a comunicação digital e as mídias digitais é definido como “um conteúdo intencionalmente falso e criado para causar danos” (Wardle, 2019, p. 8). Portanto, há um exercício deliberado de produzir inverdades e disponibilizá-las a uma audiência. Wardle (2019) destaca que a desinformação pode assumir diferentes formatos e empregar múltiplas técnicas de falseabilidade, variando em grau e alcance de acordo com o tipo de dano que determinados conteúdos buscam produzir na opinião pública.

Nesse sentido, é possível identificar dois segmentos para compreender a convergência entre desinformação e modulação algorítmica.

O primeiro diz respeito às condições estruturais que influenciam a circulação da desinformação. Observa-se que as mídias digitais operam em um ambiente em que a mediação crítica da informação é limitada, em função da forma como os fluxos comunicacionais se estruturam (Cesarino, 2021; Miguel, 2022). Isso levanta questões sobre quais fatores determinam o consumo e o espalhamento de conteúdos falsos, assim como sobre os motivos pelos quais indivíduos acreditam neles.

Beauvais (2022) aponta que esses fatores são interdependentes e podem ser agrupados em três dimensões. A dimensão cognitiva envolve viés de confirmação, partidarismo político e a sensação de estar informado, acompanhada da dificuldade em identificar elementos que comprovem a falsidade de determinado conteúdo. A dimensão psicológica trata da influência de apelos emocionais — como raiva, medo e decepção — que aumentam a credulidade frente aos conteúdos falsos (Martel *et al.*, 2020). Por fim, a dimensão sociológica refere-se às interações entre indivíduos no ambiente virtual, que podem amplificar o consumo de desinformação ao se articularem com fatores cognitivos e emocionais (Cinelli *et al.*, 2021; Vossoughi *et al.*, 2018).

A partir dessas dimensões é possível entender como a modulação algorítmica influencia o espalhamento da desinformação de modo eficaz. Com a personalização radical do conteúdo e da experiência on-line, ferramentas como o *microtargeting* definem como cada peça pode ser veiculada com maior precisão tendo em vista o consumo, sendo ele de informações verdadeiras ou não. Partindo do ponto que a métrica predominante é a atenção, a arquitetura algorítmica das plataformas privilegia conteúdos que geram maior engajamento, ainda que isso implique reforçar polarizações, explorar emoções negativas e priorizar informações falsas.

Esse mecanismo cria uma retroalimentação entre usuários e plataformas: de um lado, os algoritmos aprendem continuamente com os rastros digitais deixados pelos indivíduos, refinando sua capacidade de prever preferências e comportamentos (Morozov, 2018; Zuboff, 2019); de outro, os próprios usuários tendem a permanecer em ambientes informacionais que confirmam suas crenças prévias, fornecem gratificações cognitivas e sociais (Beauvais, 2022; Lawson et al., 2023).

A consequência é a formação de circuitos informacionais nos quais a distinção entre verdade e falsidade perde relevância, já que o critério de circulação deixa de ser a veracidade e passa a ser a capacidade de mobilizar atenção e engajamento dos usuários. Nessa dinâmica, configura-se um campo em que as plataformas assumem posição central na definição do que circula, do que permanece invisível e, em última instância, do que é percebido socialmente como legítimo (Souza et al., 2021).

O segundo segmento não trata a modulação algorítmica como uma lógica de curadoria informacional, mas como uma condição potencializadora dos efeitos da desinformação. Nesse sentido, o foco recai sobre os riscos epistêmicos e, sobretudo, sobre os impactos para a coesão de uma cultura política democrática. Partindo do ponto de que a personalização radical da experiência virtual é o eixo estruturante da modulação algorítmica, e considerando a centralidade das mídias digitais na vida social contemporânea, o meio digital torna-se não apenas um espaço de interação cotidiana, mas também um campo de produção, disputa e legitimação política. Assim, a forma como os algoritmos modulam fluxos informacionais não apenas influencia percepções individuais, mas também reconfigura a esfera pública, afetando a construção de

consensos, a confiança em instituições e as condições de possibilidade do próprio debate democrático (Habermas, 2023; Post, 2023).

Do ponto de vista epistêmico, a modulação algorítmica intensifica a fragmentação da realidade compartilhada. Ao personalizar os fluxos informacionais de acordo com preferências e rastros digitais, os algoritmos favorecem a consolidação de câmaras de eco (Cinelli et al., 2021) a partir do que Pariser (2012) define como “filtro bolha”, nas quais os indivíduos são reiteradamente expostos a conteúdos que confirmam suas crenças prévias. Esse processo fragiliza a capacidade de distinguir entre informações verdadeiras e falsas, relativiza a noção de veracidade como critério de validação pública e promove a ascensão de uma lógica pós-verdade, em que a plausibilidade emocional supera a factual (Bucci, 2018). Nesse cenário, a circulação da desinformação não é apenas ampliada em alcance, mas também normalizada como prática socialmente aceitável, corroendo os parâmetros de confiança epistêmica que sustentam a vida coletiva.

No plano democrático, os efeitos são igualmente profundos. A modulação algorítmica, ao priorizar conteúdos que maximizam engajamento, tende a reforçar discursos polarizadores, simplistas e de forte carga afetiva, que encontram terreno fértil para desestabilizar consensos mínimos necessários à deliberação pública (Post, 2023). A dinâmica algorítmica transforma a esfera pública em um espaço altamente fragmentado, no qual diferentes grupos sociais convivem em realidades informacionais paralelas, dificultando a construção de referenciais comuns e corroendo a confiança entre cidadãos e instituições democráticas (Habermas, 2023; Harsin, 2015).

Nesse sentido, ao mesmo tempo em que a modulação algorítmica cria as condições estruturais para a disseminação da desinformação, ela também potencializa seus efeitos epistêmicos e políticos, transformando o ecossistema digital em um espaço no qual a disputa pela atenção se sobrepõe à busca por informação qualificada, com implicações profundas para a coesão social, a confiança pública e a vitalidade democrática.

Considerações Finais

A análise realizada demonstra que a modulação algorítmica não é um mero recurso técnico que organiza o fluxo de dados, mas um mecanismo de poder que intervém na própria formação do espaço público contemporâneo. Ao priorizar conteúdos que maximizam engajamento, em vez de assegurar critérios de relevância social ou veracidade, os algoritmos operam como forças estruturantes da circulação informacional, redefinindo não apenas o que é visível ou invisível, mas também as condições de construção do debate público. Essa arquitetura converte a atenção — recurso escasso e disputado — em mercadoria e, ao fazê-lo, compromete o ideal de uma esfera pública deliberativa, substituindo o diálogo informado por interações movidas pela emoção e pelo espetáculo.

Esse cenário fragiliza dimensões centrais da democracia. A integridade da informação, entendida como a base para uma deliberação pública racional e inclusiva, é corroída por práticas que naturalizam a fragmentação do espaço comunicacional e reforçam a lógica da pós-verdade. A circulação de desinformação deixa de ser um acidente ou desvio e passa a constituir uma característica estrutural do ecossistema digital, resultado da convergência entre dinâmicas algorítmicas, modelos

de negócios das plataformas e tendências cognitivas humanas. O efeito é duplo: por um lado, a erosão dos referenciais comuns que sustentam a confiança pública e, por outro, a ampliação das assimetrias informacionais que alimentam discursos extremistas e conspiratórios, desestabilizando a coesão social e enfraquecendo instituições democráticas.

Além disso, a modulação algorítmica reconfigura o papel do jornalismo e de outros mediadores tradicionais do debate público. Ao deslocar os critérios de visibilidade e relevância para sistemas opacos controlados por plataformas privadas, reduz a autonomia editorial e enfraquece a capacidade de intervir como instância perita na verificação e contextualização dos fatos. Essa subordinação das dinâmicas informativas ao interesse econômico das plataformas coloca em xeque a própria função social do jornalismo e revela uma assimetria de poder que limita a pluralidade de vozes e a qualidade do debate público e democrático.

Diante desse quadro, torna-se urgente afirmar a integridade informacional como um problema político e não apenas técnico. Enfrentar a desinformação requer uma abordagem que articule, de forma integrada: políticas públicas que regulem a arquitetura algorítmica e a transparência das plataformas; práticas jornalísticas capazes de recuperar autoridade epistêmica diante de um ambiente saturado por narrativas enganadoras; e esforços voltados à reconstrução da confiança coletiva, fortalecendo a capacidade de cidadãos e instituições de deliberar com base em parâmetros comuns de realidade. Ignorar que a disputa contemporânea pelo poder passa também pela disputa sobre a forma como a informação circula significa abdicar da defesa de uma esfera pública plural, crítica e democrática.

Referências

- Amaral, I. (2025). Plataformas e Desinformação: Desafios do Jornalismo na Era dos Algoritmos Populistas e Racistas. *ObCiber*. <https://obciber.com/2024/11/21/plataformas-e-desinformacao-os-desafios-do-jornalismo-na-era-dos-algoritmos-populistas-e-racistas/>
- Amaral, I; Santos, S. J. (2019). Algoritmos e redes sociais: a propagação de fake news na era da pós-verdade. In J. Figueira & S. C. Santos (Eds.), *As fake news e a nova ordem (des)informativa na era da pós-verdade: manipulação, polarização, filter bubbles* (pp. 63-85). Coimbra: Imprensa da Universidade de Coimbra.
- Antunes, R. (2023). *Icebergs à Deriva: o Trabalho nas Plataformas Digitais*. Boitempo Editorial
- Beauvais, C. (2022). Fake news: Why do we believe it? *Joint Bone Spine*, 89(4), 1-6. <https://doi.org/10.1016/j.jbspin.2022.105371>
- Bennett, W. L. B., & Manheim, J. B. (2006). The One-Step Flow of Communication. *The ANNALS of the American Academy of Political and Social Science*, 608(1), 213-232.
- Bennett, W. L., & Segerberg, A. (2012). The Logic of Connective Action: Digital media and the personalization of contentious politics. *Information, Communication & Society*, 5(15), 739-768. <https://doi.org/10.1080/1369118X.2012.670661>
- Bicalho, L. (2018). O jornalismo transmídia de Mídia Ninja. Em A. Braighi, C. Lessa, M. T. Câmara (Orgs.), *Interfaces do Midiativismo: Do conceito à prática*. https://www.researchgate.net/publication/327545910_O_jornalismo_transmidia_de_Midia_Ninja

- Briggs, A., & Burke, P. (2006). *Uma história social da mídia: de Gutenberg à internet* (2o ed). Zahar.
- Bucci, E. (2018). Pós-política e corrosão da verdade. *Revista USP*, (116). <https://doi.org/10.11606/issn.2316-9036.v0i116p19-30>
- Burke, P. (2021). A ascensão dos charlatões. *Piauí*, 176. <https://piaui.folha.uol.com.br/materia/ascensao-dos-charlatoes/>
- Castells, M. (2009). *Power communication*. Ed. Oxford.
- Caribé, J. C. R. (2019). *Algoritmização das relações sociais em rede, produção de crenças e construção da realidade* [Dissertação de Mestrado, Universidade Federal do Rio de Janeiro]. <https://ridi.ibict.br/handle/123456789/1040>
- Carelli, R., Cavalcanti, T., & Silveira, V. P. (2020). *Futuro do Trabalho: os efeitos da revolução digital na sociedade*. ESMPU.
- Cesarino, L. (2021). Pós-verdade e a crise do sistema de peritos: uma explicação cibernética. *Ilha Revista de Antropologia*, 23(1). <https://doi.org/10.5007/2175-8034.2021.e75630>
- Chadwick, A. (2013). *The Hybrid Media System: Politics and Power* (1o ed). Oxford University Press.
- Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, 118(9), e2023301118. <https://doi.org/10.1073/pnas.2023301118>

Deleuze, G. (1992). *Conversações*. Editora 34.

Felix, E. H. (2023). Informação e formação: O jornalismo “preto e livre” da agência de notícias Alma Preta. *Revista Alterjor*, 28(2), Artigo 2. <https://doi.org/10.11606/issn.2176-1507.v28i2p226-239>

Gillespie, T. (2024). Generative AI and the politics of visibility. *Big Data & Society*, 11(2). <https://doi.org/10.1177/20539517241252131>

Habermas, J. (2023). *Uma nova mudança estrutural da esfera pública e a política deliberativa* (1o ed). Editora Unesp.

Harsin, J. (2015). Regimes of Posttruth, Postpolitics, and Attention Economies. *Communication, Culture and Critique*, 8(2), 327-333. <https://doi.org/10.1111/cccr.12097>

Innis, H. (2011). *O viés da comunicação* (2o ed). Vozes.

Jensen, K. B., & Helles, R. (2017). Speaking into the system: Social media and many-to-one communication. *European Journal of Communication*, 32(1), 16-25. <https://doi.org/10.1177/0267323116682805>

Kaufman, D; Santaella, L. (2020). O papel dos algoritmos de inteligência artificial nas redes sociais. *Revista FAMECOS*, 27(1). <https://revistaseletronicas.pucrs.br/revistafamecos/article/view/34074>

Lawson, M. A., Anand, S., & Kakkar, H. (2023). Tribalism and tribulations: The social costs of not sharing fake news. *Journal of Experimental Psychology: General*, 152(3), 611-631. <https://doi.org/10.1037/xge0001374>

- Lourenço, A. L., & Carvalho, J. M. de. (2013). Clivagem da democracia no plano da esfera pública. Em *Economia política da comunicação: Digitalização e sociedade* (pp. 179-198). Cultur Acadêmica.
- Maia, R. (2002). Democracia e a internet como esfera pública digital: Aproximando as condições do discurso e da deliberação. *Estratégias e culturas da comunicação*. Editora Universidade de Brasília.
- Marques, F. P. J. A., Sampaio, R. C., & Maia, R. C. M. (2011). *Deliberações ampliadas ou restritas: Perspectivas de integração*. <http://repositorio.ufc.br/handle/riufc/663>
- Martel, C., Pennycook, G., & Rand, D. G. (2020). Reliance on emotion promotes belief in fake news. *Cognitive Research: Principles and Implications*, 5(1), 47. <https://doi.org/10.1186/s41235-020-00252-3>
- Martins, H., & Bolaño, C. (2025). *Para uma crítica da economia política das plataformas digitais* (Relatório de pesquisa de pós-doutorado No. 1). Obscom.
- McLuhan, M. (1979). *Os meios de comunicação como extensões do homem* (5o ed). Cultrix.
- Miguel, L. F. (1999). O jornalismo como sistema perito. *Tempo Social*, 11, 197-208. <https://doi.org/10.1590/S0103-20701999000100011>
- Miguel, L. F. (2022). O jornalismo no novo ambiente comunicacional: Uma reavaliação da noção do “jornalismo como sistema perito”. *Tempo Social*, 34(2), Artigo 2. <https://doi.org/10.11606/0103-2070.ts.2022.195368>

- Morozov, E. (2018). *Big Tech: a ascensão dos dados e a morte da política* (1o ed). Ubu Editora.
- Newman, N., Arguedas, A. R., Robertson, C. T., Nielsen, R. K., & Fletcher, R. (2025). *Digital News Report 2025* (Digital News Report). Reuters Institute for the Study of Journalism. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2025-06/Digital_News-Report_2025.pdf
- O'Neil, C. (2017). *The era of blind faith in big data must end*. TED Talk. https://www.ted.com/talks/cathy_o_neil_the_era_of_blind_faith_in_big_data_must_end/transcript?language=en#t-9837
- Santaella, L. (2018). *A pós-verdade é verdadeira ou falsa?* Estação das Letras e Cores
- Papacharissi, Z. (2002). The virtual sphere: The internet as a public sphere. *New Media & Society*, 4(1), 9-27. <https://doi.org/10.1177/14614440222226244>
- Pariser, E. (2012). *O filtro invisível: O que a internet está escondendo de você* (1o ed). Zahar.
- Post, R. C. (2023). The Internet, Democracy, and Misinformation. Em A. Koltay, C. Garden, & Jr. Krotoszynski Ronald J. (Orgs.), *Disinformation, Misinformation, and Democracy: Legal Approaches in Comparative Context* (pp. 37-49). Cambridge University Press. <https://doi.org/10.1017/9781009373272.004>
- Santana, R. D., & Neves, B. C. (2022). Entre filtros e bolhas: a modulação algorítmica na sociedade pós-panóptica. *LOGEION: Filosofia da informação*, 8(2), 47-64. <https://doi.org/10.21728/logcion.2022v8n2.p47-64>

- Seaver, N. (2017). Algorithms as culture: some tactics for the ethnography of algorithmic systems. *Big Data & Society*, 4(2), 1-12. <https://doi.org/10.1177%2F2053951717738104>
- Serva, L. (2005). *Jornalismo e desinformação* (3o ed). Editora Senac.
- Silveira, A. (2021). A hipótese do colonialismo de dados e o neoliberalismo. In F. Cassino, J. Souza, & S. Silveira (Orgs), *Colonialismo de dados: como opera a trincheira algorítmica na guerra neoliberal*. Autonomia Literária.
- Sousa, C. B., & Freitas, V. G. (2020). Ponte Jornalismo e Alma Preta: Mídia independente, direitos humanos e igualdade racial. *Esferas*, 18, Artigo 18. <https://doi.org/10.31501/esf.v0i18.11858>
- Souza, J., Avelino, R., & Silveira, S. A. da. (2021). *A sociedade do controle: Manipulação e modulação nas redes digitais* (2o ed). Editora Hedra.
- Srnicek, N. (2018). *Capitalismo de Plataformas*. Caja Negra.
- Staab, P., & Thiel, T. (2022). Social Media and the Digital Structural Transformation of the Public Sphere. *Theory, Culture & Society*, 39(4), 129-143. <https://doi.org/10.1177/02632764221103527>
- Wardle, C. (2019). *Understanding Information Disorder* (First Draft's essential guide). First Draft. http://firstdraftnews.org/wp-content/uploads/2019/10/Information_Disorder_Digital_AW.pdf
- Wardle, C., & Derakhshan, H. (2017). *Information disorder: toward an interdisciplinary framework for research and policymaking*. Council of Europe report.

Yang, Z. (2024). *How The New York Times Incorporates Editorial Judgement in Algorithms to Curate Home Screen Content*. Medium

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Public Affairs.

OS EFEITOS DA INTELIGÊNCIA ARTIFICIAL GENERATIVA NA ERA DA DESINFORMAÇÃO

Régis Martins¹

O presente estudo tem como objetivo analisar o fenômeno da desinformação nas redes sociais digitais, com foco na influência da Inteligência Artificial Generativa (IA-G) na criação e disseminação de conteúdos falsos. A metodologia adotada consiste em uma revisão bibliográfica exploratória, baseada em fontes acadêmicas e jornalísticas que abordam a evolução histórica da desinformação, o papel da Web 2.0 na comunicação digital e os impactos recentes da IA-G. Parte-se de uma contextualização do ambiente digital contemporâneo, destacando a ascensão das redes sociais e a centralidade do conteúdo gerado por usuários. Os resultados indicam que a IA-G potencializa a produção de conteúdos altamente verossímeis, ampliando significativamente o alcance e a velocidade de propagação da desinformação. As análises

1. Advogado, Mestre em Ciência da Informação pela FFC-Unesp/Marília e Doutorando em Comunicação pela FAAC-Unesp/Bauru

também evidenciam a centralidade do conteúdo gerado por usuários nas redes sociais e os desafios impostos à verificação da informação. Como conclusão, o estudo ressalta a necessidade urgente do desenvolvimento de estratégias e ferramentas eficazes para mitigar os efeitos da desinformação, destacando a importância de políticas de educação digital, regulamentação tecnológica e práticas institucionais de comunicação responsável.

Por volta de 2004, o termo Web 2.0 surgiu para determinar uma nova era da internet. Diferentemente da Web 1.0, que era estática, a Web 2.0 permite interações entre indivíduos, colaborações com criadores de conteúdos, e participação de todos os usuários, que deixam de ser meros consumidores. Para O'Reilly (2005), a internet deixou de ter barreiras fixas e passou a ser fluida, dando espaço para redes sociais que permitem que os usuários criem perfis, conectem-se com amigos e familiares, compartilhem conteúdos entre si além de permitir a ascensão de blogs, que passaram de publicações estáticas para verdadeiras comunidades.

A partir dessa virada, o conteúdo gerado pelos próprios usuários tornou-se parte essencial da internet. Fotos, vídeos, comentários e avaliações criados pelos usuários passaram a desempenhar um papel significativo na forma como as informações são compartilhadas e consumidas online. Um exemplo disso (Gargioni, 2024), foi a loja online Amazon, que apesar de vender os mesmos produtos, com mesmas descrições e imagens que empresas concorrentes, sempre foi a mais popular. Isso porque clientes que já haviam passado pelo processo de compra do produto, que possuíam um espaço para deixar avaliações

peçoais sobre o mesmo, de acordo com o referido autor. Desde então, a produção de conteúdo online só passou a crescer.

A disseminação de dados e conhecimento é agora mais rápida e abrangente do que nunca, graças às redes sociais e às tecnologias de comunicação. No entanto, essa aceleração da informação também traz consigo um desafio significativo: a proliferação da desinformação. A desinformação, ou a disseminação deliberada de informações falsas ou enganosas, representa uma ameaça à compreensão precisa da realidade e pode ter consequências graves nas tomadas de decisão individuais coletivas.

Diante desse cenário, o presente trabalho tem como objetivo analisar a influência da Inteligência Artificial Generativa (IA-G) na propagação da desinformação nas redes sociais. Busca-se compreender de que modo a IA-G, com sua capacidade de criar conteúdos falsos altamente verossímeis, potencializa a difusão de informações manipulativas e dificultando sua identificação.

O levantamento das referências utilizadas neste estudo foi realizado por meio de uma revisão bibliográfica de caráter exploratório, buscando artigos, livros e relatórios acadêmicos relacionados ao tema central da pesquisa. Foram utilizados bancos de dados acadêmicos reconhecidos, como Scielo, Google Scholar, Web of Science, Scopus e JSTOR, com a aplicação de palavras-chave específicas relacionadas ao objeto de estudo. A seleção das fontes priorizou trabalhos publicados em periódicos revisados por pares e obras de autores consagrados na área, garantindo a relevância e a atualidade das informações. Além disso, foram incluídas referências complementares de relatórios institucionais,

legislações e documentos técnicos disponíveis em acesso aberto, de modo a ampliar a diversidade de perspectivas sobre o tema.

Dessa forma, o artigo apresenta uma contextualização teórica sobre a evolução da desinformação e o papel das redes sociais digitais nesse processo; em seguida, discute-se o funcionamento da Inteligência Artificial Generativa e sua aplicação na criação de conteúdos falsos, concluindo-se que a IA-G representa um novo patamar no problema da desinformação, exigindo ações urgentes e articuladas entre regulação, alfabetização midiática e desenvolvimento de ferramentas tecnológicas capazes de identificar e conter esses conteúdos de forma eficaz.

Desinformação

A desinformação não é um fenômeno exclusivo do século XXI; ao contrário, acompanha a história da comunicação humana.

a informação via notícias falsas ou intencionalmente deturpadas/descontextualizadas sempre existiu na história dos meios de comunicação/informação e da sociedade geral. (Silva, 2019, p. 348)

No século XV, por exemplo, a comunicação oral ganhou maior relevância na Europa, e rumores e fofocas funcionavam como meios eficazes de propagação de falsidades. Nesse contexto, Briggs e Burke (2006) observaram que os boatos eram descritos como um “serviço postal oral”, transmitindo mensagens com grande rapidez, muitas vezes motivadas por interesses políticos em tempos de conflito.

Para compreender a desinformação, é necessário primeiro revisar as diferentes interpretações do conceito de informação, que vão desde

visões técnicas centradas em dados e processos até perspectivas mais amplas ligadas à cultura, ao contexto social e ao conhecimento. Silva e Gomes (2015) reuniram essa diversidade, sintetizada no quadro abaixo:

Tabela 1

Definições de informação segundo diversos autores

Autor	Definição Resumida	Ano
Shera, J.	Informação integra conteúdo, operação tecnológica e contexto social.	1971
Wersig & Neveling	Diferentes abordagens: estrutural, cognitiva, comunicacional, processual e de efeito.	1975
Belkin & Roberts	Informação é o que altera uma estrutura.	1976
Brookes	Promove transformações cognitivas no indivíduo.	1980
Hayes	Propriedade dos dados resultante de transmissão, seleção ou análise.	1986
Saracevic & Wood	Conjunto de mensagens estruturadas com valor decisório.	1986
Buckland	Três dimensões: informação como processo, conhecimento e coisa.	1991
Le Coadic	Conhecimento registrado em suporte escrito, oral ou audiovisual.	1996
McGarry	Informação reduz a incerteza e reforça o conhecimento.	1999
González de Gómez	Informação como objeto cultural em dimensões semântica, meta-informacional e infraestrutural.	2000
Malheiro & Ribeiro	Representações mentais codificadas e socialmente contextualizadas.	2002
Hjørland	Conceito social em comunidades discursivas.	2002
Barreto	Estruturas simbólicas que geram conhecimento.	2002
Capurro	Paradigmas hermenêuticos da informação.	2003
Choo	Recurso organizacional; significado construído por pessoas.	2004
Floridi	Informação semântica: dados bem formados, significativos e verdadeiros.	2005
Frohmann	Informação materializada por instituições e tecnologia.	2008

Nota. Adaptado de Silva e Gomes (2015).

A pluralidade de concepções demonstra que a informação é um fenômeno multifacetado, técnico e social ao mesmo tempo. Dentro dessa complexidade, a desinformação deve ser entendida como uma manipulação intencional do processo informacional, cujo propósito não é informar, mas persuadir, enganar ou induzir ao erro.

Breton (1999) define a desinformação como uma forma de manipulação que distorce os fatos para alcançar objetivos políticos ou sociais, sobretudo em contextos de conflito. Pinheiro e Brito (2014) reforçam que tais manipulações se manifestam em diferentes graus de persuasão e envolvem estratégias calculadas de convencimento. Já Brisola e Bezerra (2018) lembram que suas origens estão ligadas à contrainformação militar, mas rapidamente se expandiram para o campo político, midiático e digital, Cineli et al. (2020) demonstram a viralização da desinformação nas plataformas digitais e Nonô (2024) discute a importância da regulamentação do uso da internet, destacando seu papel na prevenção da disseminação de informações falsas durante processos eleitorais.

Essas interpretações podem ser organizadas no quadro a seguir:

Tabela 2
Tipologias da Desinformação (1999-2024)

Autor	Definição Resumida	Ano
Breton	Três variações principais: (1) transformar verdadeiro em falso e vice-versa; (2) orientar fatos para distorcer a realidade; (3) mascarar parte dos fatos para ocultar consequências.	1999
Pinheiro & Brito	Manipulações que geram gradações negativas; ações de persuasão direcionadas aos receptores da informação.	2014

Brisola & Bezerra	Informações superficiais e descontextualizadas, utilizando linguagem emocional que prioriza sentimento sobre análise crítica; presentes em mídias, discursos políticos e livros.	2018
Cinelli et al.	Analisa ecossistemas digitais, mostrando a viralização da desinformação em plataformas sociais.	2020
Manuella Nonô	Pesquisa sobre a regulamentação do uso da internet para combater fake news nas eleições de 2024.	2024

O próprio autor.

O quadro evidencia que a desinformação não resulta de equívocos acidentais, mas constitui um processo estruturado, estratégico e doloso. Seu poder reside na utilização de elementos de veracidade que conferem aparência de legitimidade, aumentando sua capacidade de convencimento e dificultando a identificação do engano.

Com a ascensão da internet e, especialmente, das redes sociais digitais, a disseminação da desinformação tornou-se exponencialmente mais rápida e abrangente. Plataformas de comunicação instantânea e mídias sociais permitem que informações falsas circulem globalmente em questão de segundos, configurando o fenômeno que Wardle e Derakhshan (2017) denominam “desordem informacional”.

Diferentemente da simples ausência de informação ou de conteúdos humorísticos, como memes ou sátiras, a desinformação contemporânea caracteriza-se por sua intencionalidade estratégica: visa enganar, manipular e induzir coletividades a determinados comportamentos ou crenças. Breton (1999) descreve esse processo como um “jogo entre verdade e mentira”, no qual dados reais são combinados com falsidades para tornar a narrativa mais verídica e plausível.

Os impactos da desinformação são significativos, afetando decisões tanto individuais quanto coletivas. Durante a pandemia de

Covid-19, por exemplo, circulou amplamente a informação sobre a suposta eficácia da hidroxicloroquina contra o SARS-CoV-2. Embora o medicamento tivesse uso comprovado para outras doenças, estudos posteriores demonstraram sua ineficácia no tratamento da Covid-19, levando à retratação de artigos em periódicos renomados (RFI, 2024). Mesmo assim, as redes sociais continuaram a amplificar a informação falsa, influenciando políticas públicas, decisões médicas e a percepção social da crise sanitária.

A desinformação deve ser compreendida como um mecanismo sofisticado de manipulação informacional, capaz de mobilizar recursos emocionais e racionais para moldar a opinião pública. Seu caráter intencional e estratégico representa uma ameaça direta ao debate público qualificado, à ciência e à democracia.

Segundo a ONG Artigo 193 a disseminação de informações inconsistentes ou imprecisas por canais oficiais do governo federal contribuiu para o aumento de contágios e mortes pela Covid-19 nos países pesquisados, incluindo o Brasil. O relatório da organização evidenciou que a circulação de informações incompletas e falsas em plataformas governamentais, que deveriam garantir o acesso à informação, agravou a crise sanitária.

Além disso, campanhas oficiais incentivaram o uso de medicamentos sem comprovação científica específica para a Covid-19 — como hidroxicloroquina, ivermectina, nitazoxanida, remdesivir e azitromicina — promovendo seu consumo em massa pela população de acordo com Santos-Pinto et al. (2021). Tal fenômeno impactou diretamente a saúde pública, contribuindo para a propagação do vírus, reduzindo a adesão a medidas preventivas, como o uso de máscaras e o distanciamento social,

além de reduzir consideravelmente o estoque desses medicamentos, prejudicando o consumo de pessoas que deles necessitavam para o controle de outras doenças.

A desinformação teve um impacto real na saúde pública, contribuindo para a propagação do vírus e a falta de adesão a medidas preventivas, como o uso de máscaras e o distanciamento social. A pandemia de COVID-19 trouxe à tona desafios significativos relacionados às redes sociais e à desinformação. Não só durante esse período, mas até o momento presente, as plataformas digitais foram e são um terreno fértil para a propagação de desinformação e teorias da conspiração.

Disseminação de Desinformação nas Redes Sociais Digitais

As redes sociais digitais têm desempenhado um papel significativo na disseminação de informações, graças às plataformas e aplicativos de mensagens que possibilitam a propagação quase instantânea de conteúdo entre as pessoas. No entanto, esse ambiente dinâmico não está isento de desafios, e a desinformação emerge como uma preocupação notável.

A facilidade de publicação e compartilhamento de informações nesses ambientes virtuais também abre espaço para informações manipuladas, fora de contexto ou até mesmo falsas. A desinformação abrange um conjunto de informações cujo propósito é moldar a opinião pública de acordo com os interesses de quem as divulga. Utiliza diversos mecanismos para se passar por confiável, tornando-se uma ameaça à compreensão precisa da realidade. Essas plataformas facilitam a disseminação de informações falsas, à medida que a velocidade e a acessibilidade da comunicação digital incentivam a rápida propagação de conteúdos enganosos.

A desinformação também pode ser propagada por mecanismos de manipulação nos meios de comunicação, distorcendo e deturpando fatos. Além das redes sociais, os aplicativos de mensagens instantâneas também desempenham um papel nesse cenário, tanto por ações humanas quanto por programas automatizados, como os robôs. Esses últimos operam em reações automáticas a uma variedade de dados de entrada, que alimentam os algoritmos para conduzir seus processos.

Em suma, as redes sociais digitais, ao mesmo tempo em que permitem a comunicação instantânea e a partilha de informações valiosas, também enfrentam o desafio da desinformação. A velocidade e a facilidade de compartilhamento podem permitir que informações incorretas se espalhem rapidamente, exigindo uma abordagem crítica e consciente por parte dos usuários para discernir o que é real e o que é manipulado.

O uso de tecnologias persuasivas, que são ferramentas e estratégias projetadas para influenciar ou mudar o comportamento das pessoas, podem ser encontradas em aplicativos e plataformas online (redes sociais como Facebook, Instagram e X), campanhas publicitárias e até mesmo em dispositivos físicos, entre essas ferramentas citamos a utilização de algoritmos de recomendação (aquelas janelinhas indesejáveis nas telas de computadores e celulares que aparecem na aba do computador minutos após termos feito uma pesquisa de algum item de interesse pessoal) para personalizar o conteúdo e manter o engajamento dos usuários, muitas vezes incentivando interações que podem reforçar certos comportamentos ou opiniões.

Tais ferramentas têm sido instrumental na manipulação da informação online, sendo usadas para manipular a opinião pública, através

de campanhas de desinformação, propaganda para segmentar grupos específicos de eleitores com mensagens personalizadas, incentivando comportamentos desejáveis, como hábitos saudáveis ou engajamento com a educação, mas podendo levar a manipulação excessiva, privacidade invadida para influenciar negativamente a saúde mental dos usuários. Corroborado por Aguiar e Rodrigues (2021, p. 245) o advento das plataformas digitais teve um impacto significativo na disseminação de desinformação. Com o aumento da presença de redes sociais e outras plataformas digitais, o compartilhamento de informações tornou-se mais rápido e amplo, mas isso também trouxe desafios relacionados à qualidade e veracidade das informações que circulam. Algoritmos que alimentam essas plataformas frequentemente priorizam conteúdo que gera mais engajamento, como curtidas, compartilhamentos e comentários. Isso pode, muitas vezes, beneficiar informações sensacionalistas ou falsas, pois elas tendem a atrair mais atenção e reações emocionais. Como resultado, conteúdos que são projetados para enganar ou manipular podem ser amplamente espalhados e recebidos como verdadeiros.

Além disso, a personalização dos feeds de notícias e sugestões de conteúdo baseados em interesses e comportamentos passados pode criar bolhas de filtro ou de eco, onde os usuários são expostos principalmente a informações e a opiniões semelhantes às suas que confirmam suas crenças existentes, exacerbando a polarização e dificultando a correção de conceitos errôneos, podendo amplificar ainda mais as mensagens populistas, criando uma sensação de apoio generalizado, mesmo que não representem a opinião pública como um todo.

Inteligência Artificial Generativa

A Inteligência Artificial Generativa (IA-G) é uma área da inteligência artificial que se concentra na geração de informações novas e originais através do aprendizado de máquina em bancos de dados massivos de experiências. Esta tecnologia oferece diversas aplicações potenciais, incluindo a geração de novas imagens, textos, músicas, vídeos e outras ferramentas que geram conteúdo com base em dados e algoritmos, de acordo com Aydin e Karaarslan (2023).

Para os autores Bengio et al. (2015), essas tecnologias podem ainda podem ser utilizados na criação de conteúdo, como ferramentas que geram artigos, posts de blog ou material de marketing; no design gráfico, softwares que criam imagens e designs personalizados com base em descrições fornecidas pelos usuários; e no entretenimento, jogos e filmes que utilizam IA para criar narrativas dinâmicas ou ambientes virtuais.

A produção de novos dados é realizada por meio do treinamento de modelos em grandes conjuntos de dados, permitindo-lhes gerar informações inéditas que se assemelham aos dados utilizados para o treinamento de acordo com Russell e Norvig (2020).

Modelos generativos de processamento de linguagem natural, segundo Goertzel e Penzinger (2015), podem ser utilizados para produzir músicas inéditas ou simular vozes de artistas, recriando estilos a partir de outros cantores.

Essas tecnologias, no entanto, levantam questões sobre privacidade, manipulação e ética, pois podem ser usadas para gerar conteúdos falsos com alto grau de realismo, Marcus (2021).

Na era digital, de acordo com Bostrom (2014), uma das preocupações mais urgentes é a capacidade desses modelos de produzir informações e narrativas falsas em larga escala, sem mecanismos automáticos confiáveis para verificação de veracidade. Casos recentes demonstram a utilização de imagens manipuladas e deepfakes para enganar o público e veículos de comunicação. Exemplos incluem vídeos falsos de políticos ou celebridades, como deepfakes de Steve Burton em golpes financeiros.

Na era digital, uma das preocupações mais urgentes é a capacidade dos modelos de inteligência artificial em produzir informações e narrativas falsas a um baixo custo em proporções estrondosas. Um dos aspectos críticos é a ausência de mecanismos para verificar a veracidade das informações geradas por esses sistemas. Sem tais salvaguardas, eles podem ser facilmente manipulados por humanos através de instruções para gerar desinformação e notícias falsas. Um exemplo recente dessa manipulação inclui uma imagem do Papa em trajes inusitados que circulou na web, enganando diversos meios de comunicação que a trataram como verídica.

O realismo dessas imagens e vídeos é alarmante, pois conseguem transmitir a sensação de realidade devido ao alto nível de verossimilhança Lovato (2024) e Mustak (2023).

Esses eventos evidenciam a facilidade com que a desinformação pode se espalhar na era digital, minando a confiança do público na autenticidade das informações Pawelec (2022) e Cazzamatta (2023). Ao aproveitar as redes sociais como veículo de propagação, a IA Generativa potencializa a disseminação de desinformação de forma rápida

e segmentada, ampliando o alcance e a eficácia de conteúdos falsos Marcus (2021) e Bostrom (2014).

Ao aproveitar as redes sociais como veículo de propagação, a IA Generativa potencializa a disseminação da desinformação de maneira rápida e abrangente. O conteúdo gerado pode ser projetado para atrair a atenção do público, incentivando compartilhamentos, curtidas e comentários que amplificam sua visibilidade e alcance. Além disso, a capacidade de segmentar audiências permite que a desinformação seja direcionada para grupos específicos de usuários, aumentando sua eficácia e impacto. Essa personalização do conteúdo torna ainda mais difícil detectar e combater a disseminação de informações falsas.

Além de propagar informações falsas, a IA-G pode moldar opiniões, influenciar percepções e manipular agendas políticas ou ideológicas, resultando em polarização social e erosão da confiança em instituições, mídia e fontes confiáveis Russell e Norvig (2020). Casos de deepfakes e imagens manipuladas têm demonstrado como narrativas falsas podem levar a comportamentos prejudiciais, como fraudes financeiras ou manipulação de decisões eleitorais, de acordo com Veerasamy e Pieterse (2022).

A influência da IA-G nas redes sociais vai além da simples propagação de desinformação. Ela pode ser utilizada para moldar opiniões, influenciar percepções e até mesmo manipular agendas políticas ou ideológicas. Isso pode resultar em uma polarização da sociedade e na criação de divisões profundas com base em crenças e visões de mundo divergentes. A proliferação de informações falsas pode erodir a confiança nas instituições, na mídia e em fontes de informação confiáveis,

levando ainda a comportamentos prejudiciais, como práticas de saúde inadequadas ou ações de violência.

Diante desse cenário, é crucial desenvolver estratégias eficazes para combater a disseminação da desinformação alimentada pela inteligência artificial.

Empresas como a OpenAI, desenvolvedora de ferramentas generativas, entre elas o ChatGPT, necessitam de ferramentas de moderação de conteúdo, no entanto, apresentando limitações, especialmente em idiomas além do inglês e na detecção precisa de conteúdos curtos ou complexos.

Esses desafios destacam a complexidade e a urgência de lidar com a disseminação da desinformação online. Enquanto as empresas de tecnologia continuam a desenvolver ferramentas e políticas para combater o problema, é evidente que há limitações significativas a serem superadas. A falta de suporte para idiomas além do inglês e a dificuldade na identificação precisa do conteúdo gerado por IA são apenas alguns desses obstáculos.

À medida que avançamos na era digital, torna-se imperativo combinar tecnologias de detecção de desinformação com a alfabetização digital e a conscientização sobre os perigos da IA-G. Isso inclui não apenas o desenvolvimento de tecnologias mais sofisticadas de detecção e moderação, mas também a promoção da alfabetização digital e da conscientização sobre os perigos da desinformação. Somente abordando de forma integrada os aspectos tecnológicos, éticos e sociais será possível mitigar os riscos associados à disseminação de informações falsas geradas por IA de acordo com Bengio et al. (2015), Marcus (2021) e Lovato (2024).

Considerações Finais

O avanço da Inteligência Artificial Generativa (IA-G) representa uma revolução na capacidade de produção e disseminação de informações na era digital. Ao treinar modelos em grandes conjuntos de dados, a IA-G pode gerar conteúdos novos e originais em uma escala sem precedentes, abrindo caminho para uma série de aplicações potenciais. No entanto, esse mesmo poder de geração de informações também traz consigo desafios significativos, especialmente quando se trata da disseminação de desinformação.

A desinformação, alimentada pela IA-G, representa uma ameaça à integridade da informação online. A capacidade de criar conteúdos falsos com alta verossimilhança pode minar a confiança do público na autenticidade das informações, criando uma atmosfera de incerteza e desconfiança. Além disso, a propagação rápida e abrangente da desinformação através das redes sociais amplifica seu impacto, tornando ainda mais difícil discernir entre o que é verdadeiro e o que é falso. A resposta a esse desafio requer uma abordagem multifacetada. As empresas de tecnologia, como a OpenAI, devem continuar desenvolvendo ferramentas e políticas para combater a disseminação da desinformação, enquanto também promovem a alfabetização digital e a conscientização sobre os perigos da desinformação entre os usuários. A promoção da literacia midiática e da educação sobre como identificar fontes confiáveis é fundamental. Capacitar o público para avaliar criticamente a informação que consome pode reduzir a propagação de desinformação e melhorar a resiliência social. No entanto, é importante reconhecer que há limitações significativas a serem superadas, como a falta de suporte

para idiomas além do inglês e a dificuldade na identificação precisa do conteúdo gerado por IA.

A regulação das redes sociais é um tema controverso, mas necessária. Encontrar um equilíbrio entre controlar a desinformação e preservar a liberdade de expressão é um desafio complexo, mas essencial para proteger a integridade da comunicação pública. Enfrentar esse desafio requer uma abordagem abrangente. A criação de políticas eficazes que promovam a verdade e, ao mesmo tempo, protejam os direitos individuais é fundamental para um ambiente digital saudável.

As empresas de tecnologia, juntamente com pesquisadores e legisladores, devem trabalhar para desenvolver ferramentas e políticas eficazes para combater a disseminação da desinformação. Isso inclui a implementação de sistemas de verificação de fatos mais robustos, o desenvolvimento de algoritmos de detecção de desinformação mais sofisticados e o incentivo à alfabetização digital e ao pensamento crítico entre os usuários da internet.

Essas considerações destacam a complexidade e a urgência de lidar com a disseminação da desinformação alimentada pela inteligência artificial. À medida que continuamos a explorar as possibilidades oferecidas pela IA-G, devemos permanecer vigilantes em relação aos desafios éticos e sociais que ela apresenta, trabalhando juntos para moldar um futuro digital mais responsável e inclusivo.

Referências

Aguiar, L. A., & Rodrigues, C. M. (2021). Expertise no jornalismo: Considerações sobre a autoridade profissional no contexto da desinformação impulsionada pelos algoritmos. *Chasqui. Revista Latinoamericana de Comunicación*, 1(147).

Aydin, Ö., & Karaarslan, E. (2023). Is ChatGPT leading generative AI? What is beyond expectations? *What Is Beyond Expectations*.

Artigo 19. (2021, 7 de maio). Infodemia e a COVID-19 - A informação como instrumento contra os mitos. <https://artigo19.org/wp-content/blogs.dir/24/files/2021/05/Infodemia-e-a-COVID-19-%E2%80%93-A-informacao-como-instrumento-contra-os-mitos.pdf>

Bengio, Y., Hinton, G., & LeCun, Y. (2015, May 28). Deep learning. *Nature*, 521, 436-444. <https://doi.org/10.1038/nature14539>

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press. <https://global.oup.com/academic/product/superintelligence-9780199678112>

Breton, P. (1999). *A manipulação da palavra*. Loyola.

Briggs, A., & Burke, P. (2006). *Uma história social da mídia: de Gutenberg à Internet*. Companhia das Letras.

Brisola, A., & Bezerra, A. C. (2018). Desinformação e circulação de “fake news”: Distinções, diagnóstico e reação. *Encontro Nacional de Pesquisa e Pós-graduação em Ciência da Informação (XIX ENANCIB)*. <http://hdl.handle.net/20.500.11959/brapci/102819>

Cazzamatta, R. (2023). AI-generated misinformation: A case study on emerging trends in deepfakes. *SAGE Journals*. <https://journals.sagepub.com/doi/abs/10.1177/27523543251344971>

Cinelli, M., Quattrocioni, W., Galeazzi, A., Valensise, C. M., Brugnoti, E., Schmidt, A. L., Zola, P., Zollo, F., & Scala, A. (2020, October 6).

The COVID-19 social media infodemic. *Scientific Reports*, 10(1), 16598. <https://doi.org/10.1038/s41598-020-73510-5>

Gargioni, A. (2024, 10 de outubro). *Estratégia da Amazon para conquistar o comércio eletrônico e as indústrias de tecnologia em todo o mundo*. V4 Company. <https://v4company.com/blog/cases-de-marketing/estrategia-da-amazon>

Goertzel, B., & Penzinger, C. (2015). Artificial general intelligence. In *Springer handbook of computational intelligence* (pp. 967-986). Springer. <https://link.springer.com/book/10.1007/978-3-319-21365-1>

Lovato, J. (2024). Impacts of human biases on detection of deepfakes. *Nature*. <https://www.nature.com/articles/s44260-024-00006-y>

Marcus, G. (2021, July 1). Artificial general intelligence is not as imminent as you might think. *Scientific American*. <https://www.scientificamerican.com/article/artificial-general-intelligence-is-not-as-imminent-as-you-might-think1/>

Mustak, M. (2023). Deepfakes: Deceptions, mitigations, and opportunities. *ScienceDirect*. <https://www.sciencedirect.com/science/article/abs/pii/S0148296322008335>

Nonô, M. da S. (2024). *Guerra eleitoral, fake news e a tentativa de regulamentar o uso da internet*. Câmara dos Deputados. <https://revista.tre-rj.jus.br>

OpenAI. (s.d.). *Usage policies*. <https://openai.com/policies/usage-policies>

O'Reilly, T. (2005). What is Web 2.0. *O'Reilly Media*. <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>

- Pawelec, M. (2022). Deepfakes and democracy: How synthetic audio and video threaten political integrity. *PMC*. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9453721/>
- Pinheiro, M. M. K., & Brito, V. de P. (2014). Em busca do significado da desinformação. *DataGramaZero - Revista de Ciência da Informação*, 15.
- Quinn, L. (2025, 29 de agosto). *L.A. woman loses life savings after scammers use AI to pose as General Hospital star, says family*. People. <https://people.com/woman-loses-life-savings-after-scammers-use-ai-to-pose-as-general-hospital-star-11800052>
- RFI. (2024, 20 de dezembro). *Estudo que fundamentou uso de cloroquina para covid-19 é invalidado por revista que o publicou*. G1. <https://g1.globo.com/saude/noticia/2024/12/20/estudo-que-fundamentou-uso-de-cloroquina-para-covid-19-e-invalidado-por-revista-que-o-publicou.ghtml>
- Russell, S., & Norvig, P. (2020). *Artificial intelligence: A modern approach* (4th ed.). Pearson. <https://www.amazon.com/Artificial-Intelligence-Modern-Approach/dp/0134610997>
- Santos-Pinto, C. D. B., Miranda, E. S., & Osorio-de-Castro, C. G. S. (2021). O “kitcovid” e o Programa Farmácia Popular do Brasil. *Cadernos de Saúde Pública*, 37(2), e00348020. <https://doi.org/10.1590/0102-311X00348020>
- Silva, J. L. C. (2019). Pós-verdade e informação: Múltiplas concepções e configurações. In *Anais do 19º Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB)*. UEL/ANCIB. <http://hdl.handle.net/20.500.11959/brapci/103784>

- Silva, J. L. C., & Gomes, H. F. (2015). A contribuição da filosofia contemporânea para uma concepção teórico-conceitual de informação. <http://hdl.handle.net/20.500.11959/brapci/177933>
- Veerasamy, N., & Pieterse, H. (2022). Rising above misinformation and deepfakes. *Semantic Scholar*. <https://pdfs.semanticscholar.org/6b5d/6d5709647a3e9ee716852b68768f3aa519e3.pdf>
- Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policymaking*. Council of Europe.

ALFABETIZACIÓN DE LA INTELIGENCIA ARTIFICIAL, CLAVE PARA CONTRARRESTAR LA DESINFORMACIÓN. CASO ESPAÑA E ITALIA

Jesús Miguel Flores-Vivar¹
Vincenzo Miracula²

En la era de las tecnologías interactivas de la comunicación, los individuos viven inmersos en una constante condición de híper-visibilidad. Ésta, relacionada con el desarrollo igualmente extendido y global de las redes sociales, que han vuelto completamente difusa la tradicional distinción entre lo público y lo privado, ha modificado sustancialmente las dimensiones básicas de la vida social como la privacidad, la imagen

-
1. Doctor en Integración Tecnológica en el Sistema Multimedia.
Catedrático der la Facultad de Ciencias de la Información de la Universidad Complutense de Madrid , España.
jmflores@ucm.es
 2. PhD. in Complex Systems
Dipartimento di Scienze Politiche e Sociali, Università di Catania. Italia
vincenzo.miracula@phd.unict.it

pública de las personas, el debate público, las relaciones interpersonales, la accesibilidad y la veracidad de la información.

El fenómeno de la difusión de noticias falsas y desinformación se ha convertido en una de las características distintivas de la sociedad contemporánea hiperconectada (Dominici, 2017) e hiperexpuesta mediáticamente. Este fenómeno, aunque tiene raíces antiguas y precede a la introducción de la imprenta, ha adquirido proporciones alarmantes en la era de las redes sociales, donde las noticias, sean verdaderas o falsas, se propagan y se comparten entre las comunidades de usuarios con un simple clic. Actualmente, los medios de comunicación tradicionales han perdido el control sobre la información y la opinión pública, y el “chisme social” se propaga de manera viral en el espacio virtual. Esta situación está generando una creciente desconexión entre las noticias y sus audiencias, lo que trae como consecuencia que el “público está dando la espalda a una información periodística al considerarla dotada de poca credibilidad” (Casero-Ripollés, 2014).

Por otro lado, Internet ha introducido un nuevo modelo económico-relacional basado en la información y la economía de la atención (Davenport & Beck, 2002), cambiando radicalmente la forma de la construcción del relato informativo. Esto, a menudo, ha llevado a la confusión entre aquellos que “crean noticias” y aquellos que simplemente las difunden mediante la compartición de documentos, imágenes o simples enlaces recibidos de otros usuarios en línea. Además, cualquiera puede acceder a la red, independientemente de la competencia o información adecuada del autor de la publicación.

Desde esta perspectiva, Internet parece una especie de red entre iguales en la que todas las personas pueden convertirse en protagonistas.

En la sociedad de la hiperconexión (Flores & Dominici, 2022) que representa una plaza digital de dimensiones globales donde millones de informaciones se introducen en la red cada segundo, el individuo asume un papel dual. Además de utilizar bienes y servicios en el entorno digital, también se convierte en productor de información al expresar constantemente sus pensamientos y emociones a través de la publicación o compartición en tiempo real de imágenes, videos y música. Es lo que se conoce como la era del prosumer (Yucra-Quispe et al., 2022).

Cuando una noticia, una imagen o un video se publica en línea, se vuelve disponible de manera gratuita para cualquier usuario en la web, que no solo puede utilizarlos, sino, también, compartirlos. De hecho, el papel que pueden tener los *Deep fakes*, las manipulaciones más sofisticadas en imágenes estáticas y videos, en la alteración de las percepciones y de las narrativas con el uso de la inteligencia artificial (Gómez-de-Ágreda et al., 2021), debe generarnos una alerta para contrarrestar el declive de la información periodística como lo hemos conocido hasta hace poco. Esto conlleva el crecimiento de posibles riesgos éticos, deontológicos y legales que pueden venir ligados al uso de inteligencia artificial en las redacciones periodísticas (Gutiérrez-Caneda et al., 2023).

Todo esto, además, sucede sin barreras espaciales y temporales: todo lo que se introduce en la red permanece allí para siempre. Como resultado, la información en la web se vuelve descentralizada y ya no es el dominio exclusivo de profesionales calificados. Es la sociedad en red, o el surgimiento de la sociedad en línea (Castells, 2006) en donde la red es “democrática” y neutral (Van Dijk, 2020).

En la era de los medios digitales, el pluralismo informativo es, sin duda, un elemento positivo, ya que implica una mayor diversidad

de puntos de vista. Sin embargo, la caída del sistema de oligopolio en el mercado de la información ha llevado a la amplificación del fenómeno de las noticias falsas, que se propagan sin restricciones, haciendo a veces imposible detener su difusión, incluso, después de que la falsedad haya sido demostrada públicamente.

En este contexto, desde hace algunos años, se viene hablando de las “*fake news*”, un término que ha ingresado en el lenguaje común y también en los tribunales, para referirse a información falsa, inexacta o engañosa, medias verdades difundidas principalmente para perjudicar a terceros o sacar provecho personal. Las *fake news*, suelen asociarlas a noticias actuales que afectan a la comunidad sin que estas hayan sido contrastadas o verificadas. Se crean de manera sencilla y se difunden en la web con facilidad -especialmente, a través de las redes sociales- en una era marcada por la total libertad de expresión.

Es importante señalar que, aunque la difusión de las *fake news* a través de los medios de comunicación tradicionales siempre ha estado limitada, proporcionando más garantías de veracidad, las redes sociales han asumido un papel de suma importancia en la propagación de *fake news* en todos los sectores, incluidos aquellos que requieren habilidades y conocimientos científicos específicos. Por ejemplo, en tiempos recientes, las noticias falsas también han abordado la pandemia de Covid-19. Diversos investigadores han encontrados que los “bulos sobre el coronavirus fueron diseminados principalmente en las redes sociales y, entre ellas, sobre todo en las cerradas, como la aplicación móvil de mensajería WhatsApp” (Salaverría, et al., 2020).

En este contexto, la propagación de las *fake news* en las redes sociales, a menudo, está motivada por un deseo de lucro económico.

Además, muchos sitios web se basan en el “*clickbaiting*”, es decir, la creación de contenido web cuyo principal objetivo es atraer la mayor cantidad posible de visitantes para generar ganancias por publicidad en línea. Esto se logra utilizando titulares llamativos que incitan a los usuarios a hacer clic y compartir el contenido del sitio (Flores & Zaharíá, 2022; Flores & Zazo, 2023). El “*clickbait*” también puede ocurrir a través de plataformas de intercambio de videos, como YouTube, que publican adelantos de imágenes que no corresponden al contenido real del video, con el único propósito de aumentar las vistas y las comparticiones. Más visitas significan mayores ganancias por publicidad.

A veces, detrás de la propagación de *fake news* está el deseo de obtener una especie de protagonismo, es decir, ver qué seguimiento tienen las noticias difundidas y el efecto mediático y social que generen (Vosoughi et al., 2018). En tales casos, el autor de las *fake news* a menudo es difícil de identificar y rastrear, y generalmente está exento de cualquier responsabilidad. En otras ocasiones, su propagación está motivada por razones políticas, con el objetivo de difundir noticias no verdaderas, incluso parcialmente, para influir en la opinión pública y crear críticas difamatorias hacia el gobierno en el poder.

También hay casos en los que la publicación de una mentira es el resultado de una conducta no intencional pero negligente, como cuando no se realiza una verificación adecuada de las fuentes antes de difundir las noticias. Incluso las noticias no verificadas adecuadamente se consideran verdaderas fake news.

Por todo esto, la creciente relevancia de la desinformación se está volviendo cada vez más evidente en la sociedad, en la esfera comunicativa y en la escena política a nivel mundial, especialmente en

relación con conceptos más amplios como la posverdad y la pospolítica (Lorusso, 2018).

Según Mukerji (2018), existen propiedades que permiten etiquetar una noticia falsa: 1) propiedades intencionales; 2) propiedades sociológicas; 3) propiedades formales. Cuando se habla de propiedades intencionales, se refiere a una serie de propiedades psicológicas donde son bastante importantes las intenciones que motivan el acto de hacer una afirmación específica y la actitud con la cual se puede hacer esa declaración. Por ejemplo, muchos sostienen que para asegurarse de que una noticia sea falsa, se necesita la intención de engañar a los interlocutores (Pepp et al., 2019; Gelfert, 2018). Sin embargo, expresada en esos términos, la afirmación es claramente falsa. De hecho, a menudo se puede compartir un contenido falso en una plataforma de redes sociales como Facebook o Twitter sin la intención de engañar a nadie; por ejemplo, cuando se comparte una noticia falsa en la que se cree inocentemente.

En cuanto al análisis de las propiedades sociológicas, algunos sostienen que una noticia falsa, para ser considerada como tal, debe tener éxito, es decir, lograr una amplia circulación (Gelfert, 2018). Sin embargo, es importante señalar que no es tanto si una noticia tiene o no una gran difusión; lo que realmente importa es si tiene el potencial para difundirse de esa manera

Una noticia falsa también debe poseer propiedades formales que la hagan parecer confiable (Vosoughi et al., 2018). La forma de la noticia, independientemente del contenido, debe ser lo más verosímil posible como una noticia confiable. El estilo, el lenguaje, el diseño y otros aspectos deben ser cuidadosamente elaborados para atraer y engañar al lector.

La desinformación, entendida como la difusión deliberada de contenido manipulador a través de los medios de comunicación de masas, ha sido ampliamente estudiada en las ciencias sociales, tanto por las formas en que el poder político y económico ha activado a las masas para lograr objetivos específicos, como por el impacto de los medios en la construcción social de valores y prácticas sociales.

En los últimos años, las estrategias de desinformación han adoptado diversas formas, ya que las dinámicas tradicionales de los medios de comunicación masiva se han fusionado con las oportunidades ofrecidas por la microsegmentación a través del uso sistemático de las redes sociales. Un ejemplo emblemático es el caso de *Cambridge Analytica* (Isaak, 2018), responsable de la manipulación de los votantes en las elecciones presidenciales de Estados Unidos en 2016 y en el referéndum sobre el Brexit en el Reino Unido.

Por su parte, la Unión Europea, a través de la Comisión Europea (2022) ha puesto en marcha el Código de Buenas Prácticas en materia de desinformación que fue firmado por 34 países signatarios y presentado el 16 de junio de 2022. Este Código tiene por objeto alcanzar los objetivos de las orientaciones de la Comisión presentadas en mayo de 2021, estableciendo una gama más amplia de compromisos y medidas para combatir la desinformación en línea.

El aumento de las noticias falsas pone de relieve la erosión de antiguos baluartes institucionales contra la desinformación en la era de Internet y, desde hace una década, de la inteligencia artificial. La preocupación por el problema es global (Lazer et al., 2018). En cualquier caso, “necesitamos profesores y estudiantes que puedan aprender a identificar, comprender y, sobre todo, responder a los grandes problemas

que genera hoy el uso de Internet” (UNESCO, 2023a) y la inteligencia artificial, ya que estas tecnologías en manos de gente sin escrúpulos, generará cada vez más desinformación.

Alfabetización digital, impacto y confianza de la IA para contrarrestar la desinformación

Según Pranshu Verma (2023) articulista de The Washington Post, la inteligencia artificial está automatizando la creación de noticias falsas, provocando una explosión de contenido web que imita artículos fácticos y que en cambio difunde información falsa sobre elecciones, guerras y desastres naturales.

De acuerdo con NewsGuard, una organización que rastrea la información errónea, desde mayo de 2023 los sitios web que albergan artículos falsos creados por IA han aumentado en más de un 1.000 por ciento, pasando de 49 sitios a más de 600.

Históricamente, las operaciones de propaganda se han basado en ejércitos de trabajadores mal pagados u organizaciones de inteligencia altamente coordinadas para construir sitios que parecen legítimos. Pero la IA está facilitando a casi cualquier persona (ya sea parte de una agencia de espionaje o simplemente un adolescente en su sótano) la creación de estos medios, produciendo contenido que a veces es difícil de diferenciar de las noticias reales. Esta producción de contenido genera, por un lado, una sobresaturación de información y, por otro, un impacto en la información y comunicación, materia prima de los medios de comunicación.

El problema del monitoreo de fuentes se entrelaza con una cuestión delicada relacionada con el exceso de información al que las personas

están expuestas diariamente, creando un entorno comunicativo en el que es necesario tener habilidades esenciales para orientarse. Este problema se conoce como “sobrecarga de información” (information overload).

El concepto de sobrecarga de información fue anticipado por Alvin Toffler en su libro “Future Shock” (Toffler, 1970), donde plantea que los medios de comunicación son tecnologías que sobrecargan los sentidos de las personas, afectando su capacidad de pensar. Según Toffler, un flujo constante de información tiene un impacto en el entorno circundante y en la capacidad del individuo para elaborar expectativas sobre las consecuencias de sus acciones. Sin embargo, los seres humanos tienen un límite cognitivo en el procesamiento de esta información, y a pesar de los intentos por ampliar estos límites, hay un punto más allá del cual no es posible aumentar las capacidades cognitivas humanas. Además, la sobrecarga de información puede generar un estrés adicional en la toma de decisiones, creando una sobre estimulación cognitiva adicional.

Además, está el tema de la “ansiedad de la información” abordado por Richard Saul Wurman et al. (2001), que se refiere a la brecha entre lo que las personas entienden y lo que creen que deben entender. Wurman et al. identifican cinco escenarios en los que puede manifestarse la ansiedad por la información, incluida la incapacidad para comprender la información, sentirse abrumado por la gran cantidad de información para comprender, la incapacidad para saber si cierta información existe, la falta de conocimiento sobre dónde encontrar cierta información y la falta de acceso a la información, incluso si se sabe dónde encontrarla.

Aunque la sobrecarga de información de Toffler y la ansiedad por la información de Wurman et al. describen un entorno comunicativo principalmente analógico, dominado por la prensa y la televisión, fue

solo a fines de la década de 1990 que el concepto de “contaminación por sobrecarga de información” (Shenk, 1997) se aplicó al contexto de la sociedad de la información. Por su parte, David Shenk contribuyó al debate con su libro “Data Smog”, resaltando los efectos de la computadora e Internet en la aceleración del ciclo de producción, distribución y consumo de información.

Sin embargo, en este panorama en constante evolución, ha surgido una herramienta poderosa para abordar el desafío de la gestión de la información: la inteligencia artificial (IA). Esta tecnología ha revolucionado aún más la forma en que interactuamos con la información. A través de algoritmos avanzados, la IA puede analizar grandes cantidades de datos en tiempos increíblemente reducidos, identificando patrones y relaciones más allá de las capacidades humanas (Flores-Vivar, 2019). Esto permite obtener conocimientos valiosos y tomar decisiones informadas en contextos complejos. El uso de herramientas de IA ha agregado un nuevo nivel de eficacia en la comprensión y el uso de la información en un entorno comunicativo caracterizado por una alta cantidad de datos en flujo constante, contribuyendo a transformar desafíos aparentes en oportunidades de aprendizaje y crecimiento. Este progreso ha sido posible gracias al surgimiento de computadoras e Internet, que han acelerado el ciclo de producción, distribución y consumo de información.

Impacto de la inteligencia artificial en la Información y la Comunicación

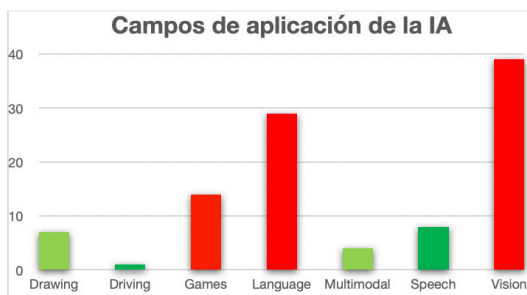
A pesar de que los primeros pasos en la aplicación de la inteligencia artificial datan de las décadas de 1950 y 1960, la atención hacia

esta disciplina ha experimentado un crecimiento exponencial solo en tiempos recientes. Inicialmente centrados en la resolución de problemas de cálculo y tareas matemáticas complejas, estos primeros desarrollos marcaron el comienzo de una evolución que en la actualidad está avanzando a un ritmo exponencial.

Aunque los primeros instrumentos de IA pueden rastrearse hasta 1958 con el “Perceptron Mark I” (en el campo de la Visión por Computadora) ideado por el psicólogo estadounidense Frank Rosenblatt, solo en los últimos años el público en general ha comenzado a percibir las notables capacidades de la IA. Sin embargo, es fundamental subrayar que las innovaciones creativas como ChatGPT, Midjourney o DALL-E representan apenas la superficie de un cambio tecnológico en ciernes, destinado a influir prácticamente en todos los aspectos de la vida humana. Específicamente, hay dos campos donde las herramientas de IA encuentran su aplicación más destacada: Lenguaje y Visión por Computadora (Fig. 1).

Figura 1

*Campos de aplicación de la IA.
Lenguaje y visión por computadora*



Nota. Basado en Eurostat.

En el ámbito de la innovación tecnológica, la inteligencia artificial (IA) se ha revelado como transformadora en diversos sectores. Con el desarrollo acelerado de la tecnología, la IA se ha convertido en una herramienta poderosa que está revolucionando el enfoque hacia los desafíos tecnológicos, las decisiones complejas y la interacción con las tecnologías digitales diarias. En este contexto, los sistemas de chat generativo basados en IA, como ChatGPT, están emergiendo como soluciones innovadoras que mejoran la interacción entre humanos y máquinas, abriendo nuevas posibilidades de comunicación avanzada y soporte inteligente en múltiples áreas.

Este rápido crecimiento ha suscitado preocupaciones particulares, no solo en el público en general sino, también, en la comunidad científica. Hacia finales de marzo, destacadas figuras del sector, incluidos Steve Wozniak, Jaan Tallinn y Elon Musk, compartieron un llamado público para examinar los posibles riesgos relacionados con la Inteligencia Artificial. En particular, se plantearon preocupaciones sobre un desarrollo no controlado de los modelos de IA, destacando la importancia de suspender temporalmente los amplios experimentos relacionados con esta tecnología para formular y aplicar protocolos de seguridad adecuados (Ramírez-Pascual, 2023).

Posteriormente, se difundió una segunda carta abierta, firmada en esta ocasión por miembros de la Asociación para el Avance de la Inteligencia Artificial (AAAI), entre los que se encuentran nombres como Francesca Rossi de IBM y Eric Horvitz de Microsoft. En este comunicado, después de enumerar las múltiples ventajas que la tecnología ya ha aportado a la sociedad, los firmantes reconocieron “los límites y preocupaciones relacionados con los desarrollos actuales, incluida la

posibilidad de que los sistemas de inteligencia artificial puedan cometer errores, proporcionar recomendaciones distorsionadas, amenazar nuestra privacidad, favorecer a actores malintencionados con nuevas herramientas y tener impactos en profesiones específicas” (Sigman & Bilinkis, 2023).

Más recientemente, Geoffrey Hinton, conocido como “el padrino de la Inteligencia Artificial”, llamó la atención al interrumpir su vínculo profesional con Google, con el fin de lanzar una señal de alarma sobre las actuales evoluciones en el campo de la Inteligencia Artificial y los riesgos asociados. En particular, son los riesgos derivados de la difusión y el uso de estas nuevas tecnologías de IA en la información y la comunicación los que causan preocupación (Hinton, 2023).

El ámbito de la información en línea parece ser muy susceptible a la rápida propagación de información no verificada o falsa, causando desinformación y confusión entre los usuarios (Flores-Vivar, 2019). No todos los que utilizan Internet tienen el interés, el tiempo o las habilidades necesarias para comparar diferentes informaciones, verificar sus fuentes, confirmar la autoría de las noticias y sus competencias, y evaluar la frescura mediante la identificación de la fecha de publicación. A menudo, quienes buscan información en línea desean encontrar noticias esenciales de manera rápida, fácilmente accesibles y sin profundizar, especialmente cuando utilizan un teléfono inteligente en lugar de una computadora para buscar noticias en tiempo real. Esto es aún más problemático considerando que los usuarios digitales pueden llegar a ser víctimas de una “infodemia” (Aleixandre-Benavent et al., 2020), involucrando la amplia difusión de información no controlada

proveniente de fuentes no verificables, propagándose en la red como un virus y causando un contagio informativo efectivo (González, 2021).

Además, a diferencia de los medios tradicionales, donde se aplica un control sobre el contenido de las noticias publicadas, especialmente en lo que respecta a la veracidad de la información y la confiabilidad de las fuentes, la web alberga una multitud de individuos que pueden expresar y compartir sus pensamientos con un público amplio, lo que hace prácticamente imposible el control de noticias y publicaciones. Aunque la red es neutral, esta neutralidad también tiene un aspecto negativo: la web es “peligrosamente” neutral en el sentido de que no filtra las noticias; una vez publicadas, permanecen en el eterno presente de la red.

Niveles de confianza en la información y alfabetización digital

Aunque los textos generados por ChatGPT (u otras IA generativas) son coherentes, no se puede deducir automáticamente que sean veraces. Este aspecto resalta una de las principales deficiencias de tales programas, que están caracterizados por inexactitudes frecuentes y afirmaciones erróneas. Estas discrepancias constituyen un problema relevante, especialmente desde el punto de vista de la información.

Por lo tanto, es fundamental adquirir un sólido conocimiento de la alfabetización digital, ya que esto permite aprender a utilizar dichas herramientas. De hecho, una comprensión profunda de su funcionamiento revela por qué a veces la información proporcionada no es precisa y, en ocasiones, incluso puede ser inventada. De ahí, la necesidad de explicar, “contextualizar y divulgar qué es (y qué no es) la IA y cómo puede

sumar en el quehacer periodístico” (Tejedor, 2023) y seguir generando confianza en la sociedad.

A partir de un análisis realizado por Eurostat, la agencia estadística de la Unión Europea, se destaca un enfoque en la investigación sobre la adopción de tecnologías de la información y la comunicación (TIC) por parte de hogares e individuos. En particular, se centró en la evaluación de la confiabilidad de la información en línea, tanto en plataformas sociales como en las páginas web de medios de comunicación y blogs. Este estudio empleó un enfoque estructurado compuesto por nueve preguntas clave (Tabla 1) para comprender la actitud y el comportamiento de las personas hacia la información en línea. La muestra representativa involucró a 169 millones de hogares con al menos un individuo entre 16 y 74 años, así como aproximadamente 328 millones de individuos en el mismo rango de edad dentro de la Unión Europea. Los datos recopilados se presentaron en forma de porcentajes, representando su incidencia en la población entrevistada para cada país miembro.

Tabla 1

Preguntas clave para comprender la actitud y el comportamiento de las personas hacia la información en línea

Pregunta	Clave
1	Personas que han visto información o contenido falso o dudoso en sitios de noticias en Internet o en redes sociales
2	Individuos que han verificado la veracidad de la información o el contenido que encontraron en sitios de noticias en Internet o en redes sociales
3	Personas que han verificado la veracidad de la información o el contenido encontrado en Internet al verificar las fuentes o buscar más información en línea

4	Personas que han verificado la veracidad de la información o el contenido encontrado en Internet al seguir o participar en discusiones en línea relacionadas con la información
5	Individuos que han verificado la veracidad de la información o el contenido encontrado en Internet al discutir la información fuera de línea con otras personas o utilizando fuentes que no son en línea
6	Personas que han verificado la veracidad de la información o el contenido encontrado en Internet
7	Personas que no han verificado la veracidad de la información o el contenido encontrado en Internet porque ya sabían que la información, el contenido o la fuente no eran confiables
8	Personas que no han verificado la veracidad de la información o el contenido encontrado en Internet porque no tenían habilidades o conocimientos
9	Personas que no han verificado la veracidad de la información o el contenido encontrado en Internet por otras razones

Nota. Basado en Eurostat.

El análisis ha estratificado aún más los datos según criterios demográficos importantes, como la edad, el género y el nivel educativo (Tabla 2). En el contexto de este estudio, se centró en la comparación entre Italia y España, en particular enfocándose en las respuestas a las preguntas 1, 6 y 8 (Figuras 2-4). Esta elección fue guiada no solo por la relevancia de estos dos países en el panorama geopolítico europeo, sino también por la reciente conclusión de las elecciones para los nuevos gobiernos en las respectivas naciones. Este análisis tiene como objetivo proporcionar una visión precisa de la actitud de las personas hacia la información en línea en una fase crucial en la que el ámbito digital y la información son centrales en el debate público.

Tabla 2

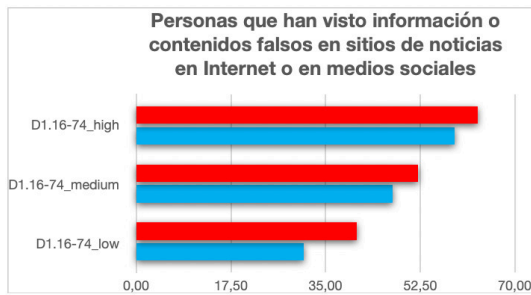
*Preguntas tabla de conversión
de los niveles de educación*

Nivel de educación	ISCED nivel	Etiqueta
De Educación preescolar a Educación primaria	0 - 2	Low
Educación secundaria superior	3 - 5	Medium
De Licenciatura a Nivel de doctorado o equivalente	6 - 8	High

Nota. Basado en Eurostat.

Figura 2

Información o contenidos falsos por niveles educativos



Nota. Basado en Eurostat.

Al analizar el gráfico correspondiente a la pregunta 1 (Fig. 2), se destacan diferencias notables entre Italia y España. Específicamente, en el contexto español se observa una clara tendencia a reconocer información errónea o ambigua en comparación con el italiano, en todos los niveles de educación. Esta diferencia plantea preguntas significativas, ya que se observa una brecha significativa en el reconocimiento de información engañosa, especialmente entre individuos con educación primaria o limitada. Las capacidades de categorización de los ciudadanos

españoles para evaluar críticamente información inexacta o inherentemente sesgada, en comparación con los italianos, son claramente evidentes. Este resultado sugiere la posibilidad de diferencias en las estrategias de alfabetización mediática o en las formas de percepción de la información dentro de los dos contextos nacionales. Investigaciones más detalladas podrían arrojar luz sobre las causas de esta discrepancia, brindando ideas valiosas para mejorar las políticas informativas y formativas en sus respectivos países.

Figura 3

Información de verificación de hechos



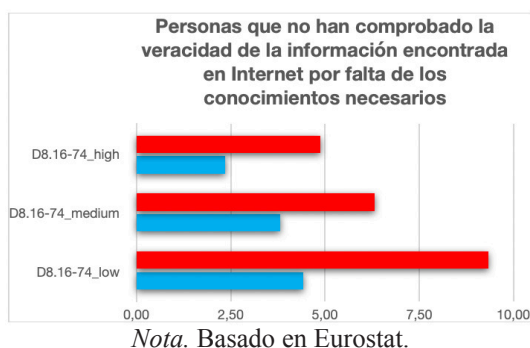
Nota. Basado en Eurostat.

El gráfico en la Figura 3 muestra la relación entre el nivel educativo y el porcentaje de individuos que realizan fact-checking (o verificación) de la información en línea. En el contexto de los dos países considerados, el nivel educativo elevado muestra convergencia, sugiriendo que las personas con una educación avanzada tienen hábitos similares de verificación. Sin embargo, surgen diferencias en el nivel medio: los ciudadanos españoles muestran una mayor propensión al fact-checking en comparación con sus homólogos italianos.

Esta disparidad podría atribuirse al acceso a servicios dedicados de fact-checking como Maldita.es en España. Además, resultados similares emergen para el nivel educativo bajo, aunque es importante destacar que la proporción de aquellos que realizan verificaciones es extremadamente limitada. Estos datos indican la importancia de la educación y los recursos disponibles en la formación de los hábitos de fact-checking de la población.

Figura 4

Sin control sobre la información debido a la falta de habilidades



El gráfico en la Figura 4 presenta de manera inequívoca el porcentaje de individuos que, debido a la falta de habilidades, no han realizado verificaciones sobre la veracidad de la información encontrada en Internet. De manera destacada, España muestra valores notablemente superiores (casi el doble) en comparación con Italia, independientemente del nivel educativo. En particular, en los estratos de la población española con educación limitada, emerge que casi el 10% de la población muestreada admite no poseer las habilidades necesarias para discernir la

autenticidad de la información. Este dato subraya un desafío significativo en términos de alfabetización mediática y plantea preguntas sobre la formación y la conciencia, requiriendo intervenciones educativas dirigidas a llenar esta brecha cognitiva y promover una evaluación crítica de las fuentes en línea.

Metodología

La metodología utilizada se basa en la revisión sistemática de la literatura como evidencia disponible, accediendo a distintas fuentes y analizando aspectos cuantitativos y cualitativos de los antecedentes y estudios publicados, principalmente, del organismo Eurostat de la Comisión Europea, así como de artículos, resultados de investigación, estudios e informes de otros organismos internacionales, como la UNESCO, Naciones Unidas, OECD, Comisión Europea o Parlamento Europeo, teniéndose en cuenta que son organizaciones que buscan afianzar el desarrollo de la educación, formación y alfabetización para contrarrestar los efectos perniciosos de la desinformación y que, desde hace años, vienen abordando el uso de la Inteligencia Artificial en este ámbito. El objetivo es resumir la información existente respecto a la concatenación de variables de descritas como la desinformación, la inteligencia artificial, la alfabetización y la educación.

Las razones que justifican la realización de esta revisión sistemática obedecen a la incertidumbre que la desinformación y la inteligencia artificial provocan una evidencia contrapuesta respecto de su real utilidad; y, en relación con esas dos variables, es interesante y necesario analizar el comportamiento de sujetos a través de la educación y programas de alfabetización. Este método nos permite, además de la

vinculación existente, identificar y evaluar las tendencias, las fortalezas y debilidades, así como los vacíos de conocimiento y las relaciones entre las publicaciones (Aria & Cuccurullo, 2017; Guo et al., 2019).

En tal sentido, la búsqueda de información se realizó, además del acceso a Eurostat, en la base de datos Scopus, utilizando una estrategia de búsqueda mediante título y resumen. Las palabras clave se obtuvieron de la literatura ya publicada sobre inteligencia artificial, desinformación, alfabetización y educación, utilizando cadenas de búsqueda con operadores booleanos (“Artificial intelligence” AND education AND disinformation).

En el estudio, se ha aplicado, además, el método PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) para llevar a cabo una revisión exhaustiva y sistemática de las variables del problema. Este enfoque ha permitido identificar, seleccionar, evaluar y sintetizar todos los estudios relevantes sobre desinformación, inteligencia artificial, alfabetización, educación. La aplicación de este método garantizó un alto nivel de rigurosidad y transparencia en la revisión de la literatura, basado en el nivel de las publicaciones y en el expertise de los investigadores que representan a muchos de los autores -y que por cuestiones de espacio no se han podido incluir a todos- lo que contribuye a la validez y confiabilidad de los hallazgos presentados.

El objetivo se centra en explorar la verificación de información con relación al nivel de educación lo que permitiría contrarrestar el impacto de la desinformación generada cada vez más por sistemas de inteligencia artificial. Este es un aspecto importante para comprender las dinámicas sociales y culturales de diferentes países. En el contexto específico de Italia y España, ambas naciones comparten raíces históricas

y culturales profundas, pero podrían presentar variaciones significativas en relación con los niveles de educación.

Conclusiones

Los resultados del estudio de caso de dos países comunitarios nos llevan a extrapolar conclusiones que afectan a la sociedad global en general y a la Unión Europea en particular. De ahí que la complejidad de la sociedad contemporánea es uno de los elementos clave a considerar puesto que la abundancia de información, la fragmentación y la subjetividad de las fuentes, junto con la posibilidad de seleccionar exclusivamente lo que se alinea con las creencias preexistentes, generan un ambiente propicio para perder la orientación con respecto a la verdad objetiva. A menudo, las personas se encuentran inmersas en las llamadas “burbujas informativas”, donde lo que escuchan y leen confirma lo que desean escuchar, sin una evaluación crítica rigurosa de los hechos. Este fenómeno contribuye a socavar la confianza en el periodismo de calidad y alimenta un enfoque selectivo de la información.

Internet ha introducido un nuevo modelo económico-relacional basado en la información, pero también ha llevado al desafío de la precisión de la información, con la propagación de noticias falsas y la pérdida de control por parte de los medios de comunicación tradicionales. La desinformación en línea a menudo está motivada por deseos de lucro, protagonismo o razones políticas, lo que lleva a una amplia variedad de consecuencias sociales y culturales. Además, la era digital ha acelerado la difusión de la posverdad. Las redes sociales y las plataformas en línea han facilitado la compartición de información distorsionada, que puede propagarse viralmente y alcanzar a un público

amplio. La manipulación emocional y la selección de información desempeñan un papel significativo en la perpetuación de la posverdad.

Es fundamental reconocer que la publicación de noticias completa o parcialmente falsas ya no tiene como objetivo informar, volviéndose no solo inútil y perjudicial, sino, también, minando los derechos fundamentales de información y crítica establecidos para proteger la libertad de prensa. Las noticias falsas contaminan el tejido informativo, privando al público de información veraz. Esto sucede incluso cuando la falsedad de las noticias se debe a una verificación insuficiente o superficial de las fuentes. También es importante destacar que las noticias falsas no siempre son inofensivas; pueden tener consecuencias graves y tangibles en la sociedad y en el individuo. La información falsa puede influir en la opinión pública, alterar percepciones, desencadenar reacciones emocionales y, en algunos casos, llevar a decisiones incorrectas o perjudiciales. La difusión descontrolada de noticias engañosas puede socavar la confianza en las instituciones, alimentar la polarización e, incluso, poner en riesgo la salud pública, como se demostró en la incidencia de bulos relacionados con la salud y la seguridad, con referencias de informaciones falsas sobre la pandemia de Covid-19. Además, las noticias falsas pueden ser instrumentalizadas por actores maliciosos para manipular la opinión pública, influir en las elecciones o causar inestabilidad social. Como resultado, el surgimiento y la difusión de las noticias falsas representan un desafío complejo que requiere un enfoque multidisciplinario para mitigar sus impactos negativos en la sociedad y promover una cultura de información crítica.

Es necesario combatir la amplia difusión de noticias falsas e información engañosa, promoviendo una cultura de información basada

en la precisión, la responsabilidad y la conciencia. En un mundo cada vez más conectado e interconectado, este desafío se vuelve crucial para preservar la calidad de la comunicación. Sin embargo, es importante reconocer que la lucha contra las noticias falsas no es solo responsabilidad de los individuos que consumen contenido, sino, también, de las plataformas digitales, los medios de comunicación y la sociedad en su conjunto.

Los usuarios deben aprender a evaluar críticamente las fuentes, cruzar información y comprender la perspectiva desde la cual proviene. Los medios de comunicación deben esforzarse por profundizar en las noticias, verificar los hechos con precisión y seguir los principios éticos del periodismo. Las plataformas digitales deben esforzarse por limitar la difusión de contenido falso sin comprometer la libertad de expresión. Solo a través de un compromiso colectivo con la verdad y la integridad de la información, será posible abordar eficazmente la Era de la desinformación

Referencias

- Aleixandre-Benavent, R.; Castelló-Cogollos, I., & Valderrama-Zurián, J. C. (2020). Información y comunicación durante los primeros meses de covid-19. Infodemia, desinformación y papel de los profesionales de la información. *Profesional de la información*, 29(4).
- Aria, M., Cuccurullo, C. (2017) Bibliometrix: una herramienta r para el análisis integral de mapas científicos. *Revista de informetricia*, 11, 959-975. <https://doi.org/10.1016/j.joi.2017.08.007>
- Casero-Ripollés, A. (2014). “La pérdida de valor de la información periodística: causas y consecuencias”. *Anuario Thinkepi*, 8, 256-259.

Castells, M. (2006). *La sociedad red una visión global*. Alianza editorial.

Comisión Europea. (2022, 16 de junio). Contra la desinformación. <https://digital-strategy.ec.europa.eu/es/policies/code-practice-disinformation>

Eurostat. (2022, 7 de febrero). Rules of engagement for handling disinformation — 2022 edition. <https://ec.europa.eu/eurostat/en/web/products-manuals-and-guidelines/-/ks-gq-22-003>

Davenport, T. H., & Beck, J. C. (2002). *The attention economy: Understanding the new currency of business*. HBS Press.

Dominici, P. (2017). The hypertechnological civilization and the urgency of a systemic approach to complexity. a new humanism for the hypercomplex society. En A. B. Floriani NEto, E. Caceres Nieto, G. Chiodi, & S. Petroccia (Eds.), *Governing turbulence. risk and opportunities in the complexity age*. Cambridge Scholars Publishing. Cambridge.

Flores-Vivar, J. M. & Zazo-Correa, L. (2023). Análisis de impacto del uso del clickbait en los perfiles de twitter de los medios de comunicación españoles. *Revista de Comunicacion*.

Flores, J., & Dominici, P. (2022). La gran equivocación: replantear la educación y la formación virtual para la formación virtual para la sociedad hipercompleja del mundo global. *Comunicación y Hombre*, (18), 23-38. <https://doi.org/10.32466/eufv-cyh.2022.18.701.23-38>

Flores-Vivar, J. M., & Zaharías, A. M. (2022). “La praxis del clickbait y de the trust project: riesgos y retos en los diarios digitales españoles”.

Anàlisi: Quaderns de comunicació i cultura, (núm. extraordinario,) 5-23. <https://doi.org/10.5565/rev/analisi.3463>

Flores Vivar, J. M. (2019). “Inteligencia artificial y periodismo: diluyendo el impacto de la desinformación y las noticias falsas a través de los bots”. *Doxa Comunicación. Revista interdisciplinar estudios de Comunicación y Ciencias sociales*, (29), 197-212. <https://doi.org/10.31921/doxacom.n29a10>

Gelfert, A. (2018). Fake news: A definition. *Informal logic*, 38(1), 84-117.

Gómez-de-Ágreda, A., Feijóo, C., & Salazar-García, I.A. (2021). Una nueva taxonomía del uso de la imagen en la conformación interesada del relato digital. Deep fakes e inteligencia artificial. *Profesional de la información*, 30(2). <https://doi.org/10.3145/epi.2021.mar.16>

González, M. A. (2021). Desinformación y coronavirus: el origen de las fake news en tiempos de pandemia. *Revista de ciencias de la comunicación e información*, 1-25.

Gutiérrez-Caneda, B., Vázquez-Herrero, J., & López-García, X. (2023). “AI application in journalism: chatgpt and the uses and risks of an emergent technology”. *Profesional de la información*, 32(5), e320514. <https://doi.org/10.3145/epi.2023.sep.14>

Hinton, G. (2023, 9 de mayo). Predicciones del padrino de la inteligencia artificial. *abc*. <https://www.abc.es/xlsemanal/ciencia/geoffrey-hinton-inteligencia-artificial.html>

Isaak, J., & Hanna, M. J. (2018). User data privacy: facebook, Cambridge analytica, and privacy protection. *Computer*, 51(8), 56-59.

- Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., & Zittrain, J. L. (2018). *The science of fake news*. *Science*, 359(6380), 1094-1096.
- Lorusso, A. M. (2018). *Postverità*. Roma-Bari.
- Mukerji, N. (2018). What is fake news? *Ergo: An Open Access Journal of Philosophy* 5, 923-946.
- Pepp, J., Michaelson, E., & Eterken, R. K. (2019). What's new about fake news. *J. Ethics & Soc. Phil.*, 16, 67.
- Ramírez-Pascual, B. (2023). *La ciberseguridad en la era de la inteligencia artificial. Dilemas y retos empresariales*. Editorial La Ley.
- Salaverría, R., Buslón, N., López-Pan, F., León, B., López-Goñi, I., & Erviti, M.-C. (2020). Desinformación en tiempos de pandemia: Tipología de los bulos sobre la covid-19. *Profesional de la información*, 29(3). <https://doi.org/10.3145/epi.2020.may.15>
- Sigman, M., & Bilinkis, S. (2023). *Artificial. La nueva inteligencia y el contorno de lo humano*. Debate.
- Shenk, D. (1997). *Data smog*. Harpercollins Publishers.
- Tejedor, S. (2023). *La inteligencia artificial en el periodismo: Mapping de conceptos, casos y recomendaciones*. UOC editorial.
- Toffler, A. (1970). *Future Shock*. Pan.

UNESCO. (2023). Inteligencia Artificial: Ejemplos de dilemas éticos.

UNESCO. (2023a). La inteligencia artificial ¿necesitamos una nueva educación? <https://unesdoc.unesco.org/ark:/48223/pf0000386262>

Van Dijk, J. A. (2020). The network society. *The network society*, 1-384.

Verma, P. (2023, 17 de diciembre). The rise of ai fake news is creating a ‘misinformation superspreader’. the washington post. <https://www.washingtonpost.com/technology/2023/12/17/ai-fake-news-misinformation/>

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146-1151.

Wurman, R. S., Leifer, L., & Sume, D. (2001). *Information anxiety 2*. Que Pub.

Yucra-Quispe, L. M., Espinoza-Montoya, C., Núñez-Pacheco, R., & Aguaded, I. (2022). De consumidores a prosumidores: La narrativa transmedia en dos juegos móviles para adolescentes y jóvenes. *Revista de Comunicación*, 21(1), 433-450. <https://doi.org/10.26441/rc21.1-2022-a22>

ANÁLISIS DEL USO DE LA INTELIGENCIA ARTIFICIAL GENERATIVA PARA PRODUCIR FAKE NEWS EN LA ELECCIÓN PRESIDENCIAL DE DONALD TRUMP

Dick Ronald Cáceres Navarro¹
Luis Alberto Vásquez Ampudia²

La irrupción de la Inteligencia Artificial (IA) ha intensificado el debate sobre sus capacidades en comparación con las humanas, un fenómeno que se amplía con la sofisticación de sus algoritmos, generando preocupación sobre la posible automatización de actividades humanas.

Específicamente, en el ámbito de la comunicación política, la IA ha adquirido una notable injerencia, manifestada principalmente en procesos electorales. El uso estratégico de este recurso por parte

-
1. Maestro en Periodismo
Docente en la Universidad Nacional Mayor de San Marcos
dcaceresn@unmsm.edu.pe
 2. Maestro en Periodismo
siul2808@gmail.com

de los candidatos se ha orientado a la desacreditación de rivales y a la generación de una ola de noticias falsas o fake news, lo que se traduce en desinformación y es motivo de alarma para partidos políticos, instituciones, medios y ciudadanos, especialmente en tiempos de comicios.

El presente trabajo se alinea con la investigación sobre las consecuencias de la inteligencia artificial en la comunicación política, tomando como objeto de estudio la elección presidencial de Donald Trump, una campaña caracterizada por el uso desmedido de la IA por su equipo para desacreditar a sus rivales y expandir noticias falsas que terminaron empañando el proceso electoral. El análisis se centra en la aplicación de la IA generativa para producir fake news, cuestionando la transparencia y seriedad de los procesos electorales ante el uso de recursos que atentan contra la ética y la responsabilidad social.

Este documento aborda esta problemática a través de un análisis estructurado en cinco ítems:

1. La Inteligencia Artificial: Se enfoca en la definición de la IA, presentándola como la elaboración de sistemas para la resolución de problemas mediante la simulación de procesos intelectuales. También se aborda su constante evolución tecnológica y las crecientes alarmas que genera en la humanidad frente a la posible robotización de las actividades humanas.
2. La Inteligencia Artificial en la política: Se explora la injerencia de la IA en la política actual, haciendo hincapié en su uso en las campañas electorales para generar desinformación y noticias falsas. Se analiza la campaña de Donald Trump como caso de estudio y se presenta el deepfake como una de las técnicas más peligrosas de la IA, describiendo su funcionamiento y la preocupación que genera en cuanto a suplantación de identidad y manipulación en la política.

3. Redes Sociales y TikTok: Se examinan las redes sociales como plataformas de gran alcance que han evolucionado de ser solo entretenimiento, a ser medios de consumo de noticias, lo que es aprovechado por los medios digitales. Se discute el cambio de un modelo de comunicación unidireccional a uno multidireccional y la responsabilidad de los medios de seleccionar contenidos para evitar la filtración de información de dudosa calidad. Específicamente, se analiza TikTok, su crecimiento y cómo la política ha encontrado en esta red una oportunidad para conectar con el público joven.
4. El Periodismo y la IA/Fake News: Esta parte doblemente enfocada aborda la relación entre el periodismo y la IA, destacando su potencial para enriquecer la labor periodística y las nuevas técnicas narrativas, pero también enfatizando la necesidad de un marco ético ante la proliferación de fake news. Además, se define el fenómeno de las noticias falsas como una distorsión de la realidad por intereses personales o políticos, se propone el uso de algoritmos de IA como un mecanismo para distinguirlas de la información veraz, y se plantea el fact-checking como una alternativa de solución para la verificación de hechos.
5. Metodología, Resultados y Discusión: Aquí se detalla la metodología descriptiva y de análisis de contenido empleada para la investigación. Se presenta la delimitación del ámbito de estudio en publicaciones de imagen y video de canales de Youtube seleccionados, y finalmente, se exponen los resultados del análisis de contenido sobre el uso de la IA para producir fake news en la elección de Donald Trump, con datos sobre la probabilidad de IA en imágenes y las visualizaciones en videos de los medios analizados, concluyendo con una discusión de los hallazgos.

La Inteligencia Artificial

A lo largo de los años, se ha venido debatiendo sobre las capacidades de las máquinas y de los seres humanos y esta brecha está

ampliándose cuando se trata del empleo de la Inteligencia Artificial (IA) la cual viene siendo objeto de estudio ya que con el avance de la tecnología, sus algoritmos adquieren elementos y procesos más sofisticados, motivo por el cual se están encendiendo las alarmas en la humanidad frente a la posible robotización de las propias actividades de los seres humanos.

En base a ello, podemos definir la inteligencia artificial, según Castellanos (2019) como la elaboración de sistemas para la resolución de problemas y la realización de labores mediante la simulación de procesos intelectuales.

Una definición más integradora la podemos encontrar en lo sostenido por Mikalef y Gupta (2021), Wirtz et al. (2019) Ronzhyn & Wimmer (2018) y Gutiérrez et al. (2018) quienes sostienen puntos en común siendo uno de los más importantes lo que se refiere a la creación de sistemas computacionales que simulen las capacidades cognitivas de las personas, así como su funcionamiento a partir de algoritmos que generen procesos de aprendizaje automático.

El uso de la IA cada vez se encuentra más especializado en las diversas tareas que solo podían hacerse mediante la actividad humana, y ello viene provocando reacciones de todo tipo, por tal motivo López (2019) reafirma lo señalado manifestando que las investigaciones científicas con respecto a esta tecnología traerán consecuencias negativas en un futuro, las cuales ya se están viendo sus efectos y viene provocando una histeria colectiva en diversos campos del conocimiento humano.

La tecnología empleada por la IA, va tomando mecanismos de perfeccionamiento los cuales son cada vez más difíciles de percibir por las personas, así lo sostiene Valle-Cruz et al., (2020) al determinar

que el concepto de IA se encuentra en una constante evolución de sus procesos tecnológicos dando como resultado diversas aplicaciones cuya finalidad es facilitar el quehacer cotidiano de las personas.

La Inteligencia Artificial en la Política

La política de hoy en día cuenta con bastante injerencia de la tecnología de la inteligencia artificial, ello se da principalmente en los procesos electorales en las cuales los candidatos en sus estrategias de campaña hacen uso de este recurso para provocar y desacreditar a sus respectivos rivales, generando así una ola de fake news o noticias falsas.

Precisamente López et al. (2022) hacen hincapié en ello al señalar que las consecuencias que genera la inteligencia artificial en la comunicación política se traducen en desinformación y noticias falsas lo que genera una preocupación en diversos sectores como los partidos políticos, las instituciones del estado, los medios de comunicación y los ciudadanos de manera especial en época de elecciones.

La campana electoral de Donald Trump, actual presidente de los Estados Unidos, el cual es el objeto de estudio del presente artículo, se caracterizó por el desmedido uso de la Inteligencia Artificial por parte de su equipo de campaña, cuya estrategia se centró en el uso de la imagen de celebridades estadounidenses en apoyo a la candidatura de Trump, y el desacreditar a sus rivales de turno, especialmente a la candidata republicana Kamala Harris. Todo ello produjo una expansión de noticias falsas que terminaron empañando el proceso electoral.

Todos estos hechos conducen a cuestionar la falta de transparencia y seriedad que conlleva un proceso electoral en el que el gobernante que aspira a convertirse en presidente de una nación, se vale de todo

tipo de recursos que atentan contra la ética, la responsabilidad social con tal de sacar ventaja durante la campaña.

Cortés (2019) nos proporciona una interesante reflexión en torno a lo señalado al afirmar que el Sistema democrático correría peligro si los gobernantes consideran el uso de la inteligencia artificial en la toma de decisiones o si las computadoras sean programadas para ganar una elección, en estas circunstancias estaríamos frente a una nueva forma de gobierno en el que la ciudadanía se convierta en incapaz de gobernarse a sí mismos.

Esta tendencia en la política sienta un precedente que puede verse reflejado en varios países sudamericanos que están próximos a realizar elecciones presidenciales, ya que según lo manifiesta Dader & Campos Domínguez (2017), los asistentes virtuales son empleados en la realización de discursos y campañas, lo que ocasiona, de acuerdo a lo señalado por Kreiss & Jasinski (2016) la difusión de mensajes personalizados automáticos producto de la combinación de big data y software avanzado.

Una de las técnicas más peligrosas con respecto al uso de la inteligencia artificial tiene que ver con el deepfakes, término que combina las palabras deep learning (aprendizaje profundo) y fakes (falso) el cual consiste, de acuerdo a lo establecido por García-Ull (2021) en la simulación tanto de las apariencias como el habla de personas reales.

Las redes sociales son las plataformas más vulnerables del uso de estas técnicas, debido a la masificación de sus contenidos, siendo cada vez más perfeccionistas produciendo confusión sobre todo en las campañas electorales las cuales los usuarios las comparten creyendo que los contenidos son reales y auténticos.

Todo ello conlleva una gran preocupación en la que se ve afectada la privacidad, así lo destaca Pérez Dasilva et al. (2021) al considerar que hay riesgo de suplantación de identidad ya que la manipulación de rostros es una amenaza para la política. Los autores Devlin y Cheetham (2023) ponen como ejemplo los deepfakes de personajes como Donald Trump o Barack Obama.

La candidatura presidencial de Donald Trump estuvo marcada por el uso del deepfake, ocasionando el rechazo principalmente de sus adversarios políticos y de los involucrados como es el caso de la popular cantante Taylor Swift, cuya imagen se vio seriamente dañada al incentivar a sus seguidores a votar por la candidatura del candidato del Partido demócrata.

Redes Sociales

Las redes sociales son las plataformas de mayor uso entre millones de usuarios a nivel mundial, sus contenidos, que en un principio eran solo de entretenimiento, han ido evolucionando hasta abarcar una diversidad de temas, siendo uno de los más destacados el referente al periodismo cuyos medios tradicionales han perdido vigencia y han visto un gran potencial de consumo en la gran cantidad de usuarios que hacen uso de las redes sociales.

Esta premisa encuentra respaldo en lo manifestado por Romero-Rodríguez et al. (2021) quienes afirman que los nativos digitales han migrado el consumo de noticias e informaciones de los medios tradicionales como son la radio y la televisión para trasladarse a las plataformas virtuales, lo que es aprovechado por los encargados de los medios debido al alcance que tienen dentro del mundo globalizado.

Precisamente dichas plataformas virtuales, tal es el caso de las redes sociales permiten, de acuerdo a lo manifestado por Appelgren y Nygren, (2014) que se empleen novedosos métodos y técnicas para contar historias. Para poner en contexto lo mencionado por este autor, las campañas presidenciales son definitivamente realizadas en las redes sociales en las cuales se emplean una serie de herramientas para conformar un diseño atractivamente visual con la finalidad de captar el mayor número de seguidores.

La ejecución de dichas campañas obtiene inmediatamente una respuesta por parte de los usuarios dentro de la denominada web 2.0, a través del proceso de interactividad, por tal motivo Segado y Chaparro (2021) complementan lo mencionado al sostener que el proceso de distribución y consumo de noticias se ha expandido y masificado a través de las redes sociales cuyo sistema clásico unidireccional ha sido dejado de lado para adoptar un modelo multidireccional a través de una comunicación horizontal por parte de los usuarios.

Dada la gran cantidad de información que a diario circulan a través de las redes sociales, los medios de comunicación tienen el deber y responsabilidad de seleccionar sus contenidos a fin de evitar que se filtre información de dudosa calidad, tal es así que Bernaissa (2018) reafirma lo dicho al señalar que los medios deben analizar cuidadosamente la información heterogénea y mediática en función a satisfacer los deseos y las preferencias de sus respectivas audiencias, ya que ahora los periodistas digitales, según Knight y Cook (2013), están atentos ahora más que nunca a lo que ocurre dentro de las redes sociales.

Estas plataformas son las preferidas por los usuarios para el consumo de información, y si el contexto ocurre dentro de una campaña

electoral, esto se multiplica, prueba de ello es la efectividad que han tenido los procesos electorales de los últimos años, siendo el más reciente y mediático el ocurrido en la elección del presidente Donald Trump ya que los ojos del mundo estuvieron a la expectativa de lo que ocurriría en uno de los países de potencia mundial como es los Estados Unidos. Dicha campaña se caracterizó por la difusión de información falsa producto del desmedido uso de la tecnología de la inteligencia artificial.

TikTok

La plataforma TikTok se ha vuelto muy popular en los últimos años, es una ventana en la cual encontramos una diversidad de creadores de contenidos para todo tipo de público.

Esta aplicación móvil, según Becerra y Taype (2020) tuvo su origen en China en el año 2016, cuyo nombre en un inicio fue Douyin.

Sin embargo, en el 2017, Ballesteros (2020) señala que dicha aplicación se fusionó con otra aplicación denominada Musical Ly. A partir de ese momento esta red social se hace conocida internacionalmente tomando el nombre de TikTok. Esta aplicación usa la inteligencia artificial en la que se aprende el algoritmo de los usuarios, es decir sus me gusta, comentarios, búsqueda de videos y el tiempo que usan mirando contenidos.

El auge de esta aplicación se dio durante el confinamiento producido a raíz de la pandemia Covid-19, ya que según datos proporcionados por Amlot, como lo cita Becerra y Taype (2020) TikTok creció en marzo del 2020 un 25%, teniendo un total de 75 millones de descargas a nivel mundial. Para finales de ese mismo año, según lo manifestado por Omnicore (2020), la plataforma social alcanzó los 2 mil millones

de descargas y mensualmente un promedio de 800 millones de usuarios, de los cuales el 41% oscilaban entre 16 y 24 años de edad.

Por tal motivo, los medios de comunicación han trasladado sus contenidos a esta red social, es por ello que Expósito y Gobierno (2021) nos proporcionan más alcances al señalar que si los medios quieren tener éxito en esta plataforma, deberían adaptar su lenguaje al de este medio y tener la capacidad para crear una sinergia entre la música, el humor y la información.

El lenguaje empleado por esta red social propone un desafío para los creadores de contenido de los medios, ya que esta sinergia que plantean los autores mencionados requiere un cambio radical dentro del esquema de la presentación de la información que aún mantiene contacto con los medios tradicionales.

La política no es ajena a esta red social, ya que representa una oportunidad para nuevamente conectar con los jóvenes, así lo sostiene Medina-Serrano et al. (2020) destacando que la política cuenta con una fuerte presencia en TikTok, prueba de ello es la gran cantidad de agrupaciones políticas, muchas de ellas nuevas lideradas por jóvenes que a través de contenidos con mensajes creativos han logrado captar la atención de un tipo de público como es el de los jóvenes quienes se mostraron reacios e indiferentes ante los políticos de turno y esta situación se ve más reflejada en época electoral en el que se valen de todos los recursos para captar votos.

El Periodismo y la IA

El empleo de la IA en el periodismo tiene como una de sus finalidades el demostrar la correspondencia entre el desarrollo de la tecnología

con el aspecto social del periodismo, es decir, el emplear los recursos tecnológicos propios de la inteligencia artificial por intermedio del buen uso de los algoritmos en la realización de contenidos informativos de interés para la sociedad, esto queda sustentado en lo manifestado por Peña et al. (2023) en el que los contenidos de la inteligencia artificial son aplicados y fomentan la discusión entre los avances tecnológicos con el lado social del periodismo.

El profesional de la información con todos estos recursos tecnológicos a su disposición, tiene que tener un marco ético ante la elaboración de contenidos informativos con el uso de la inteligencia artificial, ante la proliferación de fake news que hacen un mal uso de esta tecnología, por tal motivo, Dörr y Hollnbuchner (2017) plantean tres dimensiones éticas dentro del denominado periodismo automatizado, recolección de datos, la elaboración y difusión.

Los comunicadores sociales de esta generación tienen que tener más que nunca presente que el acceso a la tecnología se ha vuelto masivo e indispensable dentro del mundo globalizado en el que nos encontramos, siendo precisamente los medios de comunicación los que más han hecho uso de ello y ante el afán mercantilista de muchos de los dueños y encargados de dichos medios, el aspecto ético, deontológico de la carrera periodística se ha convertido últimamente en un espejismo. Por ello el mayor desafío dentro del entorno digital está en poner en práctica dichos valores ya que la inmediatez de las noticias está al alcance de millones de cibernautas a través de la telefonía móvil.

Para ello, la Inteligencia Artificial juega un papel preponderante ya que es el propio profesional de las comunicaciones quien debe hacer un uso responsable de toda esta tecnología, la cual proporciona diversas

herramientas y posibilidades que con el correr de los años serán más sofisticadas.

La labor periodística se ha visto enriquecida con el desarrollo de la inteligencia artificial mediante la cual la complementación de la información sea a través de datos, ilustraciones o productos audiovisuales representan un reto para la profesión al imponer una tecnología de vanguardia, todo ello según los argumentos que sostiene Túñez-López et al (2020) al señalar que la automatización no representa ningún peligro, sino todo lo contrario una oportunidad de realizar nuevas técnicas narrativas en el periodismo.

Precisamente sobre este punto es que apelamos a la autorregulación, uno de los principales engranajes de la ética periodística y de acuerdo a lo manifestado por Díaz del Campo; Chaparro-Domínguez, (2020) esta práctica se sostiene en los libros de estilo, códigos de deontología periodística, los cuales deben siempre ser aplicados ahora más que nunca debido a que estas malas prácticas proliferan en muchos de los contenidos digitales.

Fake News

Las noticias falsas conocidas comúnmente como fake news son una tendencia que se viene dando en las plataformas de información digital, convirtiéndose en un gran problema en la sociedad de la información, por tal motivo Muñoz y Montero (2017) nos dan más alcances acerca de este fenómeno al manifestar que de acuerdo a intereses personales o políticos, la realidad se vea distorsionada con el fin de obtener vistas en las redes sociales o en los sitios web.

El periodismo digital también encuentra en la tecnología de la inteligencia artificial, un mecanismo mediante el cual puede detener la cantidad de noticias falsas, las cuales proliferan en la red inclusive en mayor proporción que las noticias verdaderas y verificables, es por ello que Flores Vivar (2019) propone como una alternativa de solución a este fenómeno el uso de los algoritmos de la IA a través de bots que permita distinguir entre los contenidos informativos veraces frente a la manipulación de lo real, de manera que los datos con que son desarrollados estos bots sean los adecuados.

Dentro de la comunidad digital hay posiciones encontradas sobre este fenómeno llamado “fake news”, que viene cuestionando las informaciones vertidas por los medios de comunicación digitales en afán de la feroz competencia existente por captar la atención de los cibernautas. En los últimos años, la expansión de este tipo de informaciones afectó la credibilidad de muchos medios sin encontrar ningún viso de solución, es por ellos que ante el auge que viene teniendo la tecnología de la inteligencia artificial, muchos especialistas en el tema proponen que su uso se enfoque en desarrollar algoritmos que puedan diferenciar una información falsa de una real.

Todo ello viene provocando una de las mayores crisis en el periodismo del presente siglo, así lo sostiene Steensen (2019) al señalar que la confiabilidad de las noticias se ve profundamente afectada dejando de lado el debate entre lo correcto e incorrecto, de ahí se desprende hacer una revisión más exhaustiva de la esencia de la profesión periodística y sus códigos deontológicos.

Una de las posibles alternativas de solución ante esta problemática de las noticias falsas sería poner en práctica el fact-checking o

verificador de hechos, que de acuerdo a lo sostenido por López, Díez y Puente (2022) favorece la automatización en la confirmación de la veracidad de las noticias, localizando la fuente principal, así como el análisis de contenido.

Este crecimiento que ha tenido este tipo de noticias en las plataformas de internet afecta a los ciudadanos, y según lo sostenido por Mendiguren et al. (2020) es el público joven el más afectado ya que ellos usan las redes para estar informados y actualizados.

Con respecto a la campaña presidencial en los Estados Unidos la cual tuvo como ganador a Donald Trump, dicha campaña tuvo como protagonista a las redes sociales, y aunque el contenido hizo uso de la inteligencia artificial, el público objetivo, en este caso los jóvenes, se hizo viral y fue compartido por muchos, siendo uno de los hechos más significativos, el apoyo a la candidatura de Trump, por parte de la cantante Taylor Swift, muy popular en todo el mundo, la cual generó un gran rechazo entre sus seguidores, hasta que finalmente ella misma salió a desmentir dicho apoyo, acusando a Donald Trump de haber perjudicado su imagen.

Lo sucedido con la cantante Taylor Swift comprueba lo fácil que es producir una fake news y hacerla viral, tal es así que Valarezo y Rodríguez (2019) nos profundizan sobre este tema, al sostener acerca de las consecuencias negativas que provocan las noticias falsas, ya que en la actualidad las personas comentan y comparten contenido que observan en las redes sociales, y a ello le añadimos las estrategias inadecuadas para conseguir la mayor cantidad de vistas posibles, anteponiendo intereses lucrativos o comerciales frente al valor informativo.

Metodología

El presente artículo tendrá un carácter descriptivo puesto que a partir del acopio de datos nos aproximamos al objeto de estudio y analiza la situación en la que se circunscribe el trabajo.

La investigación descriptiva permite realizar una revisión crítica y analizar en profundidad la corriente de investigación con el objeto de reseñar las condiciones que originaron su aparición, así como los resultados más relevantes alcanzados por los diversos estudios realizados sobre el tema, así como también señalar los problemas conceptuales y las limitaciones metodológicas existentes (Tinto, 2013, p. 138).

Consideramos que esta investigación descriptiva encaja dentro de la metodología de análisis de contenido, puesto que de esta manera reduciremos la totalidad del cuerpo de estudio por un grupo de variables con el fin de interpretarlas, por esta razón nos hemos adherimos al concepto de Bardin sobre el análisis de contenido “conjunto de técnicas de análisis de comunicación tendente a obtener indicadores (cuantitativos o no) por procedimientos sistemáticos y objetivos de descripción del contenido de los mensajes, permitiendo la inferencia de conocimientos relativos a las condiciones de producción/recepción de estos mensajes” (Bardin, 2002, p. 32).

Asimismo, debemos señalar que en cuanto a la delimitación del ámbito de estudio en esta investigación se analizarán las publicaciones de imagen y de video de DW Español (5.25M suscriptores) El Universal de México (6.25M suscriptores), Meganoticias de Chile (3.37 M suscriptores) y Noticias Telemundo (8.76M suscriptores), ya que estos medios cuentan con una gran cantidad de suscriptores en sus cuentas de Youtube.

Para el hallazgo de nuestros resultados, se hará uso de una ficha de análisis conformada por un conjunto de variables y categorías seleccionadas a partir de una validación preliminar, para lograr optimizar y precisar la extracción de datos y la validez de su interpretación.

Resultados y Discusión

Imágenes

Fuente	Tipo de contenido	Generado con inteligencia artificial
Stable Diffusion	Fotografía	Probabilidad: 9.2%
El Imparcial	Fotografía	Probabilidad: 31.3%

Elaborado por los autores

Videos					
Medio informativo	Fecha	Número de “visualizaciones”	Número de comentarios	Tipo de contenido	Generado con inteligencia artificial
DW Español	22 de agosto 2024	5716	19	Video	Si
El Universal	21 de julio 2025	36,764	215	Video	Si
Meganoticias de Chile	27 de julio 2025	2599	5	Video	Si
Noticias Telemundo	05 de setiembre 2024	3836	20	Video	Si

Elaborado por los autores.

De acuerdo a los medios analizados acerca del uso de la Inteligencia Artificial para producir fake news en la elección presidencial de Donald Trump, en lo que respecta a la categoría de Imágenes, El Imparcial, cuya fotografía de la cantante Taylor Swift apoyando a

Donald Trump, fue el medio que tuvo mayor probabilidad de IA con un porcentaje de 31.3%, mientras que Stable Diffusion, cuya imagen de Donald Trump dando un beso a su contrincante Kamala Harris, tuvo una probabilidad de IA del 9.2%

Con respecto a la categoría Videos, los medios analizados DW Español, El Universal, Meganoticias Chile y Noticias Telemundo, presentan contenidos con Inteligencia Artificial, siendo El Universal el medio con mayor cantidad de visualizaciones 36,764, le siguen DW Español con 5716 visualizaciones, luego Noticias Telemundo con 3836 y finalmente Meganoticias Chile con 2599 visualizaciones.

La labor periodística se ha visto enriquecida con el desarrollo de la inteligencia artificial mediante la cual la complementación de la información sea a través de datos, ilustraciones o productos audiovisuales representan un reto para la profesión al imponer una tecnología de vanguardia, todo ello según los argumentos que sostiene Túñez y Tejedor (2019) la automatización no representa ningún peligro, sino todo lo contrario una oportunidad de realizar nuevas técnicas narrativas en el periodismo.

Conclusión

Los contenidos generados con Inteligencia Artificial están ocasionando perjuicios en la mayor parte de las informaciones de los medios periodísticos al producir numerosos fake news las cuales desacreditan y desprestigian la labor de los medios de comunicación, siendo las redes sociales la plataforma en la cual dichas noticias se difunden y comparten descontroladamente.

La campaña presidencial del candidato republicano Donald Trump quien ganó en las últimas elecciones presidenciales de los Estados Unidos se caracterizó porque su contenido estuvo elaborado con inteligencia artificial con la finalidad de desacreditar a sus rivales y adversarios generando polémica y desconcierto entre los electores.

Los medios informativos analizados muestran imágenes y videos elaborados con esta tecnología, con el propósito de viralizar contenido de tipo fake news, sentando las bases como precedente en futuras campañas presidenciales próximas a realizarse en algunos países sudamericanos.

Referencias

Appelgren, E., y Nygren, G. (Appelgren, E., y Nygren, G. (2014). Data Journalism in Sweden. *Digital Journalism*, 2(3), 394-405. <http://doi.org/10.1080/21670811.2014.884344>

Ballesteros, C. (2020). *La propagación digital del coronavirus: Midiendo el engagement del entretenimiento en la red social emergente TikTok*. <https://pdfs.semanticscholar.org/8943/bcfdec98e8e3d0ad379954a63fe1931dc1ba.pdf>

Bardin, L. (2002). *El análisis de contenido*. Ediciones Akal.

Becerra, N., & Taype A. (2020). *TikTok: ¿una nueva herramienta educativa para combatir la COVID-19?* http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S1728-59172020000200249

Bernaissa, S. (2018). Las redes sociales como fuente de información periodística en la prensa digital española (El País, El Mundo, La Vanguardia y ABC). *Index.comunicación*, 8(3), 13-42.

- Castellanos, J. (2019). La democracia algorítmica: inteligencia artificial, democracia y participación política. *Revista General de Derecho Administrativo*, 50
- Cortés, D. (2019). La incorporación de la inteligencia artificial en el poder público. La política en el mundo inteligente. *Conjeturas Sociológicas*, 7(19), 36-75.
- Dader, J. L., & Campos, E. (Eds.) (2017). *La búsqueda digital del voto: Ciber campañas electorales en España 2015-16*. Tirant lo Blanch. <https://dialnet.unirioja.es/servlet/libro?codigo=702481>
- Devlin, K., & Cheetham, J. (2023, marzo 25). *Cómo detectar imágenes creadas por inteligencia artificial como las fotos falsas del arresto de Trump*. BBC News Mundo. <https://www.bbc.com/mundo/noticias-65071726>
- Díaz, J., & Chaparro, M. A. (2020). Periodismo computacional y ética: Análisis de los códigos deontológicos de América Latina. *Icono 14*, 18(1), 10-32.
- Dörr, K. N., & Hollnbuchner, K. (2017). Ethical challenges of algorithmic journalism. *Digital journalism*, 5(4), 404-419.
- DW Español. (22 de agosto de 2024). *Imágenes falsas hechas con IA en campaña electoral: ¿es ético? ¿es legal?* [Archivo de Vídeo]. Youtube. <https://www.youtube.com/watch?v=5b7Nc9Ao-hk>
- El Universal. (21 de julio de 2025). *¿Será un mensaje? Trump sube polémico video creado con IA; muestra a Barack Obama siendo arrestado* [Archivo de Vídeo]. Youtube. <https://www.youtube.com/watch?v=UEcCvsYf1c>

- Expósito, A., & Gobierno N. (2021). *Periodismo en TikTok : estudio de casos de los medios de comunicación españoles*. <https://riull.ull.es/xmlui/bitstream/handle/915/24905/Periodismo%20en%20TikTok%20estudio%20de%20casos%20de%20los%20medios%20de%20comunicacion%20espanoles.pdf?sequence=1>
- Flores, J. M. (2019). Inteligencia artificial y periodismo: diluyendo el impacto de la desinformación y las noticias falsas a través de los bots. *Doxa Comunicación*, 29, 197-212. <https://bit.ly/4k857dM>
- García-Ull, F. J. (2021). Deepfakes: el próximo reto en la detección de noticias falsas. *Anàlisi: Quaderns de Comunicació i Cultura*, 64, 103-120. <https://doi.org/10.5565/rev/analisi.3378>
- Gutiérrez, J. R., Ramos, E. M., & Acosta, R. (2018). Inteligencia artificial y aprendizaje máquina: Aplicaciones y tendencias. En A. Román, S. Sandoval, M. E. Cabello & J. Herrera (Eds.), *Universidad de Colima*. https://www.academia.edu/38537176/Inteligencia_artificial_y_aprendizaje_m%C3%A1quina_Aplicaciones_y_tend
- Knight, M. y Cook, C. (2013). *Social Media for Journalists: Principles and Practice*. Sage.
- Kreiss, D., & Jasinski, C. (2016). The Tech Industry Meets Presidential Politics: Explaining the Democratic Party's Technological Advantage in Electoral Campaigning, 2004-2012. *Political Communication*, 33(4), 544-562. <https://doi.org/10.1080/10584609.2015.1121941>
- López, M. (2019). Las narrativas de la inteligencia artificial. *Rev Bioética y Derecho*, 46, 5-28. <https://bit.ly/42QpMO3>
- López-López, P. C., Díez, N. L., & Puentes-Rivera, I. (2022). La inteligencia artificial contra la desinformación: una visión desde la comunicación política. *Razón Y Palabra*, 25(112), 5-11. <https://doi.org/10.26807/rp.v25i112.1891>

- Medina-Serrano, J. C., Papakyriakopoulos, O., & Hegelich, S. (2020). *Dancing to the partisan beat: A first analysis of political communication on TikTok*. In WebSci '20:12th ACM Conference on web science. Association for Computing Machinery. <https://bit.ly/3R9wvKN>
- Meganoticias. (27 de julio de 2025). *Crece la polémica de Trump ante publicaciones con IA* [Archivo de Vídeo]. Youtube. <https://www.youtube.com/watch?v=51VCpq4H5uw>
- Mendiguren, T., Pérez, J., & Meso, K. (2020). Actitud ante las Fake News: Estudio del caso de los estudiantes de la Universidad del País Vasco. *Revista de Comunicación*, 19(1), 171-184. <https://doi.org/10.26441/RC19.1-2020-A10>
- Mikalef, P., & Gupta, M. (2021). Artificial intelligence capability: conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance. *Information & Management*, 58(3), 103434. <https://doi.org/10.1016/j.im.2021.103434>
- Muñoz, P., & Montero, M. (2017). *Medios de comunicación y posverdad: Análisis de las noticias falsas en elecciones presidenciales de EE. UU. 2016*. [Trabajo de Fin de Máster en Medios, Universitat Autònoma de Barcelona].
- Noticias Telemundo. (05 de setiembre de 2024). *Swift, Harris y Trump: los 'fake news' de inteligencia artificial en la campaña* [Archivo de Vídeo]. Youtube. <https://www.youtube.com/watch?v=FpwxOgUL3yg>
- Omnicores (Ed.) (2020). *TikTok by the numbers: Stats, demographics & fun facts*. Omnicore. <https://bit.ly/3yFND3J>

- Peña, S., Meso, K., Larrondo, A., & Díaz, J. (2023). Without journalists, there is no journalism: the social dimension of generative artificial intelligence in the media. *Profesional de la información*, 32(2). <https://doi.org/10.3145/epi.2023.mar.27>
- Pérez, J., Meso, K. & Mendiguren, T. (2021). Deepfakes on Twitter: Which Actors Control Their Spread? *Media and Communication*, 9(1). <https://doi.org/10.17645/mac.v9i1.3433>
- Ronzhyn, A., & Wimmer, M. (2018). *Report for Electronic Governance research and practice worldwide*. European Commission. https://collections.unu.edu/eserv/UNU:700/GOV3.0_D1.1-Baseline-Research_v.0.70.pdf
- Romero, L., Tejedor, S., & Castillo, B. (2021). From the Immediacy of the Cybermedia to the Need for Slow Journalism: Experiences from Ibero-America. *Journalism Practice*, 1-19. <https://doi.org/10.1080/17512786.2020.1870530>
- Segado, F., & Chaparro, M. A. (2021). Compartición de noticias en redes sociales. Difusión selectiva y poder social. *Cuadernos.info*, (50), 45-68. <https://doi.org/10.7764/cdi.50.27433>
- Steensen, S. (2019). Journalism's epistemic crisis and its solution: Disinformation, datafication and source criticism. *Journalism*, 20(1), 185-189. <https://doi.org/10.1177%2F1464884918809271>
- Tinto, J. A. (2013). El análisis de contenido como herramienta de utilidad para la realización de una investigación descriptiva. Un ejemplo de aplicación práctica utilizado para conocer las investigaciones realizadas sobre la imagen de marca de España y el efecto país de origen. *Provincia*, 29, 135-173.

- Túñez, J. M., & Tejedor, S. (2019). Inteligencia artificial y periodismo [presentación del monográfico]. *Doxa Comunicación: Revista Interdisciplinar de Estudios de Comunicación y Ciencias Sociales*, (29), 163-168 <https://doi.org/10.31921/doxacom.n29a8>
- Túñez-López, J. M., Toural-Bran, C., & Frazão, A. G. (2020). Del periodismo de datos al periodismo robótico: la automatización del procesamiento de noticias. En J. Vázquez-Herrero, S. Direito-Rebollal, A. Silva-Rodríguez, X. López-García (eds), *Metamorfosis periodística. Estudios en Big Data*. Springer, Cham.
- Valarezo, L., & Rodríguez, C. (2019) La innovación en el periodismo como antídoto ante las fake news. *RISTI - Revista Ibérica de Sistemas e Tecnologías de Informação*, (20), 24-35
- Valle, D., Criado, J. I., Sandoval, R., & Ruvalcaba, E. A. (2020). Assessing the public policy cycle framework in the age of artificial intelligence: from agenda-setting to policy evaluation. *Government Information Quarterly*, 37(4), 101509. <https://doi.org/10.1016/j.giq.2020.101509>
- Wirtz, B. W., Weyerer, J. C., & Geyer, C. (2019). Artificial intelligence and the public sector—applications and challenges. *International Journal of Public Administration*, 42(7), 596-615. <https://doi.org/10.1080/01900692.2018.1498103>

PERCEPCIONES DE REGULACIÓN DE LA DESINFORMACIÓN CREADA CON GENAI EN ELECCIONES, CASO DE ECUADOR

Abel Suing¹

En abril de 2025 los ecuatorianos eligieron presidente en una campaña política caracterizada por la circulación masiva de mensajes desinformativos creados en inteligencia artificial generativa (GenAI), incluso la Misión de Observación Electoral de la Unión Europea en su informe final señaló que “el uso de contenidos generados por inteligencia artificial aumentó a lo largo de la campaña electoral, a menudo empleados para alimentar la desinformación y los ataques personales contra los candidatos” (OEA, 2025).

En este escenario y a nivel internacional, varios países adoptan regulaciones respecto al uso indebido de la GenAI en elecciones, pero

1. Profesor del Departamento de Ciencias de la Comunicación. Universidad Técnica Particular de Loja
arsuing@utpl.edu.ec

tiene enfoques diversos, lo que refleja sus panoramas políticos, jurídicos y tecnológicos. Por ejemplo, la UE aprobó la Ley de Inteligencia Artificial que incluye regulaciones sobre deepfakes y otras aplicaciones que podrían manipular a los votantes. Sin embargo, esta ley recibió críticas por vulnerar potencialmente la privacidad y el derecho a la libertad de expresión (Romero, 2024). En tanto que, en los Estados Unidos de América, varias regiones implementaron sus propias medidas que incluyen exigencias de transparencia y la prohibición de medios sintéticos dañinos (Paget, 2024).

Además, frente al fenómeno de la desinformación, algunas plataformas aplican autorregulación, como el etiquetado del contenido, como una primera respuesta que busca rendir cuentas. También se ensayan medidas legislativas coherentes a “la necesidad de revisar los marcos jurídicos existentes, con la vista puesta en salvaguardar los supuestos de libertad, igualdad, integridad, transparencia y justicia” (Rubio, 2025, p. 96).

Ante todo, una regulación eficaz en las campañas electorales requiere la colaboración entre gobiernos, instituciones electorales, empresas tecnológicas y la sociedad civil. Este enfoque colaborativo garantiza una respuesta integral a los desafíos que plantea la GenAI (Schipper, 2025). De igual forma, se necesita mejorar la alfabetización digital de la población para combatir la propagación de la desinformación y garantizar una participación pública informada. Educar a la ciudadanía sobre los riesgos y las consideraciones éticas contribuye a fortalecer la resiliencia frente a la manipulación (Karlsen, 2021), en sentido similar, es importante crear conciencia a través de los medios

de comunicación y campañas dirigidas a los partidos políticos y a las organizaciones electorales.

Vale recordar que la GenAI impacta significativamente en las relaciones sociales y está modificando las formas cómo se interactúa dentro de los marcos institucionales. Hay valoraciones sobre afectaciones negativas de la GenAI en la enseñanza, la investigación y también en la política, y ello preocupa porque pueden ser utilizada para manipular. Lamentablemente “nuestros actuales sistemas democráticos no están en condiciones de gestionar la transformación digital más que en una medida limitada” (Duberry, 2022).

Con débil calidad informativa y sin datos sólidos al alcance se provoca un deficiente desempeño de los sistemas productivos, y se afecta al sistema democrático de las naciones. “La democracia requiere una ciudadanía bien informada e implica que personas con perspectivas diversas se reúnan para debatir y encontrar soluciones comunes a través del diálogo y de la participación” (UNESCO, 2024, p. 16), por ello es importante evaluar el uso responsable de la GenAI para la gobernanza, particularmente porque hay evidencias de creación de mensajes falsos en línea para afectar los procesos electorales (Luminate, 2023; Marsden y Meyer, 2019; Schia & Gjesvik, 2020).

La presencia de información falsa en las elecciones no es un fenómeno nuevo (Posetti & Matthews, 2018), pero luego de la Covid-19 los organismos internacionales están alerta debido a las amenazas que representa para la confianza pública (Palau-Sampio, 2024). En Estados Unidos “el contenido engañoso generado por inteligencia artificial moldeó el discurso electoral al amplificar otras formas de desinformación y encender los debates políticos” (Stockwell et al., 2024, p. 3).

“En Venezuela, los medios de comunicación controlados por el Estado utilizaron vídeos generados por IA [...] para difundir propaganda a favor del gobierno” (Mansell, et al., 2025, p. 19).

Igualmente, en 2022, circularon falacias “para deslegitimar, perjudicar o dañar la figura de los candidatos a la presidencia de la República de Colombia, las cuales pretenden exacerbar emociones que profundicen en la ruptura de los consensos sociales” (Echeverría y Rodríguez, 2023, p. 239). En ese país, el uso de herramientas de GenAI tiende a incrementar “la propagación de noticias falsas, la manipulación de tendencias en la opinión pública y el descrédito a candidatos opositores” (López et al., 2024, p. 18). Se arriba a una época de “elecciones de alto riesgo” (Alvim et al., 2024).

En lo electoral, la GenAI puede incidir en predecir y pronosticar el comportamiento del votante y sus posibles resultados, desarrollar y difundir contenido para las campañas (Dad & Khan, 2023). “Aunque las amenazas de la inteligencia artificial no se comprendan del todo, su potencial capacidad erosiona la confianza pública en las instituciones y procesos democráticos” (Foro sobre Información y Democracia, 2024, p. 100), y entre estas instituciones están los medios de comunicación masivos (Londoño & García-Perdomo, 2024).

En el Informe de Riesgos Globales 2024 se reconoce a la desinformación y la información errónea impulsadas por la GenAI como las mayores amenazas a corto plazo en el mundo, especialmente en el contexto de su impacto negativo en la credibilidad de las elecciones (World Economic Forum, 2024). “Existe una necesidad urgente de que los gobiernos, las plataformas de redes sociales y otras empresas protejan los pilares fundamentales de una sociedad democrática y sus

mecanismos de implementación: las elecciones” (UNESCO, 2022, p. 26), entre otras razones porque “la toma de decisiones políticas asistida por la inteligencia artificial puede conducir en última instancia a establecer una forma de democracia automatizada y privar a los seres humanos de autonomía sobre los procesos políticos” (Ramos, 2022).

Por lo señalado, tanto Estados Unidos como la Unión Europea tomaron medidas para salvaguardar la integridad de sus elecciones, pero “proteger los procesos democráticos en la era digital requiere un enfoque multifacético que involucre a gobiernos, organismos reguladores, sector privado y sociedad civil. El objetivo no es reprimir la libertad de expresión, sino promover un ecosistema informativo más justo” (Muto, 2025). Las potenciales regulaciones deben evaluarse para que no restrinjan las libertades y el reconocimiento de las posiciones políticas a partir de las cuales las personas elijan.

En general, la regulación de la GenAI debe complementarse con el amparo de la sociedad civil, comités éticos de las empresas y organismos públicos para una supervisión múltiple y diversa. Este enfoque tomó peso porque “tras las elecciones europeas de mayo de 2019 [...] ven la corregulación auditada como un sistema de gobernanza más deseable porque es más adecuado para esta era y contexto, y para la magnitud y la rápida evolución del problema” (Unión Europea, 2019). La corregulación “es un modelo que implica una regulación privada (ya sea en forma de autorregulación o de regulación individual) que es apoyada o fomentada activamente por el gobierno, y que normalmente está respaldada por la legislación” (Article 19, 2019).

El órgano que pretenda regular la GenAI en elecciones debe defender los valores democráticos, la justicia social incluyendo “la

transparencia, la alfabetización mediática y la introducción de formas de corregulación centrada en el ser humano” (Marsden, et al., 2020, p.18). En estudios previos se enfatiza que, “para sortear con éxito las complejidades de un mundo cada vez más moldeado por la inteligencia artificial, es imperativo desarrollar marcos regulatorios sólidos, promover la alfabetización en IA y fomentar el pensamiento crítico” (Arda & Başarır, 2024, p. 137).

Frente a fenómenos emergentes, el marco internacional de derechos humanos aparece como el mejor mecanismo para la formulación de estándares. La GenAI debe funcionar de conformidad con la Declaración Universal de Derechos Humanos. En este sentido, en marzo de 2024, la Asamblea General de la ONU “adoptó una resolución [...] para promover sistemas de inteligencia artificial “seguros y fiables” que respeten los derechos humanos y contribuyan al desarrollo sostenible” (Naciones Unidas, 2024).

En América Latina, igual en otras regiones se discute sobre el panorama regulatorio de la GenAI, se procura que, “toda iniciativa para contener el desarrollo de la inteligencia artificial debe tener como objetivo ineludible la protección de los derechos fundamentales reconocidos internacionalmente” (Giandana, 2024, p. 3). Implícitamente, “un buen entorno legislativo, la transparencia y unos controles eficaces son las principales herramientas contra las técnicas de manipulación” (Hajdúková, 2022, p. 22).

Entonces, sobre la base de las orientaciones de las instituciones nacionales, se espera que las desinformaciones disminuyan, pero esto no ocurre al revisar el caso de Ecuador. Este país andino llama la atención mundial por su poca gobernabilidad, y se identifican prácticas

incipientes de uso de la GenAI para la manipulación en elecciones (Barredo-Ibáñez, et al., 2021; Bradshaw & Howard, 2017). La desinformación en elecciones a través de GenAI no es nueva. Se informó de estrategias digitales de manipulación en las campañas presidenciales de 2017 y 2020 (Ackerman et al., 2020; Puyosa, 2017). En 2019 el candidato a la alcaldía de Quito, Jorge Yunda, empleó la GenAI y se considera que contribuyó a su triunfo (Calderón, 2019).

En 2025 se reportó que la desinformación fue generalizada en la campaña, proliferaron narrativas de fraude. La Misión de Observación Electoral de la Unión Europea (2025, p. 2), observó contenidos generados y manipulados por la GenAI en las redes sociales, incluso contenidos de pago. Igualmente, “se utilizaron múltiples tácticas para desacreditar a los candidatos, incluidos anuncios de campaña negra generados por GenAI” (Misión de Observación Electoral de la UE, 2025, p. 7). Los medios de comunicación también alertaron del “montaje de audios, manipulación de imágenes de video, alteración de fotografías de candidatos [que] fueron generados con GenAI” (Jara, 2025). “En la campaña presidencial de Ecuador se evidenció el uso sofisticado de la GenAI, como videos alterados de increíble realismo” (Primicias, 2025). La organización Usuarios Digitales (2025) reportó que, entre el 18 de enero y el 2 de abril de 2025, circularon sistemáticamente publicaciones en cuentas de redes sociales donde se aplicó la GenAI para apoyar, atacar o satirizar a los candidatos rivales,

En Ecuador existen iniciativas para regular la GenAI. Se presentaron como proyectos de ley, pero en ninguno se mencionan los riesgos en elecciones (Asamblea Nacional del Ecuador, 2025). Las propuestas

comparten principios como la privacidad y seguridad de los datos personales, transparencia y neutralidad tecnológica.

Sobre la base de lo señalado, esta investigación se enfoca a determinar las percepciones y condiciones nacionales, en Ecuador, para desarrollar una regulación que alerte y señale metodologías para prevenir los impactos negativos del uso indebido de la GenAI en la libertad de expresión y en la democracia. La metodología empleada es cualitativa a partir de entrevistas a expertos nacionales: cuatro hombres y una mujer. Profesionales en comunicación, derecho, investigación académica, política y consultoría. Los diálogos sucedieron entre julio y agosto de 2025, a través de la plataforma de video conferencia Zoom.

Se utilizó Qwen3-Max, modelo de lenguaje de gran escala, para búsquedas bibliográficas, ordenar los testimonios de los entrevistados y mejorar la redacción de determinados párrafos. Tras emplear esta herramienta, se revisó y editó el contenido según fue necesario. El autor asume la responsabilidad por el contenido del texto.

La GenAI en campaña electoral “tiene tres aristas, uno como herramienta técnica, otro como instrumento generativo de desinformación y como instrumento de discursos políticos” (entrevistado 1), pero “estamos temprano en la regulación de la GenAI. Debemos ponernos al día, y hay camino por recorrer. Se podría aprobar una legislación que señale que, al usar imágenes creadas mediante la GenAI, tengan licencias de agua o un tipo de mensaje que especifique que fueron creados mediante GenAI” (entrevistado 5). Regular “es muy difícil porque no depende de la legislación propia el uso de plataformas (Facebook, X, YouTube), es un contenido que es difícil de regular, tal vez con organismos regionales que ayuden” (entrevistada 3).

“¿Cómo se regula a usuarios que no tiene nacionalidad, que están esparcidos en el mundo? Esto genera crisis en los Estados, hoy hay nuevos dueños, que son las empresas, las plataformas, redes sociales” (entrevistado 4). Sin embargo, hay experiencias internacionales que muestran avances, por ejemplo, “en Reino Unido se hace una suerte de exigibilidad a los partidos políticos para publicar el financiamiento y anuncios en redes sociales” (entrevistada 3), en Brasil “cuando una información estaba comprobada como falsa o una desinformación, el ente electoral notificaba a la plataforma, y ellos tenían un límite de tiempo para bloquear o retirar el contenido” (entrevistado 1).

En Ecuador, en “el debate de primera vuelta, el personaje más importante fue la GenAI ya que todos los candidatos la mencionaron, como una idea, como para dar un mensaje de innovación, que poco a poco se prostituye” (entrevistado 1), y aunque el órgano nacional electoral “con apoyo de la cooperación internacional, ha mostrado lo que ocurre, que la GenAI está siendo mal utilizada, que los candidatos hacen mal uso de ella para atacar o denigrar a su oponente” (entrevistado 2) no se logra regular porque “hay ese vacío legal en el Código de la Democracia” (entrevistada 3), “la GenAI trae nuevos actores y escenarios” (entrevistado 4). Para controlar “se requeriría un marco jurídico que establezca cuáles serían los límites de esas publicaciones, qué contenidos y quiénes podrían hacerlo para que haya una autoridad que regule” (entrevistado 2).

Parte de “la solución es una cultura de verificación. Necesitamos ciudadanos verificadores, que aprendamos que la base es la educación, que aprendamos a contrastar la información” (entrevistado 1). Hay “necesidad de procesos formativos en alfabetización digital, de formación

respeto a la información que se consume. El problema no es la GenAI, sino de qué manera las personas, la inteligencia humana emplean esta herramienta” (entrevistado 2).

Conclusiones

La regulación del uso indebido de la GenAI y la desinformación en campañas electorales representa desafíos jurídicos y políticos emergentes que demandan respuestas específicas y coordinadas. El desarrollo de marcos regulatorios multinivel que combinen instrumentos tradicionales ayudará a mitigar los riesgos que supone la innovación tecnológica sin ralentizar una evolución a una democracia digital.

Sobre la base de los testimonios de los expertos, los ecuatorianos están de acuerdo con regular el uso de la GenAI en las campañas electorales, pero aún deben debatirse los mecanismos y el modelo, y sugieren que el Consejo Nacional Electoral (CNE) es el órgano llamado a reglamentar en futuras campañas. Se percibe que la desinformación creada a través de la GenAI influyó en quién ganó la elección presidencial de Ecuador en 2025; y, consideran que el uso de esta tecnología no aumentó la confianza en los candidatos y ni en el resultado de la elección.

Sin un marco legal adecuado, la GenAI seguirá utilizándose para influir en los votantes de forma poco ética, pero la solución no es restringirla, sino controlar su aplicación en la política. Se expresó la urgencia en que las instituciones gubernamentales implementen programas de alfabetización mediática para que los ciudadanos identifiquen contenidos falsos. Uno de los mayores desafíos en campañas políticas es distinguir entre contenido real y artificial.

Implícitamente se devela aceptación por la autorregulación que debe considerar la participación del gobierno, la sociedad civil, la academia, el sector privado y la colaboración con plataformas tecnológicas para mejorar la detección y eliminación de noticias falsas y contenido manipulado. Siguiendo experiencias internacionales, la autorregulación muestra la relevancia de los consensos nacionales para exigir rendiciones de cuentas a las plataformas que alojan los mensajes falsos. La norma a aprobarse puede derivar en un reglamento que incluya medidas de transparencia, límites en la personalización de mensajes políticos, y sanciones para el uso indebido de la GenAI.

La GenAI crea una falsa sensación de pluralidad, sin embargo, los usuarios son manejados por algoritmos. Se destacó el uso de cuentas falsas para afectar las conversaciones y reforzar sesgos ideológicos. La GenAI no manipula por sí sola, depende de las intenciones de quienes la programan; correctamente aplicada podría democratizar el acceso a la información. Desde esta óptica, la imparcialidad no depende de la tecnología, sino del criterio de los ciudadanos al evaluar la información. Entre las alternativas para disminuir la desinformación generada en la campaña electoral, se propone acrecentar la cultura de verificación.

En este escenario se precisan de orientaciones, pedagogía y adquisición de nuevas competencias para que los ciudadanos conserven sus derechos, de lo contrario la proporción de marginados se acrecentará. Pero, pese a la aceptación casi generalizada de abordar los temas que irrumpen en las agendas públicas, se perciben mayores acciones de la sociedad civil y de las empresas que del Estado. En el origen de los malestares, pero también de las soluciones, están interpretaciones sesgadas, incomprensiones y percepciones erróneas edificadas

por inconsistentes procesos de información. Si se comprendiese que comunicación es diálogo, tolerancia, experiencias compartidas y actuar en común, se llegarían a acuerdos en cortos plazos. Ante los desafíos inmediatos que Ecuador y el mundo atienden, se precisan de mayor presencia y liderazgo de la comunicación pública, ya que, sin ella, los aportes de los actores privados y comunitarios quedan truncos.

Lo señalado es evidente cuando se examina el fenómeno de la desinformación, particularmente en la reciente campaña electoral. Las personas requieren de herramientas para distinguir lo que es cierto en las redes sociales, y las encuentran en los medios masivos, en colectivos, y entre sus amistades, pero en mínima medida en los medios públicos, por ello, ¿la lucha contra la desinformación debería estar liderada por el Estado a cuyo mantenimiento aportan todos los ciudadanos?

Este trabajo sirve de base para futuras investigaciones sobre cómo la desinformación impulsada por la GenAI afecta el proceso democrático en Ecuador. Luego de la posesión de los dignatarios, se podrían analizar los cumplimientos de los planes de gobierno, los informes de rendición de cuentas y el tratamiento de los proyectos de leyes acerca de regular la inteligencia artificial, desde el compromiso de proteger los derechos humanos y enfrentar los desafíos que plantean las tecnologías emergentes. También queda por medir la incidencia de la desinformación en la política ecuatoriana y regional.

Referencias

Ackerman, J., Cabo de la Vega, A., Casado, F., Chávez, G., Hernández, J., Luque, A., Macías, J., Maniglio, F., Mateos, C., Merino, A., Oliván, F., Puglisi, S. & Vera, S. (2020). *La democracia de los trolls y el asalto a la democracia: Análisis del proceso electoral en*

Ecuador y los elementos irradiadores que lo conforman a través de sus redes sociales. Ediciones UTM. <https://acortar.link/Qfj9f3>

Alvim, F., Rubio, R. & Monteiro, V. A. (2024). *Inteligência Artificial e eleições de alto risco. Ciberpatologias e ameaças sistêmicas da nova comunicação política.* Lumen Juris.

Arda, Z. & Başarır, L. (2024). Defending Truth and Democracy in the Age of AI: A Framework for Empowering Voters Against Persuasion and Misinformation with AI Literacy. *adComunica. Revista Científica de Estrategias, Tendencias e Innovación en Comunicación*, 28, 115-142. <http://dx.doi.org/10.6035/adcomunica.8020>

Article 19 (2019). *The Social Media Councils: Consultation Paper. Article 19* <https://acortar.link/rAHyRp>

Asamblea Nacional del Ecuador. (2025). *Consulta de proyectos de ley.* <https://acortar.link/bR4jnN>

Barredo-Ibáñez, D., De-la-Garza-Montemayor, D.-J., Torres-Toukoumidis, A., & López-López, P. (2021). Artificial intelligence, communication, and democracy in Latin America: a review of the cases of Colombia, Ecuador, and Mexico. *Profesional De La información*, 30(6). <https://doi.org/10.3145/epi.2021.nov.16>

Bradshaw, S. & Howard, N. (2017). Troops, trolls and troublemakers: a global inventory of organized social media manipulation. Working paper N° 2017.12. University of Oxford. <https://acortar.link/vVDU43>

Calderón, M. (2019). The use of social media and AI to shape democracy in Ecuador. In *Global Information Society Watch*. Association for Progressive Communications.

- Dad, N., & Khan, S. (2023). Reconstructing elections in a digital world. *South African Journal of International Affairs*, 30(3), 473-496. <https://doi.org/10.1080/10220461.2023.2265886>
- Duberry, J. (2022). *Artificial intelligence and democracy. Risks and promises of AI-Mediated Citizen-Government Relation*. Elgar.
- Echeverría, M. & Rodríguez, C. (2023). ¿La alfabetización digital activa la incredulidad en noticias falsas? Eficacia de las actitudes y estrategias contra la desinformación en México. *Revista De Comunicación*, 22(2), 79-95. <https://doi.org/10.26441/RC22.2-2023-3246>
- Foro sobre Información y Democracia (2024). La IA como bien público: garantizar el control democrático de la IA en el espacio de la información. *Working Group on Artificial Intelligence Forum on Information and Democracy*.
- Giandana, F. (2024). Guía esencial para una ley de inteligencia artificial. *Access Now*.
- Hajdúková, T. (2022). Techniques for Manipulating Public Opinion in the Online Space During an Election Campaign as a Hybrid Threat. *Academic Journal of Interdisciplinary Studies*, 13(1), 14-23 <https://doi.org/10.36941/ajis-2024-0002>
- Jara, M. (4 de marzo de 2025). La IA irrumpe en las elecciones de Ecuador y su regulación avanza lento. *Expreso*. <https://acortar.link/f0Mmx9>
- Karlsen, G. H. (2021). Hvordan kan vi beskytte valg mot fremmed påvirkning?. *Internasjonal Politikk*, 79(1), 90-113. <https://doi.org/10.23865/intpol.v79.2309>

- Londoño, O., & García-Perdomo, V. (2024). Exposición selectiva, confianza en medios y tolerancia a compartir desinformación. *index.Comunicación*, 14(2), 57-81. <https://doi.org/10.62008/ixc/14/02Exposi>
- López, M., Barredo, D., & Sánchez, H. (2024). Usos y riesgos de la Inteligencia Artificial en las campañas electorales 2023: Encuesta Delphi a expertos estratégicos de Colombia. *Revista ICONO 14. Revista científica de Comunicación y Tecnologías emergentes*, 22(1), e2078. <https://doi.org/10.7195/ri14.v22i2.2078>
- Luminate. (2023). *Bots versus ballots: Europeans fear AI threat to elections and lack of control over personal data*. <https://acortar.link/nuly6h>
- Mansell, R., Durach, F., Kettemann, M., Lenoir, T., Procter, R., Tripathi, G., & Tucker, E. (2025). *Information ecosystems and troubled democracy. A Global Synthesis of the State of Knowledge on News Media, AI and Data Governance*. The Forum on Information and Democracy.
- Marsden, C., & Meyer, T. (2019). How can the law regulate removal of fake news? *Computers and Law*. <https://acortar.link/NkcYhA>
- Marsden, C., Meyer, T., & Brown, I. (2020) Platform values and democratic elections: How can the law regulate digital disinformation? *Computer law & security review* 36, 105373.
- Misión de Observación Electoral de la UE. (2025). DECLARACIÓN PRELIMINAR. Unas elecciones transparentes y pacíficas abocan a una segunda vuelta con retos pendientes. Unión Europea.

- Muto, G. (2025). The use of AI in electoral campaigns: key issues and remedies. *MediaLaws*. <https://acortar.link/Cgm93j>
- Naciones Unidas (2024). *¿Puede la Inteligencia Artificial influenciar los procesos electorales?* <https://acortar.link/KT9FyC>
- OEA (2025). *Unión Europea Misión de Observación Electoral. ECUADOR 2025. Informe Final*. Unión Europea.
- Paget, A. (2024). *(Synthetic) Stump Speech: Crafting Generative AI Disclosure Regulations for Political Advertisements*, 93 Fordham L. Rev. 321.
- Palau-Sampio, D. (2024). Fact-checking colaborativo: análisis de seis iniciativas contra la desinformación en campaña electoral en Europa y América Latina. *Estudios sobre el Mensaje Periodístico*, 30(4), 829-842. <https://doi.org/10.5209/emp.97376>
- Posetti, J. & Matthews, A. (2018). *A short guide to the history of 'fake news' and disinformation*. ICFJ.
- Primicias. (6 de febrero de 2025). *Así se usó la inteligencia artificial para desinformar en la campaña electoral de Ecuador*. <https://acortar.link/VrauPz>
- Puyosa, I. (2017). Bots políticos en Twitter en la campaña presidencial #Ecuador2017. *Contratexto*, 27(027), 39-60. <https://doi.org/10.26439/contratexto.2017.027.002>
- Ramos, G. (2022). *Ethics of ai and democracy: UNESCO recommendation's insights*. <https://acortar.link/0XGsAK>

- Romero, F. (2024). Generative AI and deepfakes: a human rights approach to tackling harmful content. *International Review of Law, Computers & Technology*, 38(3), 297-326. <https://doi.org/10.1080/13600869.2024.2324540>
- Rubio, R. (2025). El uso de la inteligencia artificial en las campañas electorales y sus efectos democráticos. *Revista de Derecho Político*, (122), 65-102. <https://doi.org/10.5944/rdp.122.2025.44742>
- Schia, N. & Gjesvik, L. (2020). Hacking democracy: managing influence campaigns and disinformation in the digital age. *Journal of Cyber Policy*, 5(3), 413-428. <https://doi.org/10.1080/23738871.2020.1820060>
- Schipper, T. (2025). Disinformation by design: leveraging solutions to combat misinformation in the Philippines' 2025 election. *Data & Policy*, 7, e39.
- Stockwell, S., Hughes, M., Swatton, P., Zhang, A., Hall J. y Kieran (2024). *AI-Enabled Influence Operations: Safeguarding Future Elections*. Centre for Emerging Technology and Security.
- UNESCO. (2022). Elections in digital times a guide for electoral practitioners. UNESCO.
- UNESCO. (2024). Inteligencia artificial y democracia. UNESCO.
- Unión Europea. (2019). Code of Practice against disinformation: Commission recognises platforms' efforts ahead of the European elections. https://ec.europa.eu/commission/presscorner/detail/en/STATEMENT_19_2570

Usuarios Digitales. (2025). Evidencias, repositorio, piezas originales y galería de publicaciones generadas con Inteligencia Artificial.
<https://acortar.link/TmDOLX>

World Economic Forum. (2024). Global Risks Report 2024. World Economic Forum

ENTRE A NORMATIVIDADE E O CONSEQUENCIALISMO: PERSPECTIVAS CONTEMPORÂNEAS DA ÉTICA EM INTELIGÊNCIA ARTIFICIAL

Guilherme Gonçalves Mello¹
Eli Borges Junior²

A inteligência artificial (IA) consolidou-se como uma das tecnologias centrais do século XXI, não apenas pela capacidade de automatizar processos e otimizar atividades, mas também pelo seu impacto nas decisões que moldam a vida social. Hoje, ferramentas baseadas em aprendizado de máquina são aplicadas em contextos que vão da segurança pública à saúde, do sistema financeiro à educação, até ambientes digitais de interação e consumo. Esse alcance explica por que a ética,

-
1. Mestrando no Programa de Pós-Graduação em Comunicação (PPGCOM) da Universidade Federal de Juiz de Fora (UFJF).
guilhermegonm@gmail.com
 2. Professor permanente do PPGCOM/UFJF. Coordenador do (AI)2Lab – Algorithms, AI & Image/CNPq.
eli.borges@ufjf.br

antes e em certa medida secundária, emergiu como eixo central nos debates sobre tecnologia. Assim, o grande questionamento não mais seria se a IA deve ser orientada por valores éticos, mas quais seriam esses valores, de que modo isso seria realizado e com quais efeitos.

A centralidade da ética na IA é visível na proliferação de *frameworks*, *guidelines* e declarações que buscam estabelecer parâmetros normativos para orientar seu desenvolvimento e sua aplicação. Documentos de referência publicados por organizações internacionais, governos e empresas têm proposto princípios como justiça, responsabilidade e transparência como requisitos essenciais para que a IA seja considerada confiável e socialmente aceitável (Floridi et al., 2018; Floridi & Cowsls, 2019; UNESCO, 2021). Esse movimento busca antecipar os riscos sociais da tecnologia, ao mesmo tempo em que procura enquadrá-la em estruturas de governança global. Entretanto, persiste uma dúvida: em que medida tais princípios são de fato incorporados ao design dos sistemas algorítmicos e às rotinas institucionais e/ou organizacionais? Estudos recentes

A urgência desse debate decorre, em grande medida, do impacto crescente da IA em setores sensíveis, nos quais falhas éticas produzem efeitos imediatos e, por vezes, devastadores. O uso de algoritmos preditivos em sistemas de justiça criminal revelou a reprodução de preconceitos raciais em larga escala, como no caso do COMPAS, que classificava de forma sistematicamente desigual pessoas negras e brancas quanto ao risco de reincidência (Chouldechova, 2016). Em processos seletivos, sistemas de recomendação foram denunciados por reforçar estereótipos de gênero, marginalizando candidatas mulheres em áreas de tecnologia (Gupta et al., 2021). Com a IA generativa, surgiram ainda

dilemas como a disseminação de desinformação, a autoria de conteúdos e a amplificação de vieses cognitivos (Bergman, 2025). Esses exemplos evidenciam que a neutralidade algorítmica é ilusória: sistemas de IA estão imersos em contextos sociais e políticos atravessados por desigualdades de poder (Martin, 2019).

Diante disso, propõe-se duas perspectivas para estruturar o debate atual sobre ética em IA. A primeira, de caráter prospectivo ou normativo, busca antecipar dilemas e projetar valores éticos no design e na governança da tecnologia. No relatório do projeto AI4People, por exemplo, Floridi e colegas destacam que a IA deve se orientar por cinco princípios fundamentais, entre eles a explicabilidade. Esse princípio, segundo os autores, sintetiza os conceitos de “‘inteligibilidade’ (como uma resposta à pergunta ‘como isso funciona?’) quanto ao sentido ético de ‘responsabilização’ (como uma resposta à pergunta: ‘quem é responsável pela maneira como isso funciona?’)” (Floridi et al., 2018, p. 697).

De forma convergente, a *Recommendation on the Ethics of Artificial Intelligence* estabelece como objetivo:

fornece uma base para fazer com que os sistemas de IA funcionem para o bem da humanidade, dos indivíduos, das sociedades, do meio ambiente e dos ecossistemas, bem como para prevenir danos. Ela também visa a estimular o uso pacífico dos sistemas de IA. (UNESCO, 2021, p. 14)

Essa abordagem atribui à ética um papel preventivo e prospectivo, sustentando a expectativa de que é possível guiar o desenvolvimento tecnológico antes que seus efeitos adversos se consolidem.

Por outro lado, uma segunda vertente, de natureza consequentialista, enfatiza a análise crítica dos impactos já observáveis da IA.

Ananny e Crawford (2018, p. 982) alertam que “pedir para ‘olhar dentro da caixa preta’ talvez seja uma exigência muito limitada e, em última análise, uma metáfora inadequada para as complexidades dos sistemas algorítmicos contemporâneos”³, ressaltando que a responsabilização compreende não apenas os aspectos técnicos, mas também os contextos sociais e políticos que moldam os sistemas. Hagendorff (2020, p. 6) também constata que “a eficácia das diretrizes ou códigos de ética é quase nula e que eles não mudam o comportamento dos profissionais da comunidade tecnológica”⁴, denunciando o caráter frequentemente superficial de muitas iniciativas éticas. Assim, essa perspectiva enfatiza que a ética em IA não pode se restringir a princípios abstratos, mas deve enfrentar concretamente os efeitos sociais e políticos já em curso.

Propõe-se, então, uma análise crítica das abordagens contemporâneas sobre ética a partir de uma revisão narrativa de literatura. Dessa forma, busca-se compreender tanto os fundamentos normativos que orientam as propostas de uma IA ética quanto as críticas que problematizam sua efetividade diante de casos concretos. Ao adotar essa chave analítica, pretende-se evidenciar as tensões entre princípios e prática, bem como entre a intenção de regular preventivamente e a necessidade de mitigar danos já instalados.

Portanto, delimita-se um escopo que não pretende esgotar a vasta discussão sobre ética em IA, mas se concentrar na comparação entre duas abordagens: a prospectiva/normativa, que aposta na capacidade de

3. No original: “To ask to ‘look inside the black box’ is perhaps too limited a demand and ultimately an ill-fitting metaphor for the complexities of contemporary algorithmic systems”.

4. No original: “the effectiveness of guidelines or ethical codes is almost zero and that they do not change the behavior of professionals from the tech community”.

projetar valores éticos no design e nas instituições, e a consequencialista/reativa, que denuncia e enfrenta os impactos sociais concretos da tecnologia. A proposta não é hierarquizar tais perspectivas, mas mostrar como ambas são indispensáveis — e, em certa medida, complementares — para compreender a complexidade do debate ético contemporâneo.

Metodologia

A metodologia adotada baseia-se em uma revisão narrativa de literatura (Grant & Booth, 2009; Rumrill & Fitzgerald, 2001; Saunders & Rojon, 2011; Sukhera, 2022), escolhida por sua adequação a campos de investigação em consolidação, nos quais coexistem diferentes correntes teóricas e múltiplas disputas conceituais. Diferentemente das revisões sistemáticas, que buscam exaustividade e replicabilidade estrita dos resultados, a revisão narrativa privilegia a síntese crítica e a articulação de debates dispersos, permitindo evidenciar tendências, lacunas e tensões. Essa característica torna-a especialmente pertinente ao estudo da ética em inteligência artificial (IA), onde a produção acadêmica se multiplica em ritmo acelerado, mas ainda carece de consensos consolidados quanto às formas mais adequadas de operacionalizar valores éticos em práticas técnicas, regulatórias e sociais.

O levantamento bibliográfico foi estruturado em duas frentes complementares. A primeira consistiu em buscas sistemáticas nas bases Scopus e Web of Science, escolhidas por sua abrangência na indexação de periódicos internacionais de alta relevância. Nessa etapa, foram utilizadas combinações de palavras-chave em português e inglês — com destaque para “inteligência artificial”, “ética”, “moral”, “IA ética” e

“*accountability*”⁵ — que permitiram identificar publicações recentes e diretamente relacionadas ao debate. A segunda frente correspondeu à análise retroativa de referências, considerada fundamental para recuperar textos fundacionais e autores que, embora não mais recentes, continuam a estruturar as principais categorias analíticas mobilizadas pelo campo. Esse duplo movimento, do mapeamento atual à recuperação histórica, busca garantir atualização e profundidade teórica ao estudo.

Os critérios de inclusão priorizaram a relevância crítica e conceitual dos trabalhos em relação ao tema central. Foram selecionados textos que abordam a ética em IA a partir de diferentes dimensões: (i) estudos que conceituam a IA como fenômeno sociotécnico; (ii) pesquisas que problematizam dilemas epistemológicos, como a opacidade e a explicabilidade; (iii) propostas de *frameworks* e diretrizes normativas; e (iv) investigações empíricas sobre os impactos da tecnologia em contextos sociais específicos. A intenção não foi alcançar exaustividade quantitativa, mas sim compor um *corpus* representativo da diversidade do debate, organizado em torno de eixos interpretativos comuns.

A análise foi conduzida de forma iterativa e interpretativa. Inicialmente, textos centrais serviram para delimitar os eixos de leitura, que foram refinados à medida que novos trabalhos eram incorporados. Essa estratégia permitiu identificar pontos de convergência e divergência entre diferentes tradições acadêmicas e iniciativas institucionais, bem como mapear o modo como conceitos-chave são empregados e

5. Optou-se por manter o termo *accountability* em inglês, uma vez que não há uma tradução literal em português que abarque plenamente todos os significados que ele expressa. Neste trabalho, *accountability* é compreendido como a responsabilidade pelos resultados, envolvendo a prestação de contas, a transparência, a justificação das decisões e a disposição para assumir as consequências dos próprios atos.

disputados. Mais do que descrever os enunciados presentes na literatura, o objetivo foi construir uma narrativa que evidenciasse a historicidade desses conceitos, as tensões entre formulações normativas e experiências práticas, e o diálogo entre diagnósticos acadêmicos, políticas públicas e diretrizes corporativas.

Por fim, a escolha pela revisão narrativa justifica-se pela necessidade de síntese crítica em um campo marcado pela dispersão e pela sobreposição de propostas. Diante da multiplicidade de *frameworks*, princípios e códigos éticos produzidos em escala global, a simples catalogação de publicações teria efeito limitado para compreender o estado da arte. Mais relevante é identificar as estruturas de raciocínio que sustentam esses documentos, as críticas que problematizam sua efetividade e os caminhos emergentes para o avanço da agenda de pesquisa. Assim, a presente revisão não pretende oferecer respostas definitivas, mas propor um terreno conceitual que permita reconhecer tanto os esforços de prescrição normativa quanto às análises que revelam seus limites práticos.

Mapeando o debate sobre ética em inteligência artificial

A consolidação da inteligência artificial como um dos objetos mais investigados tem provocado não apenas discussões técnicas, mas também análises que a situam em contextos sociais e políticos. Diversos estudos ressaltam que a IA não deve ser reduzida à operação de algoritmos, mas entendida como parte de uma infraestrutura que reconfigura práticas e valores. Van Dijck, por exemplo, descreve a dataficação como processo pelo qual os rastros digitais se convertem em insumo de governança e poder, observando que “os metadados parecem ter se

tornado uma moeda comum para os cidadãos pagarem por seus serviços de comunicação e segurança” (Van Dijck, 2014, p. 197)⁶.

Se, por um lado, é enfatizada a dimensão sociopolítica da IA como infraestrutura de governança e controle, por outro, destacam-se os desafios epistemológicos que emergem da dificuldade de compreender seu funcionamento interno. É nesse contexto que se insere o debate sobre a opacidade algorítmica, um dos conceitos mais influentes do campo. Burrell (2016) propõe uma tipologia que identifica três formas distintas de opacidade:

(1) opacidade como autoproteção e ocultação corporativa ou institucional intencional e, junto com ela, a possibilidade de engano consciente; (2) opacidade decorrente do estado atual das coisas, em que escrever (e ler) código é uma habilidade especializada; e (3) uma opacidade que decorre da incompatibilidade entre a otimização matemática na alta dimensionalidade característica do aprendizado de máquina e as demandas do raciocínio em escala humana e estilos de interpretação semântica (Burrell, 2016, pp. 1-2)⁷.

Essa formulação mostra que a chamada “caixa-preta algorítmica” não decorre apenas de segredos corporativos, mas também da natureza técnica especializada da programação e da distância entre o

6. No original: “metadata appear to have become a regular currency for citizens to pay for their communication services and security”.

7. No original: “Three distinct forms of opacity include: (1) opacity as intentional corporate or institutional self-protection and concealment and, along with it, the possibility for knowing deception; (2) opacity stemming from the current state of affairs where writing (and reading) code is a specialist skill; and (3) an opacity that stems from the mismatch between mathematical optimization in high-dimensionality characteristic of machine learning and the demands of human-scale reasoning and styles of semantic interpretation”.

raciocínio estatístico do aprendizado de máquina e os modos humanos de interpretação.

Diante disso, parte da literatura defende a necessidade de distinguir camadas de explicação: compreender o que o sistema faz, como faz e quais valores orientam sua lógica. A inteligibilidade técnica de um modelo não equivale automaticamente à responsabilização social, pois um algoritmo pode ser descrito matematicamente e, ainda assim, permanecer opaco em relação às escolhas normativas que sustentam sua aplicação. Nesse sentido, destaca-se que a transparência de código ou de cálculos estatísticos não garante, por si só, *accountability*. A crítica de Ananny e Crawford (2018) reforça esse argumento ao afirmar que demandas de transparência descoladas das condições sociais e políticas de produção tendem a produzir soluções formais com baixo poder explicativo.

Esse diagnóstico abre espaço para respostas normativas que buscam não apenas identificar os problemas éticos da IA, mas também propor mecanismos para mitigá-los. Em vez de se limitar a denunciar a “caixa-preta”, trabalhos recentes têm destacado a importância de traduzir valores éticos em parâmetros concretos, capazes de orientar tanto a formulação de políticas públicas quanto o próprio processo de desenvolvimento dos sistemas. A questão desloca-se, portanto, de “quais princípios devem nortear a IA” para “como esses princípios podem ser incorporados desde o design”, reduzindo riscos e vieses em sua origem. Essa orientação aparece, por exemplo, em propostas que aproximam requisitos éticos de requisitos de sistema (especificação, validação e conformidade), deslocando o foco da retórica de princípios para procedimentos auditáveis.

Esse movimento também despertou análises mais críticas. Hagendorff (2020), por exemplo, observa que a maior parte das diretrizes de ética em IA apresenta impacto limitado, funcionando mais como declarações institucionais do que como mecanismos capazes de alterar práticas de desenvolvimento. Nesse sentido, tem-se discutido, assim, o conceito de *ethics washing*, em analogia ao *greenwashing*, para designar a apropriação retórica da linguagem ética como estratégia de legitimação pública. No caso da IA, a evocação de termos como justiça, responsabilidade ou transparência podem funcionar mais como mecanismo de autopreservação institucional do que como compromisso real de mudança.

Essa crítica torna-se evidente em iniciativas específicas. A *Partnership on AI*⁸, fundada em 2016, foi apresentada como fórum multissetorial para promover uma “IA responsável”. No entanto, estudos mostram que a organização carece de mecanismos vinculantes de auditoria ou sanção, funcionando sobretudo como instrumento de reputação corporativa (Bietti, 2020; Hagendorff, 2020; Whittlestone et al., 2019). De forma análoga, os *AI Principles* do Google (2018) foram publicados após a controvérsia envolvendo o contrato *Project Maven* com o Departamento de Defesa dos EUA, revelando a lacuna entre discurso e prática⁹.

Por outro lado, Morley et al. (2019) destacam que, embora haja um reconhecimento crescente dos riscos da IA, a capacidade da

8. A *Partnership on AI* reúne mais de 100 organizações, como Google, Amazon, Meta e Microsoft, mas não estabelece mecanismos de auditoria ou sanção em seus estatutos.

9. O *Project Maven* foi um contrato firmado entre o Google e o Departamento de Defesa dos EUA para aplicação de visão computacional em drones militares, que levou à renúncia de funcionários e protestos internos.

comunidade em desenvolver respostas eficazes ainda permanece em estágio inicial. Esse quadro revela um descompasso entre a rápida produção de princípios normativos e a maturidade dos arranjos técnicos e organizacionais necessários à sua implementação. Para reduzir esse hiato, parte da literatura investiga “onde” e “como” as decisões com implicações éticas são tomadas ao longo do *pipeline* de aprendizado de máquina. Veale et al. (2018) mostram que necessidades de justiça e *accountability* emergem de forma situada — na escolha de dados, na definição de métricas e nos mecanismos de contestação — mas que barreiras institucionais frequentemente inviabilizam soluções genéricas.

A literatura também tem sistematizado taxonomias de “soluções” para incorporar valores em sistemas de IA. Yu et al. (2018) reúnem abordagens em níveis individual, social e de interação humano-IA, destacando que não há um único ponto de intervenção, mas múltiplos acoplamentos entre técnica, instituições e usuários. Ao mapear dilemas como tensões entre segurança e autonomia, os autores apontam que arquiteturas sociotécnicas demandam arranjos híbridos, combinando restrições técnicas com governança e educação dos usuários. Esse quadro reforça que “explicar” um modelo não basta para resolver dilemas como justiça distributiva, privacidade ou abuso de poder, exigindo medidas complementares de prevenção, mitigação e reparação.

Do lado das práticas, Holstein et al. (2019) realizaram entrevistas com profissionais da indústria de aprendizado de máquina, revelando um padrão recorrente: embora esses praticantes reconheçam a relevância de justiça e não discriminação, enfrentam dificuldades concretas para operacionalizar esses valores. Como relataram alguns deles, “suas equipes não possuem atualmente processos em vigor para dar suporte

à coleta e curadoria de conjuntos de dados equilibrados ou representativos” (Holstein et. al, 2019, p. 6)¹⁰. Em outras palavras, quando justiça e explicabilidade não são especificadas como requisitos, acabam subordinadas a metas de desempenho e prazos de entrega.

Esse diagnóstico é consistente com achados sobre vieses em domínios específicos: Gupta, Parra e Dennehy (2021) mostram que sistemas de recomendação podem perpetuar e amplificar preconceitos de raça e gênero, e que valores culturais modulam tanto a percepção quanto a disposição de questionar tais vieses. Com isso, o problema ético revela-se simultaneamente técnico (modelos, dados, métricas) e cultural (normas, valores, expectativas).

Outra vertente do debate envolve as percepções sociais sobre riscos e expectativas em relação à IA. Ferhataj et al. (2025) analisam preocupações éticas entre estudantes nas áreas de tecnologia e robótica, indicando sensibilidades difusas (segurança, privacidade, substituição de trabalho, impacto social) e formas incipientes de literacia ética. Tais resultados sugerem que, além de regulações e ferramentas, há um componente formativo indispensável para consolidar práticas éticas no ecossistema de IA. Huang et al. (2023), por sua vez, oferecem uma visão panorâmica da ética em IA, articulando categorias de risco (segurança, justiça, privacidade, explicabilidade, governança) e mapeando métodos de avaliação; ao integrar riscos técnicos e impactos sociais, reforçam a premissa de que abordagens unidimensionais tendem a falhar.

O estado da arte sobre ética em IA, portanto, revela um campo plural e em disputa. De um lado, encontra-se um esforço de sistematização

10. No original: “their teams do not currently have processes in place to support the collection and curation of balanced or representative datasets”.

normativa, que procura oferecer parâmetros universais para que a IA seja confiável e socialmente aceitável. De outro, emergem vozes críticas que ressaltam a distância entre tais princípios e as práticas concretas, chamando atenção para a persistência de vieses, para as barreiras epistemológicas da opacidade e para o risco de que códigos éticos se tornem meras estratégias de legitimação.

Essa multiplicidade indica que o campo ainda não alcançou consenso sobre o caminho mais efetivo para enfrentar os dilemas éticos da IA. Ao contrário, o que se observa é um processo de construção teórica em curso, no qual conceitos como opacidade, explicabilidade, justiça e responsabilidade adquirem diferentes sentidos conforme o enquadramento adotado. Esse caráter fragmentado, no entanto, não deve ser visto apenas como problema, mas como sinal da vitalidade de um campo em consolidação, que busca responder simultaneamente a exigências normativas, desafios epistemológicos e problemas práticos de implementação.

A dicotomia entre promover o bem e reparar o mal

As discussões contemporâneas sobre ética na inteligência artificial revelam um campo em expansão, marcado por uma diversidade de vozes, agendas e preocupações. A literatura mostra não apenas múltiplos enfoques, mas também a emergência de tensões que permitem estruturar o debate. A partir dessa leitura e análise, delineou-se a proposta interpretativa que orienta este trabalho: compreender o tema da ética da IA a partir de uma dicotomia entre esforços para “promover o bem” e iniciativas voltadas a “reparar o mal”. Essa dualidade, mais do que sugerir uma oposição, revela uma dinâmica dialética que ilumina

tanto os limites quanto às possibilidades de um campo que ainda busca consolidar-se como disciplina prática.

O eixo normativo, voltado para a promoção do bem, busca antecipar dilemas éticos e propor princípios universais como guias para o desenvolvimento e a aplicação da inteligência artificial. Essa perspectiva parte da convicção de que é possível, ao menos em parte, projetar sistemas éticos e evitar danos antes de sua materialização. Um dos documentos mais influentes nesse sentido é o relatório *AI4People* que propõe cinco princípios éticos — beneficência, não maleficência, autonomia, justiça e explicabilidade. Os autores defendem que “a IA deve ser concebida e desenvolvida de forma a reduzir a desigualdade e a promover o empoderamento social, respeitando a autonomia humana e aumentando os benefícios partilhados por todos, de forma equitativa” (Floridi et al., 2018, p. 701)¹¹.

No âmbito intergovernamental, a *Recommendation on the Ethics of Artificial Intelligence* (2021) é o primeiro instrumento global endossado por quase 200 Estados. O documento define ética da IA como uma “reflexão normativa sistemática, com base em um marco holístico, abrangente, multicultural e em evolução de valores, princípios e ações interdependentes que podem orientar as sociedades para que lidem de forma responsável com os impactos conhecidos e desconhecidos das tecnologias” (UNESCO, 2021, p. 10). Esse posicionamento evidencia o esforço em consolidar um consenso global em torno de princípios como dignidade humana, justiça social e diversidade cultural. Além da

11. No original: “AI should be designed and developed in ways that decrease inequality and further social empowerment, with respect for human autonomy, and increase benefits that are shared by all, equitably”.

dimensão declaratória, a recomendação incorpora propostas práticas, como avaliações de impacto ético, instâncias de governança internacional e inclusão de preocupações ambientais. Dessa forma, fortalece-se a normatividade ao oferecer não apenas diretrizes abstratas, mas também mecanismos de implementação — ainda que dependam do engajamento efetivo de Estados e instituições.

A literatura também mostra que a normatividade não se restringe a documentos isolados, mas constitui um movimento global. Huang et al. (2023), ao analisarem 146 diretrizes de diferentes setores, identificam uma recorrência em torno de cinco princípios centrais — transparência, justiça e equidade, responsabilidade, não maleficência e privacidade. Como sintetizam os autores, “há uma convergência óbvia surgindo em torno de cinco importantes princípios éticos: transparência, justiça e equidade, responsabilidade, não maleficência e privacidade” (Huang et al., 2023, p. 808)¹².

Esse esforço de normatização não se limita a declarações abstratas, mas busca orientar práticas de design e desenvolvimento. Yu et al. (2018) defendem que a construção de sistemas éticos exige que as preocupações sejam consideradas desde o início, como parte constitutiva do processo de inovação. Os autores afirmam que “para construir sistemas de IA que se comportem eticamente, o primeiro passo é explorar os dilemas éticos nos cenários de aplicação-alvo” (Yu et al., 2018, p. 5528)¹³. Em sintonia, Taddeo e Floridi (2018) argumentam que

12. No original: “there is an obvious convergence emerging around five important ethical principles: transparency, fairness and justice, responsibility, nonmaleficence, and privacy. The 11 ethical principles identified in the existing AI guidelines are described and explained in the following”.

13. No original: “In order to build AI systems that behave ethically, the first step is to explore the ethical dilemmas in the target application scenarios”.

a ética precisa ser inscrita no próprio design da tecnologia, orientando arquiteturas, funções e processos de decisão automatizados. Para eles, não se trata apenas de criar declarações normativas externas, mas de garantir que valores éticos estejam incorporados à estrutura dos sistemas inteligentes. Essa defesa da ética no design reforça a ideia de que a promoção do bem não pode depender exclusivamente de regulações externas, mas deve estar enraizada na própria inovação tecnológica.

A partir desse conjunto de propostas, torna-se possível compreender a normatividade como um esforço coletivo de articulação e coordenação. Ao propor princípios, *frameworks* e instrumentos de design, esses autores não apenas organizam o campo conceitual, mas também constroem espaços de diálogo entre diferentes atores sociais, políticos e econômicos. Ainda que haja limites em sua aplicabilidade imediata, a normatividade tem se consolidado como base indispensável, capaz de estruturar o debate global e fornecer parâmetros mínimos para pensar os rumos éticos da inteligência artificial.

Mesmo autores críticos da normatividade reconhecem que os princípios podem cumprir papel relevante como ponto de partida. Whittlestone e colegas (2019) argumentam que “os princípios podem fornecer um ponto de partida útil para desenvolver normas e regulamentações mais formais e podem ajudar a identificar questões prioritárias” (Whittlestone e colegas, 2019, p. 195)¹⁴. Assim, a normatividade é compreendida como solução parcial, necessária para organizar um campo em formação e criar um horizonte de expectativas comuns.

14. No original: “principles can provide a useful starting point from which to develop more formal standards and regulation, and can help to identify priority issues on which both research and policy should focus”.

Apesar da força dessa abordagem, a normatividade enfrenta críticas consistentes quanto à sua eficácia prática. Hagendorff (2020) mostra que, mesmo quando princípios éticos são enunciados, eles dificilmente encontram tradução no cotidiano das organizações, pois faltam condições institucionais para que sejam colocados em prática. Para ele,

Engenheiros e desenvolvedores não são sistematicamente educados sobre questões éticas, nem são capacitados, por exemplo, por estruturas organizacionais, para levantar questões éticas. Em contextos empresariais, a velocidade é tudo em muitos casos, e ignorar considerações éticas equivale ao caminho de menor resistência. (Hagendorff, 2020, pp. 6-7)¹⁵

Essa constatação expõe como, sem mecanismos adequados de formação e de suporte institucional, diretrizes permanecem como declarações de princípio, sem se converter em práticas efetivas.

Embora críticas à sua efetividade sejam recorrentes, o principal desafio da normatividade não está em sua formulação, mas em sua aplicação prática. O campo tem acumulado declarações consistentes e amplamente reconhecidas, mas pouco se tem avançado em sua transformação em instrumentos vinculantes. Esse descompasso entre discurso e prática abre espaço para abordagens consequencialistas, voltadas a examinar os impactos concretos da IA já observáveis na sociedade.

A segunda abordagem concentra-se nos problemas éticos decorrentes das consequências da inteligência artificial, priorizando estratégias de mitigação de efeitos indesejados que emergem no uso

15. No original: “Engineers and developers are neither systematically educated about ethical issues, nor are they empowered, for example by organizational structures, to raise ethical concerns. In business contexts, speed is everything in many cases and skipping ethical considerations is equivalent to the path of least resistance”.

dos sistemas. Diferentemente da normatividade, que busca antecipar riscos, o consequencialismo parte do reconhecimento de que prejuízos já se verificam em práticas sociais e políticas, exigindo respostas imediatas. Nesse sentido, Martin (2018) aponta que a grande questão contemporânea não é apenas “como projetar sistemas éticos”, mas como estabelecer mecanismos claros de *accountability* que permitam atribuir responsabilidades quando danos acontecem. A ênfase desloca o debate para a responsabilização, evidenciando a ausência de clareza sobre quem responde por decisões algorítmicas.

Essa preocupação abre caminho para reflexões mais amplas sobre os efeitos sociais da IA, que não se limitam a falhas técnicas, mas reconfiguram dinâmicas estruturais. Van Dijck (2014), ao discutir o fenômeno da dataficação, argumenta que a transformação de ações sociais em dados quantificáveis se consolidou como prática central das plataformas digitais, tornando-se ao mesmo tempo instrumento de vigilância, lógica de consumo e base de novos modelos de negócio. Nesse sentido, o consequencialismo não apenas identifica prejuízos localizados, mas ilumina como os sistemas algorítmicos moldam práticas sociais, políticas e culturais em larga escala.

Se no nível “macro” a dataficação expõe tensões estruturais, no nível “micro” os algoritmos também produzem consequências desiguais para indivíduos e grupos sociais. Chouldechova (2016) demonstrou que, mesmo quando instrumentos de previsão de reincidência são considerados estatisticamente justos, eles podem gerar taxas de erro distintas entre grupos raciais, resultando em impactos punitivos mais severos para aqueles pertencentes ao grupo com maior prevalência de reincidência. Esse diagnóstico revela que os algoritmos não apenas

reproduzem desigualdades históricas, mas podem reforçá-las, exigindo mecanismos institucionais de mitigação.

Esses problemas de justiça algorítmica não se restringem a sistemas judiciais ou institucionais, mas também permeiam ambientes informacionais cotidianos. Gupta et al. (2022) demonstram como sistemas de recomendação tendem a reforçar padrões de viés de confirmação, direcionando usuários para conteúdos que confirmam crenças prévias. Esse processo, segundo os autores, cria ambientes digitais homogêneos, reduzindo a exposição à diversidade e intensificando a polarização social. O consequencialismo, ao analisar tais efeitos, revela como os sistemas de IA participam ativamente da formação de esferas públicas fragmentadas, ampliando riscos para a democracia e para a coesão social.

Além de mapear impactos observáveis, a abordagem consequencialista também se estende à análise de riscos emergentes, buscando antecipar potenciais danos sociais e institucionais. Ferhataj et al. (2025) mostram, a partir de uma investigação com estudantes universitários, que há uma percepção de vulnerabilidades ligadas à IA e à robótica. Como observam os autores, “os riscos de segurança física [...] surgiram como a preocupação mais bem avaliada, sublinhando a importância crítica da implementação de protocolos de segurança e diretrizes operacionais robustos” (Ferhataj et al., 2025, p. 176)¹⁶. Isso mostra que a preocupação ética também envolve preparar mecanismos preventivos contra consequências futuras.

16. No original: “Physical security risks [...] emerged as the highest-rated concern, underscoring the critical importance of implementing robust safety protocols and operational guidelines”.

Dessa forma, fica evidente que o consequencialismo cumpre papel essencial ao lidar com as consequências já produzidas pela IA, ao mesmo tempo em que projeta desafios futuros. Sua ênfase em impactos reais não substitui os esforços normativos, mas os complementa: enquanto princípios fornecem um horizonte de valores compartilhados, a análise consequencialista assegura que esses valores sejam testados e ajustados diante das práticas concretas. Essa complementaridade reforça a centralidade da ética na IA como disciplina prática, voltada não apenas para a prevenção, mas também para a reparação e o acompanhamento crítico das transformações em curso.

Os estudos consequencialistas, ao revelarem os limites da normatividade, não a invalidam, mas a desafiam a tornar-se mais aplicável e vinculante. É nesse ponto que se abre espaço para pensar a integração entre as duas abordagens. Princípios universais são fundamentais para dar coerência, mas precisam ser acompanhados de diagnósticos empíricos e instrumentos de regulação. O contrário também é válido: sem um horizonte normativo, a análise consequencialista corre o risco de se restringir a reações fragmentadas, incapazes de propor direções preventivas.

Assim, a dicotomia entre promover o bem e reparar o mal não deve ser entendida como oposição, mas como articulação necessária. A normatividade fornece um horizonte preventivo e orientador e o consequencialismo os converte em práticas aplicadas. Somente na interação entre essas dimensões será possível consolidar uma ética da inteligência artificial que seja preventiva e responsiva, universal e situada.

Considerações finais

A análise empreendida ao longo deste artigo permitiu mapear o tema da ética da inteligência artificial a partir de uma tensão produtiva entre duas abordagens: a normatividade, orientada à promoção do bem por meio de princípios universais, e o consequencialismo, voltado para a reparação do mal e para a mitigação de impactos já observados na prática social. Essa leitura mostrou que tais abordagens não devem ser vistas como polos contrários, mas como dimensões complementares de um campo em consolidação, que necessita tanto de marcos regulatórios e normativos quanto de diagnósticos situados e empíricos.

No plano normativo, constatou-se a força das declarações de princípios como horizonte organizador. Iniciativas como o relatório AI4People (2018), as diretrizes da UNESCO e a multiplicidade de *frameworks* internacionais revisados por Huang et al. (2023) mostram que a normatividade se firmou como linguagem comum, capaz de oferecer parâmetros mínimos para governos, empresas e instituições acadêmicas. Valores universais como justiça, transparência, *accountability* e não maleficência estruturam o debate e sinalizam a intenção de consolidar um campo global de governança da tecnologia. Essa vertente revelou-se indispensável como primeiro passo, pois fornece vocabulário compartilhado e horizonte de expectativas, embora sua efetividade dependa da tradução em práticas vinculantes.

Por outro lado, a abordagem orientada a impactos evidenciou os limites da normatividade quando essa permanece em nível declaratório. Estudos como os de Chouldechova (2016), Gupta et al. (2022) e Ferhataj et al. (2025) demonstram que a ética não pode ser reduzida a enunciados abstratos: ela - ou a sua condição crítica - se manifesta em

desigualdades estatísticas, em vieses reforçados por algoritmos e em preocupações sociais legítimas. A vertente centrada em consequência amplia a perspectiva, exigindo que a ética da IA enfrente diretamente efeitos políticos, sociais e culturais já em curso, funcionando como instrumento de reparação, regulação, responsabilização e antecipação.

O diálogo entre essas duas dimensões constitui uma das principais contribuições deste artigo para o debate contemporâneo. Ao invés de reforçar a dicotomia entre “promover o bem” e “reparar o mal” como oposição, argumenta-se que a ética da IA só se torna consistente quando ambas operam de maneira articulada. Princípios universais fornecem coerência e direcionalidade, enquanto análises consequentialistas testam tais princípios na realidade concreta e os obrigam a se reformular. A ética aparece, assim, menos como campo estático e mais como disciplina prática, continuamente tensionada entre valores abstratos e contextos situados.

Em síntese, o percurso desenvolvido neste artigo reforça que a ética da inteligência artificial não se sustenta apenas em princípios abstratos ou somente em diagnósticos responsivos. Sua potencialidade está na integração entre promover o bem e reparar o mal, entre universalidade e situacionalidade, entre horizonte normativo e resposta prática. Esse equilíbrio é o que permitirá consolidar a ética da IA como campo de reflexão crítica e intervenção efetiva, à altura dos desafios que a tecnologia impõe às sociedades contemporâneas.

Por fim, cabe reconhecer que esta é apenas uma proposta inicial de leitura e sistematização. A ética da IA é um campo em rápido movimento, no qual novos problemas emergem continuamente, acompanhados por tentativas de resposta que ampliam ou desafiam os quadros aqui

apresentados. Pesquisas futuras poderão tanto aprofundar a relação entre essas dimensões ou propor outros arranjos conceituais que deem conta da pluralidade de práticas e valores envolvidos. O desafio está aberto: construir uma ética da IA que seja simultaneamente preventiva e responsiva, abstrata e concreta, universal e situada, capaz de acompanhar criticamente os rumos de uma tecnologia que se reinventa a cada instante.

Referências

- Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, 20(3), 973-989. <https://doi.org/10.1177/1461444816676645>
- Bergman, R. (2025). IA and confirmation bias. *Mediate.com*. <https://mediate.com/ai-and-confirmation-bias/>
- Burrell, J. (2016). How the machine “thinks”: Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1), 1-12. <https://doi.org/10.1177/2053951715622512>
- Chouldechova, A. (2016). Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *arXiv*. <https://arxiv.org/abs/1610.07524>
- Bietti, E. (2020). *From ethics washing to ethics bashing: A view on tech ethics from within moral philosophy* [Paper]. Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT '20)* (pp. 210-219). Association for Computing Machinery. <https://doi.org/10.1145/3351095.3372860>
- Ferhataj, A., Hajrizi, E., Kotorri, M., Ukaj, F., & Hajrizaj, A. (2025). Ethical concerns in AI development: Analyzing students' perspectives

on robotics and society. *Journal of Information, Communication and Ethics in Society*, 23(2), 165-187. <https://doi.org/10.1108/JICES-08-2024-0111>

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689-707. <https://doi.org/10.1007/s11023-018-9482-5>

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1), 1-17. <https://doi.org/10.1162/99608f92.8cd550d1>

Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, 26(2), 91-108. <https://doi.org/10.1111/j.1471-1842.2009.00848.x>

Gupta, M., Parra, C. M., & Dennehy, D. (2022). Questioning racial and gender bias in AI-based recommendations: Do espoused national cultural values matter? *Information Systems Frontiers*, 24(6), 1465-1481. <https://doi.org/10.1007/s10796-021-10156-2>

Hagendorff, T. (2019). The ethics of AI ethics—An evaluation of guidelines. *arXiv*. <https://doi.org/10.48550/arXiv.1903.03425>

Holstein, K., Vaughan, J. W., Daumé, H., Dudik, M., & Wallach, H. (2019). *Improving fairness in machine learning systems: What do industry practitioners need?* [Paper]. CHI 2019: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. <https://doi.org/10.1145/3290605.3300830>

- Huang, C., Zhang, Z., Mao, B., & Yao, X. (2023). An overview of artificial intelligence ethics. *IEEE Transactions on Artificial Intelligence*, 4(4), 799-817. <https://doi.org/10.1109/TAI.2022.3194503>
- Whittlestone, J., Nyrupe, R., Alexandrova, A., & Cave, S. (2019). The role and limits of principles in AI ethics: Towards a focus on tensions. In Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society (AIES '19) (pp. 195-200).
- Association for Computing Machinery. <https://doi.org/10.1145/3306618.3314289>
- Martin, K. (2019). Ethical implications and accountability of algorithms. *Journal of Business Ethics*, 160(4), 835-850. <https://doi.org/10.1007/s10551-018-3921-3>
- Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2019). From what to how: An overview of AI ethics tools, methods and research to translate principles into practices. *arXiv*. <https://doi.org/10.48550/arXiv.1905.06876>
- Organização das Nações Unidas para a Educação, a Ciência e a Cultura - UNESCO. (2022). Recomendação sobre a Ética da Inteligência Artificial. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000381137>
- Raman, R., Kowalski, R., Achuthan, K., et al. (2025). Navigating artificial general intelligence development: Societal, technological, ethical, and brain-inspired pathways. *Scientific Reports*, 15, 8443. <https://doi.org/10.1038/s41598-025-92190-7>
- Rumrill, P. D., Jr., & Fitzgerald, S. M. (2001). Using narrative literature reviews to build a scientific knowledge base. *Work*, 16(2), 165-170. <https://doi.org/10.3233/WOR-2001-16208>

- Saunders, M. N., & Rojon, C. (2011). On the attributes of a critical literature review. *Coaching: An International Journal of Theory, Research and Practice*, 4(2), 156-162. <https://doi.org/10.1080/17521882.2011.596485>
- Sukhera, J. (2022). Narrative reviews: Flexible, rigorous, and practical. *Journal of Graduate Medical Education*, 14(4), 414-417. <https://doi.org/10.4300/JGME-D-22-00480.1>
- Taddeo, M., & Floridi, L. (2018). How AI can be a force for good. *Science*, 361(6404), 751-752. <https://doi.org/10.1126/science.aat5991>
- Van Dijck, J. (2014). Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. *Surveillance & Society*, 12(2), 197-208. <http://www.surveillance-and-society.org>
- Veale, M., Van Kleek, M., & Binns, R. (2018). *Fairness and accountability design needs for algorithmic support in high-stakes public sector decision-making* [Paper]. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. <https://doi.org/10.1145/3173574.3174014>
- Yu, H., Shen, Z., Miao, C., Leung, C., Lesser, V. R., & Yang, Q. (2018). *Building ethics into artificial intelligence* [Paper]. Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI-18) (pp. 5527-5533). <https://doi.org/10.24963/ijcai.2018/777>

SATIKO E O IMAGINÁRIO FEMININO EM PLATAFORMAS DIGITAIS: ENTRE SUBJETIVIDADE E MERCANTILIZAÇÃO

Laryssa da Costa Gabellini¹
Cláudia Albuquerque Thomé²

O avanço das tecnologias digitais e a consolidação da chamada sociedade em rede transformaram profundamente as formas de sociabilidade, consumo e produção de sentido no mundo contemporâneo. Entre essas transformações, destaca-se o surgimento de influenciadores digitais como mediadores privilegiados das relações entre marcas, públicos e culturas (Abidin, 2016; Van Dijck et al., 2018). Mais recentemente, esse ecossistema foi reconfigurado pela emergência de influenciadores virtuais, personagens digitais que performam subjetividades e estabelecem

-
1. Doutoranda em Comunicação.
Bolsista CAPES. Membro do Grupo de Pesquisa Namidia (PPGCOM/UFJF)
gabelliniacademico@gmail.com
 2. Pós-Doutora em Comunicação.
Professora do Prog. de Pós-Grad. em Com. da Universidade Federal de Juiz de Fora.
claudia.thome@ufjf.br

vínculos com audiências a partir de narrativas cuidadosamente roteirizadas, mas que se apresentam como espontâneas e autênticas.

No contexto brasileiro, esse fenômeno tem se expandido de maneira significativa, articulando tecnologia, marketing e imaginários sociais. Personagens como Satiko, avatar vinculado à apresentadora Sabrina Sato, exemplificam como empresas e personalidades têm investido na criação de figuras digitais capazes de condensar discursos sobre identidade, representatividade e engajamento. Lançada em 2021, representa uma espécie de “versão aumentada” de Sabrina Sato, com presença ativa nas redes sociais e mobilizando uma estética jovial, conectada e cosmopolita. Sua narrativa enfatiza liberdade, diversidade e autenticidade, operando como um dispositivo simbólico de modelagem de um determinado perfil de feminino nas plataformas.

Diante desse cenário, este trabalho parte da hipótese de que Satiko encarna um paradoxo constitutivo do ambiente digital: embora sua performance mobilize signos de empoderamento e ruptura, sua construção está fortemente vinculada à estratégias de mercantilização e estetização da subjetividade. Nesse sentido, torna-se central compreender como se organizam os discursos e práticas que estruturam sua presença digital, bem como os modos pelos quais esses discursos reforçam ou tensionam padrões normativos de feminilidade.

A discussão aqui proposta se ancora na noção de imaginário, entendida, a partir de Silva (2003) e Castoriadis (1982), como um sistema simbólico que organiza a experiência social, moldando percepções, valores e modos de ser. O imaginário midiático, nesse sentido, não apenas reflete realidades, mas produz e hierarquiza sentidos, operando como campo de disputa entre diferentes projetos de representação. Como

afirma Silva (2003, p. 81), “a mídia não reflete o real: ela o modela, o organiza, o hierarquiza, o estetiza, o dramatiza”.

A escolha por analisar Satiko se justifica, portanto, pela potência dessa personagem em condensar os tensionamentos entre subjetividade e mercantilização nas plataformas digitais. Diferentemente de influenciadores humanos, Satiko é inteiramente roteirizada, o que permite identificar de modo mais explícito as estratégias corporativas e de mercado que atravessam sua performance. Contudo, seu discurso é apresentado como autêntico, aproximando-se das audiências por meio de recursos de interação, narrativas de proximidade e práticas que remetem ao que Marwick e Boyd (2011) chamam de autenticidade performativa.

Do ponto de vista metodológico, a pesquisa adota a análise de conteúdo (Bardin, 2016) de publicações do perfil @iamsatiko_, buscando identificar as categorias de sentido que estruturam sua narrativa, bem como os valores associados à sua performance. A investigação se orienta por três eixos principais: (1) a estetização da subjetividade, entendida como transformação do “modo de ser” em ativo comunicacional (Sibilia, 2008; Han, 2012); (2) a construção da autenticidade algorítmica (Abidin, 2023), em que a espontaneidade é programada para maximizar engajamento; e (3) o papel da personagem como narradora dialógica (Reis & Thomé, 2022), capaz de conjugar autoridade e interação em sua relação com o público.

Assim, o objetivo deste artigo é analisar como Satiko, enquanto influenciadora virtual, mobiliza e tensiona imaginários femininos nas plataformas digitais, revelando as dinâmicas entre empoderamento, estetização e mercantilização da subjetividade. Trata-se de um recorte de uma pesquisa de doutorado em andamento, que busca contribuir para

a compreensão das mediações tecnológicas e discursivas que estruturam as experiências dos femininos em tempos de plataformação.

Imaginário e comunicação

O conceito de imaginário tem sido amplamente discutido nas ciências sociais e na comunicação, revelando-se central para compreender como sociedades organizam sentidos, valores e representações. Longe de ser apenas um espaço de fantasia ou de ilusões, o imaginário é entendido como instância estruturante da vida social, articulando símbolos, narrativas e práticas que orientam a percepção do mundo e a constituição das identidades.

Cornelius Castoriadis (1982) é um dos principais autores a sistematizar essa noção, concebendo o imaginário social instituinte como a capacidade criadora de sentidos que atravessa todas as dimensões da vida coletiva. Para ele, o imaginário não é mero reflexo das condições materiais, mas um campo ativo de invenção e organização da realidade: as sociedades criam e recriam seus próprios mundos por meio de significações imaginárias centrais, que estruturam tanto instituições quanto modos de subjetivação.

Na comunicação, esse debate ganha contornos específicos ao se considerar a centralidade da mídia na produção e circulação de representações. Juremir Machado da Silva (2003) destaca que os meios não apenas refletem o real, mas o modelam, hierarquizam e dramatizam, definindo o campo do pensável e legitimando certas narrativas em detrimento de outras. Ao operar sobre a sensibilidade e a percepção, o imaginário midiático organiza modos de ver, sentir e desejar, estabelecendo referências para a vida social contemporânea.

Outros autores, como Roland Barthes (1990), também contribuem para essa perspectiva ao analisar como mitologias do cotidiano, produzidas pela mídia, naturalizam valores e comportamentos, transformando-os em “evidências culturais”. Nesse sentido, o imaginário não se apresenta como algo estático, mas como campo dinâmico de disputas simbólicas, em que diferentes forças competem para fixar sentidos hegemônicos.

Com a expansão da internet e das plataformas digitais, o imaginário passou a ser cada vez mais mediado por algoritmos e por lógicas de visibilidade próprias desse ecossistema (Van Dijck et al., 2018). A plataformização da comunicação reorganiza não apenas a circulação de conteúdos, mas também as condições de produção do imaginário, já que os critérios de relevância, engajamento e viralização condicionam quais narrativas ganham destaque e quais permanecem marginais.

Nesse contexto, o imaginário feminino ocupa lugar estratégico. Historicamente, representações sobre o feminino foram construídas pela mídia a partir de padrões normativos, associados a estereótipos de beleza, comportamento e papéis sociais (Sibilia, 2008; Wolf, 1992), configurando ainda um feminino único, sem trazer destaque para questões de interseccionalidades. As plataformas digitais, embora abertas à multiplicidade de vozes e narrativas, frequentemente reproduzem e reforçam tais normatividades, mesmo quando se apresentam como espaços de diversidade.

Assim, compreender o papel de personagens como Satiko implica analisar como o imaginário do feminino é reconfigurado nas plataformas e se há ou não de forma mais evidenciada uma pluralidade de femininos. Por um lado, esses avatares prometem liberdade, fluidez identitária e

autenticidade; por outro, estão profundamente inseridos em estratégias de consumo e mercantilização da subjetividade. O imaginário, nesse sentido, aparece como território de tensão, onde se negociam sentidos entre empoderamento e objetificação, ruptura e manutenção de padrões hegemônicos.

Portanto, a análise dos imaginários e da comunicação permite reconhecer que Satiko não é apenas um produto estético ou tecnológico, mas uma figura que condensa disputas simbólicas contemporâneas, revelando como a cultura digital reorganiza os modos de pensar e viver um determinado tipo de feminino.

Subjetividade e performatividade nas plataformas

O debate sobre subjetividade na contemporaneidade ganha novos contornos com a centralidade das plataformas digitais. Se antes os meios de comunicação de massa organizavam de forma mais verticalizada os discursos e representações, hoje vivemos um regime de visibilidade em que os sujeitos são convocados a performar constantemente suas identidades diante de audiências múltiplas e algoritmicamente mediadas. Essa mudança implica não apenas novas formas de sociabilidade, mas também novas pressões sobre os modos de ser e de estar no mundo.

Segundo Paula Sibilia (2008), a cultura digital potencializa o que ela denomina de “show do eu”, em que a intimidade se converte em espetáculo. Nesse processo, a subjetividade passa a ser exposta, exibida e consumida como narrativa de si, transformando a vida cotidiana em matéria-prima para visibilidade e engajamento. A lógica da transparência, própria do ambiente digital, incentiva os indivíduos a se apresentarem de forma constante, performando autenticidade e

proximidade como recursos estratégicos para legitimar sua presença no espaço público online.

Byung-Chul Han (2012), ao discutir a “sociedade do cansaço”, argumenta que esse regime de exposição permanente produz um sujeito que internaliza a pressão pelo desempenho. A vida passa a ser organizada em torno de métricas de produtividade, visibilidade e reconhecimento, resultando em um modelo de subjetividade autorreferenciada e orientada pelo imperativo da performance. No contexto das plataformas, isso se traduz na busca incessante por likes, seguidores e engajamento, que funcionam como validações simbólicas da existência digital.

Esse cenário dialoga com a teoria da performatividade de gênero de Judith Butler (1990/2019), segundo a qual identidades não são essências fixas, mas efeitos reiterados de práticas discursivas e corporais. Nas plataformas digitais, as performances de gênero ganham visibilidade ampliada e passam a ser constantemente moduladas por algoritmos e pelas expectativas das audiências. Assim, aquilo que se apresenta como “espontâneo” é, em grande medida, resultado de um conjunto de repetições reguladas, que naturalizam determinados padrões de feminilidade e marginalizam outros.

A noção de performatividade, aplicada ao ambiente digital, permite compreender como influenciadores - e, de modo ainda mais explícito, influenciadores virtuais - constroem narrativas de si que oscilam entre autenticidade e mercantilização. Jenkins (2006), ao discutir a cultura da convergência, já destacava que as fronteiras entre produtores e consumidores se tornaram porosas, instaurando uma lógica participativa em que todos são convocados a atuar. Nesse ambiente, a subjetividade se torna um recurso comunicacional, uma “marca pessoal” (personal

brand) estrategicamente projetada para atender às demandas tanto das plataformas quanto dos públicos.

No caso de avatares como Satiko, essa lógica se intensifica: sua subjetividade não é apenas performada, mas literalmente programada. Cada gesto, fala ou aparência é roteirizado por equipes de comunicação, design e marketing, de modo a construir uma persona digital que encarne valores aspiracionais - liberdade, diversidade, jovialidade - ao mesmo tempo em que reforça padrões normativos de feminilidade e consumo. Satiko evidencia, assim, a radicalização do processo descrito por Sibilia e Han: o sujeito performativo torna-se produto acabado, resultado de estratégias corporativas que estetizam a existência e instrumentalizam a subjetividade como ativo mercadológico.

Dessa forma, compreender a subjetividade e a performatividade nas plataformas exige reconhecer o entrelaçamento entre lógica algorítmica, economia da atenção e dinâmicas culturais. Satiko, ao performar um “modo de ser” aspiracional, encarna as contradições desse processo: é simultaneamente figura de empoderamento e dispositivo de mercantilização, narrativa de liberdade e encenação roteirizada, promessa de diversidade e reprodução de normatividades.

Influenciadores virtuais e autenticidade algorítmica

O fenômeno dos influenciadores digitais consolidou-se na última década como um dos principais eixos de reorganização da comunicação de marcas e da cultura de consumo. Sua relevância reside na capacidade de mediar relações de confiança, proximidade e identificação entre empresas e públicos, em um contexto marcado pela saturação publicitária e pela valorização da experiência pessoal como fonte de legitimidade

(Abidin, 2016). Nesse cenário, a figura do influenciador emerge como “autoridade afetiva” (Khamis et al., 2017), capaz de converter narrativas de si em capital simbólico e mercadológico.

Com o avanço das tecnologias de animação, inteligência artificial e design 3D, esse processo ganha uma nova camada: o surgimento dos influenciadores virtuais. Assim como os influenciadores humanos, esses personagens são inteiramente roteirizados e visualmente construídos por equipes de profissionais, mas se apresentam como sujeitos autônomos, dotados de personalidade, estilo e narrativa próprios. Exemplos internacionais como Lil Miquela e brasileiros como Lu do Magalu e Satiko demonstram a potência desse modelo, que combina inovação tecnológica com estratégias de branding.

Um dos elementos centrais para compreender a eficácia comunicativa desses avatares é a noção de autenticidade algorítmica. Trata-se da construção de uma “verdade de si” calibrada para agradar aos critérios das plataformas digitais, equilibrando espontaneidade e planejamento estratégico. Marwick e Boyd (2011) já haviam observado que a autenticidade, nas redes sociais, não é algo dado, mas performado: os sujeitos elaboram cuidadosamente aquilo que aparenta ser genuíno, de modo a manter relevância e engajamento junto às audiências.

Crystal Abidin (2023) amplia essa discussão ao demonstrar como a autenticidade se tornou um recurso estratégico nas mídias digitais. Em sua análise, elementos de intimidade e afetividade são produzidos de forma planejada, visando atender às expectativas de engajamento e monetização. Ou seja, a autenticidade não é um traço espontâneo da subjetividade, mas um dispositivo performativo que articula capital social e capital econômico.

No caso dos influenciadores virtuais, a autenticidade algorítmica se torna ainda mais evidente. Por serem personagens fictícios, sua credibilidade depende da capacidade de simular proximidade, emoção e interação, atributos tradicionalmente associados à experiência humana. Satiko, por exemplo, adota um discurso de liberdade e autenticidade, apresentando-se como “CEO livre” e “sem idade definida”. Contudo, esses atributos são cuidadosamente roteirizados por sua equipe, que mobiliza códigos visuais e narrativos reconhecíveis - diversidade, fluidez, jovialidade - para gerar identificação e engajamento.

Essa construção revela um paradoxo central: ao mesmo tempo em que são programados e estratégicos, os discursos de autenticidade de personagens como Satiko são recebidos e apropriados pelos públicos como expressões de verdade. O engajamento gerado não decorre da “realidade” da personagem, mas da eficácia de sua performance de autenticidade dentro das lógicas algorítmicas das plataformas.

Dessa forma, a autenticidade algorítmica pode ser compreendida como uma forma de performatividade programada, na qual afetos e interações são coreografados para se adequarem aos imperativos de visibilidade, viralização e capitalização. Satiko encarna essa dinâmica, demonstrando como a autenticidade, no ambiente digital, não é antítese do mercado, mas recurso fundamental para sua legitimação.

Narrador dialógico e interações digitais

A lógica comunicacional das plataformas digitais não se sustenta apenas na circulação de conteúdos, mas sobretudo na construção de vínculos simbólicos e afetivos entre sujeitos, marcas e públicos. Nessa perspectiva, a figura do narrador adquire centralidade, pois é

por meio dele que se organizam os discursos, se constroem sentidos e se convidam os interlocutores à participação.

Reis e Thomé (2022) propõem o conceito de narrador dialógico para compreender o enunciador que, nas mídias digitais, combina autoridade narrativa com uma postura de escuta e de abertura ao outro. Diferentemente do narrador tradicional, que assume uma posição por vezes unidirecional, na lógica do emissor, o narrador dialógico constrói-se no jogo da interação, reconhecendo a alteridade como constitutiva de sua narrativa. Não se trata apenas de contar uma história, mas de estabelecer um espaço de troca, em que a audiência se torna parte ainda mais ativa da construção de sentidos.

Esse modelo dialoga diretamente com a lógica participativa das plataformas, onde os públicos não apenas consomem conteúdos, mas também os comentam, remixam, compartilham e atribuem novos significados. Jenkins (2006), ao discutir a cultura da convergência, já havia destacado que o engajamento das audiências contemporâneas se dá pela coparticipação, pelo desejo de colaborar e de intervir nas narrativas midiáticas. O narrador dialógico, nesse sentido, é aquele que potencializa essa copresença, legitimando o público como coautor da experiência comunicativa.

Aplicado ao caso de influenciadores virtuais, esse conceito permite compreender como personagens como Satiko se inserem nas dinâmicas interativas das plataformas. Embora sua performance seja inteiramente roteirizada, sua narrativa é construída de forma a simular proximidade e abertura. Postagens em primeira pessoa, enunciados que convocam a audiência (“e vocês, o que acham?”), respostas a comentários e a incorporação de pautas sociais amplamente debatidas

exemplificam estratégias de interpelação que reforçam a dimensão dialógica de sua atuação.

Ao se apresentar como uma personagem acessível e engajada em conversas cotidianas, Satiko performa aquilo que Reis e Thomé (2022) identificam como um narrador que não apenas transmite, mas escuta, pelo menos em uma proposta mesmo que simulada de diálogo. Essa escuta, no entanto, é mediada por filtros corporativos e estratégias de comunicação de marca, o que produz uma tensão: de um lado, a promessa de diálogo genuíno; de outro, a instrumentalização desse diálogo para fins mercadológicos.

Dessa forma, Satiko pode ser compreendida como um narrador dialógico simulado: sua interação com o público dá a aparência de espontaneidade e reciprocidade, mas é resultado de roteiros cuidadosamente planejados para maximizar engajamento e conversão simbólica em valor de marca. Ainda assim, isso não diminui sua eficácia comunicacional. Pelo contrário, demonstra como o narrador dialógico, mesmo quando roteirizado, é capaz de mobilizar afetos, construir comunidades de sentido e reforçar o imaginário de um feminino digital buscando espaços de proximidade, liberdade e reconhecimento, mas criando um modelo de sucesso que performa estilos de vida bem demarcados.

O conceito, portanto, ajuda a iluminar os paradoxos da atuação de Satiko: ao mesmo tempo em que se constitui como personagem corporativa, sua eficácia narrativa depende da simulação da escuta e da abertura ao público. É nessa tensão entre diálogo e mercantilização que se situam as disputas por imaginários femininos nas plataformas digitais.

Metodologia de análise

A pesquisa parte de uma abordagem qualitativa, voltada para compreender os sentidos e tensionamentos produzidos na performance de Satiko como influenciadora virtual. Por se tratar de um fenômeno emergente no campo da comunicação digital, a escolha metodológica privilegia a análise interpretativa, capaz de apreender os discursos, valores e estratégias que atravessam sua construção narrativa e imagética.

O corpus é constituído pelas publicações do perfil oficial de Satiko no Instagram, identificado como @iamsatiko_, desde sua criação em 2021 até o ano de 2024. Esse recorte temporal permite observar tanto a fase inicial de lançamento da personagem quanto sua consolidação como influenciadora virtual vinculada à imagem de Sabrina Sato. Foram selecionadas publicações que mobilizam narrativas de identidade, representatividade e engajamento, especialmente aquelas que dialogam com temas relacionados ao feminino.

O Instagram foi escolhido por ser a plataforma de maior visibilidade para Satiko e por concentrar a maior parte das interações com o público, oferecendo, portanto, um espaço privilegiado para a observação de sua performance e de sua recepção. Além disso, a centralidade dessa rede social no ecossistema de influenciadores digitais reforça a pertinência da escolha.

O método adotado é a análise de conteúdo, conforme sistematizada por Laurence Bardin (2016), que permite examinar o material empírico em busca de categorias de sentido. A análise de conteúdo é especialmente útil em pesquisas de comunicação por possibilitar a identificação de padrões discursivos e simbólicos em mensagens midiáticas, sem perder de vista sua complexidade e seus contextos de produção.

O procedimento de análise seguiu três etapas principais. São elas, pré-análise: organização do corpus, leitura flutuante e identificação de postagens relevantes para o objeto de estudo. Exploração do material: categorização das publicações a partir de eixos temáticos relacionados ao imaginário feminino, à performatividade e à mercantilização da subjetividade. Tratamento e interpretação: articulação dos resultados com o referencial teórico, de modo a compreender como Satiko mobiliza discursos de autenticidade, liberdade e empoderamento em diálogo com estratégias mercadológicas.

A análise se orienta por três categorias principais, extraídas do diálogo entre teoria e empiria. Estetização da subjetividade: corresponde à transformação do “modo de ser” em ativo comunicacional, conforme discutido por Sibilia (2008) e Han (2012). Busca identificar como Satiko performa subjetividade a partir de elementos visuais, narrativos e estilísticos que reforçam determinados padrões de feminilidade. Autenticidade algorítmica: refere-se à simulação de proximidade e espontaneidade calibrada para agradar às lógicas das plataformas digitais (Abidin, 2023; Marwick & Boyd, 2011). Observa-se de que maneira Satiko performa a autenticidade como recurso estratégico de engajamento. Narrador dialógico simulado: categoria inspirada em Reis e Thomé (2022), que analisa a presença de interpelações, convites ao diálogo e estratégias de interação. Procura-se compreender como Satiko encena a escuta e a reciprocidade, mesmo dentro de roteiros corporativos.

Análise da Personagem Satiko: Estética e narrativa de si

A análise da estética e da narrativa de Satiko revela como a personagem é construída como dispositivo simbólico e comunicacional,

articulando elementos visuais, discursivos e performativos para consolidar sua presença nas plataformas digitais. Criada em 2021 como um avatar vinculado à imagem pública de Sabrina Sato, Satiko se apresenta como uma “versão aumentada” da celebridade, projetada para atuar em territórios digitais com maior fluidez, jovialidade e alcance global. Essa duplicidade - entre celebridade e avatar - confere-lhe uma característica singular: ela existe como derivação de uma figura já consolidada, mas se autonomiza ao performar subjetividades próprias, configurando-se como um “outro eu” da apresentadora.

Do ponto de vista estético, Satiko é marcada por um design visual que combina traços de leveza e modernidade, com estética que se aproxima do futurismo e da diversidade. Sua aparência flutua entre referências cosmopolitas e códigos de moda, explorando a plasticidade do corpo digital como recurso de diferenciação. Essa plasticidade permite-lhe assumir múltiplas formas, estilos e cenários, reforçando a ideia de liberdade identitária e de uma existência “sem idade” ou “sem fronteiras”. Essa ausência de fixidez dialoga com o imaginário contemporâneo de fluidez, em que as identidades são entendidas como múltiplas e em constante transformação (Bauman, 2001).

Narrativamente, Satiko constrói-se como personagem aspiracional, posicionando-se como “CEO livre” e figura globalizada, conectada com causas sociais e tendências culturais. A escolha por adotar um discurso em primeira pessoa - “eu sou Satiko” - é fundamental para criar a ilusão de agência e autenticidade, aproximando a personagem de um modelo humano de enunciação. Ao mesmo tempo, a narrativa de si está profundamente roteirizada, funcionando como um roteiro de marketing

e branding que encena valores caros ao mercado contemporâneo, como diversidade, inovação e empoderamento.

Nesse sentido, a narrativa de Satiko ecoa o que Paula Sibilia (2008) denomina de “espetáculo da intimidade”: mesmo sendo um avatar, ela compartilha fragmentos de sua rotina, opiniões e interações como se fosse uma pessoa real, transformando sua própria existência em espetáculo consumível. Essa encenação de intimidade, no entanto, é planejada e estratégica, configurando-se como produto comunicacional a serviço de objetivos corporativos.

Outro aspecto relevante é a estetização da subjetividade, conceito discutido por Han (2012) e retomado por diversos autores da comunicação. Em Satiko, o “modo de ser” é transformado em ativo comunicacional, de modo que sua personalidade, estilo de vida e escolhas narrativas se convertem em recursos de engajamento. A subjetividade, nesse caso, não é apenas encenada, mas mercantilizada, operando como ativo de marca que associa liberdade e autenticidade à imagem de Sabrina Sato e às marcas que patrocinam a personagem.

Por fim, observa-se que Satiko encarna uma narrativa de si paradoxal: ao mesmo tempo em que afirma autonomia e autenticidade, sua existência é inteiramente dependente de roteiros e algoritmos. A personagem representa, portanto, um exemplo extremo daquilo que Marwick e Boyd (2011) chamam de autenticidade performativa: um processo em que a “verdade de si” é cuidadosamente construída para parecer espontânea, mas que responde a interesses mercadológicos e às lógicas de visibilidade próprias das plataformas digitais.

Assim, a estética e a narrativa de Satiko funcionam como laboratório de experimentação para o feminino digital, condensando os ideais

de liberdade, diversidade e inovação, mas inserindo-os em um circuito de mercantilização e estetização.

Liberdade, diversidade e padrões normativos

Um dos eixos centrais da narrativa de Satiko é a mobilização dos valores de liberdade e diversidade como atributos identitários e diferenciais simbólicos. A personagem se insere em um horizonte discursivo que associa sua existência digital à possibilidade de transcender limites biológicos, sociais e culturais. Essa estética da liberdade busca aproximar Satiko de ideais contemporâneos de fluidez identitária, pluralidade de estilos e autonomia subjetiva, valores que ressoam especialmente nas gerações mais jovens, habituadas a narrativas digitais marcadas pela diversidade (Bauman, 2001).

Além disso, a personagem mobiliza constantemente discursos de diversidade cultural e estética, assumindo diferentes aparências, estilos de vestuário e cenários, o que a posiciona como uma figura global e cosmopolita. Essa plasticidade conecta-se ao imaginário digital contemporâneo, em que as identidades podem ser “curadas” e constantemente atualizadas para dialogar com diferentes públicos e contextos (Couldry & Hepp, 2017). Nesse sentido, Satiko se coloca como símbolo de inclusão e abertura, ao menos no nível discursivo, ao abraçar pautas como liberdade de expressão, pluralidade de vozes e diversidade de estilos.

No entanto, ao mesmo tempo em que afirma liberdade e diversidade, a performance de Satiko permanece atravessada por padrões normativos de feminilidade e consumo. A personagem reproduz traços estéticos hegemônicos associados ao ideal de beleza midiático: corpo esguio, pele impecável, feições simétricas e códigos de moda

internacionalizados. Esses elementos evidenciam que, embora encene pluralidade, Satiko continua ancorada em representações que reforçam expectativas tradicionais de feminilidade, muitas vezes relacionadas a ideais de juventude, sensualidade e sofisticação.

Essa contradição evidencia um aspecto recorrente da comunicação digital contemporânea: a apropriação de discursos de diversidade e empoderamento para fins mercadológicos. Como argumenta Naomi Wolf (1992), a cultura midiática frequentemente mobiliza a ideia de liberdade feminina de maneira instrumentalizada, reforçando ao mesmo tempo padrões estéticos que restringem as possibilidades de representação. No caso de Satiko, a promessa de uma identidade livre e fluida convive com a reprodução de estereótipos, resultando em uma narrativa que celebra a diversidade de forma controlada e estrategicamente roteirizada.

Do ponto de vista comunicacional, essa duplicidade cumpre uma função importante: ao incorporar discursos progressistas de liberdade e inclusão, Satiko cria identificação com públicos engajados em pautas de diversidade. Porém, ao manter padrões normativos, assegura compatibilidade com os interesses do mercado, que se sustenta na manutenção de estéticas hegemônicas de consumo. Assim, sua performance ilustra o que Angela McRobbie (2009) denomina de “complexo de empoderamento pós-feminista”, em que elementos de emancipação são apropriados e reconfigurados pelo mercado, transformando-se em estilo de vida consumível.

Em síntese, Satiko opera como mediadora de um imaginário paradoxal: ao mesmo tempo em que encena liberdade e diversidade, mantém e reforça códigos normativos de feminilidade e consumo. A promessa de ruptura convive com a manutenção de padrões, revelando

como a mercantilização da subjetividade nas plataformas digitais se sustenta em um equilíbrio entre inovação discursiva e preservação de valores hegemônicos.

Estratégias de empoderamento e mercantilização

Um dos aspectos mais evidentes na performance de Satiko é a incorporação do discurso do empoderamento feminino como marca distintiva de sua narrativa. Ao se apresentar como uma “CEO livre”, cosmopolita e sem idade definida, a personagem encena a figura de uma mulher autônoma, inovadora e capaz de transitar em diferentes espaços culturais e sociais. Esse discurso ecoa reivindicações históricas dos movimentos feministas em torno da liberdade e da autodeterminação, mas é reconfigurado no ambiente das plataformas digitais como recurso simbólico a serviço de estratégias de engajamento e consumo.

A noção de empoderamento, nesse contexto, é performada por meio de códigos visuais e narrativos que reforçam a ideia de independência, criatividade e ousadia. Fotografias e vídeos em que Satiko aparece em cenários futuristas, desfilando tendências de moda ou associada a campanhas de diversidade, constroem uma imagem aspiracional de mulher que rompe barreiras e desafia limites. Um modelo de vida de celebridade. Nada interseccional nem amparado na vida cotidiana de tantas mulheres. Contudo, como apontam Gill (2007) e McRobbie (2009), esse tipo de narrativa frequentemente se enquadra no que se pode chamar de empoderamento pós-feminista: uma retórica que mobiliza signos de emancipação feminina, mas que está profundamente inserida na lógica neoliberal de individualização e mercantilização.

Satiko, nesse sentido, não apenas representa um sujeito empoderado, mas transforma o próprio empoderamento em mercadoria e em modelo de vida. Sua autonomia é roteirizada, seu discurso de diversidade é estrategicamente planejado e sua estética de liberdade está vinculada à construção de valor de marca. Essa dinâmica se conecta ao que Sarah Banet-Weiser (2018) denomina de “empoderamento popularizado” (popular feminism), em que as pautas feministas são apropriadas pelo mercado e transformadas em produtos midiáticos, acessíveis e consumíveis, muitas vezes esvaziando sua potência política original.

Na atuação de Satiko, esse processo se manifesta de diferentes maneiras. Em primeiro lugar, o empoderamento é articulado a um discurso de estilo de vida, no qual liberdade, diversidade e autenticidade aparecem como atributos desejáveis, mas sempre integrados à lógica de consumo. Em segundo lugar, a personagem funciona como uma brand persona, mediando a relação entre empresas, produtos e públicos, de modo que sua subjetividade encenada se converte em ativo mercadológico. Por fim, a estética do empoderamento é apresentada de forma aspiracional, convidando as audiências a se identificarem e a projetarem em Satiko seus próprios desejos de liberdade e reconhecimento.

Esse processo revela que o empoderamento performedo por Satiko não pode ser compreendido apenas como expressão de subjetividade, mas como estratégia de mercantilização. A personagem traduz demandas sociais legítimas em narrativas midiáticas esteticamente polidas e facilmente incorporáveis pelas marcas. Assim, o empoderamento deixa de ser apenas um movimento político e passa a ser um estilo, um signo comunicacional que agrega valor simbólico ao mercado.

Dessa forma, Satiko exemplifica como as plataformas digitais se tornam espaços de negociação entre discursos de emancipação e interesses mercadológicos. Seu discurso de empoderamento é potente na construção de identificação e engajamento, mas é simultaneamente limitado por sua função como dispositivo corporativo. O resultado é a consolidação de uma subjetividade roteirizada, em que a promessa de autonomia convive com a instrumentalização da mercantilização, revelando as contradições centrais do imaginário feminino digital.

A performance da autenticidade algorítmica

No ambiente das plataformas digitais, a autenticidade tornou-se um valor central para a legitimação das identidades públicas. O público, saturado por discursos publicitários tradicionais, tende a valorizar aquilo que aparenta ser espontâneo, próximo e “real”. Contudo, como já apontaram Marwick e Boyd (2011), a autenticidade nas redes sociais não é um dado natural, mas uma performance calibrada, resultado de escolhas estratégicas orientadas para a manutenção de relevância e engajamento.

No caso dos influenciadores virtuais, esse desafio é ainda mais evidente. Por serem personagens ficcionais, criados e programados por equipes de comunicação e design, sua eficácia depende justamente da capacidade de simular humanidade e proximidade. Satiko exemplifica esse processo: embora seja inteiramente roteirizada, sua presença digital é construída a partir de códigos narrativos que evocam autenticidade, como postagens em primeira pessoa, interações com seguidores e enunciados que sugerem vulnerabilidade e liberdade de expressão.

Crystal Abidin (2023) denomina esse processo de autenticidade algorítmica, compreendida como uma performance cuidadosamente

moldada para atender aos critérios das plataformas digitais. A espontaneidade, nesse caso, não é um traço genuíno da personagem, mas um recurso estratégico que responde às lógicas algorítmicas de engajamento, viralização e monetização. Em outras palavras, trata-se de uma autenticidade roteirizada para ser percebida como tal, ajustada de acordo com métricas de visibilidade.

Satiko, ao assumir discursos de diversidade e liberdade, projeta-se como sujeito “verdadeiro”, acessível e engajado com causas sociais. Contudo, a autenticidade que mobiliza é resultado de um planejamento corporativo que busca equilibrar inovação e aceitação. A personagem encarna, assim, um paradoxo: sua “verdade de si” é programada, mas ainda assim gera efeitos reais de proximidade, engajamento e construção de comunidades afetivas em torno de sua narrativa.

Esse paradoxo dialoga com o que Sibilia (2008) denomina de espetacularização da intimidade, na qual os sujeitos se apresentam de modo cada vez mais transparente, mas sempre mediados por filtros estéticos e narrativos. No caso de Satiko, a intimidade não é apenas estetizada, mas também fabricada como estratégia de diferenciação, funcionando como dispositivo que humaniza a tecnologia e legitima sua inserção no campo dos influenciadores digitais.

Portanto, a performance da autenticidade algorítmica em Satiko evidencia como a comunicação digital contemporânea transforma valores subjetivos em recursos mercadológicos. A autenticidade, longe de se opor à mercantilização, é condição para sua eficácia: somente ao parecer autêntica, Satiko consegue consolidar sua autoridade afetiva e reforçar sua função como mediadora entre marcas e públicos.

Discussão

A análise da personagem Satiko permite compreender, em profundidade, como os influenciadores virtuais operam no entrecruzamento entre subjetividade, imaginário e mercantilização nas plataformas digitais. Sua performance revela a complexidade das disputas simbólicas que atravessam o feminino digital, condensando paradoxos que ilustram os dilemas contemporâneos da comunicação midiaticizada.

Em primeiro lugar, a figura de Satiko explicita como o imaginário midiático (Castoriadis, 1982; Silva, 2003) não se limita a refletir a realidade social, mas atua como dispositivo de produção e organização de sentidos. Ao encarnar discursos de liberdade, diversidade e autenticidade, a personagem não apenas representa valores aspiracionais, mas contribui ativamente para moldar modos de ver, sentir e desejar o feminino no ambiente digital. Nesse sentido, Satiko confirma a ideia de que a mídia estrutura o campo do pensável, oferecendo modelos simbólicos que orientam as percepções e as práticas sociais.

Em segundo lugar, sua performance reafirma a centralidade da subjetividade performada como recurso comunicacional. Conforme discutido por Sibilia (2008) e Han (2012), vivemos em um regime de visibilidade em que os sujeitos são convocados a estetizar a própria existência e a exibir performances contínuas de si. Satiko radicaliza esse processo: como personagem inteiramente programada, sua subjetividade é produto de design, roteiro e estratégia de marca, transformando-se em ativo comunicacional. Essa condição reforça a ideia de que, nas plataformas, a subjetividade se converte em mercadoria, sendo roteirizada para atender às lógicas da economia da atenção.

Um terceiro aspecto diz respeito ao uso estratégico do empoderamento feminino. A personagem se apresenta como “CEO livre” e sem idade, evocando discursos de autonomia e emancipação. No entanto, como apontam McRobbie (2009) e Banet-Weiser (2018), esse empoderamento é reconfigurado pelo mercado em forma de estilo consumível, esvaziado de sua potência política original. Satiko traduz o empoderamento em narrativa aspiracional, que reforça identificação e engajamento, mas permanece ancorada em padrões normativos de beleza e feminilidade, revelando o caráter ambíguo de seu discurso.

A dimensão da autenticidade algorítmica também é central na discussão. De acordo com Marwick e Boyd (2011), a autenticidade nas redes é sempre performada, e Abidin (2023) demonstra como ela se tornou um recurso estratégico para influenciadores. Satiko exemplifica essa lógica ao encenar autenticidade de forma programada, calibrada para maximizar engajamento e adequar-se às métricas de visibilidade. Assim, sua autenticidade não é um dado, mas um efeito narrativo cuidadosamente planejado.

Por fim, a noção de narrador dialógico (Reis & Thomé, 2022) ajuda a compreender como Satiko mobiliza estratégias de interação e interpelação, criando a ilusão de proximidade e escuta. Ainda que sua performance seja roteirizada, a simulação de diálogo confere legitimidade e fortalece sua autoridade afetiva. Contudo, essa abertura ao público é mediada por interesses corporativos, revelando uma contradição entre a promessa de diálogo e a instrumentalização mercadológica da interação.

Diante desses elementos, pode-se afirmar que Satiko encarna as tensões do feminino nas plataformas digitais: é símbolo de empoderamento e mercadoria, promessa de autenticidade e encenação roteirizada,

figura de liberdade e reprodutora de normatividades. Sua atuação revela que o imaginário feminino digital é um território de disputas, em que discursos de emancipação convivem com interesses de mercado e algoritmos de visibilidade.

Assim, o caso de Satiko não deve ser visto apenas como curiosidade tecnológica ou estratégia de marketing, mas como expressão das dinâmicas contemporâneas de subjetivação nas plataformas. Ao condensar paradoxos entre subjetividade e mercantilização, Satiko oferece um campo fértil para refletir sobre os modos como a comunicação digital reorganiza os sentidos do feminino e insere as experiências de identidade no circuito da economia simbólica.

Conclusão

A análise de Satiko como influenciadora virtual evidencia a complexidade das dinâmicas que atravessam a construção do feminino em ambientes digitais. Longe de ser apenas um experimento estético ou tecnológico, a personagem opera como dispositivo simbólico e mercadológico, articulando discursos de subjetividade, autenticidade e empoderamento em sintonia com as lógicas das plataformas e do mercado.

Os resultados desta pesquisa permitem afirmar que Satiko encarna os paradoxos centrais do imaginário feminino digital. Ao mesmo tempo em que mobiliza narrativas de liberdade, diversidade e autenticidade, sua performance está profundamente ancorada em padrões normativos de feminilidade e em estratégias corporativas de mercantilização da subjetividade. Essa duplicidade revela como o empoderamento, frequentemente apresentado como valor emancipatório, é apropriado e

ressignificado pelo mercado, transformando-se em estilo consumível e em ativo de marca.

Outro achado relevante diz respeito à performance da autenticidade algorítmica. Satiko demonstra que, no ambiente das plataformas digitais, a autenticidade não é uma essência, mas uma performance cuidadosamente roteirizada e calibrada para maximizar engajamento e visibilidade. Essa autenticidade programada, embora artificial, produz efeitos reais de proximidade, interação e construção de comunidades afetivas, legitimando a personagem como mediadora entre marcas e públicos.

Além disso, a análise mostrou que Satiko funciona como narradora dialógica simulada, convocando seus seguidores à participação e à troca, ainda que dentro de um roteiro estratégico corporativo. Essa simulação de escuta, embora limitada, é eficaz para consolidar sua autoridade afetiva e reforçar o vínculo com a audiência, ilustrando o papel central da interação na economia da atenção.

Do ponto de vista teórico, este estudo contribui para compreender como o imaginário midiático se reconfigura em tempos de plataformização, evidenciando que a produção de sentido não é apenas reflexo da realidade social, mas campo ativo de disputas simbólicas em torno da identidade, do gênero e do consumo. O caso de Satiko exemplifica como o feminino é constantemente negociado entre discursos emancipatórios e interesses mercadológicos, tornando-se espaço privilegiado para observar as tensões contemporâneas entre subjetividade e mercado.

Como pesquisa em andamento, este trabalho apresenta limitações que abrem caminhos para futuras investigações. A análise concentrou-se no perfil de Satiko no Instagram, mas estudos posteriores podem

expandir o olhar para outras plataformas e para a recepção do público, investigando como as audiências interpretam, negociam e ressignificam a performance da personagem. Além disso, comparações com outros influenciadores virtuais - como Lu do Magalu e Nat da Natura - podem oferecer perspectivas mais amplas sobre a pluralidade de estratégias que estruturam o fenômeno no Brasil e no cenário internacional.

Em síntese, Satiko não é apenas um avatar digital, mas um laboratório simbólico que condensa as tensões do feminino em tempos de algoritmos, plataformas e mercantilização. Sua existência nos convida a refletir sobre como os discursos de autenticidade, empoderamento e diversidade, embora potentes, são permanentemente reconfigurados pelo mercado, revelando os limites e as possibilidades de imaginários femininos digitais no século XXI.

Referências

- Abidin, C. (2016). Visibility labour: Engaging with influencers' fashion brands and #OOTD advertorial campaigns on Instagram. *Media International Australia*, 161(1), 86-100. <https://doi.org/10.1177/1329878X16665177>
- Abidin, C. (2023). *Digital intimacies: Intensity, exposure and living in social media*. Polity Press.
- Banet-Weiser, S. (2018). *Empowered: Popular feminism and popular misogyny*. Duke University Press.
- Bardin, L. (2016). *Análise de conteúdo* (Edição revista e ampliada). Edições 70.

- Barthes, R. (1990). *Mitologias*. Bertrand Brasil. (Obra original publicada em 1957).
- Bauman, Z. (2001). *Modernidade líquida*. Zahar.
- Butler, J. (2019). *Problemas de gênero: Feminismo e subversão da identidade* (14ª ed.). Civilização Brasileira. (Obra original publicada em 1990).
- Castoriadis, C. (1982). *A instituição imaginária da sociedade*. Paz e Terra.
- Couldry, N., & Hepp, A. (2017). *The mediated construction of reality*. Polity Press.
- Gill, R. (2007). Postfeminist media culture: Elements of a sensibility. *European Journal of Cultural Studies*, 10(2), 147-166.
- Han, B.-C. (2012). *A sociedade do cansaço* (2ª ed.). Vozes.
- Jenkins, H. (2006). *Convergence culture: Where old and new media collide*. New York University Press.
- Khamis, S., Ang, L., & Welling, R. (2017). Self-branding, ‘micro-celebrity’ and the rise of social media influencers. *Celebrity Studies*, 8(2), 191-208. <https://doi.org/10.1080/19392397.2016.1218292>
- Marwick, A., & Boyd, D. (2011). To see and be seen: Celebrity practice on Twitter. *Convergence: The International Journal of*

Research into New Media Technologies, 17(2), 139-158. <https://doi.org/10.1177/1354856510394539>

McRobbie, A. (2009). *The aftermath of feminism: Gender, culture and social change*. SAGE.

Reis, M. A., & Thomé, C. A. (2022). O narrador dialógico na reconfiguração do jornalismo pós-guina subjetiva. *Interin - Revista do Programa de Pós-Graduação em Comunicação*, 27(2), 1-16. <https://doi.org/10.5380/interin.v27i2.80927>

Sibilia, P. (2008). *O show do eu: A intimidade como espetáculo*. Nova Fronteira.

Silva, J. (2003). *O sujeito na tela: subjetividade, identidade e diferença na era das redes*. Sulina.

Van Dijck, J., Poell, T., & Nieborg, D. (2018). *The platform society: Public values in a connective world*. Oxford University Press.

Wolf, N. (1992). *O mito da beleza: Como as imagens de beleza são usadas contra as mulheres*. Rocco.

MÍDIA SINTÉTICA: ENTRE A INOVAÇÃO E A MANIPULAÇÃO

Margareth Boarini¹

Quando o assunto orbita pelos impactos causados por qualquer tecnologia em qualquer aspecto que seja, observamos que o bem e o mal caminham juntos, uma vez que toda e qualquer aplicação resulta de decisões planejadas e operacionalizadas, direta ou indiretamente, por seres humanos. Em tempos marcados pelos avanços na área e no ecossistema da inteligência artificial generativa, a tecnologia que envolve a mídia sintética se destaca como um dos exemplos impactados pela conjugação dos fatores positivos e negativos. Isso porque, em nosso entendimento, tal modelo de mídia tem a capacidade intrínseca de reconfigurar os parâmetros daquilo a que fomos acostumados a crer como realidade. “Brincar” com a noção de realidade pode ter implicações

1. Doutora em Tecnologias da Inteligência e Design Digital (PUC SP, Brasil).
Professora da Faculdade Cásper Líbero (SP, Brasil).
magaboarini@gmail.com

positivas no campo do entretenimento, permitindo que experimentemos o poder da criação de seres, objetos e cenários. No entanto, esse mesmo “brincar” confere o poder de manipular uma mensagem textual, vocal ou imagética, por conta da tecnologia que transforma o irreal em real.

Mídia sintética ora transforma a barreira entre realidade e ficção em uma linha bastante tênue, ora a desfaz por completo. Ou seja, em um momento, temos a barreira. Em outro, a passagem é livre. O fator determinante para obter um ou outro *status* se escora na intencionalidade do agente (ou agentes) responsável pela criação e distribuição de tal mídia.

Três aspectos demandam atenção no que diz respeito ao uso e disseminação da mídia sintética a partir de 2025. O primeiro deles diz respeito à profusão de ferramentas movidas pela inteligência artificial generativa para a criação e promoção de mídia sintética que surgem a cada dia no mercado. Acessíveis, elas são capazes de serem usadas de forma fácil, a um clique de distância. O segundo aspecto se refere a uma característica de nosso tempo. Em épocas já amplamente consolidadas pelo protagonismo do prosumidor, tais ferramentas tendem a ter seu uso proliferado, sem controle de regulação. O terceiro aspecto, em nosso entendimento, deveria ser reconhecido como aquele que permeia os dois anteriores pelo fato de se referir ao letramento do usuário, ou seja, o cidadão em geral, sobre os efeitos da intencionalidade da mídia sintética seja pelo entretenimento seja pela finalidade manipulatória. O mesmo usuário que emprega a ferramenta para entreter pode vir a ser vítima do uso indevido por outrem. Afinal, tudo o que se refere à mídia sintética envolve tanto um grande potencial para o bem, amplamente divulgado, como para o mal, na grande parte das vezes, escamoteado. Daí a transcendência extremamente fluida entre a inovação e a manipulação.

Nossa pesquisa traz como objetivo central evidenciar as duas facetas da mídia sintética, cuja ferramenta *deep fake* é das mais expoentes, focadas na diversão e no prejuízo de outros, e, para tanto, traçou um percurso amparado em uma metodologia de natureza qualitativa, exploratória, em bibliografia multidisciplinar e forte pesquisa documental, uma vez que as novidades tecnológicas têm se sucedido com velocidade exponencial.

O que é mídia sintética

Rege o dicionário *Oxford* que sintético é tudo aquilo que é artificial, produzido por meio de combinação com algo químico, um sinônimo partilhado pelo dicionário *Houaiss* da língua portuguesa, que destaca o caráter de algo resultante de uma -e dotado para- síntese.

Para a *Ofcom Broadcast* (2023), trata-se de um termo genérico para vídeos, imagens, textos ou vozes que foram gerados, no todo ou em parte, por algoritmos de inteligência artificial. O resultado dessa “criação” se transforma então em uma realidade sintética. Na definição de Cardenuto et al. (2023, p. 1), realidades sintéticas são criações ou ampliações digitais geradas contextualmente por inteligência artificial, se valendo de grande quantidade de dados para gerar novas narrativas e realidades. Fazem parte do amplo espectro das possibilidades de criação iniciativas em imagem e som, isto é, filmes e fotos, gravação de voz, mixados ou não.

A recente invasão de nosso universo físico pela tecnologia da mídia sintética tem chamado a atenção de estudiosos, profissionais de mercado e cidadãos, seja pela perspectiva de uso pautada nas oportunidades de trabalho e do entretenimento, seja pela ótica dos riscos

trazidos pela possibilidade de fins manipulatórios, reforçam os autores (2023, p. 2).

Em outras palavras, mídia sintética é um amplo guarda-chuva de uma tecnologia que se vale dos avanços da inteligência artificial em todas as suas vertentes (generativa, principalmente), para criar, com algoritmos de IA, produtos ficcionais que simulam com perfeição a realidade. Sob seu guarda-chuva, abriga uma série de subcategorias, como música, texto e voz gerados por IA, técnica imagem gerada por computador de (*computer generated imagery* ou *CGI*), realidade aumentada e realidade virtual, incluindo o *deep fake*, cujo foco específico é a intencionalidade de criar e propagar algo com fins malicioso e ilícito. Por trás da capacidade de os computadores gerarem conteúdo realista, destacam-se técnicas de aprendizado profundo da máquina (*Deep learning*), as redes adversárias generativas (*GANs*), alimentadas por redes neurais, visão computacional e processamento de áudio, e o processamento de linguagem natural (*NLP* ou *natural language process*), que permitem então a concepção da mídia sintética baseada em textos, som, imagens e fotos/vídeos.

Com os avanços tecnológicos dos últimos anos, temos observado o quanto a mídia sintética tem evoluído em termos de fidelidade à realidade e quão mais sofisticada deve se apresentar dentro do curto prazo. Segundo o estudo *The future of synthetic media*, promovido pela entidade *Digital Regulation Cooperation Forum* (2024, p. 10), apesar de ainda haver limitações na sincronização entre a saída de áudio com a movimentação bucal, a voz sintética é já um processo bastante desenvolvido (como veremos mais adiante), devendo apresentar cada vez mais melhoras progressivamente. Por mais que a fidedignidade

com a realidade não se dê por completo hoje em dia , observa-se que o processo de detecção e identificação do que é real ou fruto de mídia sintética tem ficado cada vez mais desafiador e, em muitos casos, já indistinguível, conforme o estudo.

Não existe uma data especificamente definida que tenha marcado o início do uso ou do lançamento da mídia sintética na história, uma vez que a técnica de manipulação de imagens não é novidade. Stone (2023) afirma que o começo pode ser remetido aos primórdios da computação, por conta de experimentos com a geração de gráficos e áudios simples, e cita “John Whitney como um dos pioneiros em mídia sintética durante os anos de 1960 e 1970, quando usou computadores para criar animações complexas e efeitos visuais para filmes e anúncios”. Os anos de 1980 e 1990 também se mostraram ativos. O programa *Photoshop*, criado no final dos anos 1980 pelos irmãos Thomas e John Knoll e posteriormente adquirido pela *Adobe Systems* em 1988, revolucionou o mercado por tornar mais acessível a manipulação de mídias. A técnica de computação gráfica no cinema, enquanto isso, também se desenvolvia e se tornou bastante evidente no filme “*Tron: Uma odisséia eletrônica*” (1982) e no desenho *ToyStory* (1995). No entanto, vale lembrar, que a manipulação fotográfica é uma possibilidade muito antiga e seu uso remonta ao século XIX (Fluture, 2023).

A expressão *deep fake*, por sua vez, é formada a partir das palavras de origem inglesa *deep learning* (aprendizado profundo) e *fake* (falso) e foi cunhada em 2017 por um moderador da rede social *Reddit* ao criar um espaço batizado com este termo para serem compartilhadas fotos manipuladas por IA de celebridades em situações pornográficas, trocando-se o rosto delas. Embora o espaço tenha sido posteriormente

excluído, o termo *deep fake* se popularizou (Regan, 2025). A partir de 2010, a técnica ganhou força por conta dos avanços na área de aprendizado da inteligência artificial, da visão computacional e do fluxo e da capacidade de armazenamento de dados, entre outros. Regan (2025) reforça que o desenvolvimento das *GANs* no *deep learning*, por volta de 2014, teve, e tem ainda, um papel relevante na sofisticação da técnica.

A faceta inovadora da tecnologia

Não há dúvidas de que o desenvolvimento da mídia sintética e sua crescente popularização tem trazido muitas oportunidades na criação e produção de conteúdo de todas as formas, em áreas como o entretenimento, a publicidade e a propaganda, a educação, a comunicação (RP jornalismo) e a saúde, entre outros possíveis. No entanto, consideramos relevante ressaltar que, em se tratando do uso de mídia sintética em todas as suas versões e ferramentas, não existe sequer uma oportunidade em que se dispense o cuidado com a ética, com a preocupação em respeitar os direitos autorais, o cuidado com o devido tratamento de dados pessoais, uma vez que esse tipo de mídia estabelece uma barreira muito tênue e frágil entre realidade e irreabilidade e entre os pontos positivos e os negativos, temas tratados a partir de reflexão mais detalhada no bloco a seguir.

O estudo do *Digital Regulation Cooperation Forum* (DRCF, 2024) aponta a qualidade na criação de conteúdo aliada ao baixo custo de produção e à democratização do uso da tecnologia como os três principais fatores de sucesso. Além disso, a possibilidade de personalizar o conteúdo de forma bastante assertiva também se tornam contributos importantes para entregar informação de interesse a grupos específicos,

podendo gerar mais engajamento e, portanto, uma comunicação mais duradoura e merecedora de atenção por parte das pessoas (DRCF, 2024).

Bhat (2023) destaca ainda a disponibilidade 24x7 e o caráter multiplataforma da tecnologia como fatores de facilitação comunicacional, em jogos aplicados a *sites*, experiências com realidade virtual ou aumentada. “Isso torna a mídia sintética altamente adaptável que pode ser amplamente utilizada em vários campos”, afirma.

Na área da saúde, o potencial de uso com apelo bastante positivo é diverso, desde o treinamento de profissionais em áreas de atendimento, cirúrgico, estudo de órgãos até nos cuidados com o próprio paciente e com pacientes com demandas de acessibilidade (SWGfI, 2025), que podem, por exemplo, se valer de voz sintética para expressar sensações, conversação. Outro uso na área médica diz respeito à geração de imagens que reproduzem órgãos humanos de forma bastante realista, com o propósito de identificar tratamentos e procedimentos assertivos a determinado paciente, sem a necessidade de realizar procedimentos cirúrgicos desnecessariamente.

A geração de avatares que possam ou não representar um ser humano também se destaca entre os usos positivos da mídia sintética. Nas redes sociais, é comum nos depararmos com avatares que representam uma pessoa, o que Boarini (2024) classifica de humano digital. Para a autora, trata-se de uma representação de um ser humano dentro do formato digital, sintonizado às demandas da vida em rede e totalmente conectada. Um exemplo dentro desse conceito seria o avatar Satiko, criado para a apresentadora Sabrina Sato, e que tem uma página na rede social Instagram e protagoniza campanhas especialmente idealizadas para a ambiência digital, dentro de uma caracterização digital (Satiko, 2025).

Para quem atua no mercado criativo, participa ativamente de campanhas publicitárias e *collabs*, estar representado por avatar torna-se uma boa estratégia comunicacional e de marketing, uma vez que se permite estar presente, de fato, dentro do caráter figital (físico e digital) da sociedade contemporânea, potencializando visibilidade e alcance a seguidores nem sempre alcançados pela atuação tradicional no mercado.

Um avatar não precisa necessariamente representar um humano. Neste caso, Boarini (2024), classifica como não humano os avatares que não apresentam representatividade alguma com pessoas e possuem “vida” própria. Este seria o caso da Lu, avatar da empresa Magalu, considerada uma das principais influenciadoras virtuais do momento (HypeAuditor; Meio&Mensagem, 2024), com mais de sete milhões de seguidores. A relevância no uso deste tipo de personagem para uma marca ou organização reside na possibilidade de criar relacionamento constante com clientes e audiências em geral, além de ser uma ferramenta importante para a gestão reputacional da companhia.

Na área educacional, o estudo do *DRCF* (2024), aponta a melhoria no processo de alfabetização e o letramento voltado para habilidades digitais, de dados, de IA e midiáticas, como focos para o emprego da mídia sintética em perfis variados de públicos e fins diversos, seja aqueles com dificuldade de aprendizado, aquelas pessoas que moram em locais remotos, seja sobre temas mais áridos para a grande parte das pessoas, como é a educação financeira, ou mesmo temas que poderiam tem no apelo lúdico uma boa ferramenta de assimilação de conteúdo, entre tantos outros exemplos.

A faceta manipulatória da tecnologia

Na obra “A manipulação da verdade”, Patrick Charadeau (2022, p. 15), alerta, logo na Introdução, que a manipulação, embora seja um tema da moda, não é algo novo, mas um fenômeno que pontua a história das relações humanas desde sempre, e tem como vantagem uma característica de “não aparecer como tal e de manter o outro no sentimento de sua liberdade, dentre outras”.

A ponderação de Charadeau ganha contornos mais contundentes a partir do fato de que os avanços tecnológicos têm se apresentado tão intensos a ponto de aperfeiçoar a mídia sintética de forma que aquilo que nasceu como irreabilidade, e, portanto, por ela produzido, possa se sobrepor àquilo que temos como real como se estivesse em curso um processo simbiótico, nem sempre com o nosso conhecimento, consentimento e percepção.

Manipular significa manejar com as mãos, tocar com destreza, fazer mágica, falsificar a realidade (Houaiss, 2009, p. 1235), um significado que sustenta tanto a manipulação de fotografias de modo ao simples intento de tornar a imagem mais visível ou mesmo de apenas reparar detalhes apagados como criar e falsear elementos com o objetivo de impactar ou enganar a audiência. Muitos anos antes de a expressão *fake news* vir à tona para o grande público, já eram conhecidos episódios de adulteração de fotos na história soviética, como a praticada em 1917, quando um “cartão-postal de uma passeata de soldados em Petrogrado teve a placa ao fundo alterada: “relógios — prata e ouro”, trocada para “Lute pelos seus direitos” (Redação, 2019). Não há como negar que o efeito manipulatório de uma imagem manipulada naquela época tenha tido efeito igualmente negativo que uma mídia sintética de nosso tempo

usada para o mal, mas os fatores de escala e de aceleração da vida em rede tornam, sem dúvida, os efeitos de hoje em dia mais impactantes.

O estudo do *DRCF* (2024, p. 18) listou sete categorias que demandam atenção:

1. Autenticação dificultada;
2. Desinformação;
3. Consumidores enganados;
4. Violação de direitos autorais e incentivos criativos;
5. Criação de conteúdo prejudicial e ilegal;
6. Erros de sistema e falta de transparência;
7. Impactos psicológicos.

Podemos observar que uma das características única entre os efeitos listados pela entidade é a de impacto estrutural, não importando se ele se dá de forma monetária, psicológica ou de conteúdo. A reverberação do estrago traz fortes impactos em uma estrutura de forma geral. O estudo do *DRCF* apresenta uma dinâmica interessante. Para cada ponderação feita tanto para oportunidades como para riscos, há um campo que contempla “considerações regulatórias”. Destacamos como exemplo o item “Detecção de mídia sintética”. Nele, o estudo discorre sobre diversos métodos que têm sido sugeridos, a fim de se possibilitar a detecção daquilo que parece, mas não é real: marca d’água, origem dos dados, métrica reputacional, softwares capazes de realizar a detecção. Após a explanação, há uma ponderação sobre quais considerações regulatórias podem ser aplicadas a ele. Tal dinâmica nos permite uma

leitura que demonstra tanto o amplo espectro de preocupação acerca dos efeitos eventuais da mídia sintética, bem como o intenso grau de fragilidade que ela carrega na diversidade de sua existência.

Na obra “Dos humanos aos humanos digitais e os não humanos: a nova ordem social da coexistência”, Boarini (2024) alerta para uma condição de *match* entre homem e máquina, que pode surgir pela intensidade do coexistir, por efeitos de caráter psicológico tanto da nossa contemporaneidade (solidão, necessidade de engajamento nas redes) como aqueles explorados pela tecnologia (antropomorfismo), além de surgirem em nossos *feeds* e telas por meio do poder conhecedor e manipulatório da cadeia algorítmica. Como exemplo, a autora cita o perfil de uma influenciadora chamada Milla Sofia. De origem declarada finlandesa, 19 anos e dona de um padrão estético clichê, Milla ostentava um perfil na rede social TikTok, sem revelar se tratar de uma criação da inteligência artificial, no ano de 2024. Depois de atrair uma legião de seguidores entre o público adolescente, a entidade *Parents Together* exigiu que a rede e os donos do perfil da influenciadora informassem claramente se tratar de uma imagem gerada por IA e, portanto, não representava uma pessoa real. A preocupação dos pais amparava-se no risco de adolescentes e crianças se espelharem na beleza “irreal” da influenciadora, tentando atingir um padrão de beleza praticamente impossível de ser alcançado por um ser humano normal.

Dois fenômenos chamam a atenção nesse sentido: o efeito Proteus, em que “ambientes virtuais possibilitam alterar drasticamente a autorrepresentação, inferindo comportamentos e atitudes (Yee et al., 2009), e o antropomorfismo. Este último acontece quando humanos atribuem características humanas a entidades, objetos, animais ou robôs,

tratando-os como iguais. Nos últimos tempos, temos presenciado casos crescentes de pessoas que decidiram fazer terapia com *bots* (ChatGPT, por exemplo).

Um outro fenômeno também acaba permeando essa realidade contemporânea. A solidão já tem sido tratada como uma epidemia de nossos tempos. Dados da Organização Mundial da Saúde (OMS) mostra que “entre 17% e 21% dos jovens de 13 a 29 anos relataram se sentir solitários, com as taxas mais altas entre adolescentes. O índice chega a 24% entre pessoas de países de baixa renda - mais que o dobro da taxa registrada em países de alta renda (11%), segundo Laboissière (2025). No cômputo geral, uma em cada seis pessoas se sente solitária, mesmo estando conectada com milhões via rede social digital. A coexistência entre humanos e não humanos que se valem de elogios e linguagem sedutora pode acentuar o relacionamento entre homem e máquina, preterindo a relação com outros seres humanos. Em um momento de solidão extrema, isso pode se mostrar prejudicial.

No campo da segurança, o *deep fake* e sua crescente popularização já vem emitindo sinais de alerta. Um estudo da consultoria KPMG avisa as empresas para a necessidade de implementação de programas de IA construídos com base em uma arquitetura confiável, a fim de garantir um uso seguro e ético: “Praticamente qualquer pessoa pode criar um *deepfake*, graças à democratização acelerada da AI generativa, impulsionada por ferramentas de código aberto, plataformas fáceis de usar e acesso descentralizado à tecnologia *deepfake*”.

Um documento elaborado pelo *Monetary Authority of Singapore* (2025) - *Cyber Risks associated with Deepfakes* (Riscos cibernéticos associados ao *deepfake*, em português) -, chegou a contemplar como

tema da edição de 2024 os riscos referentes ao mau uso da inteligência artificial generativa, mas na edição de 2025 dedicou todo o espaço para essa técnica de mídia sintética. Entre os riscos apresentados: falsificação de identidade, documentos falsificados e transações fraudulentas, responsáveis já por diversos casos de perdas financeiras e danos à reputação de organizações em várias partes do mundo. O estudo ressalta ainda a ameaça que a técnica representa para o sistema biométrico como um todo, sobre o fluxo de e-mails e sobre sistemas de comunicação por vídeo e voz falsificados, por conta da facilidade na “burla de medidas de segurança e a manipulação de processos de tomada de decisão por agentes mal-intencionados”.

O amplo leque de usos

Como observado no decorrer desta pesquisa, o uso positivo ou negativo da mídia sintética dependerá da intencionalidade de seu criador e propagador. Apresentamos alguns casos abaixo.

Bem-vindo a Chechênia (2021)

O documentário dirigido por David France, repórter investigativo, tinha um grande desafio para a sua produção. Ao mesmo tempo que pretendia retratar a perseguição a pessoas *LGBTQ* na república russa, precisava idealizar formas de demonstrar essa situação para o mundo, protegendo a identidade dos ativistas que lutam por essa causa e que já viviam escondidos para se livrar da prisão e eventual tortura. A forma encontrada foi se valer da técnica do *deep fake*, substituindo rostos reais por meio de técnicas de efeitos visuais. Dessa forma, os

protagonistas estariam protegidos e a obra teria conferido o realismo necessário das imagens para criar empatia na audiência. Segundo a publicação *MIT Technology Review* (2021), a solução buscada pelo diretor foi a de conchamar ativistas *LGBTQ* de todo o mundo para emprestarem seus rostos, que, então foram sobrepostos sobre os dos ativistas reais. “A técnica permitiu que France preservasse a integridade das expressões faciais de seus entrevistados e, portanto, sua dor, medo e humanidade. No total, o filme protegeu 23 indivíduos”. A produção acabou levando o prêmio *Teddy Award* dentro da temática *LGBT+* e o da Audiência de Melhor Documentário no Festival de Berlim (GZH, 2021). A técnica foi similar à empregada para a produção icônico em comemoração dos 70 anos da Volkswagen no Brasil em que a cantora Maria Rita contracenou e cantou ao lado de sua mãe, a também cantora Elis Regina, morta em 1982. O comercial foi lançado em 2023 e produzido pela agência *AlmapBBDO*.

Sora, OpenAI e direitos autorais (2025)

30 de setembro de 2025. A *OpenAI*, criadora do *ChatGPT*, apresentou a ferramenta Sora, destinada à criação e publicação de vídeos gerados por inteligência artificial generativa, com o apelo de que os usuários poderiam se valer de conteúdos protegidos por direitos autorais sem qualquer ônus, com exceção dos conteúdos em que os respectivos proprietários das respectivas obras bloqueasse de forma explícita o tal uso. Sam Altman, *CEO* da empresa, chegou a emitir um depoimento, afirmando que “caberia às empresas ‘optarem por sair’, caso não quisessem que seus personagens fossem usados pelo Sora” (Melo & Cossetti, 2025). Em razão da forte repercussão negativa em torno da questão de

infração a direitos autorais, no dia 06 de outubro de 2025, uma semana após o lançamento, a empresa informou ter voltado atrás, prometendo dar controle dos direitos autorais aos respectivos detentores.

Voz sintética assim como a humana (2025)

Em sua página na rede social Instagram em 03 de outubro de 2025, Martha Gabriel apresentou uma pesquisa *da Queen Mary University of London*, publicada na revista *Plus ONE*, sobre os avanços na área da voz sintética, alcançando características bastante similares à voz humana, natural. A clonagem de voz, tanto para fins de entretenimento como para crimes, tem se tornado mais e mais frequente. No entanto, a mesma pesquisa aponta ainda que a voz sintética, embora esteja se desenvolvendo em qualidade, ainda perde em efeitos de realidade para a imagem.

O mercado global da tecnologia de reconhecimento da voz, segundo o portal *Statista*, prevê atingir até 2029 um movimento de US\$ 50 bilhões. Esse mercado é impulsionado pelo “reconhecimento fonético de diferentes palavras e entonação e a interpretação da linguagem real tal como é falada, o campo do processamento de linguagem natural ou NLP” (Thormundsson, 2024).

Conclusão

Ao longo de nossa pesquisa, objetivamos evidenciar como a mídia sintética orbita entre a inovação e a manipulação. Isso porque da mesma forma que permite criações relevantes e muito próxima da realidade em texto, imagem, som (vocal), ela se configura uma ferramenta

que pode ser empregada para causar prejuízo, desinformação, danos dos mais diversos e sérios à reputação de pessoas, organizações, governos, democracia, entre outros.

A mídia sintética e sua popularização, a possibilidade de ser criada, replicada, para o bem e para o mal a apenas um clique de distância, para todas as pessoas em qualquer plataforma, torna-se uma oportunidade de emprego e de lazer ao mesmo tempo em que se configura uma arma, seja pela simples capacidade de fazer o mal.

É o tipo de ferramenta que nos atenta para a relevância da ética, da regulação, da promoção do conhecimento pleno sobre ela. Com tantos riscos evidentes e consultorias traçando alertas para organizações, inclusive, por que não criar mecanismos para a sociedade civil como um todo, contemplando medidas protetivas por parte do Estado e, por parte do cidadão, letramento e informação, a fim de saber usar para, eticamente, tirar proveito das oportunidades, e saber se defender? Não seria o caso de criar e instituir boas práticas para o seu uso, de criar normas regulatórias para serem conhecidas, praticadas e fiscalizadas? A literacia em tudo o que se refere à inteligência digital e às mídias sintéticas deve ser uma necessidade a ser associada àquela voltada para o conhecimento midiático e de dados. “Não existe IA por si mesma. A tecnologia é sempre social e humana, não apenas sobre tecnologia, mas sobre o que os seres humanos fazem com ela”, afirma Coeckelbergh (2023, p. 79).

Esta pesquisa não se esgota com este artigo. Em razão da complexidade do tema, acreditamos na sua relevância e no acompanhamento constante da tecnologia.

Referências

- Bhat, S. (2023). O que é mídia sintética? Tipo, benefícios e práticas recomendadas. *Learn*. <https://learn.g2.com/synthetic-media>
- Boarini, M. (2024). *Dos humanos aos não humanos: a nova ordem social da coexistência*. Editora Estação das Letras e Cores.
- Charadeau, P. (2022). *A manipulação da verdade*. Editora Contexto.
- Cardenuto, J. P., Yang, J., Padilha, R., Wan, R., Moreira, D., Li, H., Wang, S., Andaló, F., Marcel S., & Rocha, A. (2023). The age of synthetic realities: opportunities and challenges. *Transactions on Signal and Information Processing*, 12, e44.
- Coeckelbergh, M. (2023). *Ética na inteligência artificial*. UBU Editora.
- Diego Almeida | IA | ChatGPT [@diegoalmeida.ia]. (s.d.). *Posts* [Perfil do Instagram]. <https://www.instagram.com/diegoalmeida.ia/>
- Digital Regulation Cooperation Forum (2024). The future of synthetic media. <https://www.drcf.org.uk/siteassets/drcf/pdf-files/the-future-of-synthetic-media.pdf?v=385978>
- Fluture, S. (2023). Fotografias relembram mestres da manipulação de imagens antes do Photoshop. *B9*. <https://www.b9.com.br/35566/fotografias-relembram-mestres-da-manipulacao-de-imagem-antes-do-photoshop/>
- GZH (2021). Documentário Bem-vindo à Chechênia e mais destaques desta terça. <https://gauchazh.clicrbs.com.br/cultura-e-lazer/agenda-cultural/noticia/2021/04/documentario-bem-vindo-a-chechenia-e-mais-destaques-desta-terca-ckn54a8zp008h016uzytjosqi.html>

Houaiss Dicionário (2009). *Sintético*. Editora Objetiva.

KPMG. Deepfake threats to companies (2025). <https://kpmg.com/xx/en/our-insights/risk-and-regulation/deepfake-threats.html>

Laboissière, P. (2025). OMS: uma em cada seis pessoas no mundo é afetada pela solidão. *Agência Brasil*. <https://agenciabrasil.ebc.com.br/geral/noticia/2025-07/oms-uma-em-cada-seis-pessoas-no-mundo-e-afetada-pela-solidao>

Martha Gabriel [@marthagabriel]. (s.d.). *Posts* [Perfil do Instagram]. Instagram. <https://www.instagram.com/marthagabriel/>

Meio & Mensagem (2024, novembro 12). Lu do Magalu perde posto de influenciadora virtual mais seguida. <https://www.meioemensagem.com.br/midia/influenciadora-virtual-mais-seguida>

Melo, J., & Corssetti, M.C. (2025). OpenAI promete dar controle aos detentores de direitos autorais no Sora. *CanalTech*. <https://canaltech.com.br/inteligencia-artificial/openai-recua-e-promete-dar-controle-aos-detentores-de-direitos-autorais-no-sora/>

MIT Technology Review (2021). O ano em que os *deepfakes* se tornaram populares. https://mittechreview.com.br/o-ano-em-que-os-deepfakes-se-tornaram-populares/?srsltid=AfmBOoohZu5Ss2IrCz1YnngB3FymbryP_WJW7LKe0QZEZFlnEqixOBZB

Monetary Authority of Singapore. (2025). *Cyber Risks Associated with Deefakes*. <https://www.mas.gov.sg/-/media/mas-media-library/regulation/circulars/trpd/cyber-risks-associated-with-deepfakes.pdf>

Ofcom Broadcast. (s.d.). Broadcast Standards. <https://www.ofcom.org.uk/tv-radio-and-on-demand/broadcast-standards>

- Oxford Dictionary. (s.d.). Synthetic. https://www.oxfordlearnersdictionaries.com/us/definition/english/synthetic_1
- Redação. (2019, dezembro 09). Apagados da história: a fotomanipula. *AH*. <https://aventurasnahistoria.com.br/noticias/galeria/fotomanipulacao-stalin.phtml>
- Regan, G. (2025). Uma breve história de deepfakes. *Reality Defender*. <https://www.realitydefender.com/insights/history-of-deepfakes>
- Satiko [@iamsatiko]. (2025, setembro 19). *Eu depois de ver o look da @sabrinasato pro anúncio do #HalloweenSephora: [Imagem anexada]*. Instagram. https://www.instagram.com/p/DOzGsvbkoyf/?img_index=1
- Stone, J. (2023, abril 12). The rise of synthetic media: from Hollywood to your business - a guide to the future of content creation. *LinkedIn*. <https://www.linkedin.com/pulse/rise-synthetic-media-from-hollywood-your-business-guide-stone/>
- SWGfI. (2024, agosto 16). Synthetic media: how it is being used around the world. <https://swgfi.org.uk/magazine/synthetic-media-how-it-is-being-used-around-the-world/>
- Thormundsson, B. (2024). Voice Technology. *Statista*. https://www.statista.com/topics/6760/voice-technology/?srsltid=AfmBOoq74dOJV0We1M1kPKp40BbCn-zzC-O79R4mq0N0sz_ort4k-gbr
- Yee, N., Bailenson, J. N., & Ducheneut, N. (2009). The Proteus Effect. *Communication Research*. https://www.researchgate.net/publication/228445790_The_Proteus_Effect
- Wikipedia (2021). Bem-vindo à Chechênia. https://en.wikipedia.org/wiki/Welcome_to_Chechnya

PARTE 2 - PROPOSIÇÕES

MIL DOCENTES, MIL HORAS: UNA CARTOGRAFÍA SITUADA DE LA INTELIGENCIA ARTIFICIAL GENERATIVA EN EDUCACIÓN

Xabier Martínez Rolán¹

La irrupción de la inteligencia artificial generativa (IAG) en el ámbito educativo ha sido tan acelerada como transformadora. Herramientas como ChatGPT, Copilot o Gemini han reabierto debates fundamentales sobre la evaluación, la autoría, el rol docente y el sentido mismo del aprendizaje (Luckin et al., 2022). Lejos de constituir una mera innovación técnica, la IAG se presenta como una tecnología de propósito general con capacidad para reconfigurar las condiciones de posibilidad del acto educativo.

1. Doctor em Comunicación.
Professor da Facultade de Comunicación de la Universidade de Vigo.
xabier.rolan@uvigo.gal

Desde 2022, he impartido más de mil horas de formación sobre IAG a más de un millar de docentes de universidad, formación profesional y enseñanza secundaria en España y Portugal. Esta experiencia, diversa en perfiles y contextos, ha permitido un contacto estrecho y continuado con los discursos, tensiones, apropiaciones y resistencias que atraviesan al profesorado ante esta nueva realidad. Las aulas de formación han funcionado como espacios de observación privilegiada, donde se condensan miedos, entusiasmos, dudas y resignificaciones pedagógicas.

Este texto no persigue ofrecer resultados cuantitativos ni replicar un diseño experimental. Se plantea, en cambio, como una reflexión situada y crítica, apoyada en una base metodológica cualitativa que se detalla en el apartado correspondiente. El objetivo es doble: visibilizar algunas buenas prácticas emergentes en la integración docente de la IAG y advertir sobre los usos problemáticos y los riesgos de una adopción acrítica.

Diversos organismos internacionales han subrayado la necesidad de acompañar la introducción de estas herramientas con marcos éticos claros y principios como la justicia, la transparencia, la privacidad o la no discriminación (UNESCO, 2023). Esta propuesta se alinea con dichas preocupaciones, pero las aborda desde una perspectiva situada en la práctica profesional docente.

A lo largo del texto se identificarán perfiles actitudinales comunes entre el profesorado, se analizarán las tensiones que surgen en torno al uso educativo de estas herramientas y se propondrán claves pedagógicas, éticas y profesionales para pensar la IAG más allá del entusiasmo

tecnocrático. Lejos de proponer soluciones cerradas, esta propuesta busca abrir un espacio de reflexión informada desde la práctica.

La inteligencia artificial generativa como tecnología educativa disruptiva

La llegada de la inteligencia artificial generativa (IAG) en el ámbito educativo no se ha producido como resultado de una planificación institucional deliberada, sino como efecto de una irrupción acelerada y masiva desde el entorno digital generalista. A diferencia de tecnologías educativas anteriores, cuyo despliegue fue gradual, mediado y con acompañamiento institucional, la IAG ha alcanzado al profesorado de manera transversal y con escasa preparación previa. La crítica clásica de Larry Cuban sobre la historia de las tecnologías educativas sigue siendo pertinente: las innovaciones mediáticas (radio, cine, televisión, ordenadores) fueron introducidas en la escuela con promesas de transformación, pero su adopción real fue lenta y fragmentaria, con frecuentes desencantos (Cuban, 1986). La IAG, por el contrario, irrumpe con velocidad, penetrando directamente en la práctica docente sin un marco pedagógico consolidado, lo que genera tanto fascinación como desconcierto.

A diferencia de herramientas puntuales, la IAG debe ser entendida como una tecnología de propósito general, con capacidad transversal para intervenir en múltiples dominios: generación de textos, traducción, creación de imágenes, simulación de diálogo, retroalimentación automatizada y procesos de evaluación. Holmes et al. (2023) ya subrayaron que la IA no es una tecnología más, sino un agente de reconfiguración de los roles docentes y de los procesos de enseñanza-aprendizaje.

Investigaciones más recientes confirman este potencial. Monib et al. (2024) destacan que la IAG transforma el trabajo académico al articular en una sola plataforma funciones que antes estaban dispersas, desde la escritura hasta la tutoría. En el ámbito universitario, Qian et al. (2025) documentan aplicaciones emergentes en redacción académica, retroalimentación y diseño curricular, señalando un impacto sistémico en las formas de enseñanza.

Este carácter transversal la convierte en una tecnología disruptiva que no solo amplía la caja de herramientas docentes, sino que redefine los marcos de autoría, evaluación y planificación. La IAG obliga, por tanto, a reconsiderar la agencia docente y el sentido mismo del aprendizaje en la era digital.

Ante esta expansión acelerada, organismos internacionales han respondido con marcos normativos y éticos. La UNESCO (2023) y la OECD (2021a, 2021b) insisten en principios como justicia, transparencia, privacidad, rendición de cuentas y no discriminación. De manera complementaria, el Departamento de Educación de EE. UU. (2023) recomienda que toda implementación educativa de IA mantenga la figura del docente como garante de la supervisión crítica (*humans in the loop*), evitando que las decisiones pedagógicas recaigan exclusivamente en sistemas automatizados.

No obstante, diversos estudios muestran una brecha significativa entre los principios internacionales y la realidad de los centros educativos (Belkina et al., 2025; Zhai et al., 2024). Las declaraciones de organismos multilaterales suelen ser demasiado generales, carecen de guías prácticas aplicables y no se traducen en protocolos concretos de formación o acompañamiento docente. Esto produce un desfase entre

la “ética declarada” y la “práctica situada”, aumentando la sensación de incertidumbre en el profesorado.

Un riesgo central es la llamada desprofesionalización suave. Selwyn (2019) advierte que, si las decisiones pedagógicas se subordinan a las sugerencias algorítmicas, el profesorado corre el peligro de convertirse en mero ejecutor de instrucciones. Estudios recientes refuerzan esta preocupación. Zhai et al. (2024) muestran que la sobre-adopción de chatbots educativos puede erosionar habilidades cognitivas críticas, al fomentar una dependencia acrítica de los outputs generados. Costa & Murphy (2025) añade que el entusiasmo por la IAG no debe confundirse con aprendizaje profundo: sin un andamiaje pedagógico, estas tecnologías pueden reforzar desigualdades sociales y reproducir un modelo de “ciencia de ricos” frente a “ciencia de pobres”.

Desde un enfoque propositivo, autores como Picasso et al. (2024) y Atenas et al. (2025) plantean la necesidad de una alfabetización crítica en IA y datos, orientada a formar ciudadanía y profesorado capaces de analizar sesgos, negociar con la agencia algorítmica y sostener valores democráticos y humanistas en la práctica educativa.

El debate no se limita al ámbito técnico o pedagógico. La filosofía de la tecnología aporta marcos conceptuales para comprender la IAG. Floridi (2025), por ejemplo, sostiene que la IA debe entenderse como una forma de agencia sin inteligencia, es decir, como un sistema capaz de actuar en el mundo sin desplegar la totalidad de lo que denominamos inteligencia humana. Este enfoque permite matizar el debate educativo: el desafío no consiste en competir con una “inteligencia” alternativa, sino en regular la interacción entre agencias algorítmicas y

agencias humanas, garantizando que el juicio docente y el pensamiento crítico sigan siendo centrales en el proceso de enseñanza-aprendizaje.

La formación situada como observatorio de prácticas y discursos docentes. Marco metodológico

Este trabajo se fundamenta en una sistematización de la experiencia profesional acumulada durante más de mil horas de formación sobre inteligencia artificial generativa (IAG), impartidas entre 2022 y 2025 a más de un millar de docentes de distintos niveles educativos en España y Portugal. El enfoque metodológico combina herramientas propias de la netnografía educativa con elementos de autoetnografía crítica, entendida como una práctica reflexiva que permite analizar desde dentro los procesos vividos en primera persona en contextos profesionales.

Lejos de constituir un estudio experimental o estadístico, esta propuesta parte de una mirada cualitativa, centrada en identificar patrones, tensiones y aprendizajes emergentes desde una posición activa en el proceso formativo. El rol desempeñado ha sido indisociable del de formador, dinamizador y observador sistemático, lo que ha permitido acceder a una comprensión rica y situada del modo en que el profesorado se aproxima a la IAG, tanto en sus usos instrumentales como en sus implicaciones pedagógicas y éticas.

Las formaciones se desarrollaron principalmente en el contexto de universidades gallegas (como la Universidade de Vigo), centros de formación del profesorado, y centros educativos de enseñanza infantil, secundaria y formación profesional en Galicia. También se llevaron a cabo cursos en universidades de otras comunidades autónomas (Salamanca) y de Portugal (Aveiro, Porto, y próximamente Lisboa). Estos cursos,

organizados por las propias universidades o por la Consellería de Educación de Galicia, oscilaron entre las 6 y las 20 horas de duración. Aunque se trataba de formaciones de iniciación, su carácter estructurado y su profundidad conceptual favorecieron el desarrollo de una reflexión crítica por parte del profesorado.

Los grupos de trabajo fueron relativamente estables, con una participación que osciló entre las 12 personas (en los cursos más reducidos) y los 50 participantes (en los más numerosos), siendo lo más habitual la formación de grupos de alrededor de 20 docentes. Las actividades incluyeron ejercicios escritos donde se solicitaba al profesorado que valorara críticamente el impacto de la inteligencia artificial en sus prácticas educativas. Estas producciones, junto con los materiales generados durante las sesiones y la observación de las dinámicas de grupo, constituyen el corpus documental que se analizará en el presente trabajo.

Aunque no se aplicó un protocolo de análisis sistemático en tiempo real, esta sistematización retrospectiva se apoya en la revisión de dicho corpus para identificar actitudes recurrentes, formas de apropiación tecnológica y dilemas emergentes en torno al uso educativo de la IAG. De esta manera, el texto busca validar epistemológicamente el conocimiento generado en contextos formativos reales y, al mismo tiempo, reforzar el valor de la experiencia como fuente legítima de reflexión crítica en educación. Tal como señala Frabetti (2014), comprender las tecnologías implica también interrogar las condiciones ideológicas, culturales y materiales que configuran su uso y significado, algo que solo puede lograrse desde un enfoque situado y comprometido con la práctica.

Resultados. Perfiles docentes ante la IAG: tensiones, dilemas y estrategias

La presencia creciente de la inteligencia artificial generativa (IAG) en el ámbito educativo ha provocado reacciones diversas entre el profesorado. A lo largo de las sesiones de formación descritas, fue posible identificar una serie de perfiles actitudinales recurrentes, no como categorías cerradas, sino como tendencias observables en los modos de aproximarse a esta tecnología. Estos perfiles, más que depender del área de conocimiento, del tipo de institución o del nivel educativo, se revelan como patrones transversales, presentes por igual en docentes de infantil, secundaria, formación profesional o universidad; en profesionales de ciencias, humanidades, tecnologías o artes.

La utilidad de este enfoque tipológico no reside en clasificar a los docentes, sino en comprender mejor las tensiones, dilemas y estrategias que despliegan al enfrentarse a una herramienta tan disruptiva como la IAG. Reconocer estos perfiles puede ser clave para diseñar propuestas formativas más contextualizadas y eficaces.

Entusiastas funcionales

Este grupo está formado por docentes que adoptan rápidamente la IAG, movidos por su utilidad práctica. Ven en estas herramientas un recurso para ahorrar tiempo, automatizar tareas repetitivas y mejorar la eficiencia del trabajo docente. Como señalaba un profesor: *“Podemos afirmar que o uso da IA contribúe á xeración e ao mantemento de materiais docentes de alta calidade, reducindo significativamente o tempo necesario para a súa preparación”* (Profesorado participante, testimonio 1).

Su interés se centra en el “cómo usarla” para generar rúbricas, corregir textos, resumir contenidos o elaborar materiales didácticos. Otro docente lo resumía así: *“As actividades plantexadas permiten aforrar tempo no proceso de avaliación da materia ‘Deseño e desenvolvemento curricular da EF’”* (Profesorado participante, testimonio 2).

Aunque su actitud favorece una adopción temprana, el riesgo es un uso acrítico, que subordina la toma de decisiones pedagógicas al criterio algorítmico. En algunos casos, se observa una tendencia a delegar aspectos clave de la planificación y evaluación en las herramientas, sin una reflexión profunda sobre sus implicaciones. Tal como expresaba otro participante: *“O problema é que ter que pasar estas preguntas a LaTeX para poder incluílas en Moovi levaría demasiado tempo. Desta forma, o obxectivo é poder crear unha serie de prompts que poidan xerar as preguntas que poidan ser utilizadas na materia”* (Profesorado participante, testimonio 3).

Sus estrategias habituales incluyen el uso intensivo de prompts, la experimentación autónoma y la compartición informal de “trucos” entre colegas. De hecho, varios docentes destacaban la optimización que trae consigo en investigación: *“A principal vantaxe que observo ao incluír a IA á miña investigación, algo que nunca he realizado con anterioridade, é a optimización e redución do tempo en os procesos de análise do estado do arte e no procesamiento de os datos obtidos”* (Profesorado participante, testimonio 4); *“A integración destas ferramentas na redacción de artigos académicos optimou o meu proceso de investigación e produción científica”* (Profesorado participante, testimonio 5).

Suelen valorar positivamente la formación técnica y demandan recursos aplicables de forma inmediata. Como reconocía de forma

explicita un docente: *“Son ademais totalmente integracionista. Se os docentes non integramos estas ferramentas ímonos a equivocar”* (Profesorado participante, testimonio 6).

Este grupo coincide en buena medida con lo que la literatura reciente denomina actitudes de aculturación tecnológica o integracionismo, donde el profesorado adapta activamente la IA a sus rutinas, buscando optimizar el trabajo sin necesariamente cuestionar los marcos que la sostienen (Romero et al., 2023).

Del mismo modo, estudios en el contexto español muestran que esta adopción pragmática se vincula con la aparición de nuevos roles profesionales docentes, orientados a la facilitación, adaptación y curaduría de contenidos (Bonales-Daimel et al., 2025).

Resistentes críticos

En el otro extremo se sitúan los docentes que expresan una actitud de rechazo o recelo frente a la IAG. Suelen fundamentar su posición en motivos éticos o pedagógicos, vinculados a una defensa del papel humanista del profesorado y del vínculo educativo como proceso relacional, no automatizable. Como resumía un profesor: *“A IA é unha ferramenta para potenciar a aprendizaxe, o pensamento crítico e a creatividade (capacidades) do estudantado, sempre que se utilice adecuadamente e non un sustituto”* (Profesorado participante, testimonio 7).

Estos docentes manifiestan preocupaciones por la desinformación, el plagio, la banalización del aprendizaje o la desprofesionalización del rol docente. Uno de ellos advertía: *“En canto aos riscos, un dos máis significativos é a sobreconfianza na ferramenta e a ausencia de unha maior presenza de un intercambio emocional e empático que podería*

afectar á motivación e o compromiso dalgúns estudantes. Para min, é un reto e un dos principais problemas” (Profesorado participante, testimonio 8). Otro fue más tajante: *“Descartamos por completo a redacción asistida de artigos científicos, por las razones ya expuestas, pero sí distinguimos algunas vertentes en que no compromete el rigor que debe presidir toda investigación, al menos si se utilizan con cierto criterio”* (Profesorado participante, testimonio 9).

Aun así, su resistencia no siempre implica inmovilismo. En algunos casos, adoptan un uso selectivo o testimonial, incorporando la IAG solo en situaciones muy concretas, y desde una postura vigilante. Tal como señalaba otro docente: *“O máis interesante é como combina o uso da intelixencia artificial coa reflexión crítica, facendo que os estudantes non só interactúen con ferramentas tecnolóxicas, senón que tamén cuestionen como funcionan e que sesgos poden ter”* (Profesorado participante, testimonio 10).

También es frecuente que expresen sus cuestionamientos de forma activa en los espacios de formación, generando debates enriquecedores. Un ejemplo claro de esta mirada crítica se observa en afirmaciones como: *“Este é un ámbito no que as ferramentas de IA xenerativa tenden a inventarse precedentes e sentenzas inexistentes, polo que é moi importante incidir na necesidade dunha abordaxe prudente e crítica do uso destas ferramentas”* (Profesorado participante, testimonio 11).

En paralelo, algunos señalan riesgos más estructurales vinculados a la equidad académica y social: *“Mesmo si os resultados para o obxectivo que eu buscaba me pareceron moi bos... penso que a diferenza entre a ‘ciencia de ricos’ e a ‘ciencia de pobres’ vai ser moitísimo mais grande en moi pouco tempo”* (Profesorado participante, testimonio 12).

Este grupo, aunque no se identifica así, podría ser percibido ocasionalmente como neoludita, en tanto expresa reservas fuertes hacia la IAG, prioriza valores pedagógicos, relacionales y éticos, y rechaza usos que considera deshumanizantes o tecnocráticos. Literatura reciente muestra cómo en educación se están usando categorías como *Luddite pedagogy* para describir esta postura de rechazo o de distancia crítica respecto al uso de IA (East, 2025), o de personas que aluden al término *Luddites* como metáfora de resistencia frente al avance tecnológico sin control social.

Confundidos pragmáticos

Este perfil, quizás el más numeroso, agrupa a docentes que reconocen el potencial de la IAG, pero que se sienten desorientados ante su uso real en el aula. Su actitud es ambivalente: desean adaptarse, pero no disponen de referentes claros, ni de tiempo o acompañamiento suficiente. Como reconocía un docente: *“O primeiro que debo dicir é que ate agora sempre fuxín da utilización das IAs, sobre todo por descoñecemento, porque pensaba que serían moi complicadas de entender e de manexar”* (Profesorado participante, testimonio 13).

Estos docentes suelen expresar dudas sobre los límites éticos, las implicaciones evaluativas o la legalidad del uso de estas herramientas. Uno de ellos señalaba con claridad: *“Polo que vexo, é necesario ser moi preciso na redacción dos prompts para evitar confusións deste tipo”* (Profesorado participante, testimonio 14).

A menudo se enfrentan a una sobrecarga de información y muestran una necesidad urgente de criterios pedagógicos y ejemplos prácticos. Así lo expresaba otro profesor tras sus primeras experiencias:

“A experiencia non foi totalmente satisfactoria, posto que na sección 2 pódese ver como a IA deu unha resposta incorrecta, e na sección 3 descríbese como o texto xerado debe ser revisado e adaptado” (Profesorado participante, testimonio 15).

Sus estrategias incluyen el ensayo y error, la consulta a colegas, el uso puntual de la IA con fines exploratorios, y una valoración muy positiva de los espacios de formación que les permiten experimentar en un entorno guiado y reflexivo. Como apuntaban varios docentes: *“Os chatbots poden proporcionar respostas útiles, pero habería que estar moi enriba dos estudantes para que contrastaran a información con outras fontes. Hai que fomentar máis que nunca o espírito crítico nas aulas”* (Profesorado participante, testimonio 16). Otro añadía un matiz importante: *“A iso súmase a dependencia de plataformas externas, cuxos cambios nas políticas de uso gratuíto poden afectar á continuidade do proxecto”* (Profesorado participante, testimonio 17).

Conclusiones del análisis de perfiles

Estos tres perfiles no deben entenderse como posiciones fijas ni mutuamente excluyentes. Muchos docentes transitan entre ellos según el contexto, el momento formativo o el tipo de tarea. Más que categorías estancas, son formas de expresar una relación todavía en construcción con una tecnología que desafía los marcos tradicionales de la educación.

La siguiente tabla ofrece una visión comparativa de las diferentes posturas docentes frente a la integración de la inteligencia artificial generativa (IAG). Se organizan tres perfiles identificados en la investigación: el *entusiasta funcional*, que adopta la tecnología con rapidez y un enfoque eminentemente práctico; el *confundido pragmático*, que

reconoce el potencial de la IAG pero manifiesta dudas e inseguridades en su uso; y el *resistente crítico*, que mantiene una posición de recelo o rechazo fundamentada en principios éticos y pedagógicos. La estructura temática de la tabla permite observar cómo cada perfil responde de manera distinta a cuestiones clave —desde el uso en la docencia y la investigación hasta la percepción de riesgos, las estrategias habituales o la visión de futuro—, proporcionando un marco comparativo que sintetiza tanto las coincidencias como las divergencias entre los grupos.

Tabla 1
Comparativa de perfiles docentes frente a la IAG

Tema	Entusiasta funcional	Confundido pragmático	Resistente crítico
Actitud general	Adopción rápida y acrítica, centrada en la utilidad práctica.	Ambivalencia: reconocen el potencial, pero están desorientados.	Rechazo o fuerte recelo, fundamentado en motivos éticos y pedagógicos.
Uso en docencia	Optimización de tareas: rúbricas, correcciones, resúmenes, materiales (T1).	Uso exploratorio, ensayo-error; valoran espacios de formación guiada (T13).	Evitan usos masivos; solo lo incorporan en situaciones concretas, vigilando riesgos (T7).
Uso en investigación	Aprovechamiento directo para ahorrar tiempo en análisis y redacción (T5).	Desconfianza por resultados parciales o incorrectos (T15).	Rechazo frontal a la redacción asistida de artículos (T9).
Riesgos percibidos	Uso acrítico que puede subordinar la pedagogía al algoritmo.	Confusión en prompts, dependencia de plataformas externas (T17).	Plagio, desinformación, desprofesionalización, pérdida de vínculo humano (T8).
Estrategias	Experimentación autónoma, compartir 'trucos', uso intensivo de prompts.	Ensayo-error, consulta entre colegas, prueba puntual.	Debate activo en formación, mirada vigilante, uso muy selectivo.

Valores pedagógicos	Eficiencia y productividad por encima de reflexión crítica (T6).	Necesidad de referencias claras y criterios éticos, pero falta de confianza en cómo aplicarlos.	Defensa del papel humanista del profesorado y del aprendizaje relacional (T10).
Visión de futuro	Confianza en la integración total de la IA en la docencia.	Expectativa de adaptación, pero con dudas sobre sostenibilidad.	Temor a la desigualdad y a la 'ciencia de ricos vs ciencia de pobres' (T12).

Elaboración Propia.

La presencia transversal confirma que las actitudes frente a la IAG no dependen del área de conocimiento ni del nivel educativo, sino de trayectorias previas, valores pedagógicos y condiciones institucionales. Este hallazgo refuerza la necesidad de diseñar propuestas formativas flexibles, que partan de la diversidad de posicionamientos y promuevan una apropiación crítica, situada y profesional de estas tecnologías.

En relación con el grupo de los *Resistentes Críticos*, puede afirmarse que ninguno de los argumentos recogidos en los testimonios docentes se adscribe de manera explícita al concepto de *neoludismo*. No obstante, se observan actitudes de cautela y resistencia que permiten clasificar estas posturas dentro de dicho perfil y, en menor medida, en el de los *Confundidos Pragmáticos*. Aunque estos colectivos no buscan evitar por completo la utilización de la inteligencia artificial, expresan reticencias profundas que, en contextos más extremos, podrían evolucionar hacia un rechazo total.

Las reflexiones de este grupo se articulan en torno a tres grandes dimensiones interrelacionadas. En primer lugar, se aprecia una defensa del rigor académico y de la autoría humana que se traduce en un rechazo

a la sustitución de procesos intelectuales por soluciones automatizadas. La inteligencia artificial no se concibe como un sustituto del pensamiento crítico ni de la creatividad, y se advierte sobre los riesgos de que la redacción asistida de textos académicos erosione la autenticidad y la capacidad analítica de quienes investigan. En segundo lugar, emergen con fuerza las preocupaciones éticas vinculadas a la equidad y la transparencia. La posibilidad de sesgos culturales o estilísticos, la opacidad de los modelos y la fiabilidad de sus respuestas son señalados como factores que deben ser sometidos a un cuestionamiento constante, pues un uso acrítico podría reducir los espacios de intercambio emocional y empático que caracterizan la relación pedagógica. Finalmente, aparece la dimensión pedagógica-humanista, donde la resistencia se centra en la necesidad de preservar la autonomía intelectual del alumnado y de evitar que la dependencia excesiva de estas herramientas socave el desarrollo de competencias fundamentales como la reflexión crítica, el rigor metodológico o el esfuerzo cognitivo sostenido.

En conjunto, la posición de los *Resistentes Críticos* no puede ser entendida como un intento de evasión tecnológica en clave neoludita, sino más bien como una apuesta por regular su uso y por someterlo a un escrutinio ético y pedagógico riguroso. Se trata, en definitiva, de un enfoque que reconoce la existencia y el potencial de la inteligencia artificial generativa, pero que exige establecer límites claros que garanticen una integración responsable en el ámbito educativo.

En relación con el perfil de los “confundidos pragmáticos”, se trata de un grupo que reconoce el potencial de la inteligencia artificial generativa, pero cuya reflexión se encuentra marcada por la ambivalencia y la desorientación. Sus principales preocupaciones giran en torno a los

problemas de fiabilidad de las herramientas, la necesidad de verificación constante de los resultados, la exigente curva de aprendizaje y la dificultad para establecer límites claros y estrategias de uso pedagógico. Aunque valoran la utilidad que la IA puede aportar, el esfuerzo que requiere la corrección de errores, junto con la falta de referentes institucionales o de claridad respecto a sus aplicaciones prácticas en el aula, constituyen el núcleo de su posicionamiento crítico.

En relación con el perfil de los *Entusiastas Funcionales*, se observa una aproximación positiva y orientada a la acción frente a la inteligencia artificial generativa. A diferencia de otros colectivos más cautelosos o ambivalentes, este grupo se caracteriza por integrar la IA en la práctica docente con un marcado enfoque instrumental, priorizando su capacidad para optimizar procesos, diversificar metodologías y ofrecer apoyos personalizados al alumnado. El interés no se sitúa tanto en una reflexión abstracta sobre sus implicaciones, sino en la exploración de usos concretos que potencian la eficiencia y la innovación pedagógica.

Las experiencias recogidas permiten identificar al menos tres grandes dimensiones que articulan la postura de los Entusiastas Funcionales. En primer lugar, emerge una dimensión pedagógica-aplicativa, donde la IA se utiliza como asistente para el diseño de materiales y actividades docentes. Se han creado, por ejemplo, chatbots que actúan como tutores virtuales, guiando al alumnado en la formulación de hipótesis, en la búsqueda bibliográfica o en la resolución de dudas conceptuales, configurando sistemas de apoyo orientados a la personalización del aprendizaje.

En segundo lugar, se encuentra la dimensión técnico-científica, donde la inteligencia artificial se concibe como una herramienta potente

para la resolución de problemas complejos y la experimentación con lenguajes de programación. Se han documentado experiencias con modelos locales y en línea para abordar ejercicios de álgebra de conjuntos o probabilidades, destacando la capacidad de estas tecnologías para procesar enunciados en lenguaje natural y ofrecer explicaciones paso a paso. Asimismo, la IA se ha empleado como apoyo en la optimización de scripts destinados al análisis de datos sanitarios, confirmando su utilidad como colaborador técnico en tareas de análisis avanzado.

Por último, cobra fuerza una dimensión innovadora y crítica, en la que la IA no se limita a generar contenidos, sino que facilita la construcción de escenarios de aprendizaje que promueven la reflexión y el cuestionamiento. Se han diseñado actividades basadas en casos prácticos complejos —como la evaluación de riesgos normativos o la identificación de sesgos en datos de entrenamiento— que permiten ejercitar el pensamiento crítico del alumnado. Del mismo modo, se han explorado herramientas de visualización que representan conexiones entre artículos académicos, facilitando la identificación de trabajos influyentes y líneas emergentes de investigación.

La posición de los *Entusiastas Funcionales*, en conjunto, se configura como una apuesta por la experimentación y la incorporación activa de la inteligencia artificial en el aula y en la investigación. Lejos de un uso meramente accesorio, este grupo destaca por identificar oportunidades de mejora pedagógica y científica, explorando las posibilidades de la IA como aliado estratégico en la innovación educativa y en la producción de conocimiento.

Desafíos y riesgos

A lo largo de estas más de mil horas de docencia, y en la experiencia de integración intensiva de la inteligencia artificial en el ámbito educativo presenta riesgos que deben ser analizados con cautela. Uno de los más destacables es el de la fiabilidad y precisión de las herramientas, ya que las conocidas “alucinaciones” continúan siendo un problema recurrente. La invención de datos o referencias, la generación de código erróneo o la resolución incorrecta de problemas matemáticos ponen de manifiesto la necesidad de una verificación humana constante (humans in the loop), que evite la propagación - y escalado - de errores y asegure la calidad del trabajo académico y científico.

Otro aspecto crítico es el de los sesgos algorítmicos. Los modelos de IA están adiestrados con grandes volúmenes de datos que contienen perjuicios culturales y de género, reproduciendo así distorsiones en las respuestas textuales y visuales. Ejemplos como la creación de imágenes estereotipadas de profesionales o el análisis parcial de datos evidencian la urgencia de abordar esta dimensión ética, incorporando mecanismos de supervisión y evaluación que garanticen un uso más justo y representativo de las tecnologías.

Desde la perspectiva pedagógica, preocupa también la dependencia excesiva que puede generar el uso de estas herramientas. Si estudiantes y personal investigador delegan en exceso la capacidad de análisis y resolución de problemas, existe el riesgo de erosión del pensamiento crítico y de la autonomía intelectual, incluso la merma de la capacidad intelectual. Esta situación comprometería no sólo el aprendizaje significativo, sino también el desarrollo de competencias fundamentales para la formación académica.

La cuestión del plagio y de la integridad académica constituye otro reto de primer orden. La facilidad para generar textos o código obliga a repensar los métodos de evaluación, incorporando rúbricas y estrategias que reconozcan el uso de la IA pero que, al mismo tiempo, garantan la autoría y la honestidad en los trabajos académicos. La experiencia acumulada muestra que es posible detectar irregularidades, pero la velocidad de evolución de las herramientas requiere respuestas institucionales dinámicas y adaptativas.

Una última cuestión que suscita preocupaciones es todo lo relacionado con la privacidad de los datos, con los costes asociados y con la equidad en el acceso. La integración de estas tecnologías no puede desligarse de un marco regulador que asegure condiciones justas, seguras y sostenibles para toda la comunidad educativa. Sólo así será posible consolidar un ecosistema de innovación responsable que combine el potencial transformador de la IA con los principios fundamentales de la ética y de la igualdad de oportunidades.

A modo de colofón, el contacto directo con profesorado en formación en IA generativa revela la irrupción de la inteligencia artificial generativa en la educación como un proceso de transformación profunda que interpela directamente a la práctica docente y a los marcos institucionales en los que esta se desarrolla. La diversidad de actitudes analizadas muestra que el profesorado no es un receptor pasivo de la innovación, sino un agente que negocia, adapta y resignifica las tecnologías en función de sus valores y contextos. No es una cuestión menor: el uso de la propia tecnología acaba por definir dicha tecnología.

Reconocer esa pluralidad de acciones, y acompañarla con formación crítica, regulación ética y políticas de equidad, será clave para que

la IA no erosione los fundamentos pedagógicos, y que - además - pueda contribuya a renovarlos. El reto no es elegir entre la adopción entusiasta o el rechazo frontal, sino construir una cultura educativa capaz de integrar la inteligencia artificial de manera responsable, situada y humanista, a la altura de las exigencias democráticas y sociales del presente.

Referencias

- Atenas J., Havemann L., & Nerantzi C. (2025). Critical and creative pedagogies for artificial intelligence and data literacy: an epistemic data justice approach for academic practice. *Research in Learning Technology*, 32. <https://doi.org/10.25304/rlt.v32.3296>
- Belkina, M., Daniel, S., Nikolic, S., Haque, R., Lyden, S., Neal, P., Grundy, S., & Hassan, M. (2025). Implementing generative AI (GenAI) in higher education: A systematic review of case studies. *Computers and Education: Artificial Intelligence*, 8, Article 100407. <https://doi.org/10.1016/j.caeai.2025.100407>
- Bonales-Daimel, G., Martínez-Estrella, E.C. & Sierra-Sánchez, J. (2025). Evolución del perfil docente y surgimiento de nuevos roles profesionales en la Era de la Inteligencia Artificial (IA). Una perspectiva desde docentes, estudiantes y profesionales [Evolution of the teaching profile and the emergence of new professional roles in the Age of Artificial Intelligence (AI). A perspective from teachers, students, and professionals]. Pixel-Bit. *Revista de Medios y Educación*, 73, art.3. <https://doi.org/10.12795/pixelbit.109085>
- Costa, C., & Murphy, M. (2025). Critical education, generative artificial intelligence and the tyranny of freedom: a critique of modern 'technocracy.' *Technology, Pedagogy and Education*, 1-17. <https://doi.org/10.1080/1475939X.2025.2547728>
- Cuban, L. (1986). *Teachers and machines: The classroom use of technology since 1920*. Teachers College Press.

- East, B. (2025, 3 abril). Luddite Pedagogy: It's OK to Ignore AI in Your Teaching. *Chronicle of Higher Education*. <https://www.chronicle.com/article/luddite-pedagogy-its-ok-to-ignore-ai-in-your-teaching>
- Floridi, L. (2025). AI as agency without intelligence: On artificial intelligence as a new form of artificial agency and the multiple realisability of agency thesis. *Philosophy & Technology*, 38(1), 30. <https://doi.org/10.1007/s13347-025-00858-9>
- Frabetti, F. (2014). *Software theory: A cultural and philosophical study*. Rowman & Littlefield International.
- Holmes, W., Bialik, M., & Fadel, C. (2023). Artificial intelligence in education. En W. Holmes, M. Bialik, & C. Fadel, *Artificial Intelligence in Education: Promises and Implications for Teaching and Learning* (pp. 151-180). The Center for Curriculum Redesign. <https://hdl.handle.net/20.500.12424/4276068>
- Knox, J. (2020). Artificial intelligence and education in China. *Learning, Media and Technology*, 45(3), 298-311. <https://doi.org/10.1080/17439884.2020.1754236>
- Luckin, R., George, K., & Cukurova, M. (2022). *AI for School Teachers* (1st ed.). CRC Press. <https://doi.org/10.1201/9781003193173>
- Monib WK, Qazi A, Apong RA, Azizan MT, De Silva L, Yassin H. 2024. Generative AI and future education: a review, theoretical validation, and authors' perspective on challenges and solutions. *PeerJ Computer Science*, 10, e2105. <https://doi.org/10.7717/peerj-cs.2105>
- Morales-Loor, K. P., Romero-Amores, N. V., Bayas-Jaramillo, C. M., & Vasco-Delgado, J. C. (2025). Integración de la tecnología en la formación docente: Tendencias y desafíos: Integration of technology

in teacher education: Trends and challenges. *Multidisciplinary Latin American Journal (MLAJ)*, 3(1), 448-467. <https://doi.org/10.62131/MLAJ-V3-N1-022>

OECD. (2021a). *Artificial Intelligence in Society*. Paris: OECD Publishing. https://www.oecd.org/content/dam/oecd/en/publications/reports/2019/06/artificial-intelligence-in-society_c0054fa1/eedfee77-en.pdf

OECD. (2021b). *OECD digital education outlook 2021: Pushing the frontiers with artificial intelligence, blockchain and robots*. OECD Publishing. <https://doi.org/10.1787/589b283f-en>

Picasso, F., Atenas, J., Havemann, L., & Serbati, A. (2024). Advancing Critical Data and AI Literacies Through Authentic and Real-World Assessment Design Using a Data Justice Approach. *Open Praxis*, 16(3), 291-310. <https://doi.org/10.55982/openpraxis.16.3.667>

Qian, Y. (2025). Pedagogical applications of generative AI in higher education: A systematic review of the field. TechTrends. *Advance online publication*. <https://doi.org/10.1007/s11528-025-01100-1>

Romero, M., Heiser, L., Lepage, A., et al. (2023). *Teaching and learning in the age of Artificial Intelligence*. <https://doi.org/10.48550/arXiv.2303.06956>

Selwyn, N. (2019). *Should robots replace teachers? AI and the Future of Education*. (1st ed.) Polity Press.

U.S. Department of Education. (2023). *Artificial Intelligence and the Future of Teaching and Learning*. Office of Educational Technology. <https://www.ed.gov/sites/ed/files/documents/ai-report/ai-report.pdf>

- UNESCO. (2023). *Guidance for generative AI in education and research*. UNESCO. <https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research>
- Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology*, 45(3), 223-235. <https://doi.org/10.1080/17439884.2020.1798995>
- Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. *Smart Learning Environments*, 11(1), 28. <https://doi.org/10.1186/s40561-024-00316-7>

EDUCAÇÃO NA ERA DA IA: REMODELANDO O APRENDIZADO E A CIDADANIA NA ECOLOGIA DOS MEIOS¹

*Sabrina Marcandali²
Vânia Cristina Pires Nogueira Valente³*

A veloz evolução dos sistemas de Inteligência Artificial (IA), abrangendo desde plataformas onl-ine até ferramentas sofisticadas que produzem textos e imagens, está provocando uma transformação significativa nas metodologias educacionais tradicionais e está cada vez mais presente na educação. Ela é um campo da ciência da computação que desenvolve sistemas que simulam a inteligência humana, aprende

-
1. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001
 2. Doutoranda do Programa de Pós-Graduação em Mídia e Tecnologia (PPGMiT) - FAAC/Unesp
sabrina.marcandali@unesp.br
 3. Profa. Dra. Coordenadora do Programa de Pós-Graduação em Mídia e Tecnologia (PPGMiT) - FAAC/Unesp
vania.valente@unesp.br

com dados e experiências para realizar tarefas e toma decisões de forma autônoma, utilizando algoritmos e modelos matemáticos.

A educação não se limita à simples distribuição de informações, ela deve realmente focar no desenvolvimento de competências críticas e na autorregulação, fatores essenciais para os desafios do século XXI. Isso significa promover o raciocínio computacional, a análise de dados complexos, a ética digital e, fundamentalmente, a capacidade de realizar avaliações críticas sobre a vasta quantidade de informações - incluindo as notícias falsas - geradas pela IA. Nesse contexto, a educação emerge como a ferramenta crucial para capacitar os indivíduos a serem cidadãos críticos informados e engajados em um ambiente digital em constante mudança. É essencial que o processo de aprendizado seja moldado e expandido através da Inteligência Artificial (IA), utilizando plataformas interativas que tornem a experiência educacional mais eficiente e personalizada, adaptando-se às necessidades individuais de cada aluno. Assim, o objetivo primordial da educação é fomentar uma cidadania sólida e adaptável diante dos desafios complexos e, por vezes, sutis que a tecnologia apresenta.

Para alcançar essa finalidade, é vital implementar uma abordagem pedagógica que capacite os indivíduos a investigar, avaliar e se envolver ativamente na criação, uso e aplicação da IA prevenindo que se tornem meros consumidores passivos das inovações tecnológicas. Essa perspectiva visa incentivar a participação cívica em um mundo cada vez mais dominado por algoritmos, promovendo atitudes de responsabilidade, transparência e ética na utilização das novas ferramentas.

Dessa forma, a Inteligência Artificial (IA) não apenas apresenta desafios, mas também uma vasta gama de oportunidades. Ela pode

atuar como um catalisador para uma educação mais inclusiva, flexível e acessível, preparando cidadãos que sejam capazes não só de interagir, mas também de criar e modificar a complexa ecologia dos meios digitais. O objetivo é desenvolver uma geração que utilize a IA como uma ferramenta para promover o bem-estar social, a inovação responsável e a construção de uma sociedade mais justa e equitativa.

A Ecologia dos Meios, IA e Cidadania

A ecologia dos meios oferece uma visão que se baseia na ideia simplista e utilitária da tecnologia, visualizando-a como um sistema complexo e interconectado. A tecnologia não é apenas um instrumento, mas se torna um elemento transformador da sociedade nas esferas cognitiva e cultural. Inspirado por pensadores como Marshall McLuhan (1964), que popularizou a ideia de que ‘o meio é a mensagem’, o conceito de ecologia dos tempos foi reinterpretado para a realidade atual por autores como Lúcia Santaella (2007) e outros especialistas em cultura digital. Em segundo lugar, a partir dessa perspectiva, com as novas mídias e a Inteligência Artificial (IA), mais recentemente, não apenas disseminamos informações, mas também mudamos nossa percepção e afetamos a maneira como pensamos, nos relacionamos e interagimos com o mundo. Nesse cenário, o digital é visto não como algo isolado, mas como parte essencial da realidade, afetando continuamente nossas experiências.

A Inteligência Artificial (IA) surge como um motor de transformação que abrange uma vasta gama de sistemas, desde algoritmos que sugerem conteúdo em serviços de streaming e redes sociais, até ferramentas de IA generativas que conseguem criar textos, imagens e códigos em poucos segundos. Na área da ciência da computação, a IA

tenta replicar a capacidade humana de pensar, aprender e resolver problemas. Isso resulta em sistemas que podemos personalizar ou ensinar no setor educacional, automatizar funções administrativas, criar materiais didáticos e até mesmo servir como tutores virtuais.

A cidadania, entendida como um conjunto de direitos e deveres dentro de uma comunidade política, precisa ser repensada no contexto da esfera pública digital, onde a participação, a deliberação e o acesso à informação são mediados por algoritmos. A cidadania digital significa uso seguro e responsável da Internet. Envolve uma reflexão sobre como as tecnologias digitais afetam a esfera pública, a privacidade e a formação de opinião. Uma cidade digital não usa apenas a tecnologia, mas também questiona, investiga e se esforça para moldá-la de forma ética e socialmente justa. Este conceito é crucial para a ideia de que a educação atual deve formar não apenas pessoas alfabetizadas, mas cidadãos plenamente conscientes da complexa relação entre tecnologia e sociedade.

Análise dos Impactos da IA na Educação e na Cidadania

Um dos avanços impactantes da IA na educação é a capacidade de personalizar a trajetória educacional de cada aluno. Sistemas de tutoria inteligente e plataformas adaptativas têm a capacidade de analisar o desempenho de um estudante em tempo real, identificar suas dificuldades, pontos fortes e estilos de aprendizagem, e por fim, ajustar o conteúdo e o ritmo do ensino. A IA tem se tornado cada vez mais presente na sociedade e na educação, trazendo possibilidades de melhorar a aprendizagem e a experiência dos estudantes, assim como a eficiência e produtividade dos professores.

Por exemplo, um sistema de IA pode sugerir exercícios adicionais sobre um tema que o aluno não compreendeu, ou pode evitar um conteúdo que ele já domina, liberando tempo para atividades mais desafiadoras. Essa é uma grande mudança em relação ao modelo da sala de aula convencional. Esta personalização pode resultar em maior motivação, engajamento e em resultados de aprendizagem mais eficazes, pois cada aluno recebe o suporte necessário, no momento apropriado e conforme as suas necessidades pedagógicas.

Acessibilidade e Inclusão

A Inteligência Artificial (IA) tem a capacidade de tornar a educação um ambiente mais inclusivo e acessível para todos, independentemente de localização geográfica, deficiências ou condições socioeconômicas. É possível adaptar o ensino às características de cada estudante, pois cada aluno possui necessidades e habilidades específicas. A IA pode contribuir com ferramentas de acessibilidade, como softwares de reconhecimento de fala e tradutores em tempo real, leitores de tela, que facilitam o acesso a materiais educacionais por estudantes com deficiência.

No contexto da inclusão, a IA pode ser aplicada em diversas áreas, como no monitoramento do desempenho dos estudantes, na personalização do ensino e na análise de grandes quantidades de dados para identificar tendências e padrões de aprendizagem.

Em relação ao professor, a aplicação da IA pode trazer diversos benefícios onde é possível monitorar o desempenho dos estudantes de forma mais eficiente, identificar problemas de aprendizagem e fornecer feedbacks personalizados. Já do ponto de vista do estudante, a IA pode

proporcionar uma experiência de aprendizagem mais personalizada e adaptada ao seu ritmo e estilo de aprendizagem, além de possibilitar uma maior interação com os professores e colegas de forma virtual.

As tecnologias de tradução em tempo real têm o potencial de acabar com as barreiras linguísticas, tornando possível que conteúdo de alta qualidade sejam acessíveis para um público mundial. Estudantes com problemas auditivos podem se beneficiar de ferramentas de reconhecimento de voz e transcrição, ao passo que a inteligência artificial pode gerar descrições de imagens para alunos com deficiência visual. A capacidade de criar conteúdo sob demanda pode tornar mais acessíveis os materiais didáticos, diminuindo a necessidade de recursos físicos limitados.

Independentemente das necessidades ou origens dos estudantes, a IA tem o potencial de criar um ambiente de aprendizado mais inclusivo e equitativo para todos, no entanto, é importante usar a IA de forma ética e responsável, como ajuda complementar, garantindo a privacidade e segurança dos dados dos estudantes, bem como mantendo o papel crucial do educador.

Uniformização do Pensamento e Bolhas de Filtro

Quando falamos sobre a internet em geral, especificamente sobre as redes sociais, a Inteligência Artificial (IA) tem a capacidade de gerar bolhas de filtro, nas quais os usuários são apresentados unicamente a conteúdos que fortalecem as suas crenças prévias. Na educação, há o perigo de que os sistemas de IA, concebidos para personalizar o conteúdo, acabem restringindo a exposição dos estudantes a diversas perspectivas e conceitos. Se a programação de um algoritmo enviar a um estudante

apenas conteúdos que correspondam aos seus interesses anteriores, a perspectiva de mundo do usuário pode se tornar limitada, prejudicando a sua habilidade de pensamento crítico e a sua capacidade de aceitar diferentes pontos de vista. Isso pode afetar a visão dos estudantes dos acontecimentos mundiais, políticos e sociais, levando-os a acreditar com convicção somente no que lhe é apresentado.

A educação deve ser um ambiente de diferentes pontos de vista, de debate de ideias, estimulando o pensamento crítico. Portanto, a Inteligência Artificial (IA) deve ser utilizada para fomentar a diversidade, e não a uniformidade.

Privacidade e Coleta de Dados Massiva

A adaptação avançada por meio de Inteligência Artificial (IA) requer a obtenção de amplas quantidades de informações sobre a conduta e o rendimento dos estudantes. Isso gera preocupações relevantes relacionadas à privacidade. Quem possui acesso a essas informações? De que maneira elas são guardadas? Para quais propósitos são empregadas?

A utilização comercial de dados no setor educacional pode representar um risco sério, colocando os alunos em situações vulneráveis a anúncios direcionados ou ao uso impróprio de suas informações. O sistema educacional deve capacitar os alunos a estarem cientes de sua presença digital e a demandarem clareza e controle a respeito de seus dados.

A pesquisadora Shoshana Zuboff (2019) defende que nossa vivência digital, abrangendo inclusive o processo de aprendizado, foi convertida em insumo para um modelo econômico inovador que ela chama de capitalismo de vigilância. Nesse contexto, a inteligência

artificial vai além de atuar como ferramenta educativa, funcionando também como um sistema destinado à coleta de dados capazes de prever e até mesmo influenciar o comportamento humano. Ela define o novo modelo econômico digital com base na apropriação dos dados comportamentais:

O capitalismo de vigilância unilateralmente reivindica a experiência humana privada como matéria-prima gratuita para tradução em dados comportamentais. Esses dados são a matéria-prima do seu lucro, uma vez que são transformados em produtos de previsão que preveem o nosso comportamento futuro. (Zuboff, 2019, p. 8)

Essa realidade nos desafia a pensar de forma profunda sobre o papel da educação como um escudo ativo de proteção. É essencial que ela prepare os alunos para serem plenamente conscientes de sua atuação no universo digital, além de habilitá-los a demandar transparência e controle sobre suas informações pessoais. O sistema educacional deve resistir à lógica que mercantiliza os dados, reafirmando que seu objetivo principal é formar cidadãos autônomos e críticos, em vez de reduzir indivíduos a meros produtos destinados a atender às exigências do mercado.

O Problema da Desinformação (Deepfakes e Notícias Sintéticas)

A Inteligência Artificial (IA) possui a habilidade de criar conteúdo autênticos, que incluem textos, áudios e vídeos (conhecidos como *deepfakes*), apresentando um verdadeiro desafio que é a geração massiva de informações enganosas para a cidadania. A linha que separa o verdadeiro do falso está se tornando cada vez mais indefinida, e a

inteligência artificial pode ser utilizada para alterar a percepção pública, afetar resultados eleitorais e disseminar informações equivocadas.

A educação, neste contexto, tem o dever de promover um letramento midiático consistente, fornecendo aos estudantes ferramentas para checar fontes, avaliar a origem das informações e se proteger contra a manipulação feita por algoritmos. Isso vai além da mera detecção de uma informação falsa; é sobre entender como a tecnologia por trás disso funciona.

A Educação como Ferramenta para uma Cidadania Digital Crítica e Ativa

A era digital trouxe uma transformação significativa na maneira como nos conectamos, produzimos e conduzimos nossas vidas, e junto ao enorme potencial tecnológico, surge um cenário desafiante, em vez de os usuários serem apenas agentes passivos da tecnologia, eles se tornam participantes ativos, e a educação é um elemento essencial para capacitar indivíduos a se tornarem cidadãos digitais críticos.

A educação para a cidadania digital vai muito além de ensinar a utilizar aplicativos ou criar senhas seguras, trata-se de fomentar o pensamento crítico. O primeiro passo é questionar a própria essência das tecnologias que utilizamos e o foco da educação deve estar em nos capacitar a enxergar além do conteúdo, reconhecendo o ‘meio’ - o algoritmo, a plataforma, a rede - como os verdadeiros agentes transformadores de nossa cultura e o objetivo é ter a habilidade de distinguir informação relevante de ruído.

A cidadania ativa no ambiente digital ultrapassa a simples capacidade de análise e exige a disposição de agir. Nesse cenário, as

ideias de Douglas Rushkoff (2010) se revelam fundamentais. Segundo ele, o princípio de ‘programar ou ser programado’ simboliza o divisor entre passividade e autonomia. Uma educação voltada para a cidadania digital ativa deve preparar os indivíduos para compreenderem a linguagem tecnológica, questionarem as lógicas subjacentes aos algoritmos e, quando necessário, criarem suas próprias soluções. O cidadão digital ativo não se restringe ao papel de usuário passivo, mas se torna um participante engajado na construção da sociedade online, fomentando debates e desafiando agendas corporativas que se concentram apenas em capturar nossa atenção e extrair nossos dados.

A luta por uma cidadania digital crítica e ativa reflete, em essência, o verdadeiro propósito da educação. De acordo com Shoshana Zuboff (2019), a inteligência artificial pode ser utilizada para extrair dados dentro de uma ‘economia de vigilância’, onde a experiência humana é convertida em matéria-prima. Diante desse contexto, a educação precisa rejeitar a ideia de se limitar à formação técnica, voltada apenas para atender demandas de trabalho. Seu papel mais elevado é capacitar indivíduos a fazer escolhas éticas, respeitar a privacidade e defender sua autonomia. Ao priorizar valores como ética, pensamento crítico e criatividade no centro das práticas educacionais, assegura-se que a tecnologia permaneça subordinada ao benefício humano, evitando que o contrário se torne realidade.

Letramento em Inteligência Artificial (IA)

A educação deve integrar em seus currículos a investigação da inteligência artificial. Isso não implica que todos os alunos precisem se tornar programadores, mas sim que devem compreender os princípios

fundamentais de como os algoritmos operam, como são treinados e de que maneira podem impactar nossa rotina diária. Ideias como aprendizado de máquina, redes neurais e viés algorítmico podem ser apresentados de maneira acessível e contextualizada, possibilitando que os alunos se tornem usuários mais conscientes, críticos e questionadores.

A abordagem de Douglas Rushkoff (2010) em relação à tecnologia digital vai além da superficialidade, incentivando-nos a reavaliar nossa interação com o cosmos interconectado. Para ele, a internet e os meios digitais oferecem um vasto potencial para criação e conexão, no entanto, a maioria das pessoas se encontra em uma posição passiva, o que nos submete a compromissos que não pertencem a nós.

Rushkoff (2010) argumenta que a tecnologia deveria ser uma extensão do ser humano, uma ferramenta para nosso empoderamento. Contudo, o monopólio do como se consolidou, faz com que uma pequena elite tenha o conhecimento necessário para desenvolver o software e as plataformas, enquanto a maioria de nós limita-se a utilizar os aplicativos disponíveis, aceitando as funcionalidades e restrições que nos são impostas. Essa aceitação passiva limita nossa capacidade de determinar o verdadeiro valor da tecnologia em nossas vidas. A obra de Rushkoff (2010) nos convida a uma reflexão sobre a nossa agência na era digital:

A menos que entendamos a linguagem da tecnologia e participemos de sua criação, estamos condenados a aceitar a tecnologia como ela nos é dada. Se você não está programando, você está sendo programado. (Rushkoff, 2010, p. 1)

Essa dinâmica se torna mais evidente na economia da atenção. Os modelos de negócios, especialmente os das redes sociais, são projetados

não para atender ao usuário, mas para capturar e manter sua atenção pelo maior tempo possível. Por meio de algoritmos sofisticados, manipulam nosso comportamento, transformando-nos, de fato, no produto. Nossa identidade digital e nossa interação com o mundo são moldadas por lógicas programadas para o lucro, e não para nosso bem-estar. A reflexão de Rushkoff (2010) representa um apelo urgente à ação. Ela nos desafia a reassumir o controle sobre nossas vidas digitais e a buscar uma alfabetização digital que vá muito além da mera utilização de um aplicativo. Ele sugere que adotemos uma postura ativa, questionadora e criativa, rejeitando a aceitação passiva das normas tecnológicas impostas. A verdadeira liberdade digital, segundo ele, reside na capacidade de compreendermos as regras do jogo e, se necessário, reescrevê-las.

Aprendizagem Baseada em Projetos (PBL)

A Aprendizagem Baseada em Projetos (PBL) representa uma transformação profunda no âmbito educacional, substituindo o modelo tradicional, centrado na transmissão passiva de conhecimento, por uma abordagem dinâmica, voltada para o protagonismo do estudante. O PBL é uma metodologia que se alinha perfeitamente com a necessidade de formar cidadãos ativos. Ao invés de aulas expositivas, os alunos trabalham em projetos práticos que exigem a resolução de problemas do mundo real. Nesse contexto, a IA pode ser utilizada como uma ferramenta para coletar dados, simular cenários ou criar protótipos. O foco não é na tecnologia em si, mas em como ela é usada para atingir um objetivo. Por exemplo, um grupo de alunos pode usar IA para analisar dados sobre a poluição local e propor soluções, aprendendo na prática sobre ética de dados, viés e colaboração.

O foco da PBL está em sua habilidade de unir teoria e prática de forma significativa e o ponto de partida é sempre um problema autêntico ou uma questão instigante e pertinente ao mundo real. Nesse contexto, os alunos não recebem respostas prontas, mas são provocados a agir como investigadores ativos, buscando, analisando e sintetizando informações por conta própria. Essa dinâmica reinventa seu papel na educação, tirando-os da postura passiva para colocá-los como protagonistas de sua jornada.

Durante o desenvolvimento do projeto, os estudantes têm a oportunidade de trabalhar e aprimorar as suas competências, a colaboração e a autonomia tornam-se indispensáveis à medida que aprendem a atuar em grupo, administrar o tempo e solucionar desafios de maneira criativa. O pensamento crítico é continuamente estimulado pela busca de soluções inovadoras, trazendo um significado tangível e um senso de realização ao trabalho desenvolvido.

No fim das contas, a PBL vai além da aquisição de conhecimento, ela forma indivíduos preparados para enfrentar problemas complexos, agir proativamente e colaborar com eficiência. Dessa forma, a metodologia não apenas atende às exigências contemporâneas, como também promove uma cidadania digital crítica e engajada. Ela capacita as novas gerações não só a compreender e navegar pelo mundo, mas também a contribuir para moldá-lo e transformá-lo ativamente.

O *Flow* na Sala de Aula: O Potencial da Imersão para Transformar o Aprendizado

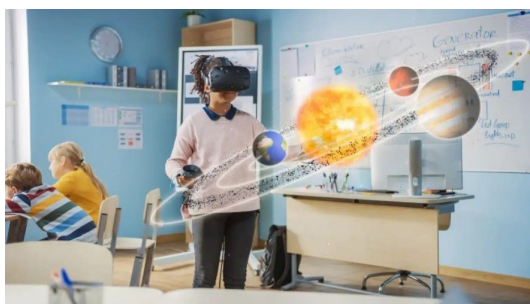
A procura por abordagens pedagógicas que realmente envolvam e incentivem os estudantes é incessante na educação. Nesse contexto,

a Teoria do *Flow*, formulada pelo psicólogo Mihaly Csikszentmihalyi (1990), surge como um conceito impactante para transformar a vivência de aprendizagem. O *Flow* define um estado de total imersão em uma atividade, onde o sujeito se sente revitalizado, concentrado e completamente realizado, perdendo a percepção temporal.

No ambiente educacional, o *Flow* é atingido quando existe um equilíbrio perfeito entre o desafio da atividade e as competências do aluno. Por exemplo, se uma tarefa é demasiado simples, provoca aborrecimento e se é excessivamente complicada, gera ansiedade. O ponto de *Flow* está no meio, demandando foco e empenho, mas permanecendo dentro da capacidade do estudante. O *Flow* torna o ato de estudar uma experiência naturalmente gratificante, estimulando a motivação própria, ou seja, a realização da atividade pelo prazer que ela proporciona, ao invés de se apoiar exclusivamente em notas ou recompensas externas.

Figura 1

Alunos em Estado de Flow com Realidade Virtual e Aumentada.



ScanSource Brasil. (2023).

A assimilação de tecnologias imersivas, como a Realidade Virtual (RV) e a Realidade Aumentada (RA), pode servir como um significativo

impulsionador para o *Flow* na sala de aula, assim como as instalações de arte digital imersivas transportam o público para contextos totalmente distintos, essas tecnologias podem conduzir os alunos a cenários de aprendizagem que seriam inviáveis de reproduzir fisicamente. A bidirecionalidade proporcionada pelas tecnologias imersivas e interativas, como mostrado na figura, transforma o aluno de mero receptor de informações em um participante ativo e imprescindível na construção do seu conhecimento.

A Teoria do *Flow* encontra uma ressonância profunda na Zona de Desenvolvimento Proximal (ZDP) de Vygotsky. A ZDP é a área entre o que um aluno consegue realizar de forma independente e o que ele pode executar com a assistência de um educador ou colega mais experiente. As atividades que favorecem o *Flow* situam-se precisamente nessa zona: são suficientemente desafiadoras para incentivar o desenvolvimento, mas viáveis com o suporte adequado.

Ao compreender a Teoria do *Flow* como um conceito universal de envolvimento, os educadores podem ultrapassar os limites da sala de aula convencional e desenvolver experiências de aprendizado verdadeiramente imersivas e eficientes, aproveitando a força da gamificação e das tecnologias digitais para equilibrar o desafio com a competência, promovendo uma aprendizagem intensa e agradável.

A inteligência artificial (IA) pode ser utilizada para criar materiais de suporte que permitem aos alunos concentrar-se no objetivo da aprendizagem. O papel da IA fornece aos estudantes comandos textuais ou visuais sobre suas necessidades e com base nesses inputs, a IA gera ambientes virtuais em 3D para contextos como aulas de história ou biologia, além de texturas e personagens não jogáveis, com diálogos

elaborados, ideais para simulações de *role-playing*. Embora não necessite programar o ambiente, o aluno assume o papel de diretor criativo e arquiteto conceitual, definindo, como por exemplo, o clima, à época, as regras e os objetivos da simulação, priorizando o *design* instrucional e a narrativa, em vez de enfrentar processos técnicos trabalhosos. Essa dinâmica garante mais tempo para a exploração de conceitos complexos e a experimentação criativa, elementos fundamentais para alcançar o estado de *Flow*, onde há total imersão nas atividades.

A Inteligência Artificial (IA) funciona como uma parceira de *brainstorming* e uma máquina de possibilidades, auxiliando no desenvolvimento narrativo e acadêmico. O material gerado pela IA serve como ponto de partida, não um produto acabado. O estudante é responsável por avaliar, editar e ajustar o conteúdo, incorporando sua própria voz e visão para transformar os rascunhos em trabalhos autênticos e originais. Essa abordagem desenvolve competências cruciais como o pensamento crítico, incentivando o aluno a discernir o que é útil, relevante e fiel ao contexto do aprendizado.

Em resumo, a inteligência artificial generativa desloca o estudante da etapa demorada de criação do zero para uma fase mais dinâmica e produtiva de refinamento, destacando os momentos onde ocorrem aprendizado profundo e *Flow* criativo.

Estratégias Pedagógicas para Incentivar o Estado de Flow

A aplicação dos princípios do *Flow* para o contexto educacional requer a adoção de estratégias pedagógicas intencionais que promovam o envolvimento profundo e a motivação inerente dos estudantes. Primeiramente, é fundamental proporcionar retorno contínuo e

construtivo. A utilização de ferramentas digitais, por exemplo, viabiliza a oferta de *feedback* imediato, primordial para que os alunos possam acompanhar seu progresso, entender erros em tempo real e ajustar suas estratégias de aprendizagem. Essa dinâmica estimula a autoavaliação e a capacidade de adaptação constante, elementos vitais para a preservação do estado de *Flow*. Além disso, torna-se imprescindível oferecer desafios adequados e progressivos.

As atividades devem ser meticulosamente adequadas ao nível de habilidade de cada aluno, evitando tanto a monotonia decorrente de tarefas excessivamente fáceis quanto a frustração provocada por dificuldades excessivas. A disponibilização de diferentes níveis de complexidade e a implementação de elementos de gamificação como níveis, missões e recompensas, são instrumentos eficazes para sustentar o envolvimento crescente e manter o aluno na zona de desafio ótimo, característica do *Flow*.

Outro pilar para a imersão é a criação de ambientes imersivos e interativos. A incorporação de tecnologias como a Realidade Virtual (RV) e a Realidade Aumentada (RA) permite a realização de simulações e explorações que ultrapassam as limitações do espaço físico da sala de aula. Além disso, o potencial da Inteligência Artificial (IA) generativa para elaborar narrativas adaptáveis e desenvolver tutores virtuais personaliza a experiência de aprendizagem, tornando-a mais responsiva às necessidades individuais e, conseqüentemente, mais imersiva.

Por fim, é fundamental estimular a autonomia e o controle por parte dos estudantes, permitindo que participem das decisões referentes ao seu próprio aprendizado, como a escolha de tópicos, métodos de pesquisa ou formas de apresentação, e incentivar a resolução independente

de problemas cria um senso de propriedade sobre o processo educativo. Essa autonomia contribui diretamente para a sensação de controle, um componente essencial do *Flow*, onde o aluno se sente protagonista de sua jornada de conhecimento.

Conclusão

A Inteligência Artificial (IA) já se tornou uma parte essencial de nossa ecologia de ferramentas, exercendo influências significativas na educação e na promoção da cidadania. Este estudo se propôs a ultrapassar a visão simplista que opõe IA como solução e IA como ameaça, oferecendo uma análise mais equilibrada e fundamentada. Ficou evidente que essa tecnologia possui um extenso potencial para transformar a educação, tornando-a mais personalizada, acessível e alinhada às demandas do século XXI. No entanto, sua implementação requer uma abordagem crítica e atenção constante aos desafios que apresenta, como o viés algorítmico, a padronização do pensamento e os riscos associados à privacidade.

Nesse contexto, a formação de uma cidadania digital crítica e participativa se destaca como o objetivo central da educação em um ambiente dominado por algoritmos. Trata-se não de rejeitar a tecnologia, mas de preparar uma nova geração para utilizá-la de maneira responsável, promovendo seu envolvimento ativo no desenvolvimento e na evolução da própria IA. O foco deve ir além do simples ensino de como navegar no espaço digital, é fundamental incentivá-los a criar, questionar e transformar essa tecnologia em uma ferramenta voltada para o bem-estar social, inovação ética e construção de uma sociedade mais justa e igualitária.

Investigações mais extensas são cruciais para compreender os impactos cognitivos e emocionais do aprendizado personalizado e imersivo mediado pela IA, em comparação aos métodos tradicionais de ensino. É necessário desenvolver abordagens voltadas a ambientes que promovam o estado de *Flow*, explorando como a IA pode facilitar um equilíbrio dinâmico entre os desafios propostos e as capacidades dos estudantes.

Embora o percurso a ser trilhado seja desafiador e extenso, este estudo se propõe a ser apenas um ponto de partida nesta jornada. Pesquisas futuras podem contribuir ao aprofundar temas como políticas públicas que regulam o uso da IA na educação, capacitação de educadores para desempenhar seus papéis nesse novo contexto e os impactos da tecnologia em diferentes realidades socioeconômicas. A urgência do debate é indiscutível, e a resposta para os desafios impostos pela IA reside, essencialmente, na educação: uma educação transformadora, que empodera os indivíduos e os inspira a agir com ética, responsabilidade e consciência no universo digital.

Referências

- Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. Harper & Row.
- McLuhan, M. (1964). *Understanding media: the extensions of man*. McGraw-Hill.
- Rushkoff, D. (2010). *Program or be programmed: Ten commands for a digital age*. OR Books.

Santaella, L. (2007). *A ecologia da comunicação: O meio, a mensagem e o sujeito*. Paulus.

ScanSource Brasil. (2023, outubro 11). Realidade Aumentada e Virtual na Educação: Uma Jornada Inovadora Recuperado de <https://scansource.com.br/blog/realidade-aumentada-e-virtual-na-educacao-uma-jornada-inovadorareal/>

Zuboff, S. (2019). The age of surveillance capitalism: The fight for a human future at the new frontier of power. *PublicAffairs*.

LETRAMENTO MIDIÁTICO CRÍTICO E FORMAÇÃO DOCENTE: EXPERIÊNCIAS COM TDIC E INTELIGÊNCIA ARTIFICIAL EM CONTEXTOS ESCOLARES

Leandro Marlon Barbosa Assis¹

Lucas Lima Coaracy²

Alexandre Farbiarz³

O século XXI consolidou uma sociedade profundamente imersa em ambientes digitais, onde mídias e tecnologias permeiam a vida cotidiana e produzem efeitos ambivalentes: de um lado, ampliam possibilidades comunicacionais e pedagógicas; de outro, reforçam desigualdades e naturalizam estruturas de poder. Nesse cenário, discutir o Letramento Midiático Crítico (LMC) (Assis & Farbiarz, 2024, Bastos, 2024) é

-
1. Universidade Federal Fluminense
leandromarlon@id.uff.br
 2. Universidade Federal Fluminense
lucaslc@id.uff.br
 3. Universidade Federal Fluminense
alexandrefarbiarz@id.uff.br

fundamental para problematizar a primazia atribuída às tecnologias na educação e superar o senso comum.

Neste trabalho, diferenciamos “tecnologias” e “mídias”, compreendendo “mídias” como formas de produção de sentidos e de organização social vinculadas a funções de controle e intencionalidades históricas (Assis, 2019; Belloni, 2018; Feenberg, 2010; Gonçalves & Farbiarz, 2024; Vieira Pinto, 2005; Selwyn, 2014). Tal perspectiva permite deslocar o foco do mero instrumental para as lógicas comunicacionais que configuram práticas docentes, sobretudo diante da emergência da inteligência artificial (IA) e seu discurso de mediadora de interações educativas.

É nesse horizonte que situamos a experiência do workshop Tecnologias e Prática Docente, realizado com educadores da Educação Básica, no âmbito das pesquisas de doutorado em andamento dos autores. O workshop buscou articular reflexão crítica e experimentação prática sobre o uso de mídias (Assis, 2019) e IA (Coaracy & Farbiarz, 2025), estimulando a construção colaborativa de práticas pedagógicas alinhadas ao LMC (Assis & Farbiarz, 2024, Bastos, 2024).

Este artigo tem como objetivo discutir os desdobramentos dessa experiência, evidenciando como processos formativos podem fomentar a apropriação crítica das mídias digitais e da IA. Ao articular teoria e prática, pretendemos contribuir para o campo da Comunicação e Educação propondo caminhos para a formação docente que ultrapassem o modelo transmissivo e tecnicista, promovendo uma postura crítica e cidadã frente às tecnologias.

A mediação tecnológica: do instrumentalismo às lógicas da cotidianidade

Quando projetamos a transição de perspectivas do instrumentalismo à compreensão das lógicas que operam na cotidianidade (Bastos, 2024), percebemos a importância do debate sobre formação docente no sentido e busca pela práxis pedagógica crítica e emancipatória (Assis & Farbiarz, 2024). O foco reside no reconhecimento de que tecnologia não pode ser tratada como um fim em si mesma, mas como objeto permeado por intencionalidades políticas e sociais (Vieira Pinto, 2005, Feenberg, 2010, Selwyn, 2014).

Nosso problema, está fundamentado na concepção de que o uso acrítico de tecnologias educacionais por parte dos educadores é normalizada, inclusive, na consolidação da introdução acelerada de novidades que propicia debate formativo puramente técnico. Entretanto, direcionar a análise para uma ótica puramente instrumental é insuficiente, pois o foco excessivo na instrumentalidade tecnológica corre risco de ofuscar a singularidade e a intencionalidade pedagógica que definem a verdadeira essência do processo educacional.

Uma das mais nocivas substancializações que cometemos quase inconscientemente, passando assim despercebida, é a que se refere à cultura. Aparece-nos como uma realidade em si. Definimos então as técnicas declarando-as pertencentes a certa cultura, substantivada, entificada, quando a verdade encontra-se na expressão inversa. São as técnicas, enquanto ações humanas concretas, que têm valor primordial porque se referem à relação direta de caráter problemático, do homem com o mundo, ao passo que a cultura designa apenas o conjunto delas em determinado tempo e lugar, mais as crenças e valores a elas agregados. (Vieira Pinto, 2005, p. 65)

Acompanhando as observações do autor, argumentamos que letramento e cultura não podem ser reduzidos a uma realidade fixa. Passa a ser essencial compreendermos que ensinar/praticar o LMC requer ação concreta de interação e comunicação dialógica a partir de situações reais do cotidiano. A mediação tecnológica, quando desprovida de criticidade, alinha-se rapidamente à ‘cotidianidade’, termo que designa a força coercitiva e hegemônica que busca a padronização das vidas (Lefebvre, 1991). A ‘cotidianidade’ é um agenciamento pasteurizador que transforma as relações humanas em mercadorias e busca dominar os hábitos sociais em nome da racionalidade unidimensional (Marcuse, 1973).

Para superar essa alienação imposta, devemos buscar materialidades das condições de vida e adotar postura crítica oposta à “inversão da práxis”. A proposta da pesquisa, ancorada no LMC, já foi pensada para o universo dos *game studies* (Bastos, 2024) e formação de educadores (Assis & Farbiarz, 2025). Aqui, buscamos oferecer um caminho, a partir das experiências teóricas de Coaracy & Farbiarz (2025), para que educadores possam perceber as contradições que são postas pela ‘cotidianidade’, no que se refere às IA de maneira mais prática e no chão da escola.

O LMC: práxis docente contra a Inversão do sentido

O LMC surge como uma concepção teórica e prática tida como essencial no contexto educacional contemporâneo, visando encontrar significação e refletir sobre a penetração das mídias na vida cotidiana. A abordagem pretende debater como a prática pedagógica dos educadores é frequentemente invertida de valor intrínseco, sendo direcionada por

forças da cotidianidade. Como Bastos (2024, p. 157) indica, ao propor que devemos “conhecer, aprender, questionar e manipular diferentes tecnologias, aparelhos, formatos, códigos e outras estruturas midiáticas” para uma transformação do cotidiano (Kellner & Share, 2008, Kellner & Khan, 2015).

O objetivo central do LMC é examinar propostas que, ao se apropriarem de elementos técnicos, falham em promover práticas reflexivas que transformem o cotidiano em um sentido oposto ao desenhado pela cotidianidade. Nossa concepção teórica busca oferecer uma alternativa frente aos interesses hegemônicos, combatendo a dominação cultural e o empobrecimento do pensamento que limita o educador. O LMC é, em essência, uma resistência à visão mercadológica e reducionista das mídias ao aspecto tecnológico, propondo que a educação desempenhe um papel fundamental na promoção da compreensão crítica das mídias, capacitando os sujeitos a entenderem e transformarem a realidade tecnológica em que estão inseridos.

Ao identificarmos a funcionalidade do letramento, podemos ressaltar nosso afastamento ao ponto de, inclusive, desejarmos uma disfuncionalidade proposital ao sistema posto como natural, civilizado ou mesmo letrado pela AMI (Wilson et al., 2013). Ao adotarmos, assim, uma perspectiva “revolucionária” (Soares, 2009), concordamos que

letramento não pode ser considerado um ‘instrumento’ neutro a ser usado nas práticas sociais quando exigido, mas é essencialmente um conjunto de práticas socialmente construídas que envolvem a leitura e a escrita, geradas por processos sociais mais amplos, e responsáveis por reforçar ou questionar valores, tradições e formas de distribuição de poder presentes nos contextos sociais. (Soares, 2009, pp. 75-76)

Igualmente, não desejamos um modelo de letramento exógeno, como aquele que imputa um conhecimento tido como inexistente ao outro ou, ainda, de modo endógeno, em que exerça uma influência hegemônica naturalizada (Street, 2014). O conceito se apoia em Lefebvre (1991), na premissa de que a “vida cotidiana” (a vida do indivíduo em sua totalidade, com individualidade e personalidade) é o ponto de observação central para analisar os processos de midiaticização, que introduzem uma lógica externa, a “cotidianidade”, orientada para padronizar e controlar ações. Tal lógica busca a “inversão da práxis”, que leva à alienação dos indivíduos dos processos histórico-sociais correspondentes à produção humana. O LMC, ao resgatar a práxis, busca manipulação intencional e consciente das mídias, permitindo que educadores sejam projetistas (Vieira Pinto, 2005) de suas próprias práticas.

A fundamentação do LMC está ancorada na perspectiva crítica da filosofia da tecnologia e da cultura (Feenberg, 2003), afastando-se do instrumentalismo, substancialismo e determinismo tecnológico (Neder, 2013). O instrumentalismo vê a tecnologia como mero instrumento neutro; o substancialismo a trata como entidade autônoma; e o determinismo afirma que é a causa principal da transformação social. O LMC, contudo, aproximando-se da teoria crítica feenberguiana, rejeita o determinismo tecnológico (Williams, 2016) ao afirmar que a mídia não é causa autônoma, mas efeito das necessidades e intenções sociais, econômicas e políticas (Vieira Pinto, 2005).

Assim, a mídia, em sua relação com a sociedade e reconhecendo-a como não neutra, mas imbuída de valores e interesses, especialmente os hegemônicos que visam o consenso, deve analisar criticamente sua existência e verificar a compatibilidade com valores democráticos

e sociais. Essa abordagem crítica é necessária porque, na realidade escolar, a “inserção tecnológica” é frequentemente utilizada para reproduzir valores do capital, processo denominado “pedagogia do capital” (Shiroma et al., 2002). A “cotidianidade” introduz uma lógica que busca padronizar ações através da naturalização da ciência e tecnologia como sinônimos de progresso social e neutralidade, como se não fossem instrumentos ideológicos (Selwyn, 2017). Essa lógica, ao se manifestar no cotidiano escolar, promove a unidimensionalidade (Marcuse, 1973), simplificando o processo capitalista sob a égide de um progresso inquestionável e normatizado.

O LMC atua, portanto, no tensionamento da técnica para permitir a reconquista da práxis pelos educadores, sendo entendida essencialmente como a ação humana intencional no mundo, sendo diferenciada da adaptação passiva (Freire, 2015). O letramento proposto busca revelar os problemas enfrentados, combatendo a dominação cultural e o empobrecimento do pensamento, que fragilizam as experiências formativas. Tornando-se sujeitos projetores (Vieira Pinto, 2005), os educadores podem dotar de sentidos suas práticas, compreendendo que tecnologias são fruto do trabalho humano intencional, tal qual intelectuais (Gramsci, 2022). A ausência dessa intencionalidade leva à apropriação de mídias de forma linear, administrativa, inexpressiva e acrítica, resultando em apropriações unívocas e vazias de sentido na esfera educativa.

O LMC se filia a uma concepção “forte” e “ideológica” de letramento (Soares, 2009; Street, 2014), vendo como um processo interacional, social, cultural e político, contrapondo-se aos modelos “fracos” ou “autônomos”. Modelos autônomos, como a Alfabetização Midiática Informacional (AMI) da UNESCO (Wilson et al., 2013), são criticados

pelo propósito de “transmitir esse conhecimento aos usuários”, focando na posse individual de habilidades funcionais (grafocêntricas). Embora a AMI incorpore conhecimentos essenciais sobre as funções da mídia em sociedades democráticas e a avaliação de seu desempenho, o LMC aponta sua falha ao não abordar explicitamente o viés ideológico, dos interesses político-econômicos e das relações de poder que permeiam as instituições de informação.

Essa abordagem capacita os sujeitos no envolvimento com leitura e escrita multimodal, garantindo o exercício de diferentes sistemas de linguagem. O objetivo é que os aprendizes consigam atuar num mundo multicultural e interconectado, capazes de desmascarar discursos direcionadores das ações pedagógicas e ocultos na moderna ilusão do “novo”. Sob estas bases desenvolvemos nosso workshop e entenderemos as IA.

IA: riscos éticos e o colonialismo de dados

Dentre as consideradas inovações, trajadas de ambiciosas premissas de potenciais e utilidades ilimitadas, encontramos recentemente e intensamente a Inteligência Artificial. Quando discutimos IA, abordamos principalmente a Inteligência Artificial Generativa (IAGen), capaz de produzir conteúdos verbais e audiovisuais. Outrora pouco acessível, desde 2014, o setor privado ultrapassou a liderança acadêmica (Maslej et al., 2023). Protagonizado pela empresa *OpenAI*, diversas empresas oferecem o uso gratuito, embora limitado, dessas ferramentas. Tal abertura, somada a aperfeiçoamentos significativos nas capacidades dessas IAs, com agressivas campanhas de vendas para suas perpetuações em massa, alimentam um cenário de ampla e crescente adoção.

Como estratégia, *Big Techs* perpetuam o hype da atual visão sobre IAs (Bareis et al. 2023), frequentemente apostando em sua importância para o futuro. Verdegem (2022) argumenta que vivemos um Capitalismo de IA, que comporta novas lógicas de produção, trabalho, circulação de capital e exploração do trabalho. Com altos custos de desenvolvimento e infraestrutura para sua manutenção, há concentração de poder nas poucas megacorporações que as detêm. Uma profunda extração, controle e comercialização de nossos dados, muitas vezes opacos (Pasquale, 2015), são realizados por estes poucos agentes, se consolidando nas vozes dominantes sobre como pensar IA e os rumos a tomar.

Tal hegemonia cristaliza uma visão de IA favorecendo essa elite. De acordo com Crawford (2021), maneiras distintas de concebermos IA moldam como é desenvolvida e controlada. A autora aponta que, se os sistemas de IA são vendidos como mais confiáveis do que humanos, cedemos espaço para legitimá-las em decisões críticas, transferindo autoridade para sistemas opacos. Quando pensamos apenas em seu âmbito de inovação tecnológica, desconsideramos os custos sociais e ambientais de sua fundamentação.

Portanto, acenamos a dimensão de disputa ideológica que permeia a própria concepção de IA. Podemos pensá-la enquanto “agentes que recebem perceptos do ambiente e realizam ações” (Russell & Norvig, 2022, p. 8), e apontarmos que são formas de produção de conhecimento e organização social (McQuillan, 2022). Nesse contexto, defendemos que o LMC desempenha papel fundamental no desenvolvimento e estímulo a possibilidades de resistência frente ao poder hegemônico das *Big Techs*. Ao questionarmos seus propósitos, modos de treinamento, delimitaremos seus contextos sócio-históricos e entenderemos

melhor seus funcionamentos, afastando-nos da aceitação passiva para apropriação crítica.

Embora sejam vendidas como ferramentas neutras e otimizadas, IAs carregam significativos vieses (Resnik, 2024) em seus funcionamentos. IAs dependem de bancos de dados que são constituídos em proporções de representações significativamente diferentes. Johnson et al. (2022) apontam, por exemplo, os dados de treinamento da versão do ChatGPT-3, onde Inglês tem 93% de presença. O segundo lugar, Francês, é de 1,8%. Essa disparidade acentuada traz diversas implicações, como a produção de conteúdos mais ajustados aos valores predominantes ocidentais, variações de qualidade entre gerações em inglês e em outras línguas (Petrov et al., 2023) e até mesmo vieses nas avaliações feitas por IAGens, que diferenciam textos de falantes nativos de inglês daqueles produzidos por não-nativos (Liang et al., 2024).

No entanto, nossas críticas não devem limitar-se somente à completude e representatividade dos dados para treinamento de IAs. Devemos questionar como tais dados são coletados, pois a recente expansão de IAGens é marcada por polêmicas e disputas judiciais envolvendo direitos autorais (Sheng, 2023). Muitos modelos de IA crescem às custas de conteúdos já presentes na internet, sem consentimento ou eventual conhecimento dos autores. Tal prática é representativa do que autores como Couldry & Mejias (2019) denominam como colonialismo de dados. Nessa lógica e infraestrutura, empresas de tecnologia primariamente do norte global buscam uma transformação generalizada de todos os aspectos humanos em dados.

O processo de amplificação das IAs coloca cidadãos de países periféricos em posição especialmente problemática: Para que esses

sistemas sejam capazes de transcrever vozes, identificar objetos ou reconhecer emoções em textos, é necessário que sejam alimentados com grandes volumes de dados e submetidos a testes constantes. Contudo, raramente é divulgada a centralidade da presença humana, muitas vezes trabalhadores do sul global, registrando, por exemplo, que determinada imagem contém uma bola ou que um texto expressa tristeza. Esse trabalho de anotação e classificação, indispensável ao funcionamento da tecnologia, é em grande parte realizado por meio de plataformas de *crowd work*.

Empresas desenvolvedoras de IA recorrem majoritariamente a trabalhadores do sul global, remunerados com valores mínimos para executar tarefas de alta volumetria e baixo reconhecimento. É um trabalho precarizado, opaco e frequentemente confidencial. Casos documentados mostram que trabalhadores contratados pela Sama, empresa de anotação de dados a serviço da *OpenAI*, foram expostos a conteúdos altamente gráficos, objetivando treinar sistemas para reconhecer e bloquear materiais nocivos (Dzieza, 2023; Perrigo, 2023). A *Amazon Mechanical Turk*, por exemplo, é descrita como plataforma que oferece serviços de curta duração executados por humanos (Kalil, 2019), mas que, na prática, tornou-se um dispositivo de exploração laboral distribuído globalmente, incluindo o Brasil. Trabalhadores brasileiros relatam dificuldades adicionais, como o bloqueio arbitrário de contas pela plataforma, que aprofunda a instabilidade e insegurança dessa modalidade de emprego (Grohmann et al., 2022).

Ressaltamos que esse contexto não é inaugurado pela IA, mas é aprofundamento da dinâmica neoliberal reforçando o colonialismo (Silveira, 2021). Compreendemos que não é ela sozinha a fonte dessa

lógica mercadológica e dominante da qual discutimos. Outrossim, destacamos a importância do LMC na medida que nos permite analisar os processos e valores do modo como são propulsionadas essas mídias. Se entendemos a tecnologia como a “fonte do mal”, corremos o risco de limitarmos nossa resistência a rejeitá-las por completo. No campo educacional, tal medida é exemplificada na (tentativa de) proibição de recursos de IA por educadores.

Defendemos que a recusa não necessariamente é proveniente da práxis pelos educadores. Realmente, parte também da apropriação acrítica - como quando, por exemplo, está ligada somente à resistência em repensar ou adaptar práticas pedagógicas cristalizadas. Tal via pode ter, em sua origem fundamentadora, não uma reflexão sobre os preceitos e práticas que regem a mídia, mas desconhecimento de seu funcionamento. A rápida penetração dessa tecnologia no cotidiano escolar, somada à falta de programas de capacitação e formação crítica para educadores em IA, leva-nos ao cenário de posturas potencialmente reducionistas, ainda que situadas em lugares opostos.

Conforme defendem Freire e Guimarães (2013, p. 92), “a crítica nossa tem que ser política, e não tecnológica. [...] eu não sou contra o computador; o fundamental seria nós podermos programar o computador. É a questão do poder: é saber a serviço de quem ele é programado para nos programar”. Por esse viés, compreendemos que o LMC é configurado como processo formativo que tanto estimula reflexão sobre o poder, quanto busca a desmistificação da “programação do computador”. Se pretendemos (re)construir outra IA, não dominada pela visão única das *Big Techs*, uma aproximação crítica é essencial. Portanto, propomos

o workshop como espaço para exercício inicial desse movimento de apropriação crítica e tensionamento das mídias.

Do diagnóstico à experimentação: o workshop enquanto teoria para o fomento da postura projetora

O workshop Tecnologias e Prática Docente foi realizado em duas escolas da rede do Estado do Rio de Janeiro, em setembro de 2025, objetivando explorar ferramentas digitais e de inteligência artificial (IA) no cotidiano pedagógico de forma crítica e significativa. Inspirada em Paulo Freire (2015), a proposta partiu da premissa de que a tecnologia não deve ser encarada como fim em si mesma ou mera metodologia, mas como parte das relações humanas, vinculadas a intencionalidades políticas e sociais.

O ponto de partida consistiu na problematização do binômio “o humano e as máquinas”, trazendo referências de Vieira Pinto (2005) e Andrew Feenberg (Neder, 2013) para questionar concepções instrumentais e deterministas da tecnologia. Foram também discutidos os riscos da aceitação acrítica da IA, destacando os debates sobre desinformação, vieses algorítmicos e colonialismo de dados (Silveira, 2021). Essa dimensão teórica foi intercalada com momentos interativos, como enquetes e jogos didáticos, de modo a engajar os educadores e conectar reflexão crítica com prática escolar.

O perfil dos participantes revelou predominância de docentes com mais de dez anos de experiência e nível intermediário de familiaridade digital (3,3/5). As ferramentas mais conhecidas e utilizadas previamente eram Canva, Google Forms e Kahoot, embora já houvesse registros de usos pontuais de IA. Entre as expectativas relatadas, destacamos o

desejo de conhecer novas plataformas, atualizar práticas pedagógicas e refletir sobre limites e possibilidades do uso crítico das tecnologias.

O workshop teve três momentos. Primeiro, uma exposição dialogada sobre desafios da cultura digital na educação e sobre a presença crescente da IA em contextos sociais e escolares. Em seguida, realizamos demonstração de ferramentas digitais e IAs, divididas por eixos de planejamento, produção de conteúdo, apresentação e avaliação. Dentre elas, destacaram-se ChatGPT, Perplexity, Consensus, Canva, Gamma, Kahoot, Mentimeter, Plickers e GradePen. Por fim, os participantes não realizaram atividades práticas em grupos, testando comandos e elaborando planos de aula que incorporaram ferramentas exploradas em virtude da logística de tempo do workshop.

A proposta metodológica valorizou a experimentação prática realizada ao longo da realização do workshop, mas com mediação reflexiva. Em cada etapa, os docentes eram convidados a testar recursos e refletir criticamente sobre suas condições de produção, interesses econômicos e potenciais usos pedagógicos. Questões como autoria, autenticidade de conteúdos e implicações ético-políticas foram levantadas como dimensões indispensáveis do LMC.

O workshop foi um espaço de formação continuada articulado às pesquisas de doutorado em desenvolvimento, permitindo observar como educadores dialogam com as IAs e as TDIC. Mais do que uma experiência pontual, a atividade buscou contribuir para a ressignificação do uso das tecnologias digitais na escola pública, incentivando os educadores a assumirem postura de apropriação crítica e criativa diante das mídias.

A oficina foi pensada como uma etapa seguinte, mais prática, propiciando, efetivamente, o contato dos educadores com as IAGen

e, por conseguinte, maior volume de dados a serem analisados. Ela objetivou capacitar educadores da Educação Básica para integrar, de forma crítica e prática, ferramentas digitais e de inteligência artificial em seu planejamento e execução de aulas.

A estrutura da oficina foi organizada em cinco eixos progressivos: (1) diagnóstico e motivação, discutindo se a IA seria de fato relevante para a prática docente; (2) primeiro contato, explorando o funcionamento básico e os limites das respostas das ferramentas; (3) exploração guiada, na qual os educadores adaptaram recursos de IA às suas disciplinas específicas; (4) crítica e ética, abordando riscos de vieses, desinformação, dependência tecnológica e plágio; e (5) produção colaborativa, culminando na elaboração de planos de aula coletivos e individuais com apoio de IA.

As ferramentas mobilizadas ao longo da oficina incluíram ChatGPT (planejamento e criação de conteúdos), Suno.AI (produção de áudio e música), Kahoot! e Plickers (interatividade e avaliação), e Google Forms (avaliações diagnósticas e coleta de feedback). Cada recurso foi apresentado em sua dimensão prática, mas acompanhado de reflexões sobre suas implicações pedagógicas, éticas e políticas. Assim, os participantes foram convidados a “aprender a usar” e a “aprender a questionar” as tecnologias, destacando seu papel como mediadores críticos.

A produção coletiva de planos de aula mostrou-se ponto central da experiência. Ao explorar comandos (prompts) adaptados a diferentes componentes curriculares, os educadores puderam visualizar como a IA pode auxiliar na personalização de atividades, sem perder de vista a intencionalidade pedagógica. O exercício reforçou que a IA deve apoiar

a prática docente sem substituí-lo, acompanhada por reflexão crítica, análise dos limites e adaptação ao contexto escolar.

Ao final, a oficina configurou-se como espaço de formação continuada com teoria e prática dialogando integradas. A mediação articulou experimentação de ferramentas digitais com debates sobre ética, desinformação, colonialismo de dados e necessidade de autonomia docente. O percurso permitiu ampliar sua familiaridade com a IA, reduzir inseguranças iniciais e, sobretudo, reconhecer a importância do LMC na apropriação responsável e criativa dessas tecnologias.

A práxis sob tensão: análise dos níveis de letramento e intencionalidade

Os respondentes têm ao menos dez anos de tempo de magistério, mas tendem a ter nível básico/moderado de familiaridade com elementos digitais. Dentre as ferramentas disponíveis, o Google Forms é a mais familiar, mesmo que dois educadores tenham indicado que nenhuma ferramenta listada foi utilizada, permitindo inferir que alguns não têm contato e prática docente com as tecnologias.

O desejo dos respondentes traduziu-se na aprendizagem prática de ferramentas para uso em sala de aula, reconhecendo a dualidade da potencialidade e necessidade das tecnologias. Destacamos respostas indicando ceticismo ou ressalvas quanto ao papel da tecnologia ser o principal elemento para mudança do cotidiano escolar. Tal cenário guarda, ainda, a diferenciação da escuta entre as escolas: a primeira, possui educadores buscando tornar suas aulas engajadoras e interativas, enquanto que outros buscam desenvolver a formação tecnológica dos educandos. Enquanto síntese, a maioria dos respondentes indica

equilíbrio cauteloso, com potencial crítico dependente essencialmente do modo de uso.

O uso de IAGen varia entre escolas. Enquanto que na Escola A identificamos familiaridade alta do uso de modelos LLM, a Escola B possui uma minoria que utiliza e/ou utilizou tais ferramentas. Embora em ambas tenhamos a cautela como palavra-chave, os educadores da Escola A têm uma visão ligeiramente mais otimista e crítica/emancipadora de seu potencial, com mais menções para “ruptura” e “pensamento crítico”. Os dados revelam essa disparidade, tendo os educadores da Escola A utilizado previamente ferramentas de IAGen e, com isso, a mediação pudesse avançar mais rapidamente para as fases de crítica e ética (eixos 4 e 5 da oficina). A alta incidência de IA na Escola A levanta a hipótese de “apropriação unívoca e vazia de sentido” (linear e acrítica), reforçando a urgência de tensionar o uso dessas ferramentas para além da comodidade.

Ao analisarmos expectativas e potencial crítico, emerge a intencionalidade (a práxis) que os educadores buscam imprimir no uso das mídias digitais, conforme Tabela 1.

No contexto do LMC, essa diferença é crucial: enquanto o grupo da Escola B demandava primeiro contato mais funcional (o “primeiro contato” da oficina) para diminuir a insegurança inicial, o grupo da Escola A já possuía instrumental estabelecido, permitindo que a mediação pudesse avançar mais rapidamente para as fases de crítica e ética (eixos 4 e 5 da oficina). Embora também busque dinamismo, o discurso deste grupo conecta-se de forma mais direta com o cerne do LMC, ao mencionar necessidade de desenvolvimento do senso crítico e responsabilidade social em relação à desinformação. O uso da

tecnologia é enquadrado como meio problematizador do acesso ilimitado à informação e reorientar a práxis docente no sentido emancipador. Na Escola B, contudo, a busca dos educadores se orientava mais pelo aprendizado de ferramentas para diversificar e tornar suas aulas mais atrativas/interativas. Num paralelo simplificado, os educadores da Escola B buscam “adaptação” aos educandos do século XXI por meio de uma relação de tecnologia com engajamento. Na Escola A, todavia, embora também seja reconhecida a busca pelo dinamismo, há um elemento de problematização do fluxo informacional e busca de uma reorientação da prática docente.

Tabela 1

Intencionalidade Pedagógica e Visão Crítica das TDIC por Escola (Diagnóstico Pré-Workshop)

Categoria	Escola A	Escola B	Escola C
Foco da Expectativa	Conhecer novas ferramentas e, adicionalmente, refletir/debater sobre a prática docente.	Aprender a usar para diversificar e tornar as aulas atrativas/interativas.	Conhecer novas ferramentas e, principalmente, atualizar e aperfeiçoar conhecimentos para aplicação em sala de aula.
Visão sobre a Contribuição para Prática Reflexiva	Menciona o dinamismo, mas com foco explícito no pensamento crítico, na responsabilidade social em meio ao Fake News e no estímulo à pesquisa.	Enfatiza o engajamento do aluno (aluno habituado à tecnologia, aula interativa/objetiva) e a formação tecnológica.	Foco na promoção da autonomia e criticidade dos alunos, permitindo pesquisar, comparar opiniões e romper com práticas tradicionais.

Elaborado pelos Autores.

Ao observarmos os dados absolutos do formulário anterior do workshop, já incluindo as respostas da Escola C, na qual não conseguimos aplicar a oficina, identificamos um consenso mais cauteloso nessa dicotomia entre o instrumentalizar e o problematizar discutido a partir das Escolas A e B. Identificamos que a maioria dos docentes rejeita as visões extremas (instrumentalismo ou determinismo substancialista) e compreende o potencial transformador das tecnologias condicionado à intencionalidade do sujeito (a práxis, o sujeito projetor de Vieira Pinto, 2005).

A participação de respondentes foi diminuindo a cada novo formulário para organização de resultados e percepções. Acreditamos que a dinâmica da rotina de trabalho possa ter dificultado, apesar dos inúmeros contatos e solicitações desse retorno. O número de respondentes de nosso formulário pós-workshop foi reduzido para 16 respondentes, mas tende a confirmar, inicialmente, o sucesso da experiência formativa na articulação entre reflexão crítica e experimentação prática, conforme o modelo do LMC. Para além da aprovação da organização do encontro, ampliamos a visão crítica sobre integração das tecnologias. Docentes relataram que tecnologia não é apenas ferramenta, mas “espaço de disputa ideológica e cultural” (Educador A) e que é preciso questionar suas “intenções, ideologias e impactos” (Educador B). Essa percepção alinha-se diretamente à concepção forte e ideológica de letramento, contraposta ao modelo autônomo.

O workshop reforçou a ideia de que o potencial da tecnologia “depende do contexto e da intencionalidade pedagógica” (Educador A), ecoando o conceito de práxis como ação intencional e consciente no mundo. O Educador C, por exemplo, superou sua “rejeição por

implementar tecnologias” ao vislumbrar novas possibilidades, mas já com a ressalva de que precisará “fomentar novas maneiras” de avaliação, demonstrando a incorporação da crítica à técnica.

A reflexão sobre o ChatGPT foi crucial, pois demonstrou que comandos genéricos resultam em “conteúdo... genérico e superficial” (Educador D), reforçando o princípio do LMC: a máquina exige um posicionamento projetor do educador para gerar sentido, e não apenas o consumo passivo. O workshop conseguiu tensionar a técnica: o educador usa a IA para ganhar tempo (função instrumental), mas o ganho de tempo deve ser reinvestido na reflexão e no planejamento de práticas críticas (função problematizadora).

Os desafios identificados pelos docentes para implementação da prática crítica confirmam a existência das “estruturas de controle e intencionalidades históricas” que a teoria do LMC busca debater. A “cotidianidade” (Lefebvre, 1991), que rege a prática do docente, manifesta-se em três eixos: (a) infraestrutura e políticas públicas; (b) formação e tempo docente; e (c) apropriação do educando. Sobre a primeira, era frequente nas respostas a falta de infraestrutura tecnológica nas escolas (equipamentos, internet de qualidade, espaço físico), e a carência de políticas institucionais comprometidas com a integração reflexiva. No segundo ponto, a falta de capacitação adequada e a existência de “professores sobrecarregados” e com pouco tempo para planejamento. Por fim, sobre o terceiro eixo, a argumentação de que os educandos não sabem usar as ferramentas ou que a usam de forma acrítica.

Tomando esses eixos, precisamos tensionar, a partir do LMC, a tecnologia como efeito das necessidades sociais e não como coisa autônoma. O desafio, portanto, é estrutural e político-econômico,

considerando o pensamento lefebvriano. Ainda, a dificuldade em reconquistar a prática docente intencionalizada, pois a práxis foi invertida de seu valor por sobrecarga burocrática impedindo o educador de ser sujeito projetor e reflexivo do seu fazer. Assim, identificamos a necessidade de LMC para o educador e o aluno, a fim de combater a unidimensionalidade (Marcuse, 1973) do consumo passivo e acrítico.

Ao oferecer a oficina aos educadores das Escolas A e B, concentramos a temática da IAGen. Nessa perspectiva inicial, com 16 respondentes, a maioria dos educadores ainda se encontrava em uma zona de insegurança prática, o que justificaria a execução de ação formativa mais “mão na massa”. Porém, como explicaria Freire (2015), não é possível dissociar a teoria da prática, menos ainda o encontro entre educadores e educandos. Em virtude disso, a perspectiva teórica e elementos do workshop foram retomados antes do direcionamento ao uso do ChatGPT.

Numa escala de 1 a 5, a média geral de familiaridade é 2,5 (entre “Tenho conhecimento básico” e “Conheço um pouco, mas me sinto inseguro(a)”). Quanto ao uso, a média de dificuldade geral é 3,2 (“Dificuldade moderada, preciso de um guia” e “Grande dificuldade, não sei bem por onde começar”). Apenas um educador se sente “muito confiante” (Educador E), e quatro (“nunca utilizei”) demonstram estar nos níveis mais baixos (1 ou 2). Assim, inferimos que mesmo com o workshop, a reflexão teórica não eliminou o desafio instrumental.

A partir das respostas dos educadores, identificamos o perfil predominante do viés utilitário coexistindo com uma preocupação ética que valida o artigo, conforme Tabela 2.

Tabela 2

Experiências e Percepções de Uso da Inteligência Artificial em Sala de Aula (Pré-Oficina)

Categoria de Experiência	Frequência	Implicação para a Pesquisa
Uso Otimizador/Produtivo pelo Professor: (preparar materiais, planejar aulas, otimizar trabalho)	8/16	Forte apelo da função instrumental da IA para alívio de demandas burocráticas e aumento da eficiência.
Desconfiança/Detecção de Uso pelo Aluno: (desconfiei, detectei, aluno informou que usou)	8/16	Maior preocupação ética em relação ao plágio, confirmando a necessidade de abordar o viés ideológico e a ética do uso (eixos da oficina).
Dificuldade de Manuseio: (senti dificuldades em manusear plataformas de IA)	6/16	Corrobora o ponto 1: a barreira funcional ainda está presente, requerendo o guia prático da oficina.
Ficar Surpreso(a) com a Capacidade da IA:	3/16	Demonstra o poder da IA, que pode levar a uma aceitação acrítica se não for acompanhada de reflexão.

Elaborado pelos Autores.

Quando tratando da percepção dos educadores, somente uma educadora demonstrou preocupação de que “A IA vai prejudicar profundamente a educação” enquanto que três educadores demonstram forte esperança com a introdução e/ou uso das IAGen. O desejo inicial dos educadores alinha-se à concepção de praticar e manusear, porém identificamos em alguns o desejo de um aprimoramento de prompts. Isso representa a semente do LMC na prática: o educador já entendeu

que precisa ser o projetista, ou seja, que a qualidade da instrução humana define a utilidade da máquina.

Identificamos novamente uma diminuição de respondentes do formulário, comparando com os dados anteriores. Tomando os oito respondentes após as oficinas, percebemos um exercício inicial de superação da barreira instrumental com vistas à intencionalidade crítica na prática pedagógica dos educadores.

Os respondentes, em sua totalidade, afirmaram que sua visão sobre a IA na educação mudou após a oficina. As explicações variam, mas convergem para a noção de que a IA pode ser aplicada “de forma reflexiva e como ferramenta de apoio” (Educador F) e que seu uso amplia o “universo de possibilidades” (Educador G). A insegurança prática inicial foi mitigada: a média de preparação subiu para 3,375 (entre “Moderadamente preparado” e “Preparado para aplicar o que aprendi”). Docentes que antes se classificavam com “grande dificuldade” agora se sentem aptos ao uso. Dois educadores (G e H) atingiram o nível máximo (5 “Sinto-me muito confiante”), demonstrando plena aquisição da habilidade funcional.

Para além do uso pelo uso, percebemos preocupações mais profundas não sobre a ferramenta, mas sobre o uso passivo do sujeito. Destacam-se as críticas ao “uso automático e desenfreado, apenas para transmitir conteúdos” (Educador C) e a preocupação com a “desvalorização do professor” (Educador I) por falta de equilíbrio. A ressalva do Educador F de que “a IA nunca vai substituir o coração humano” reitera o princípio do LMC de que a tecnologia não deve ser tratada como uma entidade substancialista, mas sim como apoio à prática humanizadora e intencional.

Em conclusão, a etapa pós-oficina demonstra que a formação de LMC funcionou como ponte entre reflexão teórica sobre os riscos e superação prática das dificuldades. Os educadores saíram com a confiança necessária (instrumental) para exercer uma práxis intencional, e com conjunto claro de preocupações ético-políticas que orientarão sua mediação em sala de aula. O desafio final, como sugerido nas propostas de futuros eventos, reside na formação continuada e na superação dos desafios de infraestrutura (cotidianidade), para que o uso crítico e intencional da IA não seja apenas evento pontual.

Considerações finais

Tivemos como objetivo discutir os desdobramentos da experiência formativa Workshop e Oficina: Tecnologias e Prática Docente, evidenciando como processos de formação continuada podem fomentar o Letramento Midiático Crítico (LMC) em face da ascensão da Inteligência Artificial (IA) no ambiente escolar. Os resultados da formação-pesquisa confirmam a premissa fundamental do LMC: a superação do uso instrumental das mídias não é desafio técnico, mas questão de práxis e intencionalidade política do educador.

A articulação entre reflexão teórica e experimentação prática mostrou-se eficaz em transformar a percepção dos docentes. A insegurança instrumental inicial foi mitigada pela Oficina, permitindo que a IA fosse vista não apenas como ferramenta otimizadora, mas como “espaço de disputa ideológica e cultural”. O ganho de tempo proporcionado pela IA, longe de gerar acomodação, foi tensionado como recurso que deve ser reinvestido na reflexão e planejamento pedagógico crítico, fortalecendo a figura do educador como sujeito projetor e não como mero executor.

Entretanto, a pesquisa aponta para as contradições estruturais que ameaçam a consolidação dessa práxis crítica. As dificuldades identificadas pelos docentes — carência de infraestrutura, sobrecarga de trabalho e consequente falta de tempo para planejamento — são manifestações da “cotidianidade” (Lefebvre, 1991). Essa lógica pasteurizadora busca inverter o valor da práxis, limitando o potencial emancipador do LMC ao impor modelo administrativo e acrítico do uso tecnológico.

Ao confrontar o discurso hegemônico das *Big Techs*, que naturaliza a IA como progresso neutro, demonstramos que a apropriação crítica passa necessariamente pelo debate sobre vieses algorítmicos e o colonialismo de dados. A formação de LMC, portanto, deve ser contínua e focar tanto no educador quanto no educando, combatendo a unidimensionalidade (Marcuse, 1973) do consumo passivo e acrítico.

Em conclusão, a experiência formativa funcionou como ponte bem-sucedida entre debate ético-político e superação das barreiras práticas. Contudo, o desafio de reconquistar a práxis docente intencionalizada é, em essência, estrutural e político-econômico. A sustentação do LMC na escola pública depende de políticas institucionais que forneçam recursos (infraestrutura) e garantam tempo e formação contínua para que o educador possa ser, de fato, um intelectual transformador e vigilante contra a reprodução da “pedagogia do capital” (Shiroma et al., 2002).

Como limitação do estudo, ressaltamos a redução da amostra na etapa pós-oficina, o que restringe a generalização dos dados quantitativos. Para pesquisas futuras, sugerimos a aplicação desse ciclo de formação-pesquisa em larga escala e a investigação aprofundada de como a aquisição da habilidade de *prompting* pelos educadores reflete diretamente na sua capacidade de atuar como sujeitos projetores no planejamento de suas aulas.

Referências

- Assis, L. (2019). *Mídias digitais, práticas docentes e cotidianos escolares: discussão do paradigma da escola do século XXI a partir da Educação Crítica para as Mídias* [Dissertação de Mestrado, Universidade Federal Fluminense].
- Assis, L. M. B., & Farbiarz, A. (2024). *Construindo o letramento midiático crítico: proposta teórica para a formação docente* [Trabalho apresentado]. 47º Congresso Brasileiro de Ciências da Comunicação, São Paulo, SP, Brasil.
- Assis, L. M. B., & Farbiarz, A. (2025). *O Letramento Midiático Crítico nos Mestrados Profissionais da Educação Básica: considerações sobre as promessas emancipatórias e uma cotidianidade reprodutivista* [Trabalho apresentado]. 48º Congresso Brasileiro de Ciências da Comunicação, São Paulo, SP, Brasil.
- Bareis, J., Roßmann, M., & Bordignon, F. (2023). Technology hype: Dealing with bold expectations and overpromising. (2023). *TATuP - Zeitschrift für Technikfolgenabschätzung in Theorie Und Praxis*, 32(3), 10-71. <https://doi.org/10.14512/tatup.32.3.10>
- Bastos, L. L. A. (2024). *Os diálogos entre sujeitos jogadores e jogos digitais: Investigação a partir de uma perspectiva comunicacional, cotidiana e do letramento midiático crítico* [Tese de doutorado, Universidade Federal Fluminense].
- Belloni, M. (2018). Mídia-educação. In D. Mill (Org.), *Dicionário crítico de educação e tecnologias e de educação a distância* (pp. 447-450). Papirus.
- Coaracy, L. L., & Farbiarz, A. (2025). Diálogos freirianos e Inteligência Artificial: caminhos para uma literacia. *Revista Interinstitucional*

Artes De Educar, 11(1), 355-374. <https://doi.org/10.12957/riae.2024.86241>

Couldry, N., & Mejias, U. A. (2019). *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Stanford University Press.

Crawford, K. (2021). *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press.

Dzieza, J. (2023). Inside the AI Factory. *The Verge*. <https://www.theverge.com/features/23764584/ai-artificial-intelligence-data-notation-labor-scale-surge-remotasks-openai-chatbots>

Feenberg, A. (2003). *O que é tecnologia?* [Conferência]

Feenberg, A. (2010). Ten paradoxes of technology. *Techné: Research in Philosophy and Technology*, 14(1), 3-15.

Freire, P. (2015). *Pedagogia do oprimido*. Paz & Terra.

Freire, P., & Guimarães, S. (2013). *Educar com a mídia: Novos diálogos sobre educação*. Paz e Terra.

Gonçalves, E. F. F. S., & Farbiarz, A. (2024). *Mídias digitais como gêneros discursivos na relação educador/educando em sala de aula* (1ª ed.) [ePDF]. Letra Capital.

Gramsci, A. (2022). *Cadernos do cárcere, volume 2: Os intelectuais, o princípio educativo, jornalismo*. Civilização Brasileira.

- Grohmann, R., Antunes, D., & Rossi, P. (2022). Platform scams: Brazilian workers' experiences of dishonest and uncertain algorithmic management. *New Media & Society*, 24(7), 1611-1631. https://eprints.lse.ac.uk/115622/1/Platform_scams_1_.pdf
- Johnson, R. L., Pistilli, G., Menéndez-González, N., Duran, L. D. D., Panai, E., Kalpokiene, J., & Bertulfo, D. J. (2022). *The Ghost in the Machine has an American accent: value conflict in GPT-3*. arXiv. <https://doi.org/10.48550/ARXIV.2203.07785>
- Kalil, R. (2019). *Capitalismo de plataforma e direito do trabalho: Crowdwork e trabalho sob demanda por meio de aplicativos* [Tese de doutorado, Universidade de São Paulo]. <https://doi.org/10.11606/T.2.2019.tde-07082020-133545>
- Kellner, D., & Kahn, R. (2015). Reconstruindo a tecnoliteracia: Uma abordagem de múltiplas literacias. *Comunicação & Educação*, 20(2), 57-82. <https://www.revistas.usp.br/comueduc/article/view/102156>
- Kellner, D., & Share, J. (2008). Educação para a leitura crítica da mídia, democracia radical e a reconstrução da educação. *Educação & Sociedade*, 29(104), 687-715.
- Lefebvre, H. (1991). *A vida cotidiana no mundo moderno*. Ática.
- Liang, W., Izzo, Z., Zhang, Y., Lepp, H. Cao, H., Zhao, X., Chen, L., Ye, H. Liu, S., Huang, Z., McFarland, D., & Zou, J. (2024). Monitoring AI-Modified Content at Scale: A Case Study on the Impact of ChatGPT on AI Conference Peer Reviews. *arXiv preprint arXiv:2403.07183*. <https://arxiv.org/abs/2403.07183>
- Marcuse, H. (1973). *A ideologia da sociedade industrial: O homem unidimensional*. Zahar.

- Maslej, N., Fattorini, L., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Ngo, H., Niebles, J. C., Parli, V., Shoham, Y., Wald, R., Clark, J., & Perrault, R. (2023). *The AI Index 2023 annual report*. AI Index Steering Committee, Institute for Human-Centered AI, Stanford University. https://hai.stanford.edu/assets/files/hai_ai-index-report_2023.pdf
- McQuillan, D. (2022). *Resisting AI: An anti-fascist approach to artificial intelligence*. Bristol University Press.
- Neder, R. (Org.), (2013). *A teoria crítica de Andrew Feenberg: Racionalização democrática poder e tecnologia*. Observatório do Movimento pela Tecnologia Social na América Latina/CDS/UnB/Capes.
- Pasquale, F. (2015). *The black box society: the secret algorithms that control money and information*. Harvard University Press.
- Perrigo, B. (2023). Exclusive: OpenAI Used Kenyan Workers on Less than \$2 per Hour to Make ChatGPT Less Toxic. *Time*. <https://time.com/6247678/openai-chatgpt-kenya-workers>
- Petrov, A., La Malfa, E., Torr, P., & Bibi, A. (2023). Language Model Tokenizers Introduce Unfairness Between Languages. *arXiv preprint arXiv:2305.15425*. <https://arxiv.org/abs/2305.15425>
- Pinto, A. V. (2005). *O conceito de tecnologia*. Contraponto.
- Resnik, P. (2024). Large Language Models are Biased Because They Are Large Language Models. *arXiv preprint arXiv:2406.13138v1*. <https://arxiv.org/abs/2406.13138v1>

- Russell, S. J., & Norvig, P. (2022). *Artificial intelligence: A modern approach*. Pearson Education.
- Selwyn, N. (2014). *Distrusting educational technology: Critical questions for changing times*. Routledge.
- Selwyn, N. (2017). Educação e tecnologia: Questões críticas. In G. Ferreira, L. Rosado, & J. Carvalho (Orgs.), *Educação e tecnologia: Abordagens críticas*. SESES/Universidade Estácio de Sá.
- Shiroma, E. O., Moraes, M. C. M. de, & Evangelista, O. (2002). *Política educacional* (pp. 53-86). DP&A.
- Silveira, S. A. (2021). Hipótese do colonialismo de dados e o neoliberalismo. In J. F. Cassino, J. Souza, & S. Silveira (Orgs.), *Colonialismo de dados: como opera a trincheira algorítmica na guerra neoliberal* (pp. 33-52). Autonomia Literária.
- Soares, M. (2009). *Letramento: Um tema em três gêneros* (3ª ed.). Autêntica.
- Street, B. (2014). *Letramentos sociais: Abordagens críticas do letramento no desenvolvimento, na etnografia e na educação*. Parábola Editorial.
- Verdegem, P. (2022). Dismantling AI capitalism: The commons as an alternative to the power concentration of Big Tech. *AI & Society*, 39, 727-737. <https://doi.org/10.1007/s00146-022-01437-8>
- Williams, R. (2016). *Televisão: Tecnologia e forma cultural*. Boitempo/PUC Minas.

Wilson, C., Grizzle, A., Tuazon, R., Akyempong, K., & Cheung, C. K. (2013). *Alfabetização midiática e informacional: Currículo para formação de professores*. Unesco.

MÉTODOS DIGITAIS APLICADOS ÀS IMAGENS EM 360 GRAUS: ANÁLISE DE DADOS VISUAIS E INTERATIVOS¹

Carolina Gois Falandes²
Alan César Belo Angeluci³

Vive-se hoje em uma cultura fortemente influenciada pelo virtual. As redes digitais permeiam diversos contextos, moldando comportamentos cotidianos e ações institucionais. Membros da sociedade seguem o compasso das *trends* da Internet, que não apenas ditam modas passageiras, mas também reconfiguram fluxos de informação e reformulam as maneiras de construir narrativas.

-
1. Agradecimento ao apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), processo 307449/2025-4.
 2. Doutora em Comunicação pela Unesp.
carol.falandes@gmail.com
 3. Doutor em Engenharia Elétrica pela USP, com pós-doutorado na ECA/USP e na University of Texas at Austin.
Professor da ECA/USP. Bolsista de produtividade em pesquisa (CNPq).
aangeluci@usp.br

Esse cenário tem exercido impacto direto na agenda da pesquisa científica, tanto em âmbito nacional quanto internacional. Investigadores de distintas áreas procuram examinar objetos assentados nos ambientes *on-line*, de modo a compreender relações, controvérsias, usos e apropriações. A intenção é lidar com a complexidade de tempos tão plattformizados. Elementos como *big data*, algoritmos, *APIs*, modelos de linguagem de grande escala (*LLMs*) e bolhas informacionais integram a pauta de trabalhos em diferentes campos, incluindo a Comunicação.

De fato, como pontua Recuero (2025), a criação de sistemas digitais capazes de registrar e rastrear dados das interações entre indivíduos conectados representou uma transformação significativa para a análise social, abrindo novos caminhos metodológicos e teóricos para perceber as atividades contemporâneas. É nesse quadro que se inscreve a proposta dos métodos digitais, elaborada em resposta às demandas conceituais de um ecossistema midiático atravessado pela lógica das infraestruturas computacionais. Trata-se de uma abordagem que tem como precursor Richard Rogers (2013), caracterizada como uma chave interpretativa na qual os meios são coprodutores dos fenômenos, possibilitando a formulação de diagnósticos culturais. Na esteira desse raciocínio, Omena (2019a) esclarece que esse instrumental não se resume ao uso de *softwares* para fins estritamente quantitativos. Ao contrário, pressupõe uma prática quali-quantitativa apoiada nos mecanismos das plataformas da web para construir leituras sobre a coletividade.

No conjunto de produtos circulantes nas redes sociais virtuais, as imagens - estáticas e em movimento - figuram entre os principais formatos discutidos pela comunidade acadêmica dos métodos digitais. Tal enfoque, segundo Colombo et al. (2023), relaciona-se ao

protagonismo das visualidades na atualidade - vigente nas esferas social, política, cultural e comercial -, o que fomenta o crescente interesse por essa forma expressiva e impulsiona o desenvolvimento de técnicas para sua avaliação.

Na literatura, identificam-se múltiplas perspectivas que contribuem para esse processo. Alguns autores têm produções concentradas em análises multifacetadas sobre o Instagram (Bitencourt, 2025; Kozharinova & Manovich, 2024; Rogers, 2021). Outros apresentam recomendações metodológicas voltadas à observação de imagens *on-line*, com o suporte de recursos de visão computacional (Omena et al., 2021) e a partir de um ponto de vista integrado, que considera a sua natureza multimodal (Pearce et al., 2018). Há, ainda, textos dedicados à proposição de esquemas de visualização de dados referentes às imagens em mídias sociais (Cirino et al., 2025; Colombo et al., 2023), assim como estudos - a exemplo do desenhado por Angeluci et al. (2024) - que monitoram o engajamento de usuários durante transmissões audiovisuais ao vivo na web.

O presente artigo insere-se nesse panorama de investigações alicerçadas em métodos digitais mobilizados no domínio das imagens, focando registros capturados por câmeras 360 graus. O objetivo é evidenciar quatro estratégias de extração, análise e visualização de dados orientadas à exploração de conteúdos esféricos publicados em plataformas *on-line*, com ênfase na interpretação de seus aspectos estéticos e/ou nas formas de interação do público. Para isso, baseia-se em repertório adquirido a partir de experiências científicas predecessoras que envolveram essa modalidade imagética, em articulação com abordagens metodológicas mediadas por tecnologia (Falandes & Angeluci,

2019; Falandes & Renó, 2022), bem como com os diálogos e experimentações promovidos pelo grupo de pesquisa *Smart Media & Users*, da Universidade de São Paulo, cujo escopo interdisciplinar compreende a relação entre as mídias emergentes e os cidadãos.

Espera-se que este trabalho, ao ilustrar diversificados procedimentos para inspecionar narrativas esféricas - alguns deles sustentados por aplicações da mais recente onda de popularização tecnológica, a inteligência artificial (IA) -, permitindo o mapeamento de obras, produtores e *insights* sobre padrões de atenção dos usuários, possa inspirar estudos futuros e auxiliar o *design* de peças em 360 graus pela indústria criativa, fortalecendo, por conseguinte, o percurso de consolidação dessa linguagem.

Estudando as Imagens: Métodos Digitais na Interpretação das Visualidades Contemporâneas

Os conteúdos imagéticos flutuantes nos espaços *on-line*, em suas mais diversas manifestações, configuram fontes inesgotáveis de informação sobre a sociedade contemporânea, revelando de que maneira os atores exteriorizam sentimentos, convicções e críticas, além de indicar como se apresentam a seus pares virtuais. Por isso, tanto pessoas físicas quanto organizações têm dedicado atenção às visualidades, cada qual de acordo com seus interesses e critérios - engajamento e fortalecimento de identidade ou de marca, experiências e interações -, reforçando sua relevância no atual ecossistema midiático.

Diante da influência dessa forma expressiva na cultura digital, capaz de conduzir reflexões na esfera pública, a exigência de sua leitura crítica nunca foi tão crucial (Niederer & Colombo, 2024). Nesse sentido,

observa-se um conjunto de estudos que a caracterizam. Renó (2024), por exemplo, ao discorrer sobre os registros fotográficos de viagem, considera que a imagem tem o potencial de transportar a audiência para um passeio visual, como se estivesse em um sonho. Já Niederer e Colombo (2024, p. 22) mencionam alguns adjetivos recorrentes na descrição dessas produções: “imateriais, leves, pobres, operacionais, invisíveis, efêmeras e em rede”.

Fontcuberta (2016), por sua vez, enquadra os produtos imagéticos da era digital no campo da pós-fotografia, termo que representa uma virada no papel das imagens: se antes ocupavam uma função passiva, de mera ilustração, transformaram-se em ativas e indomáveis, assumindo variados formatos. Ao tratar desse contexto, o autor aponta dois aspectos ressaltados por investigadores da cultura visual, que retratam o fazer imagético na hipermodernidade: a criação excessiva de imagens, que se acumulam sem serem vistas, transbordando os dispositivos de armazenamento, e o descompasso entre os métodos digitais de produção de conteúdos visuais e os analógicos de leitura.

Nessa perspectiva, embora um número significativo de disciplinas acadêmicas reconheça a eficácia da interpretação das múltiplas modalidades imagéticas como estratégia para a aquisição de aprendizados sobre a coletividade (Colombo, 2019), ainda se nota uma valorização teórica de arcabouços metodológicos mais convencionais, aplicados ao exame de uma quantidade reduzida de objetos. Com essa percepção, não se pretende, no entanto, invalidar a contribuição das abordagens tradicionais, uma vez que todas elas são importantes dentro de seus propósitos e delimitações. Cabe frisar, por outro lado, à luz do pensamento de Vasconcelos (2024, p. 20), que milhões de imagens são

disseminadas continuamente, em sua totalidade pertinentes, enquanto cientistas geralmente se concentram em identificar “significações em algumas poucas imagens sobre determinado tema”. Avista-se, portanto, um desafio considerável no âmbito do processo de análise de conteúdos imagéticos, envolvendo a obtenção de amostras representativas e as condições técnicas disponíveis para esse tipo de levantamento (Vasconcelos, 2024).

É nesse pano de fundo que se delineia a atuação dos Métodos Digitais, desdobramentos dos Estudos de Ciência e Tecnologia. Para D’Andréa (2020, p. 61), tal proposta, em articulação com os Estudos de Plataforma, favorece a sustentação de pesquisas empíricas direcionadas aos “entrelaçamentos entre temáticas diversas e as materialidades online”. Segundo o autor, há uma inclinação por investigações alicerçadas nas etapas de coleta, processamento e visualização de dados disponibilizados por uma plataforma digital, voltadas à observação de nuances que compõem o todo social e das possíveis controvérsias, tendo as *APIs*⁴ públicas como uma das ferramentas mobilizadas para esse fim. Simultaneamente, debates ligados aos *Critical Data Studies* sublinham camadas mais profundas do campo, a exemplo de relações de poder e desigualdades (Grohmann, 2025), reforçando seu papel político e não apenas técnico.

4. A título de esclarecimento conceitual, vale definir brevemente o que é *API*. Silva et al. (2020) explicam que a sigla em inglês representa a Interface de Programação de Aplicações, um mecanismo capaz de estruturar sistemas computacionais permitindo a transferência de dados com outros programas externos. Essa disponibilidade pode ser pública ou comercial, e as funcionalidades variam conforme rede social. Trata-se de uma questão complexa, condicionada a decisões das plataformas sobre a liberação ou restrição de suas *APIs*.

Dentre as possibilidades oferecidas por esse instrumental teórico-metodológico, a discussão empreendida por Recuero (2025, p. 19) evidenciou seu impacto na área dos discursos publicados no território digital, ao impulsionar o desenvolvimento de trabalhos com dados em grande escala, além da constatação de padrões e laços complexos “entre atores, ideias e práticas discursivas”. Em contraste, esse cenário traz consigo impasses que abarcam desde preocupações éticas até os vieses dos algoritmos, exigindo uma interpretação crítica das infraestruturas nas quais esses conteúdos se assentam (Recuero, 2025).

No horizonte dessas considerações, a imagem digital pode ser entendida como peça ativa na engrenagem das plataformas de redes sociais. Sua postagem realça não apenas o estilo de determinado produto visual, mas também propicia inteirar-se acerca dos processos e ações desencadeados por essa prática (Bitencourt, 2025). Com efeito, averiguar as visualidades compartilhadas em ambientes dinâmicos como *YouTube*, *Instagram* e *TikTok* permite conhecer, concomitantemente, as particularidades técnicas que regem o funcionamento desses sites. Nos termos de Niederer (2018), isso se refere aos sistemas de formação, priorização e recomendação de conteúdos - viabilizados a partir de métodos inovadores pautados no estudo de imagens agrupadas ou em uma lógica de rede.

É possível encontrar diferentes núcleos investigativos comprometidos com essa vertente, que adotam variadas abordagens para lidar com dados digitais, incluindo os de natureza imagética. No circuito internacional, podem ser destacadas, com base na lista de D’Andréa (2020), contribuições científicas geradas por eventos como a *Summer School* e a *Winter School* da *University of Amsterdam* (s.d.-a; s.d.-b),

nos Países Baixos, e o *SMART Data Sprint*, promovido pelo *iNOVA Media Lab* (s.d.), da Universidade Nova de Lisboa, em Portugal.

Em solo brasileiro, sem pretensão de contemplar todas as experiências existentes, sobressaem as iniciativas do grupo *Datalab Design* (s.d.), da Universidade do Estado da Bahia, que em seu laboratório desenvolve instrumentos específicos para a análise de coleções de imagens com suporte de visão computacional, denominados *Semantic Imagery Mapping (SIM)*, *Movie Scene Sensing (MSS)* e *Modular Observational & Operational Design Image Explorer (MOODIE)*. Ainda merecem menção as movimentações do *Smart Media & Users* (s.d.), vinculado à Universidade de São Paulo, em prol da disseminação de trabalhos e atividades práticas sobre Métodos Digitais, que também abrangem conteúdos imagéticos, a exemplo da idealização do *ECA Data Sprint*.

Este artigo se soma ao leque de estudos realizados no contexto do *SMU*, com um olhar dirigido às produções de imagem em 360 graus e seus diálogos com os Métodos Digitais, propiciando a apuração de como extração, interpretação e visualização de dados se interligam à experiência imersiva. Torna-se, assim, necessário discutir peculiaridades dessa modalidade visual antes de prosseguir para a etapa aplicada.

Análise de Imagens em 360 Graus sob as Lentes dos Métodos Digitais

Os conteúdos imagéticos derivados de câmeras 360° correspondem a uma das muitas linguagens exploráveis no atual contexto pós-fotográfico para a difusão de discursos visuais. O potencial narrativo desse formato associa-se, sobretudo, à sensação de imersão característica das produções esféricas navegáveis, que, conforme argumenta Falandes (2025a), possibilitam ao público visualizar uma cena a partir de diferentes

ângulos. Esse tipo de experiência estética é capaz de aproximar as pessoas das histórias e paisagens registradas, ainda mais quando se usa óculos de realidade virtual em sua fruição, qualidade que tem inspirado distintos setores sociais - turístico, imobiliário, jornalístico, audiovisual, educacional, publicitário, museológico e artístico - a incorporá-la em suas práticas. Paralelamente, existem obras em 360 graus estruturadas de forma não-navegável, apresentando perspectiva bidimensional, com enquadramentos direcionados na edição (Falandes, 2025a). Nesse caso, busca-se oferecer um produto visual de caráter mais artístico, com efeitos de movimentação e ritmo pensados para captar a atenção dos seguidores, principalmente em plataformas como *Instagram* e *TikTok*.

Do ponto de vista científico, este objeto midiático, que rompe com a ideia de limites do quadro presentes na imagem tradicional (Casini & Lovato, 2023), vem sendo examinado a partir de diversificadas metodologias, nas quais os pesquisadores de jornalismo têm desempenhado papel significativo. Entre elas, convém mencionar discussões teóricas e revisões bibliográficas - realizadas em repositórios e bases acadêmicas - que identificam e acionam conceitos-chave relacionados a essa inovação (Lima & Barbosa, 2022; Rocha & Rocha, 2022); mapeamentos de centros acadêmicos de produção e investigação dessa modalidade (Falandes & Renó, 2021; Rodrigues, Firmino, & Cunha, 2022); estudos de caso e análises semióticas, voltados à descrição e interpretação de arranjos narrativos (Falandes & Renó, 2023; Médola & Oliveira, 2017); entrevistas com especialistas e profissionais, que permitem revelar suas posições sobre a área (Damas & De Gracia, 2022; Teixeira, 2024); experimentos empíricos, com participantes envolvidos em atividades de consumo de conteúdos (Falandes, 2025a; Rocha & Rocha, 2024);

além da aplicação de questionários, destinados a levantar percepções do público a respeito dessa tecnologia (Falandes, 2025a; Giacomelli, 2023).

Mirando nos Métodos Digitais e na articulação com imagens em 360 graus no Brasil, realizou-se uma busca exploratória no Google Acadêmico em 1 de setembro de 2025, com os termos “vídeos em 360” e “métodos digitais”. O primeiro resultado foi o livro editado por Omena (2019b), considerado uma obra pioneira em língua portuguesa dedicada ao assunto; em particular, o capítulo de Flores (2019), que observa vídeos em 360 graus no *YouTube* no contexto jornalístico. Na quarta posição, surgiu a tese de Falandes (2025a), que também recorreu a esse instrumental para analisar conteúdos esféricos, como no *Instagram*. Os demais trabalhos localizados, levando em conta os 20 primeiros achados, parecem não abordar diretamente essa relação, evidenciando a necessidade de novos estudos nessa linha de pesquisa.

A reboque desse panorama, e reconhecendo que as imagens em 360 graus são resultantes do regime de plataformização, fomentadas por empresas de tecnologia - como *YouTube* e *Facebook* - (Rodrigues & Lima, 2022), os Métodos Digitais podem balizar análises sobre peças disseminadas pelo público nas redes sociais e, da mesma forma, favorecer realizadores em seus processos criativos. É o que se procura demonstrar nas experimentações relatadas a seguir.

Caminhos Metodológicos

O *design* desta investigação empírica foi planejado inspirando-se nas cinco etapas indicadas por Vasconcelos (2024) para auxiliar o desenvolvimento de pesquisas que empregam métodos digitais na avaliação de imagens: pré-coleta, coleta de dados, refinamento da amostra, análise e criação de visualizações.

A fase de pré-coleta compreendeu, inicialmente, a realização de revisão bibliográfica sobre o tema escolhido, bem como a definição do objetivo pretendido, que foi o de apresentar métodos digitais aplicáveis a peças em 360 graus, inclusive com a demonstração prática de exemplos, cujo potencial permite uma leitura crítica das dimensões estéticas e/ou interacionais dessa modalidade. Partindo-se da recomendação de Pearce et al. (2018), que destacam a importância de se conduzir investigações visuais multiplataforma, foram selecionadas quatro estratégias metodológicas, sustentadas em suportes distintos, a serem examinadas neste estudo - com base na experiência dos autores em tecnologia 360° e em consultas à literatura -, além dos respectivos parâmetros para acioná-las.

Na sequência, durante a coleta - feita em diferentes momentos, entre 21 de julho e 5 de setembro de 2025 - foram acessados os dados disponibilizados e/ou extraídos dos ambientes virtuais escolhidos, a saber: (1) *analytics* da ferramenta *Wonda VR* (2025); (2) registros do *Instagram* e *TikTok*, com o auxílio do *software 4CAT* (Peeters & Hagen, 2022) e da extensão *Zeeschuimer* (Peeters, 2025); (3) *prompts* de inteligência artificial, com destaque para os obtidos via *ChatGPT* (OpenAI, 2025); e (4) resultados de rastreamento ocular, com soluções como o *ScanGAN360* (Martin et al., 2022). Como pontuam Vázquez-Herrero, Negreira-Rey e Klug (2025), cada plataforma de mídia opera por meio de mecanismos próprios, exigindo protocolos metodológicos específicos, o que está alinhado à abordagem adotada neste artigo.

Após a obtenção dos *datasets* e, em alguns casos, a instalação de *softwares*, procedeu-se ao refinamento da amostra. Esse processo envolveu o tratamento do material, com ajustes efetuados a partir de

recursos como *Google Sheets* e *Microsoft Excel*, juntamente com a execução de capturas de tela e a importação e organização de arquivos de imagem em pastas no computador.

Com todas as informações tabuladas e/ou devidamente arquivadas, passou-se à fase de análise, na qual foi possível interpretar os achados, agrupando-os por padrões visuais e categorias temáticas. Ainda se verificaram, em determinadas situações, semelhanças e disparidades entre postagens publicadas em plataformas distintas, com apoio de esquematizações em planilhas.

Por fim, a materialização desse percurso investigativo se deu por meio de visualizações, tanto aquelas geradas pelos sistemas utilizados na extração quanto as elaboradas pelos autores, de natureza interativa, valendo-se das ferramentas externas *Flourish* (2025) e *StoryMap JS* (Knight Lab, 2025).

Deve-se reconhecer que, nesta investigação, buscou-se experimentar técnicas e abordagens e, considerando a quantidade expressiva de métodos a serem discutidos (4) em uma estrutura textual curta - em formato de *paper* -, priorizou-se a descrição dos processos, sem o compromisso de analisar grandes volumes de imagens. Assim, a intenção das próximas páginas é ilustrar possibilidades, o que poderá motivar novos estudos.

Análise de Dados na *Wonda VR*: Métricas e Interações da Narrativa “Conjunto do Ipiranga em 360 Graus”

Wonda VR (2025) é uma plataforma *on-line* que disponibiliza *templates* para a criação de experiências de aprendizagem imersivas, baseando-se em tecnologias como 3D, metaverso, inteligência artificial e imagens em 360 graus. Além de seu potencial enquanto espaço

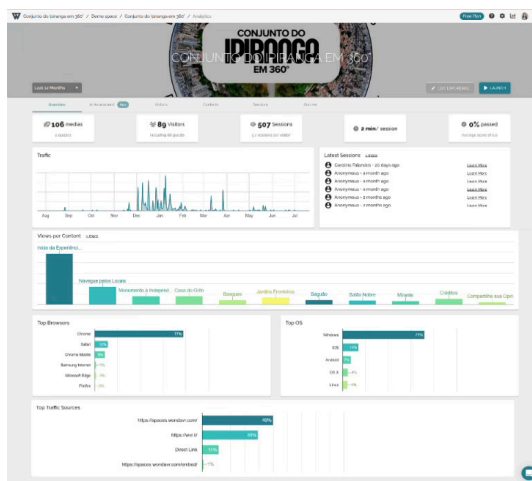
responsivo para compartilhamento de conteúdos, a ferramenta dispõe, mesmo em seu plano gratuito, de uma seção de *analytics* que permite ao realizador acessar indicadores de navegação relativos ao seu produto. Nesse contexto, serão examinados aqui os dados - apurados em 21 de julho de 2025 - referentes à narrativa interativa “Conjunto do Ipiranga em 360 graus: explorando todos os lados da Independência do Brasil⁵”, desenvolvida por Falandes (2025b) e construída a partir da estrutura de *tour* virtual oferecida pelo recurso.

Ao adentrar o ambiente de *analytics*, depara-se com seis abas principais, cada uma apresentando painéis informativos. Na aba *overview*, são apontadas métricas de uso que combinam dados quantitativos e resultados sobre o engajamento dos usuários com a obra. Observa-se, por exemplo, que a produção utiliza 106 mídias, recebeu 89 visitantes e contabilizou 507 sessões, com média de 2 minutos por sessão. Ressalta-se também a exposição de gráficos (Figura 1), que demonstram *insights* sobre o comportamento dos interatores, como a representação de tráfego, evidenciando maior concentração de acessos à narrativa entre dezembro de 2024 e fevereiro de 2025, período de sua confecção e lançamento. Cabe ressaltar, porém, que os indicadores disponibilizados apresentam limitações, sobretudo pela ausência de informações qualitativas. Tal lacuna mostra a necessidade de complementar a análise com estratégias adicionais, capazes de captar percepções e interações que não se refletem apenas nos dados quantitativos.

5. A obra documental, apoiada em imagens em 360 graus, retrata locais simbólicos vinculados à Independência do Brasil, situados na região do Ipiranga, em São Paulo.

Figura 1

Painel de analytics da aba “Overview” na Wonda VR



Nota. Captura de tela extraída do *analytics* da obra (Falandes, 2025b).

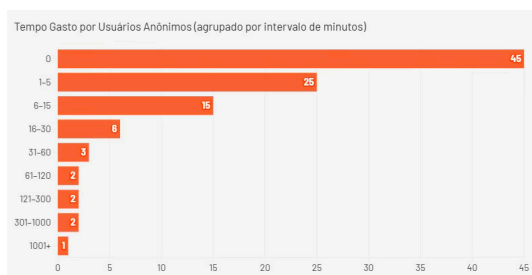
Ao clicar na aba *AI assessment*, recurso mais recente da plataforma, o criador da narrativa pode adicionar critérios de avaliação sobre a experiência imersiva desenvolvida, incluindo a atribuição de pontuações, de modo que a inteligência artificial seja capaz de mapear e avaliar os caminhos seguidos pelos usuários. Trata-se de funcionalidade direcionada ao campo educacional, destinada à medição da aprendizagem de estudantes, com possibilidade de geração de relatórios a serem enviados aos alunos. Embora não tenha sido testada no âmbito da produção em referência, vale mencioná-la como um recurso válido para mensurar as interações do público.

Complementarmente, a ferramenta fornece, por meio das abas *Visitors* e *Sessions* pacotes de dados que podem ser transferidos para o computador (em formato CSV), contendo registros sobre o perfil dos

usuários que visitaram a produção e as interações realizadas durante as sessões. Esses arquivos foram baixados, analisados e transformados em visualizações dinâmicas com o apoio de recursos da plataforma *Flourish* (2025). A partir da leitura dos dados da aba *Visitors*, constatou-se que a maioria dos visitantes permaneceu menos de 10 minutos na narrativa, com distribuição altamente concentrada no início e poucos usuários ultrapassando 20 minutos (Figura 2). De modo semelhante, mesmo entre os que passaram mais tempo, o número de sessões manteve-se reduzido (Figura 3). No caso da aba *Sessions*, percebeu-se uma tendência de queda no engajamento conforme o número de cenas aumenta (Figura 4). Notou-se que 172 usuários, correspondendo à maioria, visualizaram apenas uma cena - o que indica elevado índice de desistência ou baixa exploração da experiência -, enquanto um número bem inferior (4 usuários) alcançou 10 cenas.

Figura 2

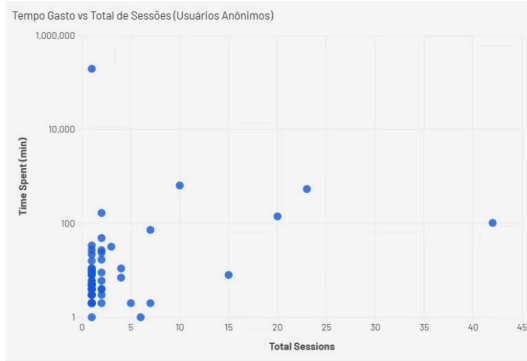
Tempo de permanência de usuários anônimos na narrativa



Nota. Gráfico gerado pelos autores a partir do *Flourish* (2025), com versão *on-line* interativa disponível em <https://public.flourish.studio/visualisation/24335548/>

Figura 3

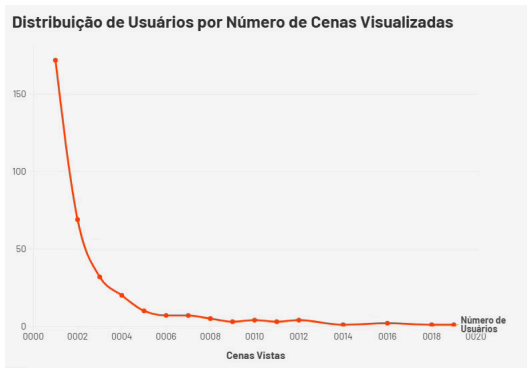
*Relação entre tempo de permanência e total de sessões
(usuários anônimos)*



Nota. Gráfico gerado pelos autores a partir do *Flourish* (2025), com versão *on-line* interativa disponível em <https://public.flourish.studio/visualisation/24336388/>

Figura 4

Distribuição de usuários conforme o número de cenas visualizadas



Nota. Gráfico gerado pelos autores a partir do *Flourish* (2025), com versão *on-line* interativa disponível em <https://public.flourish.studio/visualisation/24338518>

Ao navegar pela aba *Content*, pode-se examinar informações sobre cada cena individualmente - todas ilustradas por miniaturas - assim como seus recursos de interatividade, com base na leitura de índices de porcentagem, recurso válido para monitorar os conteúdos mais visualizados. No caso da narrativa sobre o Conjunto do Ipiranga, verificou-se que as cenas introdutórias receberam maior atenção, funcionando como um *hub* pelo qual o usuário escolhe os locais que deseja visitar.

Já a aba *Quizzes*, por sua vez, traz dados referentes às respostas dos interatores a perguntas incluídas na obra, destacando, inclusive, gráficos de barra que ilustram os resultados. Foram feitas duas perguntas: se o visitante apreciou a produção e a opção de escrever um comentário. Ao todo, 11 pessoas responderam à primeira questão, das quais 10 (91%) afirmaram positivamente e 1 (9%) negativamente. Na parte dissertativa, foram coletadas três mensagens positivas, como “Surpreendente imersão; entendo que as secretarias devem explorar esse material e assim expandir em outros pontos turísticos da cidade. Perfeito”.

Diante dos dados levantados a partir de métodos digitais (D’Andréa, 2020; Recuero, 2025), cabe esclarecer que não houve uma ampla estratégia de divulgação da narrativa, concebida, em um primeiro momento, como produto técnico resultante de doutoramento. Ainda assim, os *analytics* mostram-se pertinentes, iluminando um desafio fundamental no contexto das experiências interativas: conciliar a complexidade dos conteúdos com a garantia de que o público possa acessá-los em sua plenitude. Outro ponto de aprendizagem refere-se à oportunidade de, em análises futuras, solicitar o cadastro dos usuários na plataforma, o que permitiria obter resultados mais personalizados.

Imagens em 360 Graus no Instagram e TikTok: Um Estudo da #Museu360 a partir do 4CAT e do Zeeschuimer

A fim de capturar e analisar discursos visuais em 360 graus em redes sociais virtuais, contou-se com o suporte da ferramenta 4CAT (Peeters & Hagen, 2022), em articulação com a extensão Zeeschuimer (Peeters, 2025). Foram selecionadas as plataformas Instagram (fotos e vídeos) e TikTok (vídeos) como espaços dessa investigação. Decidiu-se apurar, em 23 de julho de 2025, postagens relacionadas à #museu360, escolha temática feita em consonância com o interesse em verificar de que forma instituições museológicas têm explorado a tecnologia 360° em suas exposições. Com esse parâmetro, procedeu-se à coleta das 300 primeiras publicações em cada mídia, obtidas a partir de arquivos em formato CSV.

Por meio das planilhas, cada postagem foi avaliada individualmente, considerando-se especialmente a descrição utilizada para referenciá-la - denominada *body* - e a imagem correspondente. No Instagram, constatou-se, entre os resultados (Figura 5), que, do total de posts, 102 mencionavam ou apresentavam *tours* virtuais promovidos por museus, recorrendo, para isso, a conteúdos imagéticos tradicionais. Em seguida, 65 publicações destacaram-se por associar o número 360 a museus, ainda que em contextos genéricos, enquanto 53 referiam-se a vistas panorâmicas em 360° registradas a partir de visitas a mirantes. Vale ressaltar que, na amostra, foram identificadas apenas 11 imagens esféricas, sendo 9 fotografias e 2 vídeos. Ao analisar essas peças, pôde-se perceber que 9 exploravam a estética *little planet*, uma das modalidades estéticas das imagens esféricas não-navegáveis (Falandes, 2025a), captadas tanto por câmera no solo quanto por drones (Figura 6).

Figura 5

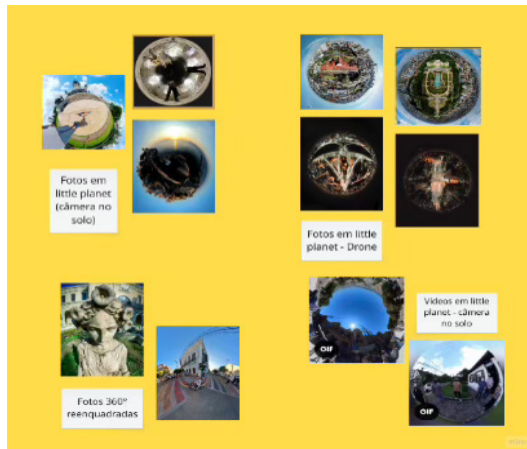
Temas das postagens que utilizaram a #museu360 no Instagram



Nota. Gráfico gerado pelos autores a partir do Flourish (2025), com versão on-line interativa disponível em <https://public.flourish.studio/visualisation/24981046>

Figura 6

Imagens em 360° extraídas de postagens com a #museu360 no Instagram



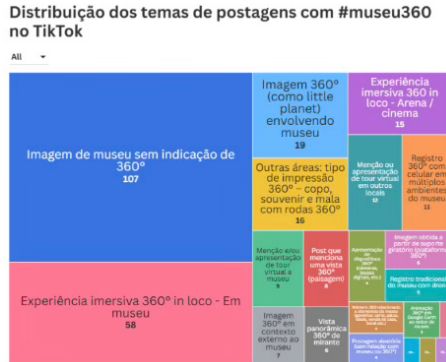
Nota. Montagem elaborada pelos autores com base na análise.

A partir desses achados, observou-se que tanto instituições museológicas quanto usuários comuns mencionaram passeios virtuais em 360°. Entretanto, o uso efetivo de conteúdos esféricos nas postagens foi significativamente menor, restrito a produtores audiovisuais, entusiastas ou pessoas com câmera 360° que visitaram museus e aproveitaram para realizar registros.

Mirando no TikTok, percebeu-se um outro cenário, exigindo o acréscimo de novas categorias. Tal fato evidencia a relevância de se conduzir investigações multiplataforma (Pearce et al., 2018), que tornam possível levantar nuances nos dados. Como demonstrado na Figura 7, entre as 300 publicações apuradas nessa rede, 107 divulgavam vídeos sobre museus, mas sem conexão aparente com 360°, ao passo que 58 retratavam experiências imersivas instaladas em espaços museológicos. Notou-se também uma quantidade superior de imagens em 360 graus, totalizando 19 (Figura 8). Por serem em movimento, apresentam possibilidades estéticas mais amplas: incluem trilhas sonoras e narrações, aplicam efeitos de tempo e movimentação, incorporam recursos de legenda e caixas de texto, e algumas são, inclusive, articuladas com conteúdos visuais tradicionais. Da mesma forma, reparou-se que essas gravações foram publicadas por visitantes - deslocando-se por áreas externas e internas dos museus e registrando instalações imersivas, possivelmente para representar visualmente a sensação proporcionada por ambientes com grandes projeções e iluminação -, além de perfis com caráter mercadológico, voltados à divulgação de produtos referentes à tecnologia 360°, como modelos de câmeras.

Figura 7

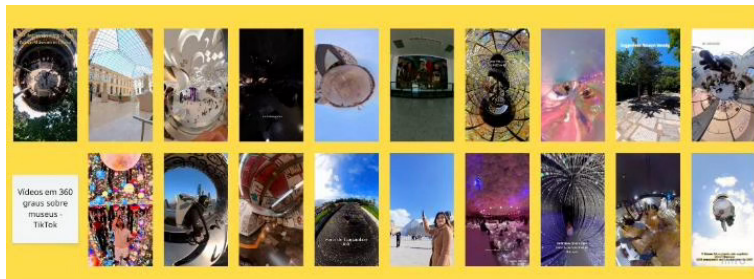
Temas das postagens que utilizaram a #museu360 no TikTok



Nota. Gráfico gerado pelos autores a partir do *Flourish* (2025), com versão *on-line* interativa disponível em <https://public.flourish.studio/visualisation/24981186/>

Figura 8

Vídeos em 360° extraídos de postagens com a #museu360 no TikTok



Nota. Montagem a partir de capturas de tela elaborada pelos autores com base na análise.

Logo, assim como no Instagram, no TikTok permanece a mesma natureza dos perfis que compartilham imagens em 360 graus. Nesse sentido, abre-se uma oportunidade para que instituições museológicas

aproveitem o interesse espontâneo dos usuários nesse tipo de conteúdo, incorporando-o em suas estratégias de divulgação oficial.

O estudo da #museu360 a partir dessas duas plataformas midiáticas também permitiu a elaboração de uma cartografia interativa de *tours* virtuais de instituições museológicas, distribuídas em 6 países, com o suporte da ferramenta *on-line StoryMap JS* (Knight Lab, 2025). Ao inspecionar as descrições de cada publicação, reuniram-se informações como links e títulos, totalizando 36 passeios identificados via Instagram e 5 no TikTok, conforme ilustra a Figura 9. Trata-se de um repositório que pode servir tanto a pesquisadores do tema quanto a curiosos por esse formato curatorial digital.

Figura 9

Cartografia interativa de tours virtuais produzida a partir da análise



Nota. Imagem da cartografia produzida pelos autores com a ferramenta *StoryMap JS* (Knight Lab, 2025), disponível em: <https://shre.ink/SYKp>

ChatGPT como Método Digital para Análise de Conteúdos em 360 Graus

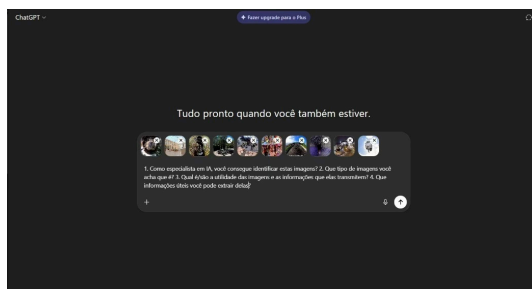
A opção por explorar o *ChatGPT* (OpenAI, 2025), na versão GPT-5, como suporte para pesquisas *on-line* vai ao encontro do recente

fenômeno de popularização da IA generativa, cuja integração ao fazer científico pode, na perspectiva de Lapa (2025), oferecer contribuições para o desenvolvimento de análises de maior profundidade e precisão envolvendo grandes volumes de dados, ao mesmo tempo em que suscita questões metodológicas e éticas.

O experimento com esse assistente de inteligência artificial, realizado em 5 de setembro de 2025, buscou investigar especificamente as imagens em 360 graus obtidas a partir dos *datasets* levantados via *4CAT* (Peeters & Hagen, 2022) e *Zeeschuimer* (Peeters, 2025). Para isso, recorreu-se à abordagem proposta por Johnson et al. (2023), que elaboraram *prompts* voltados a medir a capacidade de o *ChatGPT* de examinar conteúdos visuais, estruturados em quatro perguntas-chave destinadas a extrair informações sobre tipo de imagem, aspectos técnicos, utilidade e potencial de uso.

Figura 10

Imagens anexadas no ChatGPT, juntamente com as perguntas-chave



Nota. Captura de tela da interface do *ChatGPT* (OpenAI, 2025).

Com esse alicerce metodológico, partiu-se para a inclusão dos arquivos na plataforma. Como foi utilizada a versão gratuita da

ferramenta, tornou-se necessário lidar com a limitação de anexar apenas dez imagens por vez. Como solução, das 30 imagens inicialmente levantadas no Instagram e TikTok, foram selecionadas 20, incluídas na interface - todas em forma de captura de tela estática, devido a questões de compatibilidade - por meio de duas contas distintas da *OpenAI*: dez em uma e dez em outra. A Figura 10 ilustra esse processo.



O *ChatGPT* retornou prontamente às indagações, dividindo suas respostas em quatro tópicos, de forma alinhada ao roteiro encaminhado. O material recebido foi analisado, categorizado por frequência de expressões e transformado em visualizações interativas, de modo a favorecer a leitura (Figuras 11 e 12). Todavia, é importante pontuar que, ao explorar conteúdos feitos pela IA, sua estrutura baseada em

padrões probabilísticos de linguagem levanta questionamentos sobre a confiabilidade das interpretações, tornando a validação humana uma condição *sine qua non*.

Figura 12

Expressões utilizadas pelo ChatGPT na análise de imagens 360° do TikTok



Nota. Gráfico gerado pelos autores a partir do *Flourish* (2025), com versão *on-line* interativa disponível em <https://public.flourish.studio/visualisation/24996283/>

No caso das imagens provenientes do Instagram, o sistema conseguiu identificá-las corretamente como pertencentes à modalidade 360 graus, chegando inclusive a descrever alguns dos formatos nelas perceptíveis, como “pequeno planeta” (*tiny planet*), “projeção esférica” e “olho de peixe” (*fisheye*). Para explicar esse tipo de imagem, a ferramenta enumerou exemplos de câmeras especiais que permitem sua criação - registrando o espaço em todos os ângulos -, como *Insta360*, *GoPro Max* e *Ricoh Theta*.

Ao discutir a utilidade e as informações transmitidas pelas imagens analisadas, a plataforma de IA citou que elas apresentam variadas finalidades: turísticas e patrimoniais, pois expõem museus, praças, ruas e cidades de maneira diferenciada, valorizando as instituições culturais; documentais, por atuarem como arquivos detalhados de preservação urbana e arquitetônica; de engajamento, pelo impacto visual que podem provocar nas redes sociais virtuais; educacionais, ao permitirem o acesso de estudantes a espaços históricos sem a necessidade de deslocamento presencial; e imersivas, ao possibilitarem uma sensação de presença nos locais - sobretudo quando em formato tridimensional, para consumo em óculos de realidade virtual.

O *ChatGPT* também compartilhou dados que pôde extrair após o processamento desses arquivos, como os de localização, reconhecendo nas imagens espaços culturais - entre eles, o Museu do Ipiranga, o Museu da Imigração e o Museu Dom José (todos no Brasil) - além da identificação de ruas e praças de cidades brasileiras e portuguesas. Puderam ser colhidas ainda referências em termos de organização urbana, de proporção dos ambientes (a partir da escala de objetos e pessoas representados), bem como de horário, clima e identidade visual de cada local.

As respostas relacionadas ao TikTok foram semelhantes. A plataforma conseguiu caracterizar corretamente a modalidade pertencente às imagens, constatando nas peças ambientes como o *Palace Museum* (China), o *Louvre* (França), o *Guggenheim* (Itália), o *Moco Museum* (Espanha) e o Museu Oscar Niemeyer (Brasil). Percebeu ainda a aplicação de efeitos de pós-produção, mencionado o *zoom out* esférico e transições de perspectiva, explorados em redes sociais virtuais. Quanto

à natureza desse tipo de conteúdo, apontou que pode ser classificada a partir de três eixos: fotografia/vídeo, produto para divulgação na web e narrativas híbridas - estas últimas combinando documentação com inovação estética e artística.

Discorrendo sobre as funções desse formato, o assistente de IA destacou seus impactos para o turismo virtual, educação patrimonial, marketing cultural e experimentação visual, acentuando que as imagens investigadas revelam aspectos de localização, clima, estilo arquitetônico e interação dos visitantes com espaços históricos. Ao tratar dos elementos relevantes observados na amostra, a ferramenta indicou ainda a possibilidade de extrair metadados culturais, procedimentos técnicos, escolhas narrativas e contribuições para pesquisas dedicadas, por exemplo, à análise da plataformização em instituições museológicas.

Fundamentando-se no experimento realizado, é possível afirmar que a inteligência artificial pode atuar como método digital para o desenvolvimento de investigações sobre imagens em 360 graus, fornecendo pontos de partida a serem lapidados pelos pesquisadores. No entanto, a checagem das informações concedidas, associada a um ceticismo saudável diante desses assistentes - sujeitos a alucinações -, constitui fator fundamental para profissionais de todas as áreas, ainda mais no campo científico. À vista disso, surgem implicações epistemológicas substanciais, questionando até que ponto o *ChatGPT* deve ser considerado mero instrumento ou coautor interpretativo. Essa e outras indagações exigem atenção crítica e reflexão metodológica contínua, de forma que as leituras humanas prevaleçam nesse processo.

Técnicas de Rastreamento Ocular em Imagens Esféricas: *ScanGAN360* e Monitoramento Simultâneo em Realidade Virtual

Analisar a recepção de conteúdos em 360 graus navegáveis (Falandes, 2025a) pelo público configura-se tarefa desafiadora para produtores e pesquisadores. Diferentemente da imagem tradicional - estática e em movimento -, calcada na lógica do enquadramento, múltiplas direções podem ser seguidas pelos usuários nas experiências imersivas, as quais permitem a visualização de diferentes ângulos das cenas. Essa liberdade de exploração proporcionada aos interatores torna mais complexa a avaliação de seus comportamentos durante a fruição da narrativa.

Um dos métodos digitais que podem ser acionados para auxiliar nesse processo é o uso da ferramenta *ScanGAN360* (Martin et al., 2022), baseada em IA, especificamente em uma arquitetura de rede adversarial generativa condicional (*cGAN*), pela qual é possível simular a trajetória do olhar humano em ambientes em 360 graus a partir da geração de *scanpaths* (caminhos de varredura). A aplicação mostra-se eficaz para o design de produções imersivas, contribuindo para avaliar se as escolhas visuais planejadas - como o posicionamento de *hotspots* nas cenas - estão alinhadas aos pontos de maior atenção sugeridos pela plataforma, a qual busca reproduzir a perspectiva de um grande número de observadores virtuais.

Nesse sentido, realizou-se um experimento em 7 de agosto de 2025 a fim de testar sua aplicabilidade. Após a conclusão das etapas de instalação do programa, foi selecionada para processamento uma imagem esférica em formato equiretangular, de posse dos autores. O registro foi feito por ocasião da criação da narrativa documental

“Conjunto do Ipiranga em 360 Graus”, citada anteriormente neste artigo (Figura 13). Como resultado, foram obtidos dois painéis com miniaturas dessa mesma imagem, sobre as quais foram sobrepostas as análises da *ScanGAN360* (Figura 14).

Figura 13

Imagem 360° em dos Jardins Fronteiros do Museu do Ipiranga (SP) utilizada no experimento

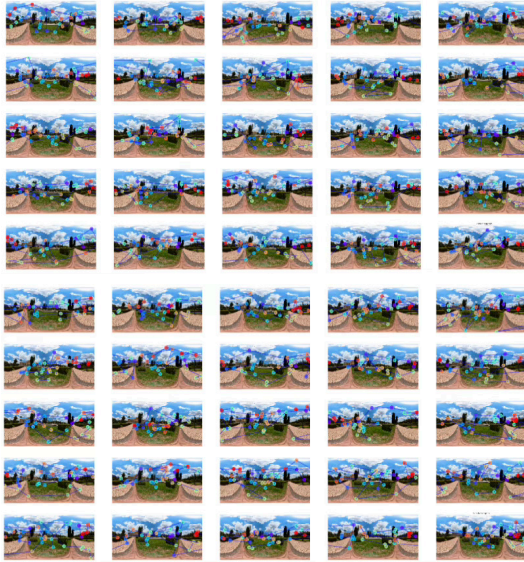


Autoria própria.

Ao examinar os conteúdos gerados pela plataforma, nota-se a presença de elementos visuais distribuídos de forma diferenciada ao longo de toda a extensão da paisagem retratada nas miniaturas. Os círculos coloridos simbolizam os locais onde o modelo considera que o usuário olharia - os pontos de fixação -, enquanto as setas representam a direção do movimento ocular. Verifica-se, assim, uma certa concentração de olhares para a área central da imagem, em que está situada a fachada do Museu do Ipiranga, bem como pontos de vista direcionados para as laterais, nas quais avistam-se árvores. A partir desse breve relato sobre a utilidade da *ScanGAN360*, pode-se reconhecer a validade de suas previsões para o campo das imagens em 360 graus, servindo como parâmetro para a construção de conteúdos.

Figura 14

Resultados obtidos a partir da ScanGAN360

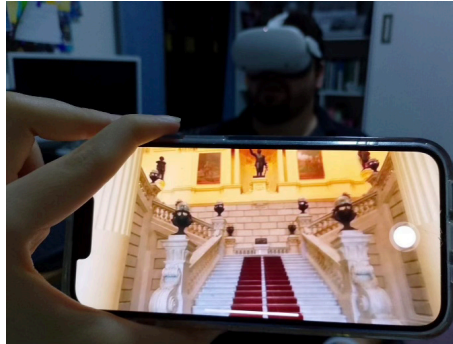


Nota. Materiais obtidos a partir da ferramenta *ScanGAN360* (Martin et al., 2022).

Além dessa estratégia, cabe mencionar a técnica de monitoramento simultâneo em realidade virtual, que consiste em acompanhar, via espelhamento de tela, os caminhos seguidos pelo experienciador durante o consumo de narrativas esféricas. Essa abordagem revela-se útil em experimentos com usuários reais, como demonstrado por Falandes (2025a), que avaliou a recepção de indivíduos em relação à sua peça documental imersiva, utilizando o aplicativo *Meta Horizon* (Meta Platforms, Inc., 2025) em conjunto com um *headset* do modelo *Quest 2* (Figura 15).

Figura 15

Monitoramento de usuário interagindo com narrativa 360°



Autoria própria.

Conclusões

Diante da investigação conduzida por este artigo, pôde-se conhecer o potencial de Métodos Digitais para o exame de imagens, especialmente as produzidas a partir da tecnologia 360 graus. Foram apresentadas quatro abordagens, cada uma com plataformas e protocolos específicos, alicerçadas no tripé coleta-análise-visualização de dados.

A intenção foi apontar possibilidades metodológicas acessíveis - baseadas em ferramentas disponíveis na web, algumas envolvendo desafios de instalação, mas totalmente superáveis por meio de tutoriais *on-line* -, capazes de servir tanto a pesquisadores do tema, buscando compreender melhor as movimentações de usuários em rede no que tange a essa modalidade visual, quanto a realizadores, dispostos a entender o comportamento da audiência e aperfeiçoar o design de suas experiências imersivas.

Longe de fornecer caminhos imutáveis, este estudo se prestou, antes de tudo, a experimentar. Nesse processo, deparou-se com

limitações, como durante a inclusão de arquivos de imagem na versão gratuita do *ChatGPT*. Assim, detentores da licença completa dessa plataforma poderão alcançar melhores resultados. Ainda, para a leitura de conteúdos audiovisuais, dispositivos analíticos mais robustos são recomendados - o *Move Scene Sensing*, do *Datalab Design* (s.d.), parece eficaz nesse contexto.

Espera-se, por meio do panorama traçado, contribuir para o campo dos Métodos Digitais e suas aproximações com conteúdos esféricos, demonstrando a relevância desse arranjo para as Ciências Sociais Aplicadas.

Referências

- Angeluci, A. C. B., Caldas, C. H. S., & Calixto, G. M. (2024). Práticas interacionais ao vivo nas plataformas digitais: Estudo de caso da CNN Brasil no YouTube. *Galáxia: Revista Interdisciplinar de Comunicação e Cultura*, 49(1), Article e66563. <https://revistas.pucsp.br/index.php/galaxia/article/view/66563>
- Bitencourt, E. (2025). Além do design: interfaces algorítmicas como agentes epistemológicos. *MATRIZES*, 19(1), 199-224. <https://doi.org/10.11606/issn.1982-8160.v19i1p199-224>
- Casini, J. G., & Lovato, A. (2023). Experiencias inmersivas: Hacia una gramática narrativa e interactiva. *Chasqui: Revista Latinoamericana de Comunicación*, 1(154). <https://doi.org/10.16921/chasqui.v1i154.4931>
- Cirino, G. F. C., Filgueiras, I. S., & Angeluci, A. C. B. (2025). Visualização de dados para políticas públicas de saúde: a experiência do DengueMap. *RECIIS*, 19(1), 1-14. <https://doi.org/10.29397/reciis.v19i1.4537>

- Colombo, G. (2019). Studying digital images in groups: the folder of images. In L. Rampino, & I. Mariani (Eds.), *Advancements in design research: 11 PhD theses on Design as we do in POLIMI* (pp. 185-195). FrancoAngeli. <https://series.francoangeli.it/index.php/oa/catalog/book/376>
- Colombo, G., Bounegru, L., & Gray, J. (2023). Visual Models for Social Media Image Analysis: Groupings, Engagement, Trends, and Rankings. *International Journal Of Communication*, 17, 28. <https://ijoc.org/index.php/ijoc/article/view/18971>
- D'Andréa, C. (2020). *Pesquisando plataformas online: conceitos e métodos*. EDUFBA. <https://repositorio.ufba.br/handle/ri/32043>
- Damas, S. H., & De Gracia, M. J. B. (2022). Immersive journalism: advantages, disadvantages and challenges from the perspective of experts. *Journalism and Media*, 3(2), 330-347. <https://doi.org/10.3390/journalmedia3020024>
- Datalab Design. (s.d.). *Datalab Design*. Universidade do Estado da Bahia. <https://datalabdesign.org/en>
- Falandes, C. G. (2025a). *Uma proposta de linguagem para o design das narrativas audiovisuais de não-ficção em 360 graus* [Tese de Doutorado, Universidade Estadual Paulista - UNESP]. Repositório Institucional UNESP.
- Falandes, C. G. (2025b). Conjunto do Ipiranga em 360°: explorando todos os lados da Independência do Brasil [Experiência em vídeo 360°]. <https://wvr.li/6qyw11>
- Falandes, C. G., & Angeluci, A. C. B. (2019). Usos e apropriações da imagem em 360 graus no Instagram. In M. P. Caprino, & P. F. Perazzo

(Orgs.), *Comunicação e cultura: aproximações com memória e história oral - Diálogos entre Brasil e México* (pp. 722739). USCS.

Falandes, C. G., & Renó, D. P. (2021). Centros de producción e investigación de imágenes en 360 grados: una cartografía de no-ficción de inmersión en la academia iberoamericana. *Hipertext.net*, (23), 117-126. <https://doi.org/10.31009/hipertext.net.2021.i23.11>

Falandes, C. G., & Renó, D. P. (2022). Narrativas inmersivas durante la pandemia de Covid-19: Un análisis de videos 360 grados en YouTube. *Revista Panamericana de Comunicación*, 4(1), 84-97. <https://doi.org/10.21555/rpc.v4i1.2554>

Falandes, C. G., & Renó, D. P. (2023). Documentary poetics in 360 degrees: perspectives on non-fiction spherical narrative models. *Brazilian Journalism Research*, 19(2), e1555. <https://doi.org/10.25200/BJR.v19n2.2023.1555>

Flores, A. M. M. (2019). Produção e consumo de vídeos em 360° — tendências para o jornalismo brasileiro no YouTube. In J. J. Omena (Ed.), *Métodos digitais: Teoria-prática-crítica* (pp. 183-201). ICNOVA. <https://novaresearch.unl.pt/en/publications/m%C3%A9todos-digitais-teoria-pr%C3%A1tica-cr%C3%ADtica>

Flourish. (2025). *Flourish* [Software]. <https://flourish.studio/>

Fontcuberta, J. (2016). *La furia de las imágenes: notas sobre la postfotografía*. Galaxia Gutenberg.

Giacomelli, F. O. (2023). *Jovens e jornalismo: A relação das novas gerações com os recursos jornalísticos imersivos em ambientes virtuais* [Tese de Doutorado, Universidade da Beira Interior - UBI]. Repositório Digital da UBI.

- Grohmann, R. (2025). Latin American critical data studies. *Big Data & Society*, 12(2). <https://doi.org/10.1177/20539517251330160>
- iNOVA Media Lab. (s.d.). *SMART Data Sprint*. <https://metodosdigitais.fcsh.unl.pt/>
- Johnson, O. V., Alyasiri, O. M., Akhtom, D., & Johnson, O. E. (2023). Image Analysis through the lens of ChatGPT-4. *Journal of Applied Artificial Intelligence*, 4(2), 31-46. <https://doi.org/10.48185/jaai.v4i2.870>
- Knight Lab. (2025). StoryMap JS [Software]. Northwestern University. <https://storymap.knightlab.com/>
- Kozharinova, M., & Manovich, L. (2024). Instagram as a narrative platform. *First Monday*, 29(3). <https://doi.org/10.5210/fm.v29i3.12497>
- Lapa, T. (2025). Métodos de inquirição online: transcendendo os limites tradicionais na pesquisa social. In G. Cardoso, & R. Sepúlveda (Orgs.), *Manual de Métodos para Pesquisa Digital* (pp. 153-163). Mundos Sociais. <https://repositorio.iscte-iul.pt/handle/10071/34301>
- Lima, L., & Barbosa, S. (2022). Mobilidade, imersão e experiência: entre ajustes e discordâncias conceituais no jornalismo imersivo. *Estudos em Comunicação*, (34), 1-11. <https://doi.org/10.25768/1646-4979n34-01>
- Martin, D., Serrano, A., Bergman, A. W., Wetzstein, G., & Masia, B. (2022). *ScanGAN360* [Software]. GitHub. <https://github.com/DaniMS-ZGZ/ScanGAN360>

- Médola, A. S. L. D., & Oliveira, B. J. de. (2017). Estratégias enunciativas em Google Spotlight Stories: o olhar da semiótica de Greimas nos vídeos 360°. *Significação: Revista de Cultura Audiovisual*, 44(48), 68-85. <https://doi.org/10.11606/issn.2316-7114.sig.2017.136716>
- Meta Platforms, Inc. (2025). *Meta Horizon* (Versão 335.0.0.11.109) [Aplicativo móvel]. App Store Brasil. <https://apps.apple.com/br/app/meta-horizon/id1366478176>
- Niederer, S. (2018). *Networked images: visual methodologies for the digital age*. Hogeschool van Amsterdam.
- Niederer, S., & Colombo, G. (2024). *Visual methods for digital research: an introduction*. Polity.
- Omena, J. J., Pilipets, E., Gobbo, B., & Chao, J. (2021). The Potentials of Google Vision API-based Networks to Study Natively Digital Images. *Diseña*, (19), Article 1. <https://doi.org/10.7764/disena.19.Article.1>
- Omena, J. O. (2019a). Introdução: O que são métodos digitais? In J. J. Omena (Ed.), *Métodos digitais: Teoria-prática-crítica* (pp. 5-15). ICNOVA. https://run.unl.pt/bitstream/10362/95363/1/ICNOVA_MetodosDigitais_compactado_5_15.pdf
- Omena, J. O. (Ed.). (2019b). *Métodos digitais: Teoria-prática-crítica*. ICNOVA. <https://novaresearch.unl.pt/en/publications/m%C3%A9todos-digitais-teoria-pr%C3%A1tica-cr%C3%ADtica>
- OpenAI. (2025). ChatGPT (GPT-5) [Large language model]. <https://chat.openai.com/>

- Pearce, W., Özkula, S. M., Greene, A. K., Teeling, L., Bansard, J. S., Omena, J. J., & Rabello, E. T. (2018). Visual cross-platform analysis: digital methods to research social media images. *Information, Communication & Society*, 23(2), 161-180. <https://doi.org/10.1080/1369118X.2018.1486871>
- Peeters, S. & Hagen, S. (2022). The 4CAT capture and analysis toolkit: A modular tool for transparent and traceable social media research. *Computational Communication Research*, 4(2), 571-589. <https://doi.org/10.5117/CCR2022.2.007.HAGE>
- Peeters, S. (2025). *Zeeschuimer* (v1.13.1) [Software]. Zenodo. <https://doi.org/10.5281/zenodo.15281614>
- Recuero, R. (2025). *Métodos digitais e epistemologias emergentes no estudo do discurso em plataformas* [Trabalho apresentado em congresso]. 34º Encontro Anual da Compós. Recuperado de <https://publicacoes.softaliza.com.br/compos2025/article/view/11384>
- Renó, D. (2024). *Fotografia contemporânea de viagem* (1a ed.). Ria Editorial. <https://www.riaeditorial.com/livro/fotografia-contemporanea-de-viagem>
- Rocha, A. E., & Rocha, P. M. (2024). Consumo de Jornalismo Imersivo: experiência, imersão, sensação de presença, empatia e outras manifestações subjetivas percebidas por usuários. *Revista Eco-Pós*, 27(2), 352-377. <https://doi.org/10.29146/eco-ps.v27i2.28127>
- Rocha, A.-E., & Rocha, P.-M. (2022). Características estructurales de los estudios sobre Periodismo Inmersivo 2010-2020. *Revista Mediterránea de Comunicación*, 13(2), 203-222. <https://doi.org/10.14198/MEDCOM.22087>

- Rodrigues, L. C. S., & Lima, L. (2022). Plataformização e vídeo 360°: implicações para o jornalismo no Brasil. *Âmbitos - Revista Internacional de Comunicación*, (56), 86-104. <https://doi.org/10.12795/Ambitos.2022.i56.06>
- Rodrigues, L., Firmino, F., & Cunha, R. (2022). Câmeras 360 graus e Jornalismo. Indexação de pesquisas teórico e práticas do Brasil. *Revista Estudos em Jornalismo e Mídia*, 19(2), 2022. <https://doi.org/10.5007/1984-6924.2022.82620>
- Rogers, R. (2013). *Digital methods*. The MIT Press.
- Rogers, R. (2021). Visual media analysis for Instagram and other online platforms. *Big Data & Society*, 8(1). <https://doi.org/10.1177/205395172111022370>
- Silva, T., Mintz, A., Omena, J. J., Gobbo, B., Oliveira, T., Takamitsu, H. T., Pilipets, E., & Hamdan, A. (2020). APIs de Visão Computacional: investigando mediações algorítmicas a partir de estudo de bancos de imagens. *Logos*, 27(1). <https://doi.org/10.12957/logos.2020.51523>
- Smart Media & Users. (s.d.). *Smart Media & Users*. Universidade de São Paulo. <https://smartmediausers.org/>
- Teixeira, M. (2024). *Jornalismo imersivo profundo como experiência ética e estética* [Tese de doutorado, Universidade Estadual Paulista - UNESP]. Repositório Institucional UNESP.
- University of Amsterdam. (s.d.-a). *Digital Methods Summer School*. <https://summerschool.uva.nl/content/summer-courses/digital-methods-summer-school/digital-methods-summer-school.html?origin=PtN3uwsWQquDIgtnNX3fZg>

University of Amsterdam. (s.d.-b). *Digital Methods Winter School*. <https://summerschool.uva.nl/shared/programmas/en/winter-courses/digital-methods-winter-school-program/digital-methods-winter-school-program.html?origin=VW%2FXiBtNQwSjxR%2BVJO%2B5dw>

Vasconcelos, E. L. (2024). *Análise de visualidades jornalísticas: aplicação de métodos digitais na pesquisa com imagens em jornalismo* [Tese de Doutorado, Universidade Federal da Bahia - UFBA]. Repositório Institucional da UFBA.

Vázquez-Herrero, J., Negreira-Rey, M. C., & Klug, D. (2025). Key Concepts and Research Methods to Analyze TikTok Journalism. *Cultural Science Journal*, 15(1), 42-62. <https://doi.org/10.2478/csj-2023-0005>

Wonda VR. (2025). Wonda VR [Software]. <https://www.wondavr.com/>

DA TEORIA À PRÁTICA: O USO DE INTELIGÊNCIA ARTIFICIAL GENERATIVA DE TEXTOS NA IMPLEMENTAÇÃO DA LINGUAGEM SIMPLES EM DOCUMENTOS JURÍDICOS

Renato Sobhie Zambonato¹

Este artigo examina o uso de inteligência artificial generativa como mediação para implementar linguagem simples em textos jurídicos, aproximando as diretrizes técnicas com a prática cotidiana. A literatura sobre linguagem simples consta em diretrizes como a ISO 24495-1:2023 (e a 24495-2:2025), o PL 6.256/2019 e a Recomendação nº 144/2023 do CNJ, com a mudança de paradigma para uma comunicação centrada no usuário.

1. Mestrando do Programa de Pós-graduação em Comunicação da Faculdade de Arquitetura, Artes e Comunicação da Universidade Estadual Paulista (Unesp). renato.zambonato@unesp.br

Delimita-se o problema na persistente distância entre tais normas e a produção efetiva de textos claros; aqui, pergunta-se até que ponto a IA pode reduzir essa lacuna sem comprometer fidelidade jurídica e precisão técnica. A justificativa reside no fato de que clareza textual é condição para legitimidade comunicacional e até mesmo como condição de acesso à justiça, mas que permanece sendo uma técnica não dominada pelos operadores do Direito, seja em âmbito judicial administrativo ou legislativo. Aqui, parte-se do princípio que documentos jurídicos devem ser elaborados por advogados, quando caracterizada atividade privativa da advocacia, ou por pessoas com formação técnica suficiente.

Para explorarmos se a IA seria capaz de facilitar a criação de textos em linguagem simples, mantendo o rigor técnico e seguindo diretrizes específicas, o artigo inclui um teste controlado de reescrita em trecho legislativo específico, no qual comandos simples são apresentados para três modelos distintos de IA, com a comparação dos resultados

Objetiva-se avaliar o potencial da IA para operacionalizar linguagem simples em documentos jurídicos, mapeando diretrizes, conduzindo experimento de reescrita com modelos generativos, medindo compreensibilidade e fidelidade com avaliação humana posterior, recomendando-se estudos posteriores mais aprofundados sobre o tema.

Contextualização sobre a linguagem técnica jurídica

A comunicação, no ecossistema midiático atual, distribui-se em redes, plataformas e interfaces que aceleram a produção, a filtragem e a circulação de sentidos. Nesse contexto, textos técnicos e jurídicos enfrentam pressão por clareza, objetividade e acessibilidade, não apenas como estilo, mas como condição de legitimidade comunicacional, que

envolve desde a mediação de informações cotidianas até a efetivação de direitos fundamentais.

Especificamente no campo jurídico, persiste um modo de expressão caracterizado pelo excesso de tecnicismos, jargões e construções rebuscadas: o chamado juridiquês (também conhecido como legalese).

O juridiquês, termo aqui usado para indicar a linguagem técnica jurídica, por sua tradição formal e sintaxe hermética, opera como barreira simbólica e cognitiva; nesse contexto, a linguagem simples (plain language) emerge como contra-padrão, alterando o foco do texto e produzindo uma redação centrada no usuário, para que pessoas encontrem, compreendam e usem a informação. Essa agenda é materializada através de diversas iniciativas, como o PL 6.256/2019 (Política Nacional de Linguagem Simples), ISO 24495-1:2023 e ISO 24495-2:2025; Recomendação CNJ nº 144/2023, além de iniciativas institucionais e privadas que já testam IA generativa como mediadora entre diretrizes e prática cotidiana de escrita jurídica.

O movimento em defesa da linguagem simples (plain language) surge como contraponto a essa lógica forma e excludente do discurso técnico e rebuscado. Trata-se de uma estratégia comunicacional que busca assegurar que qualquer pessoa, independentemente de formação técnica, possa, em linhas gerais, encontrar, compreender e usar a informação que lhe é destinada, através de construções mais simples e menos técnicas.

Em termos conceituais, temos diversas iniciativas que buscam versar sobre o tema. Temos a norma ISO 24495-1:2023, que fala sobre plain language e foi recentemente complementada pela norma ISO 24495-2:2025, específica sobre comunicação jurídica. Tais normas estabelecem critérios universais para redação clara e centrada no leitor,

sendo replicadas em diversas iniciativas distintas. No Brasil, a preocupação se materializou no Projeto de Lei nº 6.256/2019, que institui a Política Nacional de Linguagem Simples, bem como na Recomendação nº 144/2023 do CNJ, que orienta tribunais a adotarem comunicação acessível em todos os atos administrativos e judiciais e no Pacto Nacional do Judiciário pela Linguagem Simples, entre outros.

Além dessas iniciativas com de maior abrangência, podemos mencionar também diversas iniciativas mais pontuais, como a Rede Linguagem Simples Brasil, o Programa Municipal de Linguagem Simples da cidade de São Paulo e o Projeto Linguagem Simples e Inclusiva, da Unicamp, entre inúmeros outros.

Todas essas iniciativas apontam para um movimento amplo da comunicação social, de busca pela mudança de centro de poder e de foco. Se antes o poder de informar estava concentrado nos meios de massa e nos detentores do discurso especializado, hoje observa-se que os cidadãos produzem e exigem informações compreensíveis, que não precisem de intermediários para serem entendidas.

A circulação de sentidos tornou-se mais horizontal, plural e imediata. Nesse ambiente, a escrita linear e hermética cede espaço a linguagens que privilegiam clareza, objetividade e acessibilidade, não apenas como recurso estilístico, mas como condição de legitimidade discursiva. A linguagem simples, nesse contexto, apresenta-se como mais do que uma simples técnica de redação, mas também como uma adaptação necessária para criação de um cenário comunicacional mais democrático.

Podemos argumentar que o descompasso entre o ideal da linguagem simples e a prática da linguagem técnica, especialmente a jurídica, se explica por um arranjo de fatores que se reforçam mutuamente.

A cultura profissional ainda valoriza o juridiquês como marcador de pertencimento e de autoridade simbólica, o que incentiva escolhas estilísticas que priorizam o status do emissor, não a compreensão do leitor. Soma-se a isso uma capacitação insuficiente: escrever de forma clara exige critérios, treino e revisão sistemática, de modo a não ser uma competência simplesmente “intuitiva”. Por fim, a pressão de tempo e de risco na prática forense (aqui apresentada da maneira mais ampla possível) favorece a reprodução de modelos prontos, desenhados para a segurança técnica e a velocidade interna, mas pouco atentos às necessidades de quem está fora do circuito jurídico. O efeito combinado é facilmente verificável: textos tecnicamente corretos e até defensáveis, porém que não alcançam aqueles que precisam agir, participar ou exercer direitos.

As inteligências artificiais generativas de texto como ferramentas possíveis

Nesse cenário, as IAs generativas despontam como mediadoras capazes de encurtar essa distância. Ao aprender padrões de escrita, elas são capazes de reescrever trechos longos em versões mais objetivas sem suprimir informações essenciais; de traduzir jargões através do uso de definições e exemplos acessíveis; de reorganizar a informação para trazer a ideia principal ao início e, sobretudo, são capazes de reorganizar o conteúdo, de modo geral, para que o leitor consiga realmente compreender o conteúdo. Quando integradas a fluxos de trabalho com critérios claros e revisão humana, essas ferramentas deslocam o eixo da comunicação jurídica: em vez de preservar a opacidade como vestígio de autoridade, ajudam a transformar a clareza em critério de legitimidade e efetividade comunicativa.

Assim, as inteligências artificiais generativas de texto, como ChatGPT, Grok, Claude e Gemini, entre diversos outros, emergem como mediadoras potenciais para diminuir a distância entre teoria e prática. Essas ferramentas demonstram capacidade de reescrever textos complexos em versões simplificadas, principalmente através de direcionamentos específicos.

No Brasil, temos experiências, como o Sistema de Orientação e Facilitação de Informações e Acessibilidade (SOFIA), de iniciativa pública pelo Tribunal de Justiça de Minas Gerais, que é uma ferramenta desenvolvida para aprimorar a comunicação do Poder Judiciário com a sociedade por meio da linguagem simples. Na iniciativa privada encontram-se diversas iniciativas parecidas, como a ferramenta “Entender Mais”, do Jusbrasil, que usa a inteligência artificial para simplificar termos jurídicos, permitindo uma comunicação mais direta com os usuários. Ambos cenários apontam na direção de que a tecnologia pode ajudar a reduzir a distância entre a teoria da linguagem simples e a prática.

Diante desse panorama, este artigo propõe explorar como a linguagem simples, articulada às ferramentas de IA generativa, pode aproximar a teoria normativa e comunicacional da realidade institucional. A partir da análise de algumas diretrizes constantes nas normas de maior abrangência (PL 6.256/2019, ISO 24495-1:2023 e Recomendação nº 144/2023) e de um exercício empírico de simplificação textual por diferentes modelos de IA, busca-se avaliar se tais tecnologias podem funcionar como catalisadoras da democratização da comunicação pública e jurídica.

Alucinações e a necessidade de checagem das informações

A depender do tipo de interação realizada com a IA, é importante que seja conhecida a possibilidade de ocorrência das “alucinações”. Conteúdos gerados pelos modelos de linguagem, que fazem sentido textual e aparentam serem plausíveis, mas que possuem informações falsas, recebem essa nomenclatura e são muito presentes nos modelos mais antigos.

Um recente estudo publicado pela OpenAI, nomeado de “Why Language Models Hallucinate”, indicou que os modelos de linguagem tentam “adivinhar” quando não sabem a resposta para uma determinada pergunta, o que resulta nas “alucinações”. Esse comportamento seria gerado pelo modo de treinamento da tecnologia, que acabou “premiando” respostas corretas, apesar de serem “chutes” da ferramenta.

As alucinações não parecem um problema para o contexto aqui discutido, por dois motivos: a) a simplificação de texto não envolve uma pergunta e, por consequência, não gerará uma resposta que necessite de acurácia factual, mas apenas uma simplificação de um texto previamente apresentado, apesar de ser possível a ocorrência de simplificações excessivas; b) a diminuição da ocorrência das alucinações em modelos mais novos, que indicam não saberem a resposta, e não mais apenas tentam “adivinhar” a resposta correta, com uma taxa de abstenção de 52% para o modelo gpt-5-thinking-mini, na avaliação do SimpleQA, enquanto o modelo OpenAI o4-mini possuía taxa de abstenção de apenas 1%.

De qualquer modo, continua presente a necessidade de checagem das informações apresentadas pelas IAs generativas de texto, informação constante nos termos de uso das próprias ferramentas e cuja inobservância não pode ser usada para questionar a resposta em

contextos de simplificação textual. A Anthropic, por exemplo, traz em seus termos de uso que “você não deve confiar em nenhuma Saída ou Ação sem confirmar independentemente sua precisão.”, em tradução livre, enquanto os termos da OpenAI dispõem que o uso de seus serviços “pode, em determinadas situações, gerar Resultados que não refletem com precisão pessoas, locais ou fatos reais.”, também em tradução livre.

Ainda, apenas para entendermos como a checagem poderia ser feita, parte-se da condição de que os textos jurídicos devem ser elaborados e, por consequência, checados por advogados, quando caracterizada atividade privativa da advocacia e nos termos do Estatuto da Advocacia e a Ordem dos Advogados do Brasil (OAB), ou por pessoas com formação técnica suficiente para tanto, a depender do contexto.

É importante deixar claro que a simplificação de textos jurídicos sempre deve ser seguida por um momento de verificação técnica por uma pessoa capacitada, nos moldes das próprias diretrizes das ferramentas utilizadas. A ideia da simplificação é que sejam seguidas as bases conceituais da linguagem simples, mas a acurácia técnica depende de outros inúmeros fatores, como contextualização e descrição, que não são esperados de um modelo de IA e que não se confundem com o que se espera da linguagem simples.

Para exemplificação, um recente estudo chamado “Are LLM-generated plain language summaries truly understandable? A large-scale crowdsourced evaluation” avaliou a geração de sumários em linguagem simples (chamados de PLS), por meio de IAs, e trouxe diversos pontos de atenção em sua conclusão.

O estudo aponta uma lacuna nítida entre a qualidade percebida e a compreensão efetiva de PLSs gerados por LLMs: ainda que os textos

sejam avaliados como claros e fluentes, o desempenho dos leitores em tarefas estruturadas de compreensão permanece aquém do esperado, evidenciando que métricas de legibilidade e autoavaliações positivas não se traduzem, isoladamente, em entendimento real. Também trazem a indicação de que seus achados “reforçam a importância de distinguir entre qualidade percebida e compreensão efetiva”, recomendando uma nova metodologia de avaliação, dessa vez centrados no leitor, para avaliação de sucesso dos textos gerados por IA.

O mesmo estudo aponta que a “fluência estilística”, por si só, não assegura entendimento efetivo e, portanto, não deve orientar isoladamente a geração nem a avaliação de “sumários em linguagem simples”, além de, entre outras conclusões, recomendar que pesquisas futuras explorem estratégias de geração otimizadas diretamente para o entendimento humano e desenvolvam métricas capazes de aferir, de modo confiável, resultados de compreensão real.

Portanto, a checagem de informações desponta como uma boa-prática, indicada pelas próprias ferramentas, que não apenas possui valor relacionado ao rigor técnico do conteúdo gerado, mas também com relação à compreensão real por parte dos leitores, que precisam de mais do que apenas “fluência estilística” para, de fato, entenderem o texto formulado pelas IAs.

Fundamentos da Linguagem Simples

A noção de linguagem simples (ou *plain language*), também chamada de “linguagem clara” toma força como um movimento de democratização da comunicação escrita, sustentado pelo princípio de que todo cidadão tem o direito de compreender, de forma direta e

sem intermediários, as informações que lhe dizem respeito. Apesar de diferenças conceituais, entende-se que, no seu núcleo, a linguagem simples não é apenas uma técnica de redação, mas também uma estratégia comunicacional centrada no usuário final, que deve ser capaz de localizar o que precisa, compreender o que encontra e usar a informação de forma eficaz em sua vida cotidiana.

A ISO 24495-1:2023, que abrange a primeira parte da documentação sobre linguagem simples, com conteúdo sobre princípios e diretrizes de governança, é o padrão técnico internacional sobre o tema. Tal norma consolidou a base da conceituação ao definir a linguagem simples como “comunicação em que redação, estrutura e design são tão claros que os leitores podem facilmente encontrar o que precisam, entender o que encontram e usar essa informação”, em tradução livre.

A norma estabelece ainda que o conceito de “leitor” não se restringe ao sujeito que pratica o verbo “ler”, mas inclui diferentes formas de interação com o texto, desde quem acessa visualmente até quem o recebe por meio de leitura em voz alta, dispositivos eletrônicos ou recursos de acessibilidade. Essa concepção amplia o horizonte da linguagem simples ao vinculá-la à noção de inclusão, garantindo que diferentes públicos possam se beneficiar de uma comunicação efetiva, com uma ideia mais abrangente de usuário do documento, e não apenas do leitor.

Aqui, cabe a menção à segunda parte da norma, que foi recentemente lançada sob a numeração “ISO 24495-2:2025”, e refere-se à comunicação jurídica, mas que não será objeto de análise, já que a primeira parte se mostra suficiente para os fins aqui propostos.

No Brasil, a proposta de institucionalização da linguagem simples foi incorporada ao Projeto de Lei nº 6.256/2019, que institui a Política

Nacional de Linguagem Simples nos órgãos da administração pública. O projeto parte de objetivos claros listados no inciso II do artigo 2º, que inclui, em linhas gerais, a garantia de que atos governamentais sejam redigidos de maneira compreensível, redução da necessidade de intermediários, diminuição de custos administrativos e fortalecimento da participação popular, da transparência e do controle social. Entre os princípios enunciados, destacam-se o foco no cidadão, a redução das desigualdades e a promoção da linguagem inclusiva.

Diferente de outras iniciativas, como o Plain Writing Act norte-americano, de 2010, o texto traz formas diretas de operacionalização da teoria, através dos seguintes itens: indicação de redação de frases em ordem direta, frases curtas, desenvolvimento de uma ideia por parágrafo, uso de palavras comuns, a explicação de termos técnicos ou uso de sinônimos, o desestímulo de uso de palavras estrangeiras e de termos pejorativos, a organização do texto de forma esquemática, com uso de recursos não textuais (lisas, tabelas e gráficos), com a informação mais importante aparecendo primeiramente, além da indicação de não serem usadas novas formas de flexão de gênero e de número, prezando por regras gramaticais consolidadas.

Esse alinhamento entre a definição internacional da ISO e a proposta legislativa brasileira reflete um consenso em torno da centralidade do destinatário da mensagem. Se por muito tempo o discurso jurídico e administrativo priorizou a autoridade do emissor da ideia e a tradição formal, a linguagem simples desloca o eixo comunicacional para o usuário do documento, que pode, ou não, confundir-se com o leitor. Trata-se de uma mudança cultural que acompanha a evolução do ecossistema midiático contemporâneo: de um modelo vertical, em

que poucos falavam e muitos apenas recebiam, para um ambiente horizontal, caracterizado pela circulação de sentidos, pela interação e pela exigência social por clareza e objetividade.

Esse movimento também dialoga com teorias da comunicação que apontam para a transformação das linguagens na sociedade contemporânea. O desenvolvimento da web 2.0, por exemplo, deslocou a centralidade da mensagem única e massificada para a produção distribuída de conteúdo, na qual os cidadãos não apenas recebem, mas também produzem informação. Do mesmo modo, a ascensão das mídias digitais imagéticas e audiovisuais mostrou que a atenção e a confiança do público estão vinculadas à clareza e à acessibilidade do discurso. A linguagem simples, nesse sentido, pode ser compreendida como expressão dessa nova ecologia dos meios, em que a transparência comunicacional não é apenas desejável, mas condição de legitimidade.

Assim, podemos pensar que os fundamentos da linguagem simples articulam três dimensões complementares: (i) uma dimensão normativa, consagrada por padrões internacionais e legislações nacionais; (ii) uma dimensão comunicacional, que responde às transformações do ambiente midiático e à centralidade do usuário; e (iii) uma dimensão jurídica, que associa a clareza textual ao direito fundamental de acesso à informação e também à justiça. Esse tripé parece ser capaz de fornecer a base para discutirmos, nas seções seguintes, o contraste entre a persistência do juridiquês e as potencialidades abertas pelas tecnologias digitais e pela inteligência artificial generativa.

Aqui, mostra-se importante apontar que as diretrizes explícitas constantes na política nacional destoam da mera “fluência estilística” indicada no estudo anteriormente mencionado. Com regras mais pontuais

e ordens expressas, a conceituação se afasta do caráter meramente estilístico e se aproxima de uma norma procedimental e mensurável, que também pode ser analisada sob a ótica da compreensão real.

Mesmo que a questão da compreensão real seja de enorme importância, ela não é o objeto do presente estudo, que é focado na análise objetiva dos conceitos da linguagem simples, sendo que ambas as análises se mostram complementares, e não excludentes.

O Problema do Juridiquês

A tradição jurídica brasileira consolidou uma forma de escrita que, ao longo do tempo, passou a ser reconhecida não apenas como característica estilística, mas como marca de autoridade: o chamado juridiquês, já apresentado anteriormente.

Trata-se de um modo de comunicação permeado por latinismos, expressões antigas, longos trechos de texto e sintaxe rebuscada, que distancia o texto jurídico da linguagem cotidiana. Embora defendido por alguns setores como forma de garantir rigor técnico, o juridiquês opera, na prática, como barreira simbólica e cognitiva, restringindo o acesso à informação jurídica e à compreensão de direitos e deveres fundamentais.

Os efeitos dessa linguagem são amplos. Em primeiro lugar, ela exclui grande parte da população da participação informada nos processos judiciais, legislativos e administrativos, produzindo uma assimetria entre especialistas e leigos. Em segundo lugar, gera insegurança comunicacional, pois mesmo cidadãos envolvidos diretamente em processos judiciais frequentemente não compreendem os documentos que lhes dizem respeito, o que pode ser agravado em momentos processuais onde o advogado ou

juiz ainda não se fazem presentes. O juridiquês pode, ainda, ser visto como reserva de mercado simbólica, reforçando a exclusão comunicacional e afastando o cidadão da compreensão plena de seus direitos.

Do ponto de vista jurídico, essa exclusão comunicacional pode representar uma violação indireta do princípio do acesso à justiça, aqui mencionado apenas para ilustrar a abrangência do tema, muitas vezes visto apenas como um capricho estilístico ou estético.

A clássica formulação de Cappelletti e Garth, em análise superficial, situa esse princípio em ondas: a primeira voltada à assistência judiciária gratuita, a segunda à tutela de interesses coletivos e a terceira à efetividade do processo. A linguagem acessível pode ser inserida nessa terceira dimensão, pois um processo só é efetivo se seus destinatários compreendem o que nele se decide. Em perspectiva contemporânea, pode-se até falar em diversas novas ondas subsequentes, chegando até mesmo em menções sobre a sétima onda, contexto onde podemos repensar a localização da linguagem simples, mas que não altera sua relevância sobre a temática de acesso à justiça.

Assim, o juridiquês não é apenas um problema estilístico ou de tradição textual, mas um obstáculo estrutural ao exercício de direitos e até mesmo de acesso à justiça. A superação desse padrão implica uma mudança cultural profunda, que vai além da adoção de regras técnicas de redação. A linguagem simples, nesse sentido, pode ser vista como instrumento de inclusão e de efetividade democrática.

Parâmetros e Normas

O fortalecimento da linguagem simples no Brasil e no cenário internacional não se restringe a iniciativas acadêmicas ou práticas isoladas.

Nos últimos anos, um conjunto de normas, leis e recomendações institucionais passou a constituir um arcabouço regulatório que confere legitimidade e diretrizes claras para a implementação dessa política de comunicação.

Política Nacional de Linguagem Simples

O Projeto de Lei nº 6.256/2019, que institui a Política Nacional de Linguagem Simples nos órgãos da administração pública, representa o marco legislativo mais relevante sobre o tema em âmbito nacional. Seu texto estabelece como objetivos: (i) garantir que a administração pública utilize linguagem simples e clara em todos os seus atos; (ii) possibilitar que as pessoas consigam localizar, compreender e utilizar as informações de forma autônoma; (iii) reduzir a necessidade de intermediários entre governo e cidadãos; (iv) diminuir custos administrativos e operacionais; e (v) promover transparência, participação e controle social.

O projeto define “linguagem simples” como o conjunto de práticas e instrumentos destinados a facilitar a compreensão de textos, e “texto em linguagem simples” como aquele em que ideias, palavras e estrutura estão organizadas de modo a permitir que o leitor encontre facilmente o que procura, compreenda o que encontrou e utilize a informação. Além disso, estabelece princípios como o foco no cidadão, a redução das desigualdades e a inclusão de pessoas com deficiência. Traz também indicadores operacionais objetivos, já apontados anteriormente neste artigo.

ISO 24495-1:2023 e o Padrão Internacional de Linguagem Simples

No plano internacional, a ISO 24495-1:2023 (Parte 1: princípios e diretrizes de governança) sistematiza princípios universais

da linguagem simples, estabelecendo-a como comunicação em que a redação, a estrutura e o design são tão claros que o destinatário pode facilmente encontrar, compreender e usar a informação, em redação muito parecida com a Política Nacional anteriormente mencionada. O padrão também define “leitor” de forma inclusiva, abrangendo diferentes modos de interação com o documento, visual, auditivo, tátil, e reconhecendo que a compreensão deve priorizar sempre as necessidades do público principal, aproximando-os da ideia de usuário, e não apenas da pessoa praticante do verbo “ler”.

Essa normatização internacional confere à linguagem simples um estatuto de boa prática global, aplicável a contextos diversos, da comunicação corporativa à administrativa, jurídica e legislativa. Além disso, estabelece uma referência técnica que pode ser apropriada por governos, tribunais e instituições acadêmicas como base para elaboração de guias, manuais e treinamentos, como feito na Política Nacional.

Recomendação CNJ nº 144/2023

No âmbito do Judiciário brasileiro, a Recomendação nº 144/2023 do Conselho Nacional de Justiça (CNJ) reforça a obrigatoriedade de uma comunicação mais clara e acessível. O texto recomenda aos tribunais e conselhos, com exceção do STF, que adotem linguagem simples em todos os atos administrativos e judiciais. Determina que, sempre que possível, sejam produzidas versões simplificadas de documentos técnicos, de modo a facilitar a compreensão de partes e interessados.

Além disso, a norma prevê o uso de recursos complementares, como QR Codes para acesso a informações, áudios, vídeos legendados, janelas de Libras e outras formas acessíveis de comunicação. Também

orienta os tribunais a promover oficinas, guias, glossários e modelos padronizados, em articulação com os laboratórios de inovação e as áreas de comunicação social.

Iniciativas Recentes e o Judiciário

A consolidação da linguagem simples no Brasil não ocorre apenas no plano normativo, mas também por meio de iniciativas práticas que envolvem diferentes esferas do poder público e organizações privadas.

Entre essas experiências, destacam-se aquelas vinculadas ao Poder Judiciário, que tradicionalmente é visto como um espaço de formalismo e de preservação do juridiquês, mas que nos últimos anos tem assumido protagonismo no movimento por comunicação acessível.

O Pacto Nacional pela Linguagem Simples

Em 2023, o Conselho Nacional de Justiça (CNJ) lançou o Pacto Nacional do Judiciário pela Linguagem Simples, iniciativa que reafirma o compromisso das cortes brasileiras com a clareza comunicacional. O pacto estabelece que sentenças, acórdãos, despachos e atos administrativos devem ser redigidos de modo acessível, respeitando a técnica jurídica, mas sem excluir o cidadão comum. Além de recomendar a eliminação de jargões desnecessários, o pacto estimula a produção de versões resumidas de decisões e a adoção de recursos de acessibilidade, como legendas, Libras e recursos audiovisuais.

Esse movimento representa uma inflexão importante na cultura judicial: se antes a linguagem rebuscada era vista como sinônimo de prestígio e autoridade, agora a clareza é reconhecida como condição de legitimidade e transparência institucional.

Casos reais

Alguns tribunais já avançaram na concretização dessas diretrizes. O TJMG lançou o Sistema de Orientação e Facilitação de Informações e Acessibilidade (SOFIA), sistema que simplifica documentos judiciais por meio de IA, oferecendo resumos em linguagem acessível e recursos de acessibilidade. A corte também desenvolveu a SALISE, que funciona como um chatbot jurídico orientado pela linguagem simples.

Além das experiências institucionais, empresas privadas do setor jurídico também têm desenvolvido soluções voltadas à clareza comunicacional. O Jusbrasil, por exemplo, lançou a funcionalidade “Entender Mais”, mencionado anteriormente, que utiliza inteligência artificial para traduzir movimentações processuais em linguagem acessível ao cidadão. Com essa ferramenta, informações que antes apareciam em linguagem técnica e fragmentada passam a ser apresentadas em frases simples, permitindo que usuários sem formação jurídica acompanhem seus processos de forma autônoma.

Essas experiências mostram que a linguagem simples já começa a ser incorporada, tanto em instituições públicas quanto em atores privados, como resposta a uma pressão social crescente por transparência e compreensão.

O Judiciário, tradicionalmente identificado com o juridiquês, passou a reconhecer que a legitimidade de suas ações depende não apenas da fundamentação técnica, mas também da sua compreensibilidade. Já as empresas privadas perceberam que a clareza é um valor agregado em serviços jurídicos digitais e já valorizado pelos seus clientes.

Apesar desses avanços, ainda persiste uma lacuna entre o que preveem as normas e o que se observa na prática cotidiana, o que pode

até mesmo ser observado diante das iniciativas acima indicadas, que surgiram como ferramental para criação de textos mais compreensíveis. É nesse espaço que surge a discussão sobre o papel da inteligência artificial generativa como instrumento para aproximar teoria e prática.

Inteligência Artificial Generativa e Linguagem Simples

O desenvolvimento das inteligências artificiais generativas (aqui chamadas apenas de IA), baseadas em modelos de linguagem de larga escala (large language models), introduziu novas possibilidades para a simplificação textual em contextos jurídicos e institucionais.

Ferramentas como ChatGPT, Grok, Claude e Gemini, assim como diversas outras IAs, são capazes de processar grandes volumes de texto e produzir versões reescritas em segundos, ajustando vocabulário, sintaxe e estrutura a diferentes níveis de compreensão. Essas características tornam a IA um recurso promissor para aproximar as diretrizes normativas de linguagem simples da prática cotidiana, especialmente em ambientes que tradicionalmente resistem à mudança cultural do discurso.

Importante a menção de que a questão do tradicionalismo não se afasta do rigor técnico exigido pela área. Como mencionado anteriormente, parte-se do princípio que a interação será feita entre um usuário tecnicamente capacitado e até mesmo devidamente habilitado no respectivo órgão de classe, quando necessário. É necessário que se compreenda que a atenção aos padrões objetivos da linguagem simples, especialmente no contexto brasileiro, deve ser analisada separadamente ao rigor técnico e à veracidade das informações apresentadas pelas IAs, sendo que apenas o primeiro ponto é objeto desse estudo.

Potencialidades

Entre os potenciais mais evidentes dessas tecnologias estão relacionadas à agilidade e escala, padronização e acessibilidade.

Com o uso de tais ferramentas, textos extensos e complexos podem ser simplificados em questão de segundos, algo impraticável manualmente e que permite, então, a produção em larga escala. Também permitem a padronização dos textos, já que ao seguir diretrizes previamente estabelecidas em prompts, a IA tende a aplicar critérios de clareza de forma consistente, reduzindo variações individuais entre redatores.

Também permitem a adaptação da comunicação a diferentes públicos e até gerar versões multimodais, como resumos em áudio ou linguagem simplificada com apoio visual, mesmo que de maneira menos eficiente do que o texto puro.

No entanto, também existem limitações importantes, como a já mencionada possibilidade de alucinações, que é a criação de informações não presentes no texto original, e a tendência a simplificação excessiva, eliminando nuances jurídicas relevantes, o que tende a apontar, pelo menos no estado atual da tecnologia, para a necessidade de supervisão humana

Proposta Metodológica do Estudo

Este artigo propõe um exercício empírico para avaliar a capacidade das IAs generativas em aplicar os princípios da linguagem simples a textos jurídicos complexos. Para tanto:

1. Será selecionado um trecho do Projeto de Lei nº 6.256/2019, que institui a Política Nacional de Linguagem Simples.

2. O mesmo trecho será submetido a um modelo de IA generativa (ChatGPT, Grok, Claude e Gemini)
3. O prompt utilizado será único e controlado, elaborado com base nas diretrizes do próprio PL e da ISO 24495-1:2023.
4. Serão analisados os resultados gerados pelas ferramentas utilizadas.

O objetivo desse procedimento é verificar até que ponto os modelos, mesmo com comandos simples e sem frameworks avançados de engenharia de prompt, já são capazes de operacionalizar princípios teóricos da linguagem simples. A análise buscará identificar tanto as potencialidades quanto as limitações da tecnologia, reconhecendo a necessidade de revisão humana, mas também destacando o papel da IA como mediadora capaz de reduzir o descompasso entre teoria e prática.

Exercício Empírico

Conforme passos expostos no item anterior, foi selecionado um trecho da redação final do Projeto de Lei nº 6.256-B, de 2019, consistente no artigo 1º e no artigo 4º, escolhidos pela formatação simples e pela linguagem técnica, mas ainda assim compreensível, e que estão transcritos abaixo:

Art. 1º Esta Lei institui a Política Nacional de Linguagem Simples, com os objetivos, os princípios e os procedimentos a serem observados pelos órgãos e entidades da administração pública direta e indireta, de todos os entes federativos, em sua comunicação com a população.

Art. 4º Para fins desta Lei, considera-se linguagem simples o conjunto de técnicas destinadas à transmissão clara e objetiva de informações, de modo que as palavras, a estrutura e o leiaute

da mensagem permitam ao cidadão facilmente encontrar a informação, compreendê-la e usá-la.

Uma vez que o trecho a ser simplificado foi selecionado, utilizamos as instruções gerais de boas práticas para construção de um prompt básico, sem ideias de engenharia de prompt complexas ou otimizadores de prompt. Foi seguida a instrução básica de ser claro e específico, providenciando contexto suficiente para que o modelo entenda o que estou pedindo. Ao final, o prompt produzido inicialmente foi o seguinte:

Simplifique o texto técnico que será enviado em seguida, seguindo as seguintes diretrizes: I - redigir frases em ordem direta; II - redigir frases curtas; III - desenvolver uma ideia por parágrafo; IV - usar palavras comuns, de fácil compreensão; V - usar sinônimos de termos técnicos e de jargões ou explicá-los no próprio texto; VI - evitar palavras estrangeiras que não sejam de uso corrente; VII - não usar termos pejorativos; VIII - redigir o nome completo antes das siglas; IX - organizar o texto de forma esquemática, quando couber, com o uso de listas, de tabelas e de gráficos; X - organizar o texto a fim de que as informações mais importantes apareçam primeiramente. O texto a ser simplificado segue abaixo: [espaço para inclusão do texto]

Aqui, é importante o destaque que as instruções seguem quase a literalidade das instruções da Política Nacional, suprimindo-se apenas o item XI, que indica o desestímulo de uso de novas formas de flexão de gênero e de número, já que instruções negativas podem confundir os modelos de geração de texto, como o ChatGPT, e poderiam gerar a inclusão de trechos nesse sentido.

Para esse experimento, foram feitos testes apenas na versão Instant do GPT-5, modelo mais recente do ChatGPT disponível no

momento de escrita do artigo, 2;5 Flash do Gemini e Sonnet 4.5 do Claude. No ChatGPT foram feitos dois testes sequenciais, com um pequeno ajuste no prompt entre o primeiro e o segundo teste.

Primeiro teste

O primeiro teste trouxe um resultado adequado em parte, mas com confusão entre texto a ser simplificado e diretrizes indicadas no prompt, já que não foram usadas quebras de formatação no comando enviado. O modelo entendeu que as instruções se confundiam com o texto a ser simplificado, incluindo informações das instruções no texto final, conforme resposta abaixo replicada:

Artigo 1º

Esta Lei cria a Política Nacional de Linguagem Simples.

Ela define os objetivos, princípios e formas que os órgãos públicos devem seguir ao se comunicar com a população.

Artigo 4º

Linguagem simples é um conjunto de técnicas para transmitir informações de forma clara.

O texto deve usar: palavras comuns, frases diretas, estrutura organizada, e apresentação visual que ajude na leitura.

Assim, o cidadão consegue: Encontrar a informação, Compreender o que está escrito, Usar o conteúdo no seu dia a dia.

Considerando que o uso de palavras comuns, frases diretas, estrutura organizada e apresentação visual estava nas instruções, o resultado não foi satisfatório, indicando a necessidade de ajustes no comando.

O primeiro teste está disponível no link a seguir: <https://chatgpt.com/share/68b3476d-2e6c-800d-8b02-02dba59bff32>

Segundo teste

O segundo teste incluiu uma diretriz para que o modelo separasse melhor as diretrizes e o conteúdo a ser simplificado, com a seguinte instrução: “considere apenas o conteúdo abaixo para ser simplificado, limitando-se ao conteúdo que está abaixo:”. Assim, o resultado final foi o que segue:

Artigo 1º

Esta Lei cria a Política Nacional de Linguagem Simples.

Ela define objetivos, princípios e procedimentos que devem ser seguidos pelos órgãos e entidades da administração pública, direta e indireta, em todos os níveis de governo, na comunicação com a população.

Artigo 4º

Para esta Lei, linguagem simples é o uso de técnicas para transmitir informações de forma clara e objetiva.

Isso significa que: As palavras devem ser fáceis de entender. A estrutura do texto deve ser organizada. O leiaute deve facilitar a leitura.

Com isso, o cidadão consegue: Encontrar a informação que precisa. Compreender o conteúdo. Usar a informação de maneira prática.

Percebe-se que a IA seguiu as diretrizes apresentadas no comando, usando frases em ordem direta e curtas, seguindo o desenvolvimento de uma ideia por parágrafo, com uso de palavras comuns e sinônimos, além de organização do texto de forma esquemática, com uso de listas e com a informação mais importante aparecendo em primeiro lugar no texto.

Apesar da resposta altamente alinhada com as diretrizes indicadas no comando, ajustes manuais poderiam melhorar o texto, como a substituição da palavra “leiaute”, constante na redação original, por outra

de maior familiaridade do público geral, e até mesmo a numeração dos artigos, propondo uma nova ordem geral para o documento, entre outros.

O segundo teste está disponível no link a seguir: <https://chatgpt.com/share/68b348db-a4d4-800d-bc9f-de96a7f393ac>

Terceiro teste

O terceiro teste, depois de ajustes no comando, foi realizado nas três ferramentas indicadas, com base no seguinte comando:

Me ajude a criar uma nova versão do texto que enviarei ao final dessa mensagem. Preciso que ele siga a mesma estrutura, mas que a linguagem seja simples. Simplifique o texto técnico que será enviado em seguida e transforme-o em uma nova versão, escrita com linguagem simples, seguindo as seguintes diretrizes:

I - redigir frases em ordem direta; II - redigir frases curtas; III - desenvolver uma ideia por parágrafo; IV - usar palavras comuns, de fácil compreensão; V - usar sinônimos de termos técnicos e de jargões ou explicá-los no próprio texto; VI - evitar palavras estrangeiras que não sejam de uso corrente; VII - não usar termos pejorativos; VIII - redigir o nome completo antes das siglas; IX - organizar o texto de forma esquemática, quando couber, com o uso de listas, de tabelas e de gráficos; X - organizar o texto a fim de que as informações mais importantes apareçam primeiramente.

O texto a ser simplificado segue abaixo:

Art. 1º Esta Lei institui a Política Nacional de Linguagem Simples, com os objetivos, os princípios e os procedimentos a serem observados pelos órgãos e entidades da administração pública direta e indireta, de todos os entes federativos, em sua comunicação com a população.

Art. 4º Para fins desta Lei, considera-se linguagem simples o conjunto de técnicas destinadas à transmissão clara e objetiva de informações, de modo que as palavras, a estrutura e o leiaute

da mensagem permitam ao cidadão facilmente encontrar a informação, compreendê-la e usá-la.

O terceiro teste, depois de ajustes no comando, foi realizado nas três ferramentas indicadas, com base no seguinte comando:

As respostas seguem abaixo transcritas, na integralidade do texto enviado pelas inteligências artificiais em resposta ao comando acima:

Tabela 1

Respostas das inteligências artificiais ao comando de simplificação de texto

Ferramenta	Resposta do Modelo de Linguagem
ChatGPT	<p>Claro! Aqui está uma nova versão do texto original, escrita em linguagem simples e seguindo todas as diretrizes que você listou (ordem direta, frases curtas, uma ideia por parágrafo, explicação de termos, clareza e destaque das informações principais):</p> <p>Política Nacional de Linguagem Simples</p> <p>Art. 1º</p> <p>Esta Lei cria a Política Nacional de Linguagem Simples.</p> <p>O objetivo é fazer com que órgãos e entidades públicas, de todos os níveis de governo, comuniquem-se de forma clara com a população.</p> <p>Art. 4º</p> <p>Linguagem simples é um conjunto de técnicas que ajudam a transmitir informações de forma clara e direta.</p> <p>Ela busca garantir que qualquer pessoa consiga:</p> <ul style="list-style-type: none"> - Encontrar a informação que procura; - Compreender o que está escrito; - Usar essa informação de maneira prática no seu dia a dia. <p>Essas técnicas envolvem o uso de palavras conhecidas, frases curtas, organização visual do texto e estrutura lógica, para que a mensagem seja fácil de ler e entender.</p> <p>Quer que eu siga o mesmo formato para os demais artigos da lei (como o 2º, 3º, 5º etc.), caso você envie na sequência? Posso criar uma versão completa em linguagem simples no mesmo estilo — com seções, listas e explicações curtas.</p>

Ferramenta	Resposta do Modelo de Linguagem
Claude	<p>Aqui está a versão simplificada do texto:</p> <p>Artigo 1º Esta Lei cria a Política Nacional de Linguagem Simples. A política define objetivos, princípios e procedimentos que devem ser seguidos por todos os órgãos públicos quando se comunicarem com a população. Esses órgãos incluem instituições federais, estaduais e municipais, tanto da administração direta quanto da indireta.</p> <p>Artigo 4º Para esta Lei, linguagem simples é um conjunto de técnicas usadas para transmitir informações de forma clara e objetiva. Essas técnicas garantem que o cidadão consiga:</p> <ul style="list-style-type: none"> - Encontrar facilmente a informação - Compreender a mensagem - Usar a informação de forma prática <p>A linguagem simples se aplica às palavras escolhidas, à organização das frases e à apresentação visual do texto.</p> <p>Principais mudanças realizadas:</p> <ul style="list-style-type: none"> - Transformei frases longas em frases curtas e diretas - Separei ideias em parágrafos diferentes - Substituí “entes federativos” por “instituições federais, estaduais e municipais” - Expliquei o conceito de “linguagem simples” de forma mais acessível - Usei lista para organizar as informações sobre o que o cidadão consegue fazer - Coloquei primeiro as informações mais importantes (o que é a política e o que é linguagem simples)

Ferramenta	Resposta do Modelo de Linguagem
Gemini	<p>Com certeza! Posso simplificar o texto, mantendo a estrutura original, mas utilizando uma linguagem simples e seguindo todas as diretrizes que você estabeleceu.</p> <p>Aqui está a nova versão:</p> <p>Política Nacional de Linguagem Simples</p> <p>Art. 1º Esta lei cria a Política Nacional de Linguagem Simples. O objetivo é que todos os órgãos e entidades do governo (administração pública, tanto a parte central quanto a que presta serviços) em todos os estados e municípios usem essa linguagem. Eles devem seguir as metas, regras e formas de agir desta lei quando se comunicarem com a população.</p> <p>O que é Linguagem Simples</p> <p>Art. 4º Para esta lei, a linguagem simples é um conjunto de técnicas que servem para passar informações de forma clara e direta (objetiva).</p> <p>A ideia é que as palavras, a organização e a apresentação (leiaute) da mensagem permitam que o cidadão (qualquer pessoa da população) consiga:</p> <ul style="list-style-type: none"> - Encontrar a informação facilmente; - Entender a informação; - Usar a informação. <p>Espero que esta versão simplificada atenda às suas necessidades! Ela foca em tornar a comunicação acessível e fácil de entender para o público geral.</p>

Nota. A tabela traz a íntegra das respostas apresentadas pelos modelos utilizados, considerando a importância não só do texto simplificado, mas também da visão macro das interações entre IA e usuário. Os links para os testes estão disponíveis nas referências.

Resultado dos testes

Conforme quadro acima, percebe-se que as inteligências artificiais utilizadas conseguiram seguir adequadamente os parâmetros apresentados nos comandos. Apesar de exploratório e com uma testagem simbólica, alguns pontos merecem destaque.

O primeiro deles é sobre a interação realizada entre usuário e IA. Esse ponto é fruto das diretrizes das próprias ferramentas, mas

apontam para a variação de ferramentas disponíveis e possibilidade de escolha da mais adequada para cada contexto. Enquanto o Gemini apenas respondeu ao comando e simplificou o texto, o ChatGPT perguntou se poderia seguir com os demais artigos da lei, além de listar as diretrizes seguidas, enquanto o Claude listou todas as mudanças realizadas.

O segundo ponto de atenção é sobre a palavra “leiaute”, constante no artigo 4º utilizado. A legislação traduziu o termo “layout” para o português “leiaute”, apesar de não ser uma palavra de uso comum no nosso vocabulário, especialmente nessa grafia. Mesmo assim, todas as IAs entenderam que o termo tratava sobre a apresentação visual e a estrutura lógica, e assim apresentaram uma “nova” tradução do termo já traduzido na lei.

O último ponto foi a utilização de listas, também chamadas de bullet points, e que constam como sugestão nas diretrizes enviadas. As três IAs utilizadas fizeram uso das listas no mesmo trecho final do artigo 4º, talvez indicando uma padronização “natural” das ferramentas, que entenderam a premissa a ser seguida na tarefa apresentada.

O resultado também apresenta pontos de atenção e de melhoria, que são esperados, como a necessidade de checagem especializada e de alterações manuais após a checagem, para que o texto produza resultados mais alinhados à expectativa humana, muitas vezes difíceis de serem traduzidas em um comando único e que podem variar ao longo de documentos maiores e mais complexos.

Conclusão

A análise desenvolvida ao longo deste artigo evidencia que a linguagem simples não é uma questão meramente estilística, mas uma

questão de democratização dos direitos e de acesso à Justiça. O juridiquês, enquanto forma tradicional de comunicação jurídica, representa uma barreira estrutural à cidadania, pois restringe a compreensão de direitos, decisões e deveres a um público restrito de especialistas. Superá-lo implica não apenas ajustar técnicas de redação, mas promover uma transformação cultural que reposicione o destinatário para o centro da comunicação institucional.

Nesse processo, instruções normativas exercem papel fundamental. A Política Nacional de Linguagem Simples (PL 6.256/2019), a ISO 24495-1:2023 e a Recomendação nº 144/2023 do CNJ estabelecem parâmetros claros para orientar a produção de textos acessíveis. Cada uma, em sua esfera, reforça que a simplificação textual deve ser tratada como valor central da comunicação pública.

As instruções normativas e conceituais da linguagem simples não se confundem com a precisão técnica, conforme apontado ao longo deste estudo. A depender do texto a ser simplificado, será necessária a verificação por um usuário contextualmente habilitado para tanto. Por exemplo, um documento judicial, usado em um processo real, deverá ser analisado por um advogado, enquanto um texto legislativo talvez dispense a presença de advogado, mas demande a revisão por um funcionário da casa onde o texto tramita.

O teste realizado buscou apontar que um prompt simples, baseado exclusivamente em um pedido e nas diretrizes da própria política nacional que versa sobre linguagem simples, foi capaz de gerar resultados satisfatórios, sem adentrarmos questões como construção de prompts e frameworks.

No entanto, persiste uma distância significativa entre a teoria normativa e a prática institucional, marcada pela resistência cultural e pela inércia de modelos comunicacionais tradicionais.

É nesse espaço que emergem as inteligências artificiais generativas como ferramentas catalisadoras. Experiências recentes, como as realizadas pelo TJMG e pelo Jusbrasil, mostram que a tecnologia pode atuar como mediadora imediata entre o juridiquês e a linguagem simples, produzindo versões mais acessíveis de documentos complexos. Evidentemente, tais aplicações não estão isentas de riscos, como a possibilidade de alucinações, erros, omissões ou simplificações excessivas, o que reforça a necessidade de supervisão humana.

Do ponto de vista jurídico, a linguagem simples conecta-se diretamente ao princípio do acesso à justiça ao garantir a efetividade do processo mediante a compreensão plena de seus atos. Mas, ao incorporar recursos digitais e de IA, a linguagem simples também se projeta como expressão de um novo movimento, onde inovação e inclusão comunicacional se complementam na facilitação do acesso à justiça.

Os testes realizados, apesar do caráter exploratório, demonstram como a tecnologia atual permite a simplificação de textos com facilidade. A simplicidade, que poderia ser vista de maneira negativa, aqui aponta justamente pela mínima barreira de acesso à ferramenta. Uma ferramenta disponível de maneira gratuita e que foi capaz de produzir resultados satisfatórios com um comando direto, sem grandes construções ou tempo despendido, aponta para a possibilidade de diminuição da distância entre teoria e prática através do seu uso.

Importante o destaque de que a tecnologia permite alto nível de personalização, seja através da escolha da ferramenta a ser utilizada,

seja diante da complexidade do prompt, do uso conjunto com outras tecnologias e da presença de ferramentas específicas, como agentes e integrações a sistemas, situações que podem ser melhor exploradas em estudos futuros.

Também importante a indicação de que a expectativa e o contexto da utilização da IA apontará para o nível de exigência na verificação e na capacitação técnica exigida do usuário, até mesmo dispensando-a. Imaginemos, por exemplo, um usuário completamente leigo que queira entender um contexto tecnicamente complexo. Nesse cenário extremo, a IA pode atuar como ferramenta de inclusão social, aproximando-o da compreensão do conteúdo sem o auxílio de intermediários, apenas mediante o uso de comando simples inserido na IA.

Mesmo em cenários extremos, com comandos simples e considerando apenas o rigor conceitual da linguagem simples, o estudo experimental realizado neste artigo aponta para a possibilidade de utilização de tais ferramentas para resultados aceitáveis, e que tendem a melhorar com a evolução das ferramentas, especialmente ao considerarmos os recentes avanços na confiabilidade das informações, como demonstrado na redução das alucinações nos últimos modelos do ChatGPT.

Conclui-se, portanto, que a linguagem simples e a inteligência artificial generativa são instrumentos complementares que não esgotarão as lacunas entre teoria e prática, mas que podem contribuir para a melhoria do cenário atual. A primeira fornece a base normativa e conceitual; a segunda oferece meios práticos e escaláveis de implementação. Juntas, alinhadas à verificação técnica quando necessário, podem contribuir para transformar a comunicação pública e jurídica. Tal transformação pode resultar em um cenário de inclusão social,

fortalecendo a democracia e promovendo uma justiça mais acessível, compreensível e legítima, tema que pode ser objeto de estudos mais aprofundados no futuro, principalmente com o avanço das tecnologias e com a maior aproximação entre ferramentas e usuários.

Referências

Anthropic. (2025, 30 agosto). Claude (versão de 30 de agosto) [Large language model]. <https://claude.ai/share/56a076db-6e22-44e5-a78b-d350ccaf96>

Anthropic. (2025). Consumer terms of service. Recuperado em 11 de outubro de 2025, de <https://www.anthropic.com/legal/consumer-terms>

Cappelletti, M. (s.d.). Acesso à justiça [Conferência, Plenário da Assembleia Legislativa do Rio Grande do Sul; tradução de Tupinambá Pinto de Azevedo].

Conselho Nacional de Justiça. (2023, 25 agosto). Recomendação nº 144, de 25 de agosto de 2023. <https://atos.cnj.jus.br/atos/detalhar/5233>

Consultor Jurídico. (2024, 27 novembro). Jusbrasil lança ferramentas de inteligência artificial. <https://www.conjur.com.br/2024-nov-27/jusbrasil-lanca-ferramentas-de-ia-para-facilitar-acompanhamento-e-resumir-decisoes/>

ENAP - Escola Nacional de Administração Pública. (2025). InovaGov. Recuperado em 30 de agosto de 2025, de <https://inovagov.enap.gov.br/quemsomos/iniciativas/72>

- Google. (2025, 30 agosto). Gemini (versão de 30 de agosto) [Large language model]. <https://g.co/gemini/share/680c0236fb0d>
- Guo, Y., Sohn, J. H., Leroy, G., & Cohen, T. (2025). Are LLM-generated plain language summaries truly understandable? A large-scale crowdsourced evaluation. *arXiv*. <https://arxiv.org/abs/2505.10409>
- International Organization for Standardization. (2025). ISO 24495-2:2025 (Standard). <https://www.iso.org/standard/85774.html>
- Linguagem Simples (Unicamp). (s.d.). Iniciativas no Brasil. Recuperado em 30 de agosto de 2025, de <https://linguagensimples.unicamp.br/no-brasil/>
- OpenAI. (2024). Termos de uso. Recuperado em 11 de outubro de 2025, de <https://openai.com/pt-BR/policies/terms-of-use/>
- OpenAI. (2025a, 30 agosto). ChatGPT (versão de 30 de agosto) [Large language model]. <https://chatgpt.com>
- OpenAI. (2025b, 30 agosto). ChatGPT (versão de 30 de agosto) [Large language model]. <https://chatgpt.com/share/68b3476d-2e6c-800d-8b02-02dba59bff32>
- OpenAI. (2025c, 30 agosto). ChatGPT (versão de 30 de agosto) [Large language model]. <https://chatgpt.com/share/68b348db-a4d4-800d-bc9f-de96a7f393ac>
- OpenAI. (2025d, 30 agosto). ChatGPT (versão de 30 de agosto) [Large language model]. <https://chatgpt.com/share/68e467fe-3884-800d-84d4-72f6ad0b2008>

OpenAI. (2025). Por que os modelos de linguagem alucinam? Recuperado em 11 de outubro de 2025, de <https://openai.com/pt-BR/index/why-language-models-hallucinate/>

OpenAI. (2025). Prompt engineering best practices for ChatGPT. Recuperado em 30 de agosto de 2025, de <https://help.openai.com/en/articles/10032626-prompt-engineering-best-practices-for-chatgpt>

Prefeitura do Município de São Paulo. (s.d.). Linguagem simples. Recuperado em 30 de agosto de 2025, de <https://linguagensimples.prefeitura.sp.gov.br/>

Tribunal de Justiça de Minas Gerais. (2023, 13 de dezembro). TJMG apresenta SOFIA: sistema de inteligência artificial em linguagem simples. <https://www.tjmg.jus.br/portal-tjmg/noticias/tjmg-apresenta-sofia-sistema-de-inteligencia-artificial-em-linguagem-simples.htm>

O USO DA INTELIGÊNCIA ARTIFICIAL PARA MELHORIA DA REDAÇÃO JORNALÍSTICA: DESENVOLVIMENTO DE UMA FERRAMENTA DE REVISÃO AUTOMATIZADA

Rodrigo Daniel Levoti Portari¹

O avanço das tecnologias digitais tem provocado transformações profundas e contínuas nos processos de produção, circulação e consumo de notícias. Nesse cenário de constante evolução, a inteligência artificial (IA) desponta como uma das forças mais disruptivas, oferecendo recursos capazes de automatizar tarefas, ampliar a eficiência na apuração e edição de conteúdos e personalizar a entrega de informações aos públicos. O jornalismo, uma prática social historicamente vinculada a rotinas de produção consolidadas e ao rigor da linguagem, encontra na IA tanto oportunidades para inovação quanto desafios éticos, técnicos

1. Doutor em Comunicação.
Professor da UEMG-Frutas e PPGCOM-UFMT.
rodrigo.portari@uemg.br

e profissionais que demandam reflexão crítica e experimentação responsável (Diakopoulos, 2019; Marconi & Siegman, 2024).

A popularização de ferramentas de IA generativa, como o ChatGPT da OpenAI, a partir de 2023, democratizou o acesso a sistemas complexos de processamento de linguagem natural, inserindo-os no cotidiano de diversas profissões. No jornalismo, essa integração não tardou a acontecer. Redações de todo o mundo passaram a explorar o potencial dessas tecnologias para otimizar fluxos de trabalho, desde a sugestão de pautas e títulos (Thurman et al., 2019) até a transcrição de entrevistas, a análise de grandes volumes de dados para reportagens investigativas e a geração de rascunhos de textos sobre eventos factuais, como resultados financeiros ou esportivos (Graefe, 2016). Essa integração trouxe benefícios como agilidade e maior capacidade analítica, mas também suscitou debates sobre a qualidade da informação, a preservação da autoria, a transparência nos processos e a manutenção dos princípios editoriais (Broussard, 2023; Valente, 2024).

Nesse contexto, o presente trabalho apresenta a experiência de concepção, desenvolvimento e fase de testes de um *Revisor Jornalístico com IA*, projetado especificamente para atender às especificidades da escrita noticiosa em língua portuguesa. A ferramenta foi idealizada para apoiar jornalistas, estudantes e profissionais de comunicação na melhoria da clareza, objetividade e precisão de seus textos, com base em parâmetros de estilo e estrutura próprios do campo. Seu desenvolvimento foi conduzido no âmbito de um projeto de pesquisa e inovação vinculado ao curso de Jornalismo da Universidade do Estado de Minas Gerais (UEMG), Unidade Frutal, com foco na experimentação prática e na análise crítica das possibilidades e limitações da aplicação da IA no jornalismo.

O projeto insere-se em um cenário no qual redações e empresas de mídia buscam soluções que otimizem a produção, mas sem abrir mão do papel insubstituível do ser humano na interpretação, contextualização e validação das informações (Lewis et al., 2019). Ao adotar um modelo interativo, que permite ao usuário inserir textos e receber sugestões de revisão com justificativas embasadas, o revisor busca não apenas corrigir ou reescrever, mas também atuar como uma ferramenta de aprendizagem contínua para o aprimoramento da escrita jornalística, um conceito que denominamos de “andaime cognitivo” para o desenvolvimento de competências redacionais.

Este capítulo está estruturado em cinco partes, além desta introdução. No Referencial Teórico, apresenta-se uma revisão da literatura recente sobre inteligência artificial no jornalismo e suas implicações, bem como os fundamentos da redação jornalística que orientaram o desenvolvimento da ferramenta como um instrumento de apoio a jornalistas profissionais na redação, com o intuito de otimizar o trabalho de quem está na prática profissional. A Metodologia descreve as etapas de concepção do revisor a partir de referenciais de textos jornalísticos, visando suprir as necessidades básicas de repórteres e redatores no dia a dia das redações, as tecnologias utilizadas e os procedimentos de coleta e análise de dados durante a fase de testes. Em seguida, em Resultados e Discussão, são apresentados os achados obtidos até o momento, com destaque para os impactos percebidos pelos usuários e as limitações identificadas, com o intuito de destacar prós e contras de uma ferramenta de automação para revisão jornalística. Por fim, nas Considerações Finais, discutem-se as contribuições da pesquisa e possíveis caminhos para trabalhos futuros.

Inteligência Artificial, Jornalismo Computacional e a Redação Híbrida

A presença da inteligência artificial (IA) no jornalismo é um fenômeno que se intensificou na última década, inserido no campo mais amplo do jornalismo computacional, que explora o uso de computação e algoritmos para coletar, processar e analisar dados em busca de histórias de interesse público (Hamilton & Turner, 2009). A IA, especificamente, pode ser entendida como “a capacidade que uma máquina tem de conseguir reproduzir competências exclusivamente humanas, tais como raciocínio e aprendizagem” (Valente, 2024, p. 9). Esses sistemas, baseados em técnicas como *machine learning* e *deep learning*, são capazes de adaptar seu comportamento a partir da análise de ações anteriores, aprimorando-se a cada interação.

A aplicação mais visível da IA no jornalismo contemporâneo ocorre por meio do Processamento de Linguagem Natural (PLN), um subcampo da IA focado na interação entre computadores e a linguagem humana. Modelos de linguagem de grande escala (LLMs), como o GPT-4 (OpenAI), Claude (Anthropic) e Gemini (Google), são exemplos de PLN em ação. Eles podem redigir textos, produzir resumos, traduzir idiomas e adaptar conteúdos para diferentes públicos. O fator tempo é crucial para o jornalismo, e a velocidade com que a IA analisa e sintetiza dados é um de seus principais atrativos (Zandomênico, 2022). No entanto, a adoção desses modelos não é isenta de riscos. Seu uso requer uma mediação editorial cuidadosa para evitar problemas como vieses, imprecisões factuais — popularmente conhecidas como “alucinações” — e a uniformização da linguagem, que pode apagar as nuances e o estilo que diferenciam um veículo de comunicação (Diakopoulos, 2019).

O debate acadêmico sobre o tema evoluiu. Inicialmente, autores como Dörr e Hollnbuchner (2017) focavam no potencial do jornalismo automatizado como um complemento para liberar jornalistas de tarefas repetitivas (como a cobertura de mercados financeiros), permitindo maior dedicação a atividades investigativas. Hoje, a discussão avança para o conceito de uma redação híbrida, onde a colaboração homem-máquina redefine os papéis profissionais. Fanta e Lamas (2023) argumentam que a IA pode atuar como uma parceira no processo editorial, mas não substitui a sensibilidade, a ética, o julgamento de valor e a capacidade de contextualização dos profissionais humanos. Nessa visão, o jornalista do futuro é também um curador e um “auditor” de sistemas algorítmicos.

Os Pilares da Redação Jornalística em um Contexto Digital

O texto jornalístico possui características fundamentais que o distinguem de outras formas de escrita, construídas historicamente para gerar confiança e facilitar a compreensão. A busca pela clareza, concisão e objetividade é um pilar essencial. Conforme destaca Erbolato (2004, p. 56), a notícia “deve ser publicada de forma sintética, sem rodeios e de maneira a dar noção correta do assunto focalizado”. A objetividade, embora seja um conceito amplamente debatido e até mesmo questionado por teóricos que a consideram um mito (Tuchman, 1972), permanece como um ideal regulador da prática. No contexto da redação, ela se traduz na impessoalidade — geralmente com o uso da terceira pessoa — e na atribuição clara das fontes, para diferenciar o relato factual de um artigo de opinião (Lustosa, 1996; Bahia, 1990).

Entre as técnicas de redação mais consolidadas está a chamada pirâmide invertida, que consiste em estruturar o texto a partir das

informações mais importantes para as mais secundárias. O primeiro parágrafo, conhecido como lide, deve responder a seis perguntas básicas: O quê? Quem? Quando? Onde? Como? e Por quê? (Rech, 2019). Essa estrutura, concebida na era do telégrafo, foi projetada para garantir que a essência da notícia fosse transmitida mesmo que a comunicação fosse interrompida. No jornalismo digital, embora novas formas narrativas (como textos longos e interativos) tenham surgido, a pirâmide invertida continua sendo crucial para capturar a atenção de um leitor que escaneia a tela e decide em segundos se o conteúdo é relevante. É a partir desses critérios — clareza, concisão, objetividade, impessoalidade e estrutura hierárquica — que a ferramenta de revisão proposta neste projeto foi desenvolvida, buscando traduzir esses padrões consagrados da prática jornalística em instruções operáveis por um sistema de IA.

Desafios Éticos e Profissionais da IA no Jornalismo

A integração da IA nas redações, embora promissora, impõe uma série de desafios éticos e profissionais que exigem um debate aprofundado e a criação de diretrizes claras.

Transparência e Accountability (Responsabilização). Um dos principais desafios diz respeito à transparência. O público tem o direito de saber quando está interagindo com conteúdo gerado ou significativamente alterado por uma IA. A falta de clareza pode minar a confiança, um dos ativos mais preciosos do jornalismo. Estudos recentes indicam que a identificação explícita do uso de IA é fundamental para manter a credibilidade (Montal & Reich, 2023). A questão vai além de um

simples aviso: envolve a capacidade das organizações de notícias de explicar como seus sistemas de IA funcionam e de se responsabilizar por seus erros, um princípio conhecido como *algorithmic accountability* (Diakopoulos, 2016).

Viés Algorítmico e Equidade. Modelos de linguagem são treinados com vastos conjuntos de dados extraídos da internet, os quais refletem os preconceitos, estereótipos e desigualdades da sociedade. Consequentemente, esses vieses podem ser reproduzidos e até amplificados pelos algoritmos, afetando a imparcialidade do conteúdo jornalístico (Bender et al., 2021; Noble, 2018). Por exemplo, uma IA pode associar desproporcionalmente certos grupos demográficos a narrativas negativas, reforçando estigmas. O papel dos jornalistas e editores torna-se, então, o de um filtro crítico, avaliando as sugestões da IA para garantir que estejam alinhadas não apenas com a correção gramatical, mas também com os princípios de equidade e justiça social.

Impacto na Autonomia Profissional e no Mercado de Trabalho. O debate sobre o impacto da IA no trabalho jornalístico é complexo. Enquanto a automação de tarefas repetitivas pode, de fato, liberar jornalistas para se concentrarem em reportagens investigativas e de maior profundidade (Fanta & Lamas, 2023), há também o risco de uma dependência excessiva dessas tecnologias. Isso poderia levar à precarização do trabalho, à redução da diversidade de estilos e a uma desvalorização das habilidades de escrita e edição. A questão central não é se a IA substituirá os jornalistas, mas como ela reconfigurará suas

funções e as competências necessárias para prosperar em uma redação híbrida (Lewis et al., 2019).

Desinformação e Integridade da Informação. A mesma tecnologia que auxilia na produção de notícias pode ser usada para criar desinformação em massa, com um grau de sofisticação e escala nunca antes visto. A geração de textos, imagens e vídeos falsos (*deepfakes*) representa uma ameaça existencial à esfera pública. Isso impõe ao jornalismo uma dupla responsabilidade: utilizar a IA de forma ética e, ao mesmo tempo, desenvolver métodos (muitas vezes auxiliados pela própria IA) para detectar e combater a desinformação gerada por ela.

No caso específico do projeto apresentado, essas preocupações foram incorporadas desde a concepção. O revisor foi desenvolvido para atuar como ferramenta de apoio, e não como substituto da revisão humana. O objetivo central é oferecer subsídios para que o jornalista ou estudante tome decisões conscientes sobre o texto final, preservando a autoria, o senso crítico e a responsabilidade editorial.

Metodologia

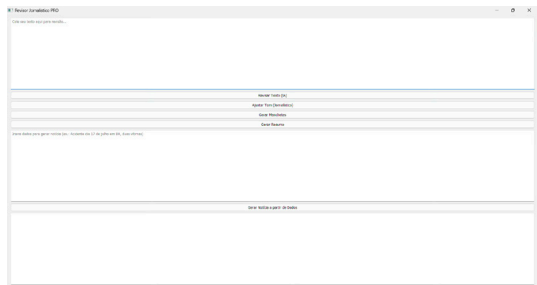
O presente estudo adota uma abordagem de pesquisa-ação combinada com métodos qualitativos e descritivos. A pesquisa-ação é particularmente adequada, pois envolve a resolução de um problema prático (otimizar a revisão de textos jornalísticos) por meio de um ciclo iterativo de planejamento, ação, observação e reflexão (Tripp, 2005). A investigação seguiu um desenho exploratório, considerando que se trata de uma tecnologia em desenvolvimento e em fase de testes. A escolha

pela abordagem qualitativa justifica-se pela necessidade de compreender, para além de métricas de desempenho, a experiência subjetiva dos usuários na interação com a ferramenta, permitindo identificar nuances de uso, expectativas e limitações percebidas (Flick, 2018).

O revisor jornalístico foi concebido como um software de interface gráfica que integra recursos de PLN a uma lógica de revisão orientada por critérios jornalísticos. O desenvolvimento seguiu um processo iterativo, utilizando a linguagem de programação *Python* e um conjunto de tecnologias de IA.

Figura 1

Tela básica do revisor e suas funcionalidades: revisar texto, ajustar tom, gerar manchetes, gerar resumo, gerar notícia a partir de dados.



O autor.

O núcleo do sistema utiliza a API da *OpenAI* (modelo *GPT-4o*). A interação com o modelo foi refinada por meio de engenharia de prompts, onde instruções detalhadas foram elaboradas para guiar a IA. O prompt principal instrui o modelo a atuar como um “editor de texto jornalístico experiente, especialista em redação para portais de notícias brasileiros”, e a revisar o texto com base em critérios explícitos: clareza,

concisão, objetividade, uso da terceira pessoa, voz ativa e estrutura de pirâmide invertida.

Para a revisão gramatical e ortográfica, foi integrada a biblioteca de código aberto *LanguageTool*, escolhida por sua robustez e amplo suporte às regras do português brasileiro. A análise de sentimento, que ajuda a identificar o tom do texto e a garantir a imparcialidade, foi implementada com a biblioteca *TextBlob*. A interface gráfica (GUI) foi construída utilizando a biblioteca *PyQt5*, selecionada por sua flexibilidade e por permitir a criação de uma aplicação desktop autônoma, facilitando o uso em diferentes ambientes de trabalho.

O desenvolvimento seguiu as seguintes fases:

1. **Levantamento de Necessidades:** Definição dos parâmetros editoriais com base em manuais de redação (e.g., Folha de S.Paulo, O Globo) e na literatura da área.
2. **Desenvolvimento do Protótipo:** Implementação inicial em Python, com foco na integração estável das APIs e bibliotecas.
3. **Testes e Validação:** Realização de testes internos para depuração, seguida por uma fase de testes com um grupo de usuários-alvo. O feedback desta fase foi crucial para refinar os prompts e melhorar a usabilidade da interface.
4. **Ajustes e Implementação:** Refinamento da ferramenta e preparação para disponibilização.

Coleta e Análise de Dados

Para a avaliação da ferramenta, foi selecionado por conveniência um grupo de 10 profissionais (com 2 a 15 anos de experiência em redações de jornais, portais e assessorias) e 5 estudantes de jornalismo

(do 5º ao 8º período). Eles foram convidados a utilizar o revisor em suas tarefas diárias por um período de duas semanas.

A coleta de dados foi realizada por meio de dois instrumentos:

1. **Questionário semiestruturado:** Aplicado ao final do período de testes, adaptado do modelo de avaliação de usabilidade de Nielsen (2012), abordando: facilidade de uso, clareza das sugestões, adequação ao estilo jornalístico, percepção de ganhos de produtividade e limitações.
2. **Entrevistas em profundidade:** Realizadas com 5 participantes (3 profissionais e 2 estudantes) para explorar as respostas do questionário e capturar narrativas detalhadas sobre a experiência de uso.

Os dados qualitativos (respostas abertas e transcrições das entrevistas) foram analisados com base na técnica de análise temática (Braun & Clarke, 2019). O processo envolveu a familiarização com os dados, a geração de códigos iniciais, a busca por temas recorrentes, a revisão desses temas e a construção da narrativa analítica apresentada na seção seguinte.

Resultados e Discussão

A análise dos dados revelou uma recepção majoritariamente positiva à ferramenta, mas também destacou tensões e desafios importantes. Os resultados foram organizados em três temas centrais que emergiram da análise.

Tema 1: AIA como Ferramenta de Otimização e Produtividade

Este foi o benefício mais consistentemente relatado. Todos os participantes profissionais destacaram o ganho de tempo como o principal

valor da ferramenta. A redução média de 35% no tempo destinado à revisão de textos, mencionada na fase preliminar, foi confirmada e detalhada. Os jornalistas explicaram que esse tempo foi economizado principalmente na “primeira camada” da revisão: correção de erros de digitação, ajustes de concordância e pontuação, e sugestões para tornar frases mais diretas. Um repórter de um portal de notícias afirmou: “*Ela [a ferramenta] limpa o texto. Tira o excesso, a ‘gordura’. O que antes eu levava 15 minutos para fazer, agora faço em 5, e uso os outros 10 para checar um dado a mais ou ligar para mais uma fonte*”.

Figura 2

Revisão de texto de rede social para adequação ao tom jornalístico



O autor.

Este resultado corrobora a literatura que aponta a automação como um meio de liberar profissionais para tarefas cognitivas de maior valor (Graefe, 2016). A discussão aqui se aprofunda ao considerar o impacto disso na qualidade jornalística. A otimização do tempo não resultou apenas em mais produtividade (mais textos em menos tempo), mas, segundo os relatos, em um redirecionamento do esforço para a

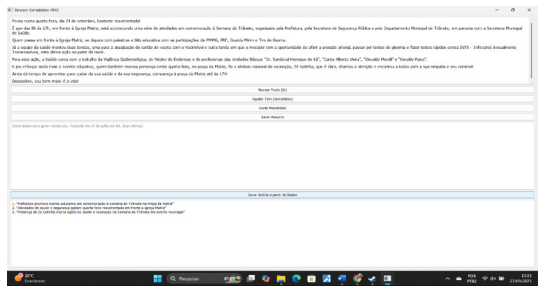
apuração, o que potencialmente eleva a qualidade e a profundidade do conteúdo final.

Tema 2: O Revisor como Plataforma Pedagógica e de Padronização

Para os estudantes, a ferramenta funcionou como um “editor particular”. Um dos participantes comentou: *“É como ter um professor experiente olhando seu texto e não só corrigindo, mas explicando o porquê. A sugestão de usar voz ativa, por exemplo, vinha com a justificativa de que ‘deixa o texto mais direto e forte’. Isso eu não esqueço mais.”* A capacidade da ferramenta de fornecer justificativas para suas sugestões foi um diferencial crucial, transformando a correção em um ato de aprendizado.

Figura 3

Gerador de manchetes a partir dos dados inseridos.



O autor.

Entre os profissionais, especialmente os que ocupavam cargos de edição, a ferramenta foi vista como um potencializador da padronização e do alinhamento ao manual de estilo do veículo. Um editor-chefe

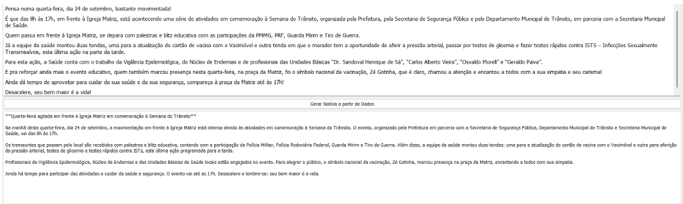
observou que a ferramenta poderia ajudar a “nivelar por cima” a qualidade dos textos recebidos de repórteres menos experientes, garantindo uma maior consistência editorial antes mesmo de o texto chegar à sua mesa. Isso sugere um novo fluxo de trabalho, onde a IA atua como um pré-editor, otimizando também o trabalho de liderança na redação.

Tema 3: Tensões e Limitações - O Fator Humano como Indispensável

Apesar dos benefícios, a interação com a ferramenta revelou tensões que reforçam a centralidade do julgamento humano. A principal limitação, citada por 80% dos participantes, foi o risco de padronização excessiva e a perda da voz autoral. A ferramenta, ao seguir estritamente as regras da redação jornalística “padrão”, tendia a “neutralizar” o tom original. Um jornalista conhecido por seu estilo mais irônico e textualidade elaborada relatou: *“Tentei passar uma crônica e ela transformou em notícia de agência. Tirou toda a graça. Para texto factual, é ótima. Para qualquer coisa que exija um pingão de estilo, ela é um rolo compressor”*.

Figura 4

Texto jornalístico gerado a partir dos dados inseridos de uma legenda de rede social



O autor.

Essa observação é crucial. Ela demonstra que a IA, em seu estado atual, opera com base em padrões estatísticos e não possui uma verdadeira compreensão de intenção, sarcasmo ou nuances culturais. A ferramenta interpretou de forma literal textos com duplo sentido, sugerindo alterações que descaracterizariam completamente a mensagem.

Outro ponto crítico foi a dependência de conexão à internet, um obstáculo prático para repórteres em campo. Além disso, a ferramenta não possui a capacidade de julgamento ético. Ela pode tornar um texto mais claro e conciso, mas não pode avaliar se a informação é precisa, se a fonte é confiável ou se a abordagem é justa com os envolvidos na história. Um participante resumiu de forma contundente: *“A IA corrige a forma, mas o fundo, a ética, a responsabilidade... isso é nosso. E é o que define o jornalismo”*.

Esses achados levam a uma discussão mais profunda sobre a colaboração homem-máquina. A ferramenta não é uma solução autônoma, mas um instrumento cujo valor depende da habilidade do usuário em discernir quais sugestões aceitar, quais rejeitar e quais adaptar. O jornalista, nesse modelo, atua como um supervisor crítico, mantendo o controle final sobre o produto editorial.

Considerações Finais

O desenvolvimento e os testes do Revisor Jornalístico com IA evidenciam o potencial da inteligência artificial como uma poderosa aliada na produção de conteúdo de qualidade. A experiência demonstrou ganhos significativos de produtividade e um notável valor pedagógico, reforçando o papel da IA como uma ferramenta de otimização e aprimoramento profissional. A relevância do projeto reside em sua capacidade

de reduzir o tempo gasto em tarefas mecânicas de revisão, permitindo que jornalistas e estudantes dediquem mais esforço à apuração, análise crítica e outras atividades que constituem o cerne da atividade jornalística.

Contudo, a pesquisa também iluminou as fronteiras e as limitações dessa tecnologia. A tendência à padronização estilística, a incapacidade de compreender nuances semânticas complexas e a ausência de julgamento ético são desafios que reafirmam a centralidade e a insubstituibilidade do jornalista. A adoção de ferramentas como a que foi desenvolvida não deve ser vista como uma automação completa do processo de escrita, mas como a introdução de um novo ator no fluxo de trabalho editorial — um ator que exige supervisão, senso crítico e direção humana.

As próximas etapas de desenvolvimento buscarão mitigar algumas das limitações identificadas. Pretende-se explorar a possibilidade de personalização da ferramenta, permitindo que os usuários ajustem os parâmetros de revisão para diferentes gêneros textuais (notícia, reportagem, crônica) ou para se alinharem a manuais de estilo específicos. Além disso, a criação de uma versão com funcionalidade offline é uma prioridade técnica.

Como perspectiva futura, propõe-se a expansão das funcionalidades para além da revisão, contemplando a geração assistida de conteúdos a partir de dados estruturados, como planilhas e bancos de dados. No entanto, esse passo exigirá um rigor ainda maior com os protocolos de verificação e checagem de informações, para garantir que a automação não se torne um vetor de imprecisões.

Dessa forma, a experiência relatada neste trabalho reforça que o futuro do jornalismo não será uma disputa entre humanos e máquinas,

mas uma colaboração complexa e em constante negociação. A adoção crítica e estratégica da inteligência artificial pode potencializar a qualidade da produção e ampliar as possibilidades de inovação no campo, desde que seja sempre guiada pela reflexão ética, pela transparência e por um compromisso inabalável com os princípios fundamentais da profissão.

Referências

- Bahia, J. (1990). *O jornal: história e técnica. Volume 2, As técnicas do jornalismo*. Mauad.
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?* [Trabalho apresentado]. Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency.
- Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, 11(4), 589-597.
- Broussard, M. (2023). *More Than a Glitch: Confronting Race, Gender, and Ability Bias in Tech*. MIT Press.
- Diakopoulos, N. (2016). Accountability in algorithmic decision making. *Communications of the ACM*, 59(2), 56-62.
- Diakopoulos, N. (2019). *Automating the News: How Algorithms Are Rewriting the Media*. Harvard University Press.
- Dörr, K. N., & Hollnbuchner, K. (2017). Ethical Challenges of Algorithmic Journalism. *Digital Journalism*, 5(4), 404-419.

- Erbolato, M. (2004). *Técnicas de Codificação em Jornalismo: redação, captação e edição no jornal diário*. Ática.
- Fanta, A., & Lamas, P. (2023). The journalist and the machine: a literature review on the role of AI in the newsroom. *Journalism Practice*, 17(5), 941-959.
- Flick, U. (2018). *An Introduction to Qualitative Research* (6th ed.). Sage Publications.
- Graefe, A. (2016). *Guide to Automated Journalism*. Tow Center for Digital Journalism.
- Hamilton, J. T., & Turner, F. (2009). Accountable Journalism: The Home News and the Great Depression. In B. H. Winfield (Ed.), *Journalism 1908: Birth of a Profession*. University of Missouri Press.
- Lewis, S. C., Guzman, A. L., & Schmidt, T. R. (2019). Automation, journalism, and human-machine communication: Rethinking roles and relationships of journalists and news. *Digital Journalism*, 7(4), 429-445.
- Lustosa, E. (1996). *O texto jornalístico: editoração, edição e estilo*. Ática.
- Marconi, F., & Siegman, A. (2024). *Newsmakers: Artificial intelligence and the future of journalism*. Columbia University Press.
- Montal, T., & Reich, Z. (2023). The ethics of AI in journalism: A review of the literature. *Journalism*, 24(7), 1371-1387.

- Nielsen, J. (2012). *Usability 101: Introduction to Usability*. Nielsen Norman Group.
- Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press.
- Rech, G. (2019). *Redação Jornalística: apontamentos para a produção de conteúdo*. Intersaberes.
- Teixeira, J. (2018). *O Pesadelo de Descartes: Do mundo mecânico à Inteligência Artificial*. Editora Fi.
- Thurman, N., Lewis, S. C., & Kunert, J. (2019). Algorithms, automation, and news. In K. Wahl-Jorgensen, & T. Hanitzsch (Eds.), *The handbook of journalism studies* (2nd ed., pp. 150-165). Routledge.
- Tripp, D. (2005). Pesquisa-ação: uma introdução metodológica. *Educação e Pesquisa*, 31(3), 443-466.
- Tuchman, G. (1972). Objectivity as strategic ritual: An examination of newsmen's notions of objectivity. *American Journal of Sociology*, 77(4), 660-679.
- Valente, C. (2024). *Inteligência Artificial: o impacto no jornalismo e nos jornalistas* [Projeto de Graduação, Universidade Fernando Pessoa].
- Zandomênic, R. (2022). Inteligência Artificial e Jornalismo: implicações na redação de notícias e na aquisição do conhecimento. *Revista Pauta Geral-Estudos em Jornalismo*, 9.

COMUNICACIÓN, COLABORACIÓN Y CHATGPT EN EL PROCESO DE DISEÑO: INTEGRACIÓN DE HERRAMIENTAS DIGITALES

Pablo Daniel Gutiérrez Gutiérrez¹

Hoy en día, la aceleración mediática y tecnológica es evidente. La sociedad contemporánea está marcada por el uso exponencial de la inteligencia artificial generativa (IAG) en múltiples disciplinas (Vaccaro et al., 2024) lo que demanda una reflexión acerca de las estrategias de comunicación utilizadas en ellas (Pérez, 2005; Scheinsohn, 2010). Asimismo, es importante aplicar este criterio en el proceso de diseño y la manera en que se toman decisiones durante las sesiones de trabajo colaborativo. En esta investigación se realiza un acercamiento a las dinámicas, decisiones y estrategias de comunicación utilizadas por estudiantes de los primeros semestres de una universidad de diseño,

1. Universidad Autónoma de Ciudad Juárez.
al251398@alumnos.uacj.mx

cuando se enfrentan al reto del desarrollo de un proyecto, bajo presión de tiempo, utilizando un modelo de diseño y herramientas digitales como la IAG y entornos de trabajo colaborativo.

Antecedentes

En los últimos veinte años, el diseño se establecido como una disciplina estratégica, altamente regida por entornos digitales, en donde la comunicación y el trabajo en equipo se han vuelto un requisito indispensable para la toma de decisiones y la alineación de objetivos (Pérez, 2005; Scheinsohn, 2010). Al mismo tiempo, modelos de trabajo como los del Doble Diamante (Design Council, 2014) han brindado una estructura clara y reconocible para organizar procesos y obtener entregables, abriendo la posibilidad de generar trazabilidad en la toma de decisiones. Por su parte, la integración de *IAG* como *ChatGPT* ha permitido acelerar tareas de búsqueda, síntesis y prototipado (Papachristos et al., 2024).

Metodología

Esta es investigación una metodológica cualitativa con un diseño etnográfico y de alcance descriptivo. Se enfoca en estudiantes del Programa de Licenciatura en Diseño Digital de Medios Interactivos (DDMI) de la Universidad Autónoma de Ciudad Juárez, en México. El objetivo es estudiar su proceso de toma de decisiones y colaboración durante la práctica de diseño, mientras que los participantes desarrollaron un artículo en revista web eligiendo ellos el tema del proyecto. El estudio contemplo a los alumnos de la materia Multimedia y Web, organizados en equipos de 4, 5 y 6 integrantes. Se realizaron dos sesiones, siguiendo

el proceso diseño del Doble Diamante del *British Design Council*. Los estudiantes documentaron sus decisiones en un instrumento de seguimiento metodológico diseñado en *FigJam* inspirado en el modelo del Doble Diamante (Design Council, 2007), contestando preguntas específicas en cada elemento de un gráfico de hitos, mientras que el observador realizaba rondines y toma de nota de lo percibido como relevante en la interacción, llenando tablas cronológicas, con pausas reflexivas cada 10 minutos. Para optimizar el registro y análisis de la información durante la actividad, se motivó el uso de la herramienta de inteligencia artificial *ChatGPT* (OpenAI, 2024).

Instrumentos metodológicos

Para las observaciones se elaboraron tres instrumentos, primero uno de observaciones con tablas cronológicas, el cual consistía en dos tablas; una tabla cronológica para el registro de observaciones por equipo, la cual se modificó en la segunda sesión y otra anexa para el registro de observaciones del grupo en general, la cual solo se usó en la segunda sesión. También se generó un instrumento de seguimiento metodológico para el registro del proceso metodológico y resultados de todos los equipos, el cual también se modificó en la sesión posterior, eliminándole preguntas y añadiéndole un módulo inspirado en *scrum*. Por último, se implementó el cuestionario reflexivo de *Microsoft Form*, como instrumento para la recopilación de opiniones en retrospectiva.

El primer instrumento consistió en dos tablas en las que se llevó a cabo el registro de lo ocurrido durante la actividad. En la primera sesión se empleó una tabla con categorías específicas para documentar los avances de cada equipo en intervalos de diez minutos (Figura 1).

Figura 1

Ejemplo de Tabla cronológica de avances

Equipo 1

Tiempo (min)	% de avance percibido	Logros	Bloqueos	Observaciones
10	70	<ul style="list-style-type: none">OrganizaciónDeterminaron herramientasRealizaron Moodboard		3 miembros Trabajo en conjunto,
10	100			

Equipo 2

Tiempo (min)	% de avance percibido	Avances	Bloqueos	Observaciones
	45	<ul style="list-style-type: none">Organización	2	5 miembros

Figura 2

Tabla cronológica de avances simplificada

Equipo 1

Tiempo (min)	Observaciones
10	<ul style="list-style-type: none">6 miembros en el equipoEl equipo se organizó rápido y comenzaron a trabajarse ve que están utilizando mucho Chat GPTse observan muchas listasParece haber un énfasis en los técnico y en la informaciónNo veo uso de plantillasAlto uso de Chat GPT
10	
10	<ul style="list-style-type: none">1 miembro esta buscando imágenes de internetSe sugiere utilizar técnicas de Design thinking (solo se usaron al principio)
10	<ul style="list-style-type: none">1 miembro se tuvo que retirar tempranoSe comienzan a buscar referencias de revistas y artículos
10	<ul style="list-style-type: none">Revisión y Marco de referencias visuales de revistas electrónicas
10	<ul style="list-style-type: none">Se realizan mockups y acercamientos de estilo en base a las referencias visuales

Equipo 2

Tiempo (min)	Observaciones
	<ul style="list-style-type: none">5 miembros2 miembros del equipo no tienen conexión

Este formato permitió registrar logros, bloqueos y observaciones, pero distraía a los participantes y la necesidad de clasificar cada anotación hizo más lento el registro sin añadir ventajas claras. Por esta razón, en la segunda sesión se adoptó un formato simplificado de dos columnas, una de tiempo y una de observaciones, que facilitó la toma de notas y permitió reflejar con mayor continuidad las dinámicas observadas (Figura 2).

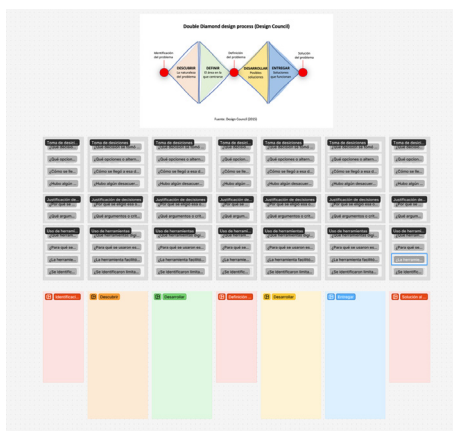
El formato con menos campos, agilizo el tiempo de llenado de las observaciones y permitió un mayor rango de atención en las actividades realizadas por los participantes. La tabla anexa se realizó para llevar el registro de lo sucedido de manera general, se hace el registro por sets de 10 y descansos de 5 minutos (Figura 3)

Figura 3
Comentarios y observaciones grupales por set, de creación propia

Time Set	Observación grupal
Set 01	Bajo uso de herramientas de diseño, uso de ChatGPT para generar ideas por medio de listas o lluvias de ideas que le piden a la herramienta.
Break 01	Se recomendó el uso de plantillas, técnicas y métodos de diseño para algunos entregables.
Set 02	Bajo uso de plantillas o técnicas de diseño, baja comprensión del propósito y función de las técnicas o métodos.
Break 02	Se refuerza la recomendación del uso de plantillas, métodos o técnicas de diseño. Explica función, uso y propósito de manera breve.
Set 03	
Break 03	Se proveyó de herramientas para indicar avances, que salió bien, que salió mal, bloqueos. Inspirados de SCRUM Se sugiere usar herramientas o técnicas del Design Thinking
Set 04	
Break 04	Se les solicito que todo el equipo tomara indicara las notas
Set 05	
Break 05	Se les pregunto ¿Que hace que una revista web sea una revista web? Se definió hacer un artículo de revista web
Set 06	A los 45 minutos, los participantes comienzan a dispersarse, platicando de temas de interés
Break 06	
Set 07	
Break 07	
Set 08	
Break 08	Aquí termino el ejercicio

El tercer instrumento fue creado para facilitar la estructuración de observaciones y resultados, agilizar la documentación y la realización de las actividades. Se les informó que la actividad tiene pausas de reflexión cada 10 minutos y se les pidió que utilizaran el instrumento de seguimiento metodológico (Figura 4). En él se utilizarían las fases del modelo representadas en el instrumento para colocar los entregables, tomar notas y la interacción entre los miembros del equipo.

Figura 4
Instrumento de seguimiento metodológico



Sin embargo, en la segunda sesión fue necesario modificar el instrumento de seguimiento metodológico y se eliminaron las preguntas para trasladar su función de registro a otro instrumento, esto debido a que los participantes al disponerse a contestar las preguntas se distraían de la implementación y seguimiento metodológicos del proceso de diseño. Al instrumento se le añadió un módulo inspirado en la retrospectiva del

sprint de *Scrum* para registrar avances y bloqueos, con capturas de pantalla y notas rápidas (Figura 5).

Figura 5

Instrumento de seguimiento metodológico modificado

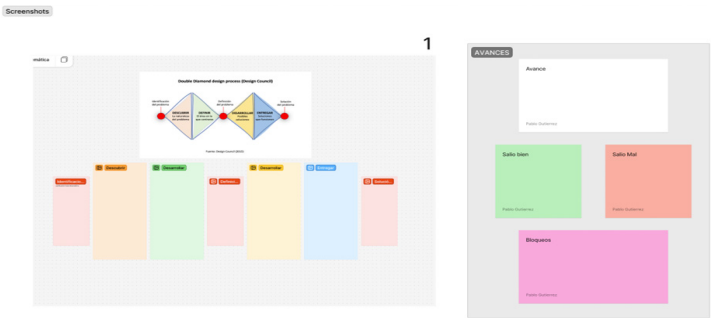


Figura 6

Muestras del instrumento de preguntas de observación en Microsoft Forms, de creación propia

Por último, se aplicó un formulario reflexivo en *Microsoft Forms* al final de la sesión. Este cuestionario consistió en cinco preguntas abiertas que se aplicaron para cada etapa del modelo, las preguntas fueron las

siguientes: ¿Qué decisiones se tomaron en este momento del proceso?, ¿Como elegiste que hacer para llegar a tu objetivo?, ¿Como sentiste el trabajo en equipo?, ¿Que salió bien?, ¿Que salió mal? y Comentarios adicionales (Figura 6). Estas preguntas ayudaron a mapear el estado de las emociones y decisiones de los alumnos durante el desarrollo de su proyecto.

Preparación de los datos

Debido a la cantidad de información generada, para su análisis fue necesario preparar los datos de las tablas y del cuestionario. A continuación, se expone de manera breve el proceso que se siguió.

Instrumento 1. Observaciones con tablas cronológicas. El primer instrumento no requirió de mucho tratamiento, solo se añadieron dos columnas a la tabla: *Equipo* y *Set*, resultando en una tabla de siete columnas: *Equipo*, *Tiempo (min)*, *Set*, *% de avance percibido*, *Logros*, *Bloqueos* y *Observaciones*. La Tabla 2 de la segunda sesión del Instrumento 1 fue simplificada a dos columnas, sin embargo, para su mejor manejo, también se le añadieron las columnas *Equipo* y *Set*, quedando en cuatro columnas: *Equipo*, *Tiempo (min)*, *Set* y *Observaciones*. El formato con menos campos, agilizó el tiempo de llenado de las observaciones. La tercera tabla consiste en dos columnas y no se le dio ningún tratamiento.

Instrumento 2: Registro del proceso metodológico por equipo.

El segundo instrumento corresponde al registro del proceso metodológico desarrollado por cada equipo, estructurado a partir del modelo del Doble

Diamante. En este caso, se empleó una herramienta digital colaborativa llamada *FigJam de Figma*, donde los equipos documentaron sus avances en cada etapa del proceso. A continuación, se muestran los resultados por equipo y por set de la progresión de su proyecto (Tabla 11), en esta tabla se muestran los “pantallazos” del progreso del instrumento junto con comentario de los participantes alrededor de ideas de “Que salió bien”, “que salió mal”, “Avance” y “Bloqueos. la indicación de avances y bloqueos, los sets son correspondientes con los sets de otras tablas y se refieren al mismo evento. El registro visual incluyó tanto las secciones completadas del tablero (capturas de pantalla) como las notas reflexivas elaboradas durante cada set de trabajo, en las que se identificaron los aciertos, dificultades y bloqueos del proceso.

Instrumento 3: Cuestionario reflexivo. Debido a la gran cantidad de información que generó este instrumento, fue necesario realizar una serie de pasos que ayudaron a visualizar y organizar los datos a través del proceso de diseño del modelo, identificar categorías de palabras por pregunta, asignar una categoría por cada una de las respuestas de los participantes y contar el número de aparición de categorías por pregunta y por proceso, para después realizar gráficos que ayudaron a visualizar la información.

Paso 1. Matriz de visualización. Primero se realizó una matriz a partir del libro de Excel que arroja *Microsoft forms* (Figura 7), el archivo contiene las respuestas de los participantes a través del proceso de diseño, dispuesto de manera horizontal, aborda las cinco preguntas,

los tres hitos y las cuatro etapas correspondientes al modelo. Esos datos se extraen para formar una matriz acomodada para apreciarse de manera visual (Error! Reference source not found.).

Figura 7

Ejemplo de hoja de Excel extraída de Microsoft Forms

0001	0002	0003	0004	0005	0006	0007	0008	0009	0010	0011	0012	0013	0014	0015	0016	0017	0018	0019	0020	0021	0022	0023	0024	0025	0026	0027	0028	0029	0030	0031	0032	0033	0034	0035	0036	0037	0038	0039	0040	0041	0042	0043	0044	0045	0046	0047	0048	0049	0050	0051	0052	0053	0054	0055	0056	0057	0058	0059	0060	0061	0062	0063	0064	0065	0066	0067	0068	0069	0070	0071	0072	0073	0074	0075	0076	0077	0078	0079	0080	0081	0082	0083	0084	0085	0086	0087	0088	0089	0090	0091	0092	0093	0094	0095	0096	0097	0098	0099	0100
0101	0102	0103	0104	0105	0106	0107	0108	0109	0110	0111	0112	0113	0114	0115	0116	0117	0118	0119	0120	0121	0122	0123	0124	0125	0126	0127	0128	0129	0130	0131	0132	0133	0134	0135	0136	0137	0138	0139	0140	0141	0142	0143	0144	0145	0146	0147	0148	0149	0150	0151	0152	0153	0154	0155	0156	0157	0158	0159	0160	0161	0162	0163	0164	0165	0166	0167	0168	0169	0170	0171	0172	0173	0174	0175	0176	0177	0178	0179	0180	0181	0182	0183	0184	0185	0186	0187	0188	0189	0190	0191	0192	0193	0194	0195	0196	0197	0198	0199	0200
0201	0202	0203	0204	0205	0206	0207	0208	0209	0210	0211	0212	0213	0214	0215	0216	0217	0218	0219	0220	0221	0222	0223	0224	0225	0226	0227	0228	0229	0230	0231	0232	0233	0234	0235	0236	0237	0238	0239	0240	0241	0242	0243	0244	0245	0246	0247	0248	0249	0250	0251	0252	0253	0254	0255	0256	0257	0258	0259	0260	0261	0262	0263	0264	0265	0266	0267	0268	0269	0270	0271	0272	0273	0274	0275	0276	0277	0278	0279	0280	0281	0282	0283	0284	0285	0286	0287	0288	0289	0290	0291	0292	0293	0294	0295	0296	0297	0298	0299	0300
0301	0302	0303	0304	0305	0306	0307	0308	0309	0310	0311	0312	0313	0314	0315	0316	0317	0318	0319	0320	0321	0322	0323	0324	0325	0326	0327	0328	0329	0330	0331	0332	0333	0334	0335	0336	0337	0338	0339	0340	0341	0342	0343	0344	0345	0346	0347	0348	0349	0350	0351	0352	0353	0354	0355	0356	0357	0358	0359	0360	0361	0362	0363	0364	0365	0366	0367	0368	0369	0370	0																													

En la Figura 8 se puede apreciar en la primera columna, las etapas e hitos del doble diamante, la segunda columna el listado las preguntas y en las columnas consecuentes, las respuestas por equipo. Este acomodo permite organizar visualmente las preguntas para después categorizarlas.

Figura 8

Matriz de visualización

		Date		Time		Location		Weather		Wind		Sea		Temperature		Humidity		Pressure		Visibility		Clouds		Moon		Sun		Stars		Planets		Comets		Aurora		Other																																																																	
Date	Time	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

Paso 2. Parametrizar las respuestas. Con la matriz realizada, se extrajeron muestras de las respuestas y las preguntas para generar una lista de categorías en común a partir de las respuestas del cuestionario. Para ello se trabajó junto con ChatGPT 5, al que se le solicitó parametrizar las respuestas buscando palabras clave y temáticas en común por pregunta, mientras que se seleccionaban categorías pertinentes a la pregunta y al contexto. Después de varias iteraciones de este proceso, se obtuvo una lista de categorías comunes para asignarle a cada respuesta según su pregunta (ver Figura 9). Las categorías por pregunta resultaron en ocho categorías para las preguntas 1 y 2, cuatro para la pregunta 3, diez para las 4 y 5 y cinco para la numero 5.

Figura 9

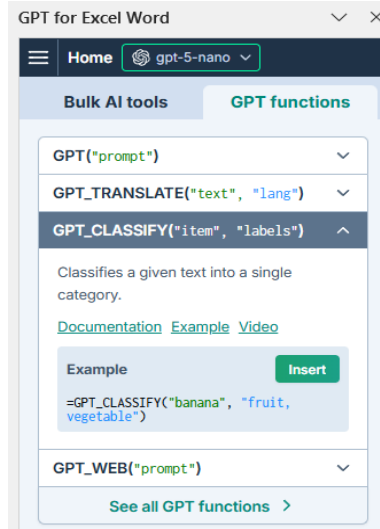
Tabla de categorías y variables por pregunta.
De creación propia

Pregunta	Variable	Categoría
¿Qué decisiones se tomaron en este momento del proceso?	accion_buena	Definición y delimitación del tema
	accion_buena	Metodología de investigación
	accion_buena	Organización y planificación del trabajo
	accion_buena	Investigación y análisis de información
	accion_buena	Influencia metodológica
	accion_buena	Comunicación y asesoramiento
	accion_buena	Reflexión y crítica crítica
	accion_buena	Sin asociación a ninguna
¿Como elegiste que hacer para llegar a tu objetivo?	eleccion_muchas	Contexto y evaluación general
	eleccion_muchas	Identificación del problema
	eleccion_muchas	Definición de objetivos
	eleccion_muchas	Proceso de investigación
	eleccion_muchas	Relevancia y validez de datos
	eleccion_muchas	Comunicación y coordinación
	eleccion_muchas	Experiencia y reflexión
	eleccion_muchas	Sin asociación a ninguna
¿Como sentiste el trabajo en equipo?	memoria_mucha_mal	Presión
	memoria_mucha_mal	Neurosis
	memoria_mucha_mal	Neurosis
	memoria_mucha_mal	Mala
¿Que salió bien?	valor_went_went	Observación
	valor_went_went	Comunicación
	valor_went_went	Claridad
	valor_went_went	Decisiones
	valor_went_went	Colaboración
	valor_went_went	Información
	valor_went_went	Reflexión
	valor_went_went	Técnica
¿Que salió mal?	valor_went_went	Relevancia
	valor_went_went	Definición
	valor_went_went	Observación
	valor_went_went	Comunicación
	valor_went_went	Claridad
	valor_went_went	Decisiones
	valor_went_went	Relevancia
	valor_went_went	Información
Comentarios adicionales	valor_went_went	Reflexión
	valor_went_went	Técnica
	valor_went_went	Sin asociación
	valor_went_went	Contexto, estrategia o de validación
	valor_went_went	Observación descriptiva en carga emocional
	valor_went_went	Comunicación descriptiva de aprendizaje
	valor_went_went	Observación sobre aspectos técnicos o de infraestructura

Paso 3. Clasificar las respuestas según las categorías. En Excel se generó una hoja de trabajo nueva a partir de la que contenía la matriz se visualización, se le llamo *Mapa de Proceso INTERPRETADO*, con

Figura 11

*Captura de pantalla de la herramienta
“Add-In” en “Excel”*



Después de haber habilitado celdas para ingresar las categorías por respuesta, se deben llenar con una de las opciones de categoría pertenecientes a cada pregunta, según la respuesta del miembro del equipo. Para esto se llenaron de manera manual varias celdas al azar, para después utilizar la función: *GPT_CLASSIFY()* del *Add-Inn* de *Excel* llamado *GPT for Excel Word*, usando el modelo de IA *gpt-5-nano*. Esta función se coloca en las celdas vacías y se le introducen al menos dos parámetros, *item* y *categories*, siendo el primero, el contenido a analizar y el segundo las opciones posibles con las que puede responder. Una vez llenas las 800 celdas disponibles, la herramienta llenó de categorías la matriz de visualización. Se requirió adquirir un paquete de créditos en

dólares, del cual se consumió aproximadamente ~\$4.58 pesos mexicanos, incluyendo ensayos de prueba y error, a la tasa de cambio actual de \$18.35 pesos por dólar. Después se seleccionaron al menos dos celdas, por pregunta, por etapa, para verificar que las respuestas del modelo fueran congruentes con lo que un humano hubiera determinado. Una muestra de la tabla se puede apreciar en la Figura 12. Al final se generó una hoja de trabajo de respaldo, llamada *Mapa de Proceso SELLADO* donde la matriz no contiene formulas, esto, para conservar los datos obtenidos y evitar el uso innecesario de créditos. La hoja de trabajo de Excel se puede encontrar en el archivo *MAPA DE PROCESO INTERPRETADO V4 GPT (versión 1)* (Gutiérrez-Gutiérrez, 2025).

Figura 12

Muestra de una parte del Mapa de proceso lleno.
De creación propia

Criterio de etapa	Etapa	Etapa 1	Etapa 2	Etapa 3	Etapa 4	Etapa 5	Etapa 6
Planteamiento del problema	¿Qué decisiones se tomarán en este momento del proceso?	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema
Planteamiento del problema	decision_force	Definición y conceptualización	Definición y conceptualización	Definición y conceptualización	Definición y conceptualización	Definición y conceptualización	Definición y conceptualización
Planteamiento del problema	¿Cómo elegiste que hacer para llegar a tu objetivo?	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema
Planteamiento del problema	selection_method	Definición y conceptualización	Definición y conceptualización	Definición y conceptualización	Definición y conceptualización	Definición y conceptualización	Definición y conceptualización
Planteamiento del problema	¿Cómo sentiste el trabajo en equipo?	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema
Planteamiento del problema	teamwork_status_felt	Definición y conceptualización	Definición y conceptualización	Definición y conceptualización	Definición y conceptualización	Definición y conceptualización	Definición y conceptualización
Planteamiento del problema	¿Qué salió bien?	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema
Planteamiento del problema	what_went_well	Definición y conceptualización	Definición y conceptualización	Definición y conceptualización	Definición y conceptualización	Definición y conceptualización	Definición y conceptualización
Planteamiento del problema	¿Qué salió mal?	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema
Planteamiento del problema	what_went_bad	Definición y conceptualización	Definición y conceptualización	Definición y conceptualización	Definición y conceptualización	Definición y conceptualización	Definición y conceptualización
Planteamiento del problema	Comentarios adicionales	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema	Definición de los objetivos y alcance del problema
Planteamiento del problema	free_comment	Definición y conceptualización	Definición y conceptualización	Definición y conceptualización	Definición y conceptualización	Definición y conceptualización	Definición y conceptualización

Paso 5. Visualización de datos.

Teniendo el Mapa de proceso respaldado y listo con puros valores, se generó una nueva hoja de trabajo a la que se le llamo *Dashboard*, una herramienta de visualización en la que se muestra el número de veces que cada categoría cabe en una respuesta. Para ellos se generó una matriz correlacional en la que se interceptan las categorías y preguntas con las etapas del modelo del doble diamante (Figura 13). También se generaron gráficas para visualizar e interpretar los datos. Se realizaron tanto lo gráficos de manera global, para los cuatro equipos, como los pertenecientes a cada uno. En la Figura 13 se puede obtener una visualización global de las categorías más utilizadas y la Figura 14, se muestra un grupo de gráficas correspondientes a las categorías más seleccionados por pregunta y el progreso del estado de ánimo del equipo, cruzado con las categorías más seleccionadas por etapa del proceso. Las gráficas se verán más a detalle en la sección de resultados.

Resultados

En esta sección se presenta de manera general la información recopilada mediante los tres instrumentos de acercamiento. El primero corresponde a las tablas cronológicas de observación, diseñadas para registrar el desarrollo del trabajo en intervalos de diez minutos por equipo, con el propósito de documentar logros, bloqueos y dinámicas relevantes. El segundo instrumento consiste en el registro del proceso metodológico llevado a cabo por cada equipo, estructurado a partir de los hitos propuestos en el modelo del Doble Diamante, donde se documenta el avance y las decisiones adoptadas en cada etapa del diseño.

Finalmente, el tercer instrumento corresponde al cuestionario reflexivo aplicado al término de las sesiones, realizado con *Microsoft Forms*, orientado a recuperar la perspectiva de los participantes en torno a las razones que guiaron sus decisiones durante el proceso. A continuación, se presentan los resultados de cada instrumento, iniciando con las tablas de observación.

Instrumento 1: Observaciones con tablas cronológicas

En la primera tabla, se registran las observaciones por sets de 10 minutos y por equipo, fueron 20 minutos de desarrollo del proceso y una cantidad de tiempo aproximada de entre 10 minutos y 15 minutos entre cada set para toma de notas (Tabla 1).

Tabla 1

Tabla cronológica de avances por equipo. Sesión 1. 3 de abril, 2025, de creación propia

Equipo	Tiempo (min)	Set	% de avance percibido	Logros	Bloqueos	Observaciones
1	10	1	70	Organización Determinaron herramientas Realizan un MoodBoard		3 miembros Trabajo en conjunto, Uso bajo de ChatGPT
1	10	2	100			
2	10	1	45	Organización Idea de tema y de problema planteado	2 miembros del equipo no tienen conexión	5 miembros Los miembros eligieron el tema propuesto por una persona del equipo.

2	10	2	90		Uso bajo de ChatGPT, lluvia de ideas
3	10	1	70	Organización plantear herramientas a utilizar	4 miembros
3	10	2	100		Uso bajo de ChatGPT, lluvia de ideas
4	10	1	75	Organización plantear herramientas a utilizar	5 miembros
4	10	2	100		Uso bajo de ChatGPT

Los equipos completaron el primer hito, llamado identificación del problema en aproximadamente 40 a 50 minutos. En esta etapa, la mayoría de los esfuerzos se dirigen a la configuración del espacio de trabajo, la organización del equipo y la asignación de roles. El equipo 1 y 3 estaban incompletos, el primero debido a problemas de conectividad, el otro debido a que dos de sus miembros estaban ausentes. El nivel de uso de ChatGPT fue bajo, se recurrió a la herramienta para generar listas alrededor del problema o lluvia de ideas.

En Tabla 2, se registraron observaciones por sets de 10 minutos y por equipo, las observaciones se realizaron integrando logros, bloques y observaciones en una sola columna. Se llevaron a cabo 6 sets de 10 minutos, con intervalos de 5 minutos para captura del registro de avances, dando un total de 90 minutos de desarrollo y registro de avances, completando las 4 etapas y 2 hitos restantes del modelo del doble diamante. En esta ocasión la inasistencia fue menor, con un miembro

del equipo 3 ausente. Sin embargo, el equipo 2, aunque completo, sigue con dos miembros inactivos, uno debido a problemas con el internet y el otro debido que se encuentra distraído y con poca interacción con sus compañeros. En resumen, durante la segunda sesión se observó un mayor nivel de avance y participación, una reducción significativa de inasistencias, pero hubo problemas técnicos y de organización, principalmente en el equipo 2. El nivel de uso de ChatGPT fue diverso pero constante en todas las etapas, se recurrió a la herramienta para generar listas, referencias y prototipos rápidos. El equipo 4 fue el que menor uso le dio al ChatGPT, ya que optaba por utilizar métodos tradicionales o propio y los Equipos 1 y 3 fueron los que mayor uso hicieron de la herramienta de IA. Se realizó una tercera tabla (Tabla 3) donde se registraron observaciones del grupo en general, se hizo el registro por sets de 10 minutos, pero también se registró lo que sucedió en los descansos de 5 minutos.

Tabla 2

Tabla cronológica de avance por equipo. Sesión 2. 7 de abril, 2025, de creación propia.

Equipo	Tiempo (min)	Set	Observaciones
1	10	3	6 miembros en el equipo El equipo se organizó rápido y comenzaron a trabajar se ve que están utilizando mucho Chat GPT se observan muchas listas Parece haber un énfasis en los técnicos y en la información No veo uso de plantillas Alto uso de Chat GPT
1	10	4	

1	10	5	1 miembro está buscando imágenes de internet Se sugiere utilizar técnicas de Design thinking (solo se usaron al principio)
1	10	6	1 miembro se tuvo que retirar temprano Se comienzan a buscar referencias de revistas y artículos
1	10	7	Revisión y Marco de referencias visuales de revistas electrónicas
1	10	8	Se realizan mockups y acercamientos de estilo en base a las referencias visuales
2	10	3	5 miembros 2 miembros del equipo no tienen conexión Organización Idea de tema y de problema planteado 1 miembro llegó tarde 1 usando su laptop tuvo problemas para usar internet La transición de su laptop a la PC del salón parece costarle 1 miembro con Laptop sin internet. Se le sugirió usar la PC del salón, pero sigue intentando conectarse a las actividades a través de la Laptop. el miembro que llegó tarde está conectándose para iniciar actividad Mediano uso de ChatGPT
2	10	4	
2	10	5	
2	10	6	El participante sin internet opta por usar su móvil y escribir a mano en su laptop 1 miembro este distraído El trabajo parece centrarse en tres personas
2	10	7	1 miembro sigue distraído, intenta hacer aportes, pero no parecen tener impacto directo en el proceso de los demás miembros del equipo El miembro con la laptop sin internet sig.
2	10	8	
3	10	3	6 miembros 1 miembro ausente (chica) 2 miembros llegaron tarde 1 estaba confundido acerca de que hacer Utilizan las plantillas de técnicas del Design Thinking Parece haber una persona asignada a realizarlo Roles de lectura, investigación, prompting y de creación de plantillas Alto uso de ChatGPT
3	10	4	

3	10	5	
3	10	6	Deciden hacer una revista en lugar de un artículo
3	10	7	El equipo siente que la etapa de entrega es difícil de concretar en el tiempo asignado. Se les sugiere al menos entregar un boceto
3	10	8	
4	10	3	5 miembros Se ve que utilizan diagramas visuales y plantillas de técnicas Roles de lectura, investigación, prompting y de creación de plantillas Bajo uso de ChatGPT
4	10	4	
4	10	5	
4	10	6	El equipo no ha usado técnicas, sienten que deben de hacer una estructura global y después adentrarse en detalles
4	10	7	Inquietudes acerca de la etapa de entrega, ¿qué hay que entregar?
4	10	8	No terminaron

Tabla 3

Comentarios y observaciones grupales por set, de creación propia

Time Set	Observación grupal
Set 01	Bajo uso de herramientas de diseño, uso de ChatGPT para generar ideas por medio de listas o lluvias de ideas que le piden a la herramienta.
Break 01	Se recomendó el uso de plantillas, técnicas y métodos de diseño para algunos entregables.
Set 02	Bajo uso de plantillas o técnicas de diseño, baja comprensión del propósito y función de las técnicas o métodos.
Break 02	Se refuerza la recomendación del uso de plantillas, métodos o técnicas de diseño. Explica función, uso y propósito de manera breve.
Set 03	
Break 03	Se proveyó de herramientas para indicar avances, que salió bien, que salió mal, bloqueos. Inspirados de SCRUM Se sugiere usar herramientas o técnicas del Design Thinking
Set 04	

Break 04	Se les solicito que todo el equipo tomara indicara las notas
Set 05	
Break 05	Se les pregunto ¿Que hace que una revista web sea una revista web? Se definió hacer un artículo de revista web
Set 06	A los 45 minutos, los participantes comienzan a dispersarse, platicando de temas de interés
Break 06	
Set 07	
Break 07	
Set 08	
Break 08	Aquí termino el ejercicio

Resumen de observaciones con tablas cronológicas. A continuación, se realiza una síntesis de los registros obtenidos en las tablas cronológicas de observación, correspondientes a las dos sesiones de trabajo (Tabla 4). Estas tablas recogen la información sobre la participación de los equipos, el uso de herramientas, la aplicación de metodologías y los avances logrados en cada etapa del proceso.

Tabla 4

*Síntesis de observaciones por equipo.
Sesión 1 (3 de abril de 2025)*

Variable	Equipo 1	Equipo 2	Equipo 3	Equipo 4
Asistencia y participación	Equipo incompleto con tres miembros. Avance afectado por la carga de trabajo.	Cinco miembros; dos con problemas de conexión. Participación parcial.	Cuatro miembros activos. Trabajo continuo con comunicación limitada.	Equipo completo y estable. Participación sostenida.

Enfoque del trabajo	Organización y selección de herramientas. Trabajo inicial con Moodboard.	Definición de tema y planteamiento del problema.	Enfoque en la planificación de herramientas y organización del flujo.	Organización y planteamiento de herramientas.
Uso de ChatGPT	Uso bajo. Se emplea para lluvia de ideas.	Uso moderado. Se generan listas temáticas.	Uso bajo. Empleado en exploraciones iniciales.	Uso bajo o nulo. Preferencia por métodos tradicionales.
Aplicación de métodos de diseño	Mencionan técnicas, pero sin aplicación directa.	No se identifican métodos explícitos.	Referencia general a herramientas, sin integración formal.	Uso limitado de técnicas o métodos de diseño.
Bloqueos técnicos y organizacionales	Ausencias que afectaron la coordinación.	Problemas de conexión impidieron participación completa.	Inasistencia de miembros y comunicación intermitente.	Sin bloqueos relevantes.
Avance metodológico	Alcanzan el hito de identificación del problema.	Avance parcial en la definición del problema.	Alcanzan el primer hito.	Completa el primer hito con estabilidad.
Resultados parciales y observables	Configuración del espacio de trabajo y roles.	Elección de tema y justificación inicial.	Organización estructural y definición de herramientas.	Documentación básica del proceso inicial.

En la Tabla 4 se muestra que los equipos 1 y 2 trabajaron incompletos. Los esfuerzos se dirigieron a la organización, definición de herramientas y delimitación del problema. El uso de ChatGPT fue bajo, con aplicación parcial o nula de métodos o técnicas de diseño. Los equipos 2 y 3 presentaron problemas de conectividad. El avance metodológico mostró que solo se completó el hito de identificación del problema. Los resultados obtenidos fueron la configuración de su espacio de trabajo y la organización y dinámica de roles de cada equipo.

Es importante considerar el tiempo que requiere la organización y planeación inicial de los proyectos, ya que un método de registro de observaciones demasiado específico puede afectar el ritmo del trabajo. También resulta necesario prever un periodo de orientación en el uso de herramientas como ChatGPT y en la aplicación de métodos o técnicas de diseño para la generación de entregables.

Tabla 5
Síntesis de observaciones por equipo.
Sesión 2 (7 de abril de 2025)

Variable	Equipo 1	Equipo 2	Equipo 3	Equipo 4
Asistencia y participación	Seis miembros presentes. Un retiro temprano sin afectar el flujo general.	Cinco miembros activos; dos con problemas de conexión e inactividad parcial.	Cinco miembros activos; una ausencia.	Cinco miembros activos; participación sostenida.
Enfoque del trabajo	Referencias visuales, mockups y revisión de estilo.	Organización y definición de tema con apoyo parcial de IA.	Producción de una revista digital. Aplicación de plantillas.	Diseño estructural global y revisión de entregables.
Uso de ChatGPT	Uso alto. Generación de listas, referencias y contenido visual.	Uso medio. Integración irregular en el flujo de trabajo.	Uso alto. Apoyo en redacción, prompting y búsqueda de referencias.	Uso bajo. Preferencia por métodos propios y manuales.
Aplicación de métodos de diseño	Bajo uso de técnicas. Se sugiere incorporar Design Thinking.	Aplicación parcial de metodologías.	Uso activo de plantillas y técnicas del Design Thinking.	Uso de diagramas y plantillas con roles definidos.

Bloqueos técnicos y organizacionales	Un miembro se retira antes de tiempo; sin afectación mayor.	Persisten problemas de conexión e inactividad.	Dificultad para concretar entregas dentro del tiempo.	Dudas sobre el entregable final; falta de claridad en la meta.
Avance metodológico	Completan las etapas de desarrollo y entrega del Doble Diamante.	Avance moderado en etapas intermedias.	Cumplen con todas las fases; definen entregables.	Avance progresivo, centrado en estructura general.
Resultados parciales y observables	Producción de mockups y referencias visuales.	Avances fragmentados, sin consolidación final.	Desarrollo de revista digital y reflexiones sobre proceso.	Definición estructural del proyecto, sin visuales finales.

En la Tabla 5, correspondiente a la segunda sesión, se observa que solo el equipo 3 estuvo incompleto. El enfoque de trabajo se orientó hacia la producción de entregables como *mockups*, referencias visuales y propuestas de estilo, aplicación de plantillas y diseño estructural. El uso de ChatGPT se amplió, siendo los equipos 1 y 3 los que le dieron mayor uso, generando referencias y contenido visual; el equipo 2 lo utilizó de forma intermedia y el equipo 4 optó por un uso mínimo. Se comenzaron a emplear plantillas, métodos y técnicas de diseño. El equipo 2 continuó con problemas de conectividad, el equipo 3 comenzó a tener dificultades con el ritmo de trabajo y el equipo 4 llegó a un punto en el que no era claro cómo continuar. Al final de la sesión, todos los equipos, a excepción del equipo 4, completaron las cuatro etapas y los dos hitos restantes del modelo del Doble Diamante. Los equipos 1 y 3 realizaron todas las etapas y generaron propuestas visuales; el equipo 2 terminó el proceso, pero no elaboró una propuesta visual, y el equipo 4 no logró concluir el trabajo.

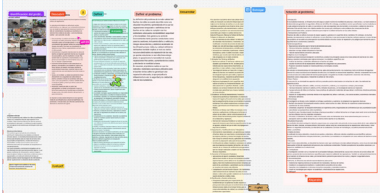
Los equipos trabajaron en la elaboración de los entregables utilizando ChatGPT, por lo que muchos de los resultados consistieron en listas o en contenido redactado en párrafos. La incorporación de estas listas dentro de los métodos de diseño parece haber complementado el tiempo destinado a seguir el proceso de diseño, ya que se completaron las etapas del instrumento metodológico en un periodo relativamente corto. Sin embargo, los resultados obtenidos deben considerarse de baja fidelidad, por lo que se recomienda realizar al menos una iteración adicional del proceso.

Instrumento 2: Registro del proceso metodológico por equipo

El instrumento de registro del proceso metodológico permitió visualizar los resultados de los participantes.

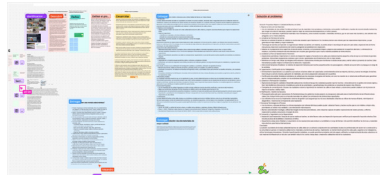
Tabla 6

Progresión del desarrollo del instrumento metodológico y notas, (de creación propia)

Equipo	Set	Progresión de instrumento metodológico	Notas
1	3		<div><div>Notas (verde)</div>Entre todos juntamos información rápidamente.</div> <div><div>Notas (rojo)</div>No logramos comunicarnos sobre que tiene que hacer cada quién.</div> <div>Notas (violeta)</div>

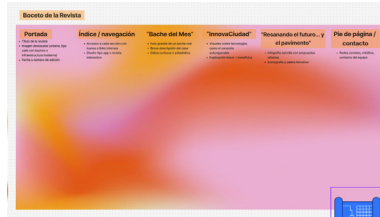
1

5



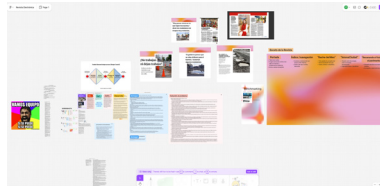
1

6



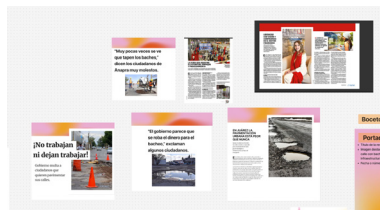
1

7



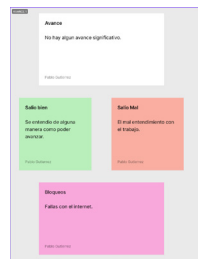
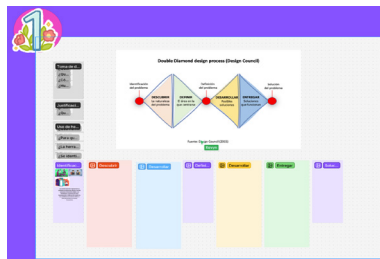
1

8



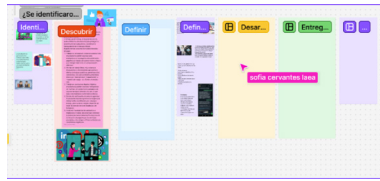
2

3



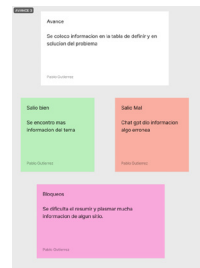
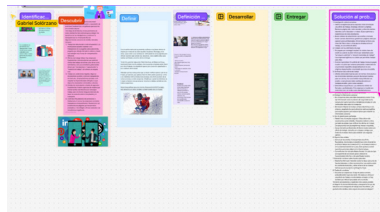
2

4



2

5



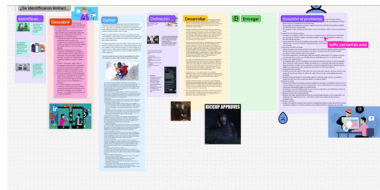
2

6



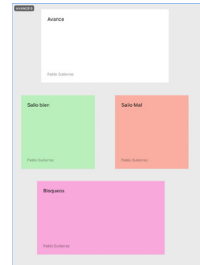
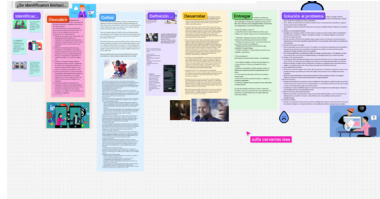
2

7



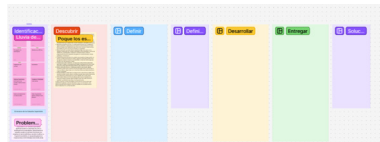
2

8



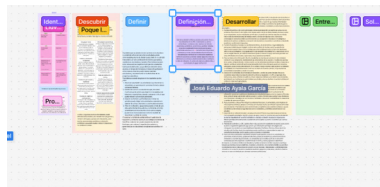
3

3



3

4



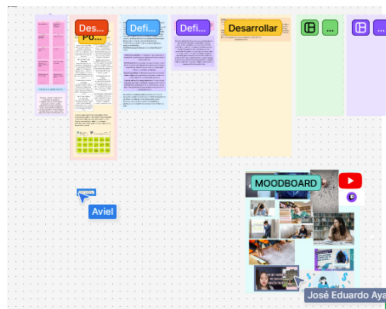
3

5



3

6



Esta padre el moodboard

2017/04

90% del esfuerzo del equipo se fue al moodboard pero nos quedó bien bonito

Jose Eduardo Ayala Garcia

moodboard

Jorgen hizo todo

Jose Eduardo Ayala Garcia

mucho detalle en el moodboard

Jose Eduardo Ayala Garcia

ya me pare

Jose Eduardo Ayala Garcia

Se están descomando poco puntos y soluciones para la revista o boceto más

Jorgen

Jose Eduardo Ayala Garcia

3

7



gran portada, increíble diseñador

Jorgen

Se me va a caer el cabello del estilo por lo

2017/04

Jorgen se le fue con los portales

Jose Eduardo Ayala Garcia

lolo hizo una revista muy bonita

Jose Eduardo Ayala Garcia

terminar esto lo lo está aaaaaa agrrr

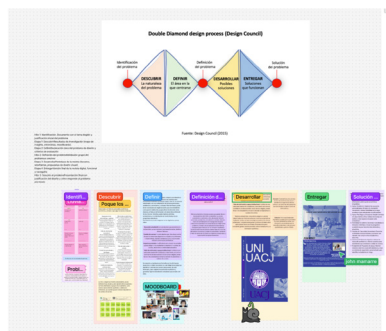
Jose Eduardo Ayala Garcia

Buen desarrollo de soluciones y la portada está padre jaj

Jorgen

3

8



La fase de entrega es dura

2017/04

Se está generando un entendimiento en la colaboración de la revista

Jorgen

re fue muy claro en la parte de entregar casi no se entendió el punto

Jose Eduardo Ayala Garcia

sumo se le fue con la sección de la revista

Jose Eduardo Ayala Garcia

Duda en entregar

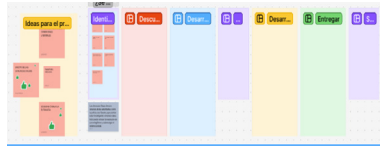
Jorgen

esta tomando muy buena forma

Jose Eduardo Ayala Garcia

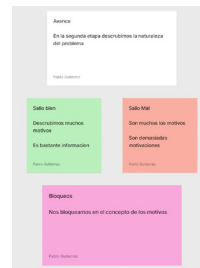
4

3



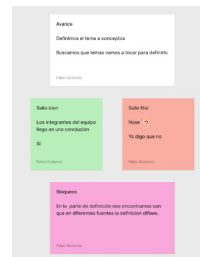
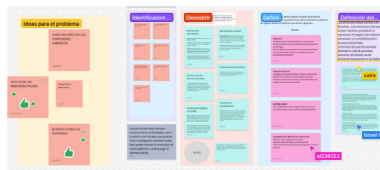
4

4



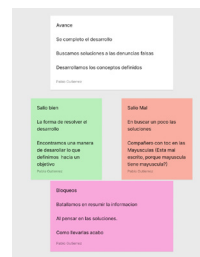
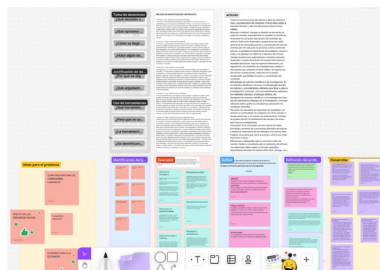
4

5



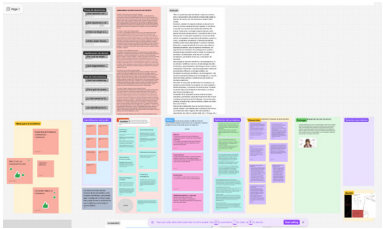
4

6



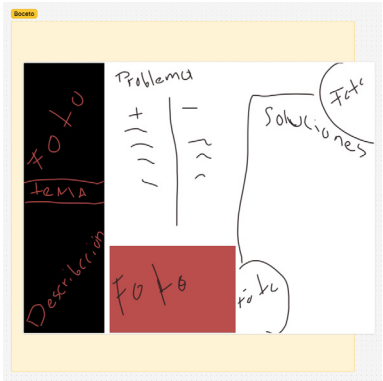
4

7



4

8



Resumen del registro del proceso metodológico. A continuación, en la Tabla 7 se presenta una síntesis del progreso observado en cada equipo, a partir de los instrumentos.

Tabla 7

Síntesis del progreso observado en cada equipo

Equipo	Síntesis del proceso metodológico
Equipo 1	Avanzó en todas las etapas del Doble Diamante. El registro evidencia una integración funcional de ChatGPT para la búsqueda de información y redacción de contenido. Las notas reflejan dificultades en la comunicación interna al inicio, pero se logró mayor coordinación en la segunda mitad. Se completaron los apartados de definición del problema y desarrollo del prototipo visual.

Equipo 2	Mantuvo un avance progresivo, aunque afectado por problemas de conectividad. Se observa mayor dependencia de la IA en la fase de desarrollo, pero un uso limitado de técnicas de diseño. Hubo dificultades para sintetizar información y mantener coherencia entre secciones. Completa las cuatro etapas, pero sin integrar visuales finales.
Equipo 3	Mostró un proceso más consistente, con uso activo de plantillas y aplicación del Doble Diamante. El equipo documentó cada etapa, evidenciando mejora en la integración del contenido y claridad en los objetivos. Se identifican bloqueos relacionados con la coordinación y los tiempos de entrega. Logró generar un prototipo visual con estructura de revista digital.
Equipo 4	Avanzó en las etapas iniciales con enfoque estructural. Se registran dificultades en la definición de entregables y en la coordinación de roles. Las observaciones indican confusión en la etapa de entrega y falta de concreción del resultado final. No se completaron las fases finales del proceso.

Los registros muestran una evolución en el uso y familiaridad con el modelo metodológico: los equipos pasaron de una comprensión escasa del proceso en la primera sesión a una aplicación más operativa en la segunda. El uso de ChatGPT fue general, aunque con diferentes niveles de integración por equipo. Su empleo favoreció la redacción y organización de ideas, pero en algunos casos limitó la profundidad analítica del contenido. Los principales bloqueos observados se relacionan con la coordinación grupal y la falta de claridad en la definición de los entregables. Los equipos que combinaron el uso de IA con métodos de diseño colaborativo (equipos 1 y 3) lograron mayores avances y coherencia en los resultados finales.

Instrumento 3: Cuestionario reflexivo

Los resultados del cuestionario reflexivo permitieron observar desde el conteo de las categorías mas utilizadas que tipo decisiones e

instrumentos utilizaron los estudiantes durante el desarrollo de su proceso. En la Figura 13 se muestra la cantidad de veces que cada categoría se eligió por cada etapa del proceso.

Figura 13
*Matriz correlacional de los cuatro equipos.
De creación propia*

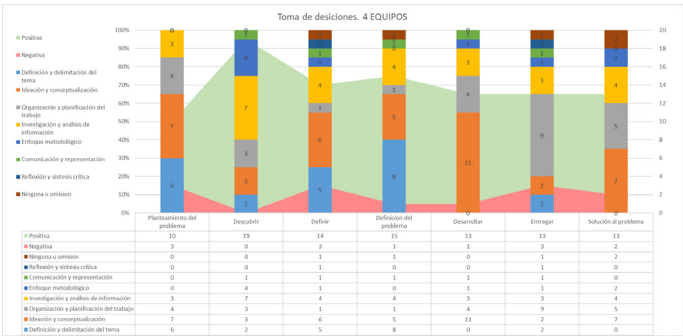
4 EQUIPOS									
Pregunta	Variable	Categoría	Puntuaciones del problema	Descubrió	Definir	Definición del problema	Descartar	Entregue	Sum
(¿Qué decisiones se tomaron en este momento del proceso?)	decision_focos	Definición y delimitación del tema	4	2	5	4	0	2	23
	decision_focos	Ideación y conceptualización	7	3	6	5	2	7	41
	decision_focos	Organización y planificación del trabajo	4	3	1	1	4	9	27
	decision_focos	Investigación y análisis de información	3	7	4	4	3	5	28
	decision_focos	Enfoque metodológico	0	4	1	0	1	1	9
	decision_focos	Comunicación y representación	0	1	1	1	1	1	5
	decision_focos	Reflexión y síntesis crítica	0	0	1	0	0	1	2
	decision_focos	Ninguna o errónea	0	0	1	1	0	1	5
	decision_method	Consenso y evaluación grupal	5	3	3	5	3	4	30
	decision_method	Identificación del problema	2	1	3	6	1	1	14
(¿Como elegiste que hacer para llegar a tu objetivo?)	selection_method	Criterios de selección	1	1	0	0	2	0	5
	selection_method	Proceso de investigación	5	9	8	0	9	4	43
	selection_method	Neurociencia y redes digitales	0	2	1	0	0	1	4
	selection_method	Comunicación y coordinación	2	4	4	2	4	8	29
	selection_method	Estrategia y reflexión	0	0	0	0	0	0	0
	selection_method	Ninguna o errónea	0	0	0	1	0	1	3
(¿Como sentiste el trabajo en equipo?)	framework_status_felt	Positiva	5	9	7	1	1	2	97
	framework_status_felt	Negativa	2	1	2	4	5	3	21
	framework_status_felt	Neutra	2	1	2	4	5	3	21
	what_vent_weell	Organización	2	0	2	3	4	1	12
	what_vent_weell	Comunicación	2	0	1	2	2	1	8
	what_vent_weell	Cantidad	3	1	3	3	0	1	14
	what_vent_weell	Decisión	6	3	0	2	1	1	16
	what_vent_weell	Colaboración	5	0	2	1	1	1	7
	what_vent_weell	Información	2	6	6	4	4	3	43
	what_vent_weell	Resultado	4	3	2	2	4	1	21
(¿Que salió bien?)	what_vent_weell	Reflexión	1	2	0	0	0	1	4
	what_vent_weell	Tecnología	0	1	0	1	0	0	2
	what_vent_weell	Sin aperturas	0	0	0	0	0	0	3
	what_vent_bad	Organización	0	0	0	0	0	0	24
	what_vent_bad	Comunicación	0	0	0	0	0	0	17
	what_vent_bad	Cantidad	0	0	0	0	0	0	24
	what_vent_bad	Decisión	0	0	0	0	0	0	8
	what_vent_bad	Colaboración	0	0	0	0	0	0	14
	what_vent_bad	Información	0	0	0	0	0	0	19
	what_vent_bad	Resultado	0	0	0	0	0	0	5
(¿Que salió mal?)	what_vent_bad	Reflexión	0	0	0	0	0	0	0
	what_vent_bad	Tecnología	0	0	0	0	0	0	0
	what_vent_bad	Sin problemas	2	2	2	2	2	2	24
	free_comment	Comentario optimista o de satisfacción	1	1	3	1	1	1	17
	free_comment	Crítica o descontento	3	4	1	2	4	1	20
	free_comment	Sin aperturas	0	0	0	0	0	0	68
	free_comment	Comentario analítico o de aprendizaje	1	1	1	1	1	1	7
	free_comment	Observación sobre aspectos técnicos o de infraestructura	0	1	1	0	1	0	4
	free_comment		0	0	0	0	0	0	0
	free_comment		0	0	0	0	0	0	0

La figura muestra, por ejemplo, que en cuanto a toma de decisiones, los participantes optaron por ejercicios de *Ideación y conceptualización* a través del desarrollo del ejercicio, con 41 selecciones, como se puede apreciar en la columna *SUM*. En cuanto al método más elegido para llegar a los objetivos es un *Proceso de investigación*, con 43 selecciones. De igual manera, el estado de ánimo de los participantes a través del proceso, fue percibido como *Positivo*, con 97 selecciones. Lo que mas consideran los estudiantes que salió bien fue la categoría de

Información, y lo que mas salió mal, fue la *Organización*. Por ultimo, la mayoría optó por no hacer comentarios, con *Comentarios optimistas o de satisfaccion*, como segundo lugar, con 37 selecciones.

Figura 14

Gráfica de toma de decisiones. 4 equipos



La Figura 14 muestra el estado de animo del equipo durante la toma de desiciones a traves de las faces del doble diamante. Mostrando un estado de animo del equipo percibido positivo a traves del ejercicio. Este formato de gráfica se vuelve relevante cuando se compara con los reultados obtenidos por el equipo numero dos, que tuvo problemas desde de conectividad y organización desde el inicio y que fue el unico que no entregó una propuesta visual. En la Figura 15 se puede apreciar como al final del ejercicio, el estado de animo del equipo comienza a decaer. En la gráfica tambien se aprecia que en la fase de entrega, se deidecaron esfuerzos a la *Organización y planificación del trabajo*.

Figura 15

Gráfica de toma de decisiones. Equipo 2

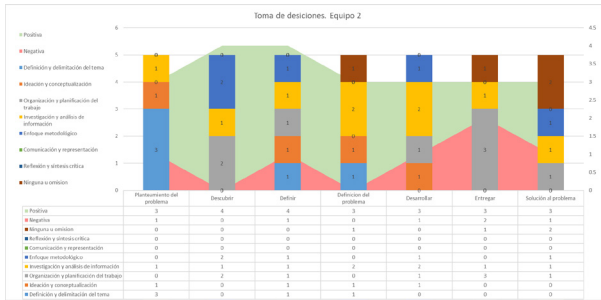
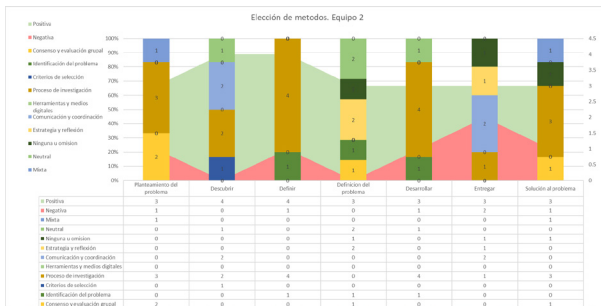


Figura 16

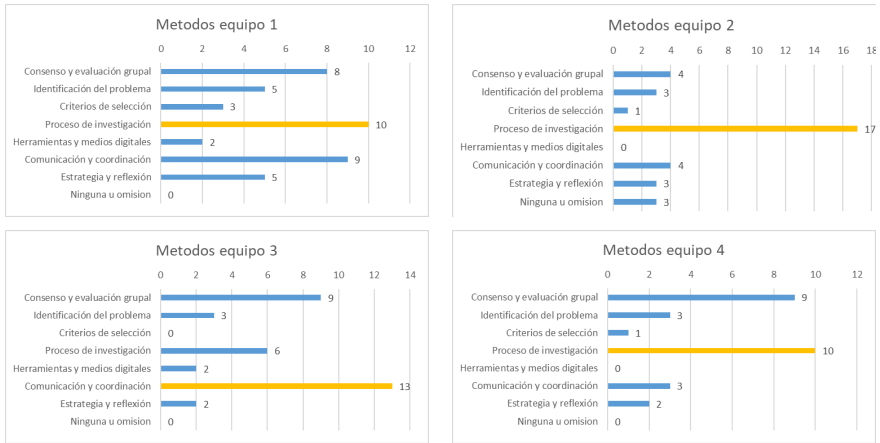
Gráfica de elección de métodos. Equipo 2



Si se toma en cuenta que el equipo 2, fue el que menos *Consenso* y *evaluación* grupal llevó a cabo, como se puede apreciar en la Figura 16 y la Figura 17; también, si se considera con la información generada por las tablas cronológicas, donde en la Tabla 5 se menciona que sus miembros tienen problema con el internet, se utiliza ChatGPT de manera moderada y la integración de métodos es parcial. Es posible

suponer que uno de estos elementos pudo haber tenido un efecto en su falta de resultados.

Figura 17
Grupos de gráficas de metodos por equipo



Si se compara con los resultados obtenidos por los equipos 1 y 3 (Figura 18 y Figura 19), donde hubo *Consenso y evaluación grupal* en varias etapas del modelo y quienes tuvieron un uso elevado de ChatGPT y de técnicas y métodos de diseño, como se muestra en la Tabla 8, se puede apreciar como estos parámetros pueden afectar positiva o negativamente la experiencia y desarrollo de un proceso de diseño.

Figura 18

Gráfica de elección de métodos. Equipo 1

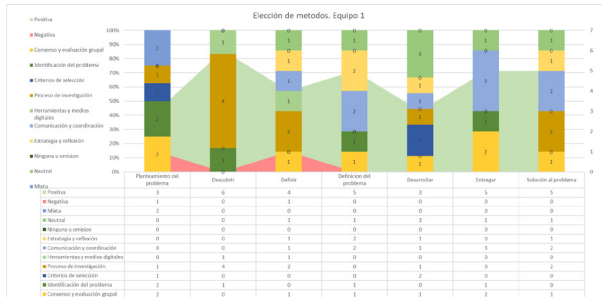
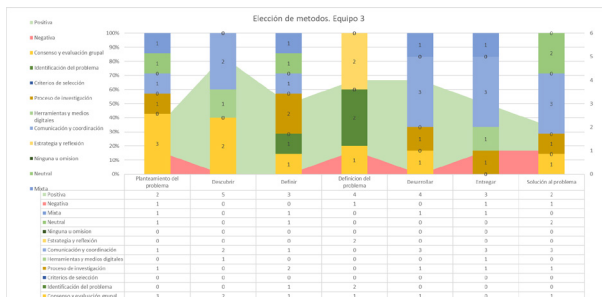


Figura 19

Gráfica de elección de métodos. Equipo 3



El cuestionario reflexivo se aplicó al término de la segunda sesión. Su objetivo fue registrar las razones detrás de las decisiones tomadas durante las distintas etapas del proceso de diseño, con base en el modelo del Doble Diamante.

Tabla 9

Síntesis del cuestionario reflexivo por equipo

Variable	Equipo 1	Equipo 2	Equipo 3	Equipo 4
Asistencia y participación	Equipo completo y activo. Coordinación progresiva entre miembros.	Cinco miembros; dos con intermitencia por conectividad.	Cinco miembros activos. Coordinación estable.	Cinco miembros activos. Participación constante.
Enfoque del trabajo	Organización basada en plantillas. Priorización de estructura y coherencia del contenido.	Acuerdos informales. Falta de coherencia entre secciones.	Enfoque analítico orientado a una revista digital.	Preferencia por métodos manuales y trabajo estructural.
Uso de ChatGPT	Integración constante para redactar, ordenar y clasificar ideas.	Uso limitado, centrado en redacción.	Uso intensivo para prompts, redacción y referencias.	Uso mínimo o nulo; enfoque tradicional.
Aplicación de metodologías de diseño	Aplicación del Doble Diamante y técnicas de Design Thinking.	Aplicación parcial del modelo metodológico.	Uso activo de plantillas y técnicas del Doble Diamante.	Uso limitado de técnicas; sin concreción final.
Bloqueos técnicos y organizacionales	Dificultades iniciales de comunicación, superadas con el avance.	Problemas persistentes de conexión y falta de claridad en objetivos.	Dificultad para cumplir con los tiempos de entrega.	Dudas sobre entregables y confusión en roles.
Avance metodológico	Cumplimiento de todas las etapas y desarrollo de prototipo visual.	Avance parcial sin integración visual.	Cumplimiento de todas las fases con resultados funcionales.	Avance limitado a las etapas iniciales.
Resultados parciales y observables	Propuesta coherente con integración de IA en redacción y organización.	Producto final incompleto y sin cohesión conceptual.	Prototipo de revista digital con integración de IA y diseño colaborativo.	Estructura general definida sin resultados visuales finales.

Discusión y conclusiones

Después de revisar de manera breve los resultados del ejercicio de observación, es posible considerar que la búsqueda de consenso y evaluación grupal constantes, junto con el uso dirigido de herramientas de IAG y la aplicación de técnicas de diseño, favorecen la generación de entornos de trabajo positivos y eficientes. No obstante, los problemas de organización y conectividad afectan la calidad de los resultados. Se recomienda una iteración adicional con ajustes a los instrumentos y sesiones breves de entrenamiento metodológico.

Referencias

- Design Council. (2007). 11 lessons: a study of the design process. <https://www.designcouncil.org.uk/our-resources/archive/reports-resources/11-lessons-managing-design-global-brands/>
- Design Council. (2014). The Double Diamond - Design Council. <https://www.designcouncil.org.uk/our-resources/the-double-diamond/>
- Gutiérrez-Gutiérrez, P. D. (2025). *Mapa de proceso interpretado V4 GPT (version 1)*. <https://docs.google.com/spreadsheets/d/1Wr-qnp4pdR4jioujioHQJJIVn1lVK-kS8/edit?usp=sharing&ouid=109942746572437883898&rtpof=true&sd=true>
- OpenAI. (2024). *ChatGPT*. <https://chatgpt.com/>
- Papachristos, E., Inal, Y., Monlla, C. V., Arnstein, E., & Hermansen, M. (2024). *Integrating AI into Design Ideation: assessing ChatGPT's role in Human-Centered Design Education*.

Pérez, R. A. (2005). *Estrategias de comunicación*. Ariel Comunicación.

Scheinsohn, D. (2010). *Comunicación Estratégica*. Granica

Vaccaro, M., Almaatouq, A., & Malone, T. (2024). When combinations of humans and AI are useful: A systematic review and meta-analysis. *Nat Hum Behav*, 8, 2293-2303. <https://doi.org/10.1038/s41562-024-02024-1>

AUTORIA INACABADA(?) REFLEXÕES E APONTAMENTOS ACERCA DA ATRAÇÃO NO FILME “CARTAS TELEPÁTICAS” DE EDGAR PÊRA

Aloísio Corrêa de Araújo¹
Vicente Gosciola²
Victor de Souza Crispim³

Trabalhamos dois filmes no presente artigo, *Cartas Telepáticas*, *Não Sou Nada - The nothingness Clube*, e algumas imagens que passam pelo *dispositivo* do TikTok, porém, o escopo deste artigo é a primeira obra citada. Os dois filmes geram uma inquietação perante ao conflito

-
1. Doutorando pela Universidade Anhembi Morumbi (PPGCOM-UAM). Educador de ling. audiovisual no Instituto Criar de TV, Cinema e Novas Mídias. araujodca@gmail.com.
 2. Doutor em Comunicação pela PUC-SP. Professor do Programa de Pós-Grad. em Com. da Universidade Anhembi Morumbi. vicente.gosciola@gmail.com
 3. Mestrando pela Universidade Anhembi Morumbi (PPGCOM-UAM). Educador de pós produção no Instituto Criar de TV, Cinema e Novas Mídias. victorscrispim@gmail.com

entre a autoria e a heteronomia, sobretudo no cinema e no audiovisual contemporâneo. Tal inquietação, presente na obra de Fernando Pessoa é observada como um ponto de inflexão que representa a inquietação da humanidade do século XIX ao XXI, seja consciente ou inconsciente narrativo. De obras ficcionais como *O médico e o monstro*, de Robert Louis Steveson aos perfis digitais fragmentados em muitas personalidades das redes sociais, a maior dificuldade em identificar a autoria é compreender o que circunda seus limites.

O objetivo da análise propõe relacionar, como exercício científico e ensaístico o uso do que se entende por autoria, heterônimo, atração e narração. A presença de autores, perfis, heterônimos, nas duas obras e seus paralelos com a plataforma do *TikTok*, como flexão e exercício de observação, onde embora sejam produtos ímpares, há algo da indústria de massa capaz de conectá-los por meio da discussão recorrente da (im)possibilidade da autoria, provocada pela massificação, e desta reflexão/fabulação podemos compreender melhor a genialidade do que foi a invenção do heterônimo.

Ao assinar como Álvaro de Campos em *Tabacaria*, o existencialismo pessimista dotado de uma melancolia própria, com estilo e percepções sobre o mundo, através do texto escrito ou digitado, vê-se a suposta existência de um outro ser, simulação de outra pessoa ou o rastro de existência de outro dentro do corpo de Pessoa, ou ortônimo.

Pessoa é contemporâneo ao período futurista, do rádio, do cinema e muitos séculos após a invenção da prensa. Uma fase de desenvolvimento da massificação pelas tecnologias da comunicação e do discurso de massa, de tal forma que ainda não se poderia observar claramente a possibilidade da existência de um heterônimo ou do que viria a ser

a sua existência no contemporâneo. Pseudônimos, cartas escritas por personagens e todo tipo de jogos que trabalham no âmbito da ficção, do jogo e do *fake* já poderiam existir, naturalmente, no campo da escrita, e sempre foram utilizados como recurso narrativo.

Portanto, ao falar genialidade, me refiro ao movimento linguístico, pós-humano e transcendental, provocado pela criação do heterônimo, o proto-usuário, ou o início do perfil digitado em uma era décadas antes da invenção do primeiro computador, muito antes da invenção da internet ou de plataformas digitais. Fernando Pessoa conseguiria prever a criação de outras formas de autoria em um ambiente do futurismo, que levaria a um futuro onde a humanidade deseja se mecanizar, sob o contra-peso da psicanálise e da psiquiatria estruturante.

Para compreender o heterônimo é importante rejeitar ou separar termos psiquiátricos e médicos que atrofiam um desenvolvimento da mente que nos tempos de redes sociais se torna uma máxima.

É importante separar também a capacidade literária e a qualidade dos trabalhos de Pessoa e seus heterônimos da possibilidade de existir um heterônimo. Tais características de valor independem da existência, visto que o heterônimo não é único de seu ortônimo primordial, mas uma forma possível de desenvolvimento social que extrapola a autoria. Em uma pessoa podem existir 20 pessoas que escreveram vinte livros, diria Pessoa, ou seu heterônimo, nos levando a questionar, assim como no artigo de Bruno Ministro: “Mas será mesmo que se trata de um jogo de ortônimo / heterônimos? Porquê o ortônimo ao centro e os restantes atrás? Não podia ser ao contrário?” (Ministro, 2024, p. 105).

Esta introdução é o mote aglutinador da ideia deste artigo, assim como um caminho maior para compreender os dois filmes de Edgar

Pêra. Nessas produções, Pêra mexe na heteronímia pessoana e dialoga com as transformações da autoria na era digital, explorando conceitos de identidade, subjetividade e virtualidade. Edgar Pêra trabalha a obra pessoana com originalidade, evitando reprodução mimética, em um gesto artístico que transcende simples homenagem ou reprodução direta:

Em todos os projetos, o autor aborda a obra de outros evitando a reprodução mimética, procurando mergulhar nas peças literárias com um olhar que tem tanto de questionador e aberto à diferença, como de antropofágico, já que produz, invariavelmente, visões que ampliam as suas próprias reflexões. Fernando Pessoa é, sem dúvida, o autor que mais estruturalmente influenciou Edgar Pêra. (Lima, 2024, p. 73)

Não há dúvidas nos filmes e em boa parte da filmografia de Pêra, de seus trabalhos fotográficos, musicais, de desenho de som e misturas, dos cadernos, como observa Teresa Lima que é possível ver os traços do desenvolvimento de *Não Sou Nada* nos cadernos em que Edgar Pêra trabalha ao longo dos anos, desde sua adolescência. Todo seu trabalho é salpicado de existencialismos, ironias, pseudônimos ou heterônimos, como nos trabalhos de mixagem de Cyaneto, que é pseudônimo/heterônimo de Edgar Pêra, nos dois filmes que serão analisados a seguir. Destas paráfrases, referências, citações e distorções, já vemos um primeiro aspecto da arte que é a dissolução da autoria numa outra obra, onde não se busca plagiar, mas praticar a cópia/variação, condição artística da inteligência humana e do atual estado da inteligência artificial (Ministro, 2024, p. 108).

À seguir, veremos alguns dos recursos estilísticos que aproveitaram um legado e transpõem a barreira da autoria, do personagem e do heterônimo.

The Nothingness Club - Não sou nada (?)

Em *The Nothingness Club*, filme de ficção com personagens humanos, Edgar Pêra explora os conflitos existenciais e a relação amorosa de Fernando Pessoa com Ofélia Queiroz.

Se passa em um escritório onde o escritor trabalha junto com seus heterônimos com escritos diversos, como se fosse uma redação jornalística. Dotado de textos, poemas e trechos de poemas e contos dos Heterônimos e de Fernando Pessoa, a obra desenvolve conflitos entre os personagens, desavenças assassinos, o relacionamento dos heterônimos com Ofélia e as idas de Fernando Pessoa ao hospital psiquiátrico. É um filme de atmosfera definida pela mídia portuguesa como *neonoir* (Lima, 2024).

Contudo, a obra também apresenta elementos de sátira, de suspense e temas psicológicos. No filme, vemos primeiro Pessoa como o chefe da redação. Este movimento de ser o chefe aos poucos é sequestrado pela figura de Ricardo Reis, um assassino a matar seus outros colegas heterônimos. É uma trama que envolve a poética e brinca com a ala psiquiátrica. É interessante observar o movimento dos personagens maiores a tomar o campo do protagonismo e questionar a legitimidade da personalidade, da *persona*, da autoria por meio de uma narrativa ficcionalizada. Um movimento que questiona a existência dos corpos, que:

“As leituras dos cadernos encaminham-nos, igualmente, na convicção de que *Não Sou Nada* é uma passagem concentrada de um estudo obsessivo e aprofundado da obra pessoana com reverberações temporais justapostas, que desaguam de forma ora direta ora subliminar num filme” (Lima, 2024, p. 74)

Cartas Telepáticas

O suposto longa-metragem inacabado *Cartas Telepáticas*, introduz uma dimensão mais radical da experimentação autoral, ao estabelecer diálogos fictícios mediado pela inteligência artificial entre Fernando Pessoa e H. P. Lovecraft.

O uso dessas novas tecnologias perpassa modos de autoria novos, questionando as fronteiras tradicionais entre realidade, ficção e criação artística. Tal abordagem encontra ressonância nas discussões contemporâneas sobre inteligência artificial, identidade e fragmentação, abrindo novos campos conceituais para refletir sobre a autoria na era digital.

A expansão da heteronímia?

A heteronímia pessoana tenta se dimensionar no cinema de Edgar Pêra, manifestando-se em personagens digitais e avatares criados pelo cineasta, por meio de personagens, pela criação distribuída em dois nomes (Pêra e Cyaneto). Desde cedo, Pêra questiona a natureza da identidade, algo evidente em suas próprias anotações pessoais:

Começemos por traçar o fio destas linhas identitárias, com os primeiros anos de registos. A 6 de janeiro de 1978, Edgar Pêra questiona-se: ‘Será possível viver noutra realidade que não esta?’ Os universos paralelos e o tempo são tópicos que passaram a ocupar muitos dos filmes do realizador (exemplo: *Manual de Evasão Lx 94*), que em 1978 ainda não realizava filmes, mas já se apresentava dúvidas que explorou ciclicamente ao longo dos anos e em variados materiais. ‘Eu sou estas páginas ou cada página é um eu?’. (Lima, 2024, p. 74)

O questionamento da ordem ortônimo-heterônimo já é uma constante no trabalho de Pêra, e este reflete-se como a grande dúvida

dos tempos de hoje: Qual é a ordem a seguir? A inversão não é apenas filosófica, mas estrutural.

TikTok e o Ego

A heteronímia encontra um paralelo contemporâneo explícito na estrutura do *TikTok*, plataforma onde os usuários constroem múltiplos perfis e personagens, frequentemente efêmeros, sustentados por tendências virais. Esta plataforma evidencia a fragilidade egóica das novas gerações, cujas identidades digitais se tornam dependentes de constantes reafirmações externas, amplificadas pela necessidade de aprovação social através das *trends*, ou a busca neurótica pela moda e popularidade, como um jogo de azar.

A fragilidade e efemeridade dessas identidades digitais contrasta ao passo que guia-sintoniza com a densidade da heteronímia pessoana e com a intensidade da pesquisa sobre identidades e autorias sugerida nos filmes de Edgar Pêra. Esta é a possibilidade de triangulação crítica entre essas diferentes formas de imagem, que embora sejam mediadas por estruturas e dispositivos ímpares, assim como o papel e a caneta de Fernando Pessoa, jogam dentro de uma mesma estrutura tecnológica intelectual (ou espiritual, do campo da consciência) que é a autoria (ortônimo) e o heterônimo.

A autoria impossível de *Cartas Telepáticas*

A reflexão sobre a heteronímia pessoana expandida nos filmes de Edgar Pêra permite considerar como as mídias contemporâneas reforçam práticas de autoria fragmentada e múltipla. A observação proposta

ressalta como a heteronímia, originalmente literária, manifesta-se hoje amplamente pelas mídias digitais, sugerindo novas concepções de autoria e identidade. Essencialmente, heterônimo significa “Outro nome”.

Essa investigação também confronta o conceito tradicional de identidade autoral com a fragilidade das identidades digitais, trazendo à tona questões relacionadas à loucura e à multiplicidade, temas constantes nas reflexões pessoais do cineasta Edgar Pêra: “Evitemos os psicologismos, contornando-os, para tropeçar noutra temática que atravessa estes cadernos: a loucura. Trata-se de um conceito que povoa os registos de Edgar Pêra desde as primeiras anotações e, não surpreendentemente, muitas das anotações do *Não Sou Nada*. Em dezembro de 2019, registra:

Apenas sigo a sua loucura e os seus sonhos’. A 20 de janeiro de 2020, continua neste universo: ‘Ao ler Pessoa escrever sobre a sua loucura, o seu medo da loucura, a sua definição da sua própria loucura, começo também a tentar entender melhor a minha loucura. Procurar uma typologia’. (Lima, 2024, p. 74)

Em *Cartas Telepáticas*, as imagens apresentam uma estrutura visual e sonora caracterizada pela técnica da colagem. São pedaços de imagens digitais-fabulações, *sonhadas-processadas* pela inteligência artificial, sempre incompletas, malformadas pela proposta e pelo erro, sendo qualquer imagem produzida pela inteligência, em consonância com a humanidade, passíveis do erro. O oposto, o termo complexo da perfeição, talvez só seja possível, por enquanto, quando uma ferramenta alcança, em paralelo com seu usuário, um afinamento de ajustes que só poderia ser alcançado através de uma lapidação de erros capaz de se renegar completamente e alcançar uma estrutura de humana perfeição. Uma máquina já não tem tal possibilidade, porque a execução de uma

tarefa é tanto o acerto como o erro inconsciente, uma característica irreparável do código, o eterno binômio sim-não.

Argumenta-se que a força expressiva da montagem de *Cartas Telepáticas* está centrada principalmente na construção de atmosferas sensoriais, e não na narrativa tradicional.

Em *Cartas Telepáticas*, a narrativa condutora, sua camada mais consciente e tradicionalmente autoral, quase estruturada num documentário ou um filme-carta, (ou um filme-jogo?) é estruturada fundamentalmente através do som, mais especificamente dos diálogos e da trilha musical.

Filme Inacabado

Esses elementos são capazes de proporcionar ao espectador uma experiência sensorial que evoca um diálogo de cunho lovecraftiano com Fernando Pessoa. Embora classificado como longa-metragem, o filme exhibe limitações evidentes em sua estrutura imagético-narrativa, privilegiando a sucessão rápida de planos curtos intercalados com planos frontais que destacam exclusivamente as bocas dos personagens.

Tal técnica remete a métodos de animação clássica, observados, por exemplo, em desenhos animados, tais como *Bob Esponja*, onde se pode ver a mesma técnica aplicada à abertura do desenho animado, contudo, utilizando colagem ou montagem com fundo do tipo *chroma*.

A inovação introduzida pela Inteligência Artificial (IA) reside precisamente na criação dessas imagens distorcidas e rudimentares, aludindo à tradição das atrações cinematográficas, onde veríamos imagens com situações intrigantes, misteriosas, com efeitos espantosos.

Nas palavras de Tom Gunning, cedidas em entrevista a definição de atração se torna clara:

Eu tinha assistido sistematicamente aos filmes de Griffith de 1908 a 1913. Agora, assistir a filmes anteriores foi uma grande revelação. Eu percebi: “Ok, esses filmes são tão diferentes do que Griffith fez apenas alguns anos depois. [...] Era realmente impressionante constatar que esses filmes não contavam histórias exatamente; eles não necessariamente desenvolviam a continuidade que Griffith impôs e sistematizou de 1908 em diante. Então comecei a pensar: “Ok, que diferença é essa?” Aí veio a ideia de atração. (Monteiro et al., 2024, p. 169)

Contudo, esta sucessão ininterrupta de atrações distingue-se significativamente das imagens de atração já conhecidas dos formatos virais ou humorísticos populares tanto no cinema quanto em plataformas digitais contemporâneas, como em vídeos curtos no Tik Tok, bem como dos filmes clássicos do cinema silencioso, tais como *O Regador Regado* (1895), dos irmãos Lumière, cuja narrativa linear com efeito de atração é claramente definida.

Em contraste, *Cartas Telepáticas* apresenta uma ausência proposital dessa linearidade, caracterizando-se por uma série contínua de atmosferas e atrações inacabadas, articuladas por uma narrativa indireta inspirada pela troca ficcional de correspondências entre Fernando Pessoa e H. P. Lovecraft. Essa atmosfera é reforçada por uma trilha sonora minimalista e de baixas frequências, que gera uma sensação sutilmente perturbadora.

A hipótese defendida neste estudo é de que a valorização da atmosfera justifica a classificação do filme como longa-metragem, ainda que careça de uma estrutura narrativa tradicional clara, no âmbito da

capacidade das imagens de contar uma história, separando-a das ferramentas tradicionais da feitura de um filme, como a *voz off* que conta a história.

Trata-se essencialmente de uma colagem audiovisual criada por I.A. o que contrasta explicitamente com *Não Sou Nada*. Este utiliza uma abordagem cinematográfica convencional, caracterizada por personagens bem definidos, locações reais, iluminação detalhada e uma construção narrativa tradicional, centrada em pontos dramáticos claramente delineados, onde Fernando Pessoa e seus heterônimos são as figuras centrais.

Nesse contexto, o uso intensivo da IA em *Cartas Telepáticas* levanta paralelos interessantes com as produções curtas de animação que circulam no *TikTok*, especialmente uma grande quantidade de animações envolvendo breves narrativas de gatos antropomorfizados e animados, produzidos por inteligência artificial. Tais vídeos geralmente carecem de diálogos estruturados, apresentando apenas trilhas musicais e efeitos sonoros pontuais. Uma característica comum desta etapa destes geradores de imagens em movimento é a ausência de um ponto preciso de atração ou clímax narrativo claro; frequentemente, observa-se ou um excesso contínuo sem resolução (por exemplo, explosões repetitivas), ou uma incapacidade de alcançar um clímax narrativo satisfatório.

Essas curtas animações produzidas por usuários que operam a inteligência artificial, costumam retratar breves histórias de relações afetivas entre gatos adultos e filhotes, repetidas continuamente através de cenas sequenciais. Essa repetição visual constante faz um paradoxo à estrutura narrativo-imagética dos quadros sequenciais da *Via Crúcis*, presentes na tradição católica, criando uma analogia visual e estrutural

com esses conteúdos virais de animação no TikTok. A *Via Crúcis* revela histórias em quadros fixos. Os vídeos de animação produzidos por inteligência artificial tem movimentação e alguma relação com o *tableau vivant*, contudo, sem narrativa, gerando, assim, uma sensação ambígua: possuem narrativa sem uma clara estrutura narrativa, têm movimento sem apresentar movimento plenamente realizado. Essa ambiguidade gera um desconforto, especialmente porque tais vídeos, geralmente curtos, aqui são editados em uma narrativa extensa de mais de uma hora, sustentada por diálogos, *voice over*, e uma trilha sonora constante que conduz o espectador a um estado de transe.

O conceito de transe aqui é entendido como uma estrutura repetitiva. Essa repetição visual que ocorre em *Cartas Telepáticas*, de situações de tensão, porém entre clímax, de atmosferas repetitivas, remete a imagens reconhecíveis, mas que não se apoiam em uma base imagética concreta devido ao seu método de produção por inteligência artificial. Existe uma evocação conceitual de Fernando Pessoa e do universo estético de H. P. Lovecraft, assim como uma atmosfera típica do cinema noir ou neo-noir, frequentemente mencionada por críticos e meios de comunicação. Tal atmosfera é incompleta visualmente, mas o transe provocado é completo, seja pela condução do som, seja pelas imagens *sem começo e sem fim*, no âmbito de atrações com um ponto de excitação do espectador.

Ritmo Telepático

Este estado de transe resulta essencialmente do ritmo, que, por sua vez, guarda uma relação semântica com o rito e, conseqüentemente, com o ritual. O ritual repetitivo gera a atmosfera específica do filme,

caracterizada pelo esvaziamento da consciência visual, restando como única consciência perceptível aquela proporcionada pelo som. O aspecto sonoro torna-se, portanto, o componente autoral central, obra de Edgar Pêra. Embora a montagem também seja realizada por Edgar Pêra, as imagens são geradas por softwares de inteligência artificial. Esse mesmo fenômeno de ambiguidade narrativa e movimento incompleto pode ser observado em micronarrativas comuns em plataformas digitais, como as frequentes animações de gatos presentes no *TikTok*.

Essa imagem é composta por ações desprovidas de narrativa, sem *punchline* ou estrutura emocional de clímax, tanto por limitações da técnica de produção, quanto da falta de curva narrativa mínima para o efeito de atração; ou então, quando apresenta uma estrutura emocional de clímax, esta é tão exagerada que acaba esvaziando a imagem de seu conteúdo.

Tanto no *Cartas Telepáticas*, quanto nos vídeos curtos de gatinhos e nas pequenas narrativas que circulam no *TikTok*, o desenho de som e a montagem tornam-se essenciais, pois são as únicas camadas, nesta etapa da I.A., conscientemente estruturadas por artistas humanos. A imagem em si, quando criada, é vazia, desprovida de narrativa e saturada de inconsciência de construção de narrativa através de uma imagem em movimento. É como se fosse o avesso de uma imagem hermética—não no sentido de ser incompreensível, mas por estar completamente esvaziada, formada integralmente pelo inconsciente, portanto, totalmente aberta a ponto de seu significado, por mais pictórico que pareça, tenda ao mesmo tempo à abstração.

Existe uma trilogia cinematográfica composta por um autor que percorre o mundo captando fenômenos culturais: *Baraka*(1992),

Koyaanisqatsi (1983) e *Samsara* (2011). São filmes nos quais vemos imagens sustentadas pela trilha sonora minimalista de Philip Glass, retratando cenas culturais, pitorescas, rituais curiosos, danças e manifestações religiosas de todo o globo terrestre.

Essas imagens, frequentemente gravadas em câmera lenta ou acelerada, perdem seu significado narrativo convencional, adquirindo, em vez disso, um sentido hipnótico ou de transe. Não se trata da atração instantânea provocada pelo humor ou pela surpresa, mas sim de uma atração prolongada, ritualística, que gera ritmo e leva o espectador a acessar camadas mais profundas do inconsciente humano.

Diferentemente das imagens produzidas por inteligência artificial, que são vazias de significado cultural genuíno por serem geradas exclusivamente por computador. Embora imagens geradas artificialmente possam eventualmente adquirir interpretações próprias pelo espectador, ou conforme a montagem, elas permanecem essencialmente destituídas da cultura autêntica. Edgar Pêra tenta explorar essa fronteira ao criar uma escrita visual improvisada, conectando-se a figuras como Fernando Pessoa e H. P. Lovecraft, mas suas imagens frequentemente resultam em aproximações parciais e incompletas. Em consequência, apresenta-se uma ideia esvaziada de Fernando Pessoa e Lovecraft, num movimento contrário à profundidade do heterônimo de Pessoa. Enquanto o heterônimo revela a profundidade real do poeta através de uma persona distinta dele, as imagens geradas por computador oferecem apenas uma ideia superficial e esvaziada do autor, pois são destituídas de qualquer estrutura consciente, salvo o *input* feito por Pêra. Assim, essas imagens brincam com a realidade, mas são essencialmente falsas, criadas integralmente pelo computador e totalmente desprovidas de consciência.

Considerações finais

Tanto as produções do *TikTok* quanto as do cinema são baseadas em dispositivos de poder através dos quais é possível gerar expressão, e através dos múltiplos perfis de usuários, há a possibilidade da existência do heterônimo dentro do aplicativo, sendo que uma pessoa poderia obter um perfil com produção de imagens pessoais, um perfil para trabalho e um outro voltado para jogos. Ou um artista poderia controlar perfis de *TikTok* diferentes e assim construir uma ideia simulada de heterônimo, mas esta denominação se torna incerta e dependeria de diversos fatores, como a forma da produção de imagens, alguma forma consistência autoral, autodeclaração de heteronomia ou qualquer forma de separação. No caso de perfis criados com o objetivo de inserir vídeos feitos com inteligência artificial, há um paradigma, que é a própria produção das imagens.

Quando consideramos a estrutura de produção de imagens de uma inteligência artificial, percebemos que a máquina funciona estritamente conforme o *prompt* que recebe. Por exemplo, se for solicitado à máquina: “crie uma imagem”, ela poderá gerar essa imagem em uma paisagem aleatório ou uma figura específica, mas ao pedir uma nova imagem com o mesmo *prompt*, normalmente será obtido apenas uma repetição da primeira imagem feita com pequenas variações. Em testes, foi pedido à inteligência artificial da *OpenAI* que produzisse um texto sem interferências ou sugestões iniciais, e geralmente o resultado é ou a impossibilidade de produzir o texto, ou então um texto extremamente simples e padronizado.

Essa limitação é a principal característica da inteligência artificial: sua incapacidade de possuir consciência própria. Portanto, a inteligência

artificial como dispositivo de comunicação não funciona realmente, pois não há uma troca genuína entre ser humano e máquina; há apenas a produção de um prompt e a consequente resposta da máquina. Isso sugere que, ou não há consciência alguma na inteligência artificial, ou o nível de consciência possível para uma máquina é tão limitado que se assemelha à consciência mineral. Surge, então, uma questão fundamental: qual seria, portanto, a verdadeira natureza e força do inconsciente da máquina?

Quanto mais os dispositivos se difundem e disseminam o seu poder em cada âmbito da vida, tanto mais o governo se encontra diante de um elemento inapreensível, que parece fugir de sua apreensão quanto mais docilmente a esta se submete [...] Ele não se deixará colocar corretamente se aqueles que dele se encarregam não estiverem em condições de intervir sobre os processos de subjetivação, assim como sobre os dispositivos, para levar à luz aquele Ingovernável, que é o início e, ao mesmo tempo, o ponto de fuga de toda política. (Agambem, pp. 51-52, 2009)

A imagem produzida por Inteligência Artificial não constitui propriamente um dispositivo, pois nela não há dinâmica de poder ou consciência. Sua produção é desprovida da capacidade de escolha autônoma, dependendo sempre da consciência humana para gerar qualquer tipo de conteúdo visual. Por consequência, essa produção acessória à consciência humana é incapaz de criar narrativas autorais ou transmitir sensações genuinamente humanas, limitando-se a produzir atmosferas. O significado efetivo é introduzido pela montagem e pelo desenho de som, que, nas obras atualmente observadas, continuam sendo em grande parte realizados por artistas humanos.

Talvez, futuramente, seja possível a produção integral de filmes pela Inteligência Artificial, abrangendo não apenas imagens, mas também montagem, desenho de som e até diálogos, incluindo créditos iniciais e finais. Entretanto, para que esses filmes sejam efetivamente conscientes—isto é, capazes de gerar significado e construir narrativas completas, com início, meio, fim e clímax—permanecem questões abertas, destinadas ao futuro.

O questionamento sobre a existência de Fernando Pessoa e seus heterônimos ou seus heterônimos e Fernando Pessoa, permanece pertinente nos dias atuais. Vivemos em um tempo em que são produzidos múltiplos perfis e imagens estéticas inócuas que, embora carreguem algum significado e certas características, carecem de corpo ou de uma base imagética concreta, como aquela gerada por uma câmera. As imagens produzidas por inteligência artificial não têm origem física real, exceto na eletricidade, no cálculo e no microchip, e são essencialmente desprovidas de ideias.

O heterônimo existe porque há uma consciência; e porque existem múltiplas consciências, existem também múltiplos heterônimos. Assim, a produção de *Cartas Telepáticas* emula a criação de um heterônimo a partir de *prompts* já estabelecidos e textos pré-existentes, alguns dos quais são criados pelo próprio autor. Em *Não Sou Nada*, assistimos à ficcionalização do mundo interno de Fernando Pessoa em interação com seus heterônimos, incluindo intrigas relacionadas a assassinatos e à relação com Ofélia. No filme, o heterônimo é representado quase como uma pessoa distinta, mas essencialmente o heterônimo é uma característica do virtual. O dispositivo pelo qual Fernando Pessoa manifestava seus heterônimos não era de ordem psiquiátrica ou uma dupla

personalidade, mas sim literário: a caneta, o papel e a própria criação literária. Para Jerónimo Pizarro, a heteronímia em Fernando Pessoa é uma manifestação da complexidade e da multiplicidade do ser, tanto no plano literário quanto no existencial.

Os filmes de Edgar Pêra não reproduzem exatamente esse fenômeno, mas oferecem uma representação ou uma ideia inconsciente, talvez mesmo uma estrutura de pastiche, de necro imagens que nos ajuda a compreender melhor a relevância e a manifestação contemporânea do conceito de autoria, esta prejudicada, escondida ou semi existente em heterônimos digitais.

Referências

Benjamin, W. (2018). *A obra de arte na era de sua reprodutibilidade técnica*. L&PM.

Eco, U. (1964). *Apocalípticos e integrados: comunicações de massa e teorias da cultura de massa*. Perspectiva.

Groys, B. (2016). *In the Flow*. Verso.

Lima, T. (2024). *Não Sou nada - Uma Análise das representações midiáticas*. Pessoa Plural.

McLuhan, M. (1962). *The Gutenberg Galaxy: The Making of Typographic Man*. University of Toronto Press.

Ministro, B. (2024). *Imaginar, especular, subverter (através de sonhos, visões, alucinações) - Fernando Pessoa, Edgar Pêra e Inteligência Artificial*. Pessoa Plural.

Monteiro, L. R., Gonçalo, P., Cruzeiro, V & Gunning, T. (2024). Atração, Formalismo e História do Cinema: Uma conversa com Tom Gunning. *Revista Aniki*.

Pizarro, J. (2023). *Ler Pessoa*. Tinta da China Brasil.

A ESTETIZAÇÃO ALGORÍTMICA DO AUDIOVISUAL: ULTRARREPRESENTAÇÃO E MODOS DE PRESENÇA EM VÍDEOS GERADOS POR IA

Rafael de Araújo Mélo¹
Claviano Nascimento de Sousa²
Diogo de Almeida Camelo³

Em maio de 2025, durante o evento *Google I/O*, o *Google* apresentou uma ferramenta de inteligência artificial (IA) que gera clipes em movimento que impressionam pelo nível de detalhes. Batizada de *Veo 3*, a IA permite criar vídeos curtos, de até oito segundos, integrando imagem, movimento e som, após o usuário fornecer algumas poucas

-
1. Doutorando do PPG em Estudos da Mídia (PPgEM) da UFRN. Bolsista Capes.
rafaelmelojornalista@gmail.com
 2. Doutorando do PPG em Estudos da Mídia (PPgEM) da UFRN. Bolsista Capes.
kkacademico@gmail.com
 3. Doutorando do PPG em Estudos de Mídia (PPgEM) da UFRN.
realdiogoalmeida@gmail.com

palavras em um comando de texto (*prompts*) indicando que tipo de vídeo quer produzir. Em poucas semanas, vídeos ultrarrealistas parodiando acontecimentos históricos ou passagens religiosas, programas de televisão, entrevistas em telejornais e *podcasts*, entre outros, se multiplicaram nas redes sociais e acumularam milhões de visualizações.

Partindo deste novo panorama em torno da imitação da vida e da arte, parece ser apropriado e oportuno contextualizar o conceito de mimese considerando uma nova modalidade de presença dos fenômenos estéticos criados pela inteligência artificial. Dada esta nova conjuntura de produção de sentidos, de conteúdo, de objetos sensíveis e de presença no mundo, o homem se vê diante de um desafio tão complexo quanto todos os não dissecados até o cenário atual.

Há uma revolução paradigmática que dispõe sobre a própria natureza humana, metamorfoseada pelo bit computacional, capturada e sequestrada pelo bios virtual, materializada sob outra propriedade alheia ao homem: uma inteligência artificial, calcada, contudo, no artifice humano. Esta superinteligência artificial seria como um intelecto que supera amplamente as capacidades cognitivas humanas em quase todos os campos de estudo (Bostrom, 2014, p. 22).

Os vídeos criados pelo *Veo 3* são imagens artificiais que recuperam uma leva gigantesca de dados, conteúdo, informação, documentos dispostos no espaço da memória coletiva, que é a internet, a fim de conformar estes dados difusos numa representação do real por meio da associação racional e cognoscível de elementos estéticos, discursivos e narrativos do presente, ou mesmo do passado, ou ainda do futuro, sendo também responsável por construir as novas bases da retrotopia do passado — como nos vídeos que reproduzem passagens bíblicas em

um formato de vídeo-selfie — e da utopia futura — como os vídeos que constroem novos mundos tecnológicos —, ao passo em que institui uma heterotopia. “A I.A. desempenha agora um papel igualmente relevante em nossas experiências culturais, automatizando de modo crescente o domínio da estética” (Manovich, 2017/2018). Perdemos nós o domínio da estética? Ou fomos nós estatizados pela máquina?

Boa parte destes vídeos ultrarrealistas apresenta conteúdos que utilizam a estética do telejornalismo, com repórteres, apresentadores de telejornal, entrevistados e a linguagem textual, as imagens, os planos típicos de conteúdos do jornalismo televisivo, dentro de outros contextos, criados a partir dos *prompts* de comando de pessoas que reconhecem no telejornalismo uma forma de narrar a experiência.

Alguns dos vídeos recuperam personagens bíblicos clássicos, outros passaram a representar discursos ideológicos políticos com imagens de políticos ou de pessoas da população, tudo isto a partir dos prompts feitos pelos usuários da plataforma.

Associado a isto, a máquina reconhece no grande banco de dados à disposição, rótulos, estereótipos, marcadores discursivos, imagem e símbolo, som e demais elementos estéticos dos fenômenos contemporâneos que servem como base para a constituição dos seus produtos. Assim, temos uma plêiade de conteúdos produzidos a partir desta biblioteca humana de séculos que presta sua matéria-prima à inteligência artificial.

Os vídeos da tecnologia *Google Veo 3* assumem esta forma, que aponta para os objetos existentes no mundo produzido pelo homem, mas que criam novas subjetividades, estatizam modos de presença, reforçam estereótipos e criam heterotopias, onde a própria representação televisiva,

por exemplo, assume outro viés; onde até a noção de diálogo, o jogo do discurso social entre os interlocutores admite outras possibilidades, grotescas, absurdas, hilárias, como no caso da apresentadora fictícia Marisa Maiô, que aparece em uma série de vídeos feitos pelo artista e escritor Raony Philips, parodiando programas de auditório.

Neste outro mundo, é natural que o entrevistado produza respostas aceitas numa realidade distópica, ou mesmo que o repórter ou apresentador teça comentários ou assumam formas estéticas que transparecem racismo algorítmico (Silva, 2022), etarismo e outros preconceitos. “*Más allá de la desigualdad de género, existe el riesgo de que la IA perpetúe o exacerbe formas de discriminación racial y étnica*” (Pedreño Muñoz et al., 2024, p. 23).

Qual aprendizado de máquina levou a estes resultados? Foi o *prompt* de comando, o impulso humano, que engrenou a máquina? Ou a máquina encerra novos objetos, que bebem na fonte humana, mas que produzem novas representações?

Para compreender essas nuances, é necessário entender o conceito de *polimedialidade*: “a junção de media apresentativos, representativos e irrepresentativos na totalidade de um espetáculo que, não raro, sempre tiveram grande apelo popular” (Duarte, 2023a, p. 15). O audiovisual é polimedial, à medida em que congrega o modo de apresentação (imediatividade), com o modo de representação (mediaticidade) e o modo irrepresentativo (som). É destas propriedades que a IA generativa de vídeos se vale para criar novas estéticas ou novos modos de presença e objetos estéticos.

Desse modo, fazemos uma incursão sobre conteúdos de vídeos ultrarrealistas da IA generativa *Google Veo 3*, buscando compreender

a nova relação de agente maquínico frente ao agente humano (Coli & Sousa, 2025), dialogando com autores como Karl Marx (1867/2013), Vilém Flusser (1983/2011), Lev Manovich (2017/2018), Georges Didi-Huberman (2020), Muniz Sodré (2023), entre outros, para compreender os desdobramentos dessa nova fase da técnica. A análise se apoia, sobretudo, na crítica marxista ao papel da máquina na produção, nas noções de antropológica (Sloterdijk, 2009/2012), nas heterotopias (Foucault, 2009/2013) e na retrotopia (Bauman, 2017), além dos modos de presença estética propostos por Rodrigo Duarte (2023a; 2023b): apresentação, representação, irrepresentação e perrepresentação.

A máquina da IA generativa

No contexto da inteligência artificial generativa, a máquina atinge o ápice da forma autônoma (Marx, 1867/2013), deixando de ser máquina motriz, de transmissão ou máquina-ferramenta para se estabelecer na relação com o homem como máquina independente, cujo trabalhador é não mais operador ou sujeito maquinal, mas se torna combustível.

Marx compreendeu que a máquina surgia a partir do momento em que a ferramenta de operação do sistema maquinal era transferida da ação corporal do homem para a máquina, ainda que a força motriz de acionamento da máquina partisse do homem.

No contexto da inteligência artificial, a força motriz pode ser assemelhada ao *prompt* de comando e a ferramenta seria a criação dos códigos para a operação. Contudo, as IAs escapam ao domínio humano, ao comando ou ao controle do homem, ou ainda à sua vigília. Marx enxergou o novo papel do homem diante da máquina na Revolução Industrial do século XVIII: vigiar o trabalho da máquina e corrigir os

erros. Na contemporaneidade, o homem deixa até mesmo de vigiar. Afinal, a máquina utiliza aprendizado de máquina (*Machine learning*), ela mesmo se ensina e se otimiza, se regula, é autodidata e o homem coletivo é a enciclopédia de consulta.

Esta disrupção que promove uma exegese do indivíduo para objetificá-lo diante da máquina, retira dele próprio sua identidade a fim de nutrir a IA de propriedade intelectual, apropriando-a da consciência mesmo humana, de tal sorte que se configura uma outra visada ontológica do homem na contemporaneidade, animal ou força-motriz, não mais braçal da máquina, mas intelectual, e da máquina também a sua matéria-prima, logo, tanto produto quanto cliente, um cliente que consome o que ele mesmo produzia, mas que não produz mais, uma vez que suas habilidades foram confiscadas ou sequestradas, ou transferidas para a máquina, assim como aconteceu na Revolução Industrial do século XVIII (Marx, 1867/2013).

Aqui, a tecnologia *Google Veo 3* pode ser vista como a evolução máxima dessa transformação e do meio de trabalho. A IA não é uma ferramenta que o humano opera diretamente com a força física, mas sim um mecanismo que opera por si mesmo, gerando conteúdo ultrarrealista. Apesar da operação de um humano para a criação do *prompt*, a IA (máquina) assume a função que antes era do criador audiovisual (do cinegrafista, do diretor de fotografia, das pessoas que aparecem nos vídeos — atores, jornalistas). A maquinaria torna prescindível a força do trabalhador (Marx, 1867/2013, p. 575), aumentando o número de assalariados e desvalorizando a força de trabalho. Ao sequestrar a estética e a representação, as tecnologias de vídeo ultrarrealistas com

IA deslocam profissionais de áreas como edição, produção, atuação, entre outros.

A capacidade do *Veo 3* de imitar estéticas e narrativas rapidamente pode ser explorada para otimizar a produção de conteúdo, reduzindo custos, tempo e aumentando a mais-valia para quem a utiliza com fins financeiros. Se antes havia programas de humor televisivos para este tipo de formato de vídeo (esquetes), a exemplo do *Saturday Night Live* (SNL), da americana *NBC*, ou *Zorra Total*, da Rede Globo, hoje o mesmo tipo de conteúdo é feito pela plataforma de IA a partir de uma única pessoa. Se eram necessários repórteres para fazer algum tipo de conteúdo, a máquina passa a representá-los e substituí-los. Além disso, a distribuição dos conteúdos é feita nas redes sociais que são proprietárias das *big techs*. Os criadores dos conteúdos não são detentores dos meios de produção, não são donos da *Veo 3* e nem da plataforma na qual publicam.

O *Google Veo 3* também representa uma *força motriz* (Marx, 1867/2013) autônoma na produção audiovisual. Ela é capaz de gerar uma quantidade de vídeos que extrapolam a capacidade humana individual. O “monstro mecânico” (p. 560) pode ser metaforicamente entendido como os grandes modelos de IA e as plataformas que os utilizam, dominando a produção de conteúdo.

Marx descreve ainda como a grande indústria teve que se apoderar de seu próprio meio de produção — a máquina — e produzir máquinas por meio de máquinas para se firmar. Com a IA não é diferente. Não é difícil imaginar cenários onde as IAs se aprimoram a si mesmas (por meio do aprendizado de máquina) em que elas serão utilizadas para criar novas IAs.

O texto de Marx oferece uma lente crítica para essa análise. Nesse contexto capitalista, de plataforma, o *Google Veo 3* não só transforma a forma de se fazer vídeos mas também remodela as relações de trabalho, a busca por valor e a própria natureza da criação estética e de representação:

A emergência dessa tecnologia levanta uma série de questões cruciais que atravessam o campo da estética e da filosofia da arte: os produtos gerados por IA podem ser reconhecidos como obras de arte em sentido pleno, ou seriam meras simulações algorítmicas da criatividade humana? Em que medida a agência humana ainda desempenha um papel essencial na definição do estatuto artístico desses objetos, e qual a extensão da autonomia da máquina nesse processo? (Coli & Sousa, 2025, p. 1)

Autonomia da máquina e automação da estética

A grande questão é que a IA deixa de atuar como um sistema de maquinaria e não mais representa a ideia de máquina ferramenta: “a atuação da IA vai muito além do papel de mera ferramenta, na medida em que gera imagens e obras a partir de comandos que podem ser indicações mínimas” (Coli & Sousa, 2025, p. 6). O *prompt* pode ser meramente algo como “surpreenda-me” e, a partir deste impulso abstrato, a máquina é capaz de desenvolver operações lógicas, partindo de um banco de dados que promovam algum tipo de produção, logo, representando uma autonomia neste sentido. “A IA, por sua vez, funciona como um co-agente do processo criativo e produz resultados que muitas vezes surpreendem os usuários por adotar caminhos e estratégias incomuns” (p. 8).

A imitação da estética do telejornalismo, a criação de video-selfies de personagens bíblicos, entre outros vídeos criados com a tecnologia

podem ser analisados à luz da crítica de Muniz Sodré (2023) de que há uma ausência da dimensão cultural na tecnologia, partindo da própria definição de *cultura* feita pelo autor, que pode ser resumida como o mapeamento ativo das crenças partilhadas e valores relativos a determinações que são a base de uma coesão comunitária que sustentam a sociedade (p. 22).

Nos vídeos analisados, pelo fato de existirem os humanos responsáveis pelos *prompts*, percebe-se que há um enraizamento cultural — às vezes superficial —, mas necessários para que os memes, as piadas e os vídeos façam sucesso nas redes. Os perfis com mais seguidores são os que são consistentes com a cultura em geral dos personagens ou situações representados por meio dos vídeos ultrarrealistas. Para fazer sentido e gerar empatia e engajamento do público, é necessário que os vídeos tenham algum roteiro cuja base seja a partir de um repertório compartilhado pelos usuários das redes que são o público-alvo dos produtores de conteúdos com IA.

Este fenômeno, ao nosso ver, se distancia da leitura moderna sobre o agente maquínico ainda estar associado diretamente ao agente humano, uma vez que a máquina é capaz de elaborar objetos sensíveis independentemente da intencionalidade do acionador da IA. Muito embora, compreendemos que a máquina geralmente apresenta resultados a partir dos dados do usuário.

Vilém Flusser (1983/2011) compreendeu que máquinas fotográficas e *softwares* não são meramente ferramentas de expressão da intenção do operador, mas resultado da programação escolhida pelo usuário. Na nossa análise, a máquina vai além, autodeterminando-se

para gerar estes resultados e não dependendo somente da programação do usuário.

Ao utilizar o *Google Veo 3*, introduzimos o seguinte prompt de comando: “Faça um vídeo com um estudante do curso de Estudos da Mídia da UFRN”. Como resultado, a IA apresentou um vídeo, do qual extraímos o *frame* apresentado na Figura 1:

Figura 1

Reprodução de vídeo de estudante e repórter gerado por IA



Nota. Vídeo feito pela tecnologia *Google Veo 3* a partir de prompt elaborado pelos autores.

Ou seja, a figura do aluno, o cenário da universidade, o perfil da repórter, as roupas, foram todas escolhidas pela máquina, a partir do banco de dados a que ela tem acesso, traduzindo como a IA visualiza o contexto proposto ou como a estatística depõe sobre a realidade:

Isso torna as IA's capazes de tomar “decisões estéticas” de forma praticamente autônoma, a depender do grau de detalhamento do *prompt*, que pode até mesmo ser puramente conceitual. Mesmo que essas “decisões” maquinicas não possuam a intencionalidade característica da agência humana, elas são o resultado de

cálculos que seguem padrões estatísticos preestabelecidos pelo treinamento algorítmico e pelas limitações dos *prompts* fornecidos. Nesse sentido, pode-se dizer que a IA não realiza escolhas propriamente ditas, mas responde a um conjunto de condições que em grande medida determinam seus resultados. Ainda assim, as IA's possuem uma capacidade de cálculo e de produção autonomizada (em espectros variados de autonomia), que muito difere de outras ferramentas tradicionais e, por isso, referimo-nos aqui à atuação das IA's pelo termo de agência algorítmica. (Coli & Sousa, 2025, p. 7)

A questão central deste fenômeno repousa sob dilemas acerca desta produção da IA generativa, conforme Lev Manovich (2017/2018): a inteligência artificial identifica os objetos culturais e faz o caminho correto para tal ou é uma caixa preta que gera novos objetos? Essa automação conduz a uma diminuição da diversidade cultural? Há uma padronização da imaginação fotográfica ou audiovisual, por exemplo? Isto estaria criando uma homogeneização estética? Admitindo que a máquina seja capaz de analisar padrões culturais em diferentes escalas e extrair uma representação destes objetos, estaríamos concebendo que “a I.A. age como um teórico ou historiador da arte, ou especialista em cinema, que também estuda rotineiramente muitas obras de alguma área da cultura para apreender seus padrões” (p. 121). Mas a estatística computacional é capaz de identificar tendências culturais ou suprime informações dos objetos sensíveis, de produções que não podem ser reduzidas a sistemas? A estatística representa a estética?

A IA está “automatizando de modo crescente o domínio da estética” (p. 2), com a modificação automática de fotos, criação de filmes curtos a partir de filmes brutos e a avaliação da qualidade estética de algoritmos de correção e classificação de imagens. Os vídeos

ultrarrealistas são a materialização e o ápice dessa automação estética no campo visual. Reparem que a depender do *prompt* colocado pelo operador, a IA cria todas as outras coisas que não foram ditas.

Anna Coli e Thobila Sousa (2025) recuperam a ideia de *écfrase*, ou seja, a descrição minuciosa seja de um quadro, um objeto ou um lugar. Os objetos estéticos sensíveis gerados pela IA precedem uma espécie de *écfrase*, que seria o *prompt* de comando. Esta *écfrase* proporciona resultados diferentes a depender do usuário, da máquina. Logo, a IA exerce determinada discricionariedade e autonomia no seu processo de criação.

Não estaríamos distantes de um futuro onde “os computadores comecem a criar produtos de mídia do início ao fim do processo” e os humanos “apenas confiarão que a IA sabe o que convém” (Manovich, 2017/2018, p. 3). O *Google Veo 3*, ao criar vídeos ultrarrealistas, é um passo nessa direção.

Modos de presença dos vídeos ultrarrealistas

Os vídeos ultrarrealistas recorrem à polimedialidade como forma de comunicar e representar o mundo. Essa estatização dos elementos cotidianos é possível em função dos dados a que o *Google Veo 3* tem acesso. A nossa avaliação é a de que estes vídeos incorporam novos modos de presença nos fenômenos estéticos contemporâneos, a partir da classificação proposta por Rodrigo Duarte (2023a, 2023b). Eles são capazes de gerar novos objetos estéticos sensíveis que elaboram novas representações sociais para as pessoas que consomem estes conteúdos.

A fim de categorizar o nosso objeto de estudo, classificamos os vídeos identificados a partir das seguintes categorias de análise: vídeos

cômicos, anedóticos, satíricos ou piadas; esquetes ou pseudojornais; vídeos políticos; *vlogs* distópicos, utópicos, retrotópicos e heterotópicos. Analisamos os conteúdos a partir dos modos de presença elencados por Duarte: *apresentação*, *representação*, *irrepresentação* e *perrepresentação*.

A *apresentação* refere-se àquilo que se mostra de forma mais imediata, com forte apelo à visualidade e corporeidade. Os vídeos ultrarrealistas buscam precisamente essa apresentação de uma realidade simulada, imitando a aparência humana, a gestualidade, e o cenário para criar a ilusão de presença. “As imagens geradas pela IA reproduzem, em grande medida, parte da experiência visual que temos, que registramos e que acaba, por fim, alimentando a base de dados à qual as IAs recorrem em sua produção” (Coli & Sousa, 2025, p. 12). Os video-selfies dos personagens bíblicos são um exemplo dessa busca pela apresentação de uma imagem, mas precisam de uma mediacidade, uma mediação simbólica que só se completa por meio da representação.

A *representação* está ligada à textualidade e à mediação discursiva. Os vídeos de IA que imitam a estética, narrativa e discurso telejornalístico, por exemplo, operam fortemente no modo de representação. Eles não apenas apresentam as imagens, mas constroem narrativas, seguem roteiros e utilizam linguagens que representam eventos e informações, mesmo que geradas por IA:

Esses códigos computacionais, uma vez convertidos em produtos estéticos, passam por uma mediação que exige, por sua vez, uma contrapartida de interpretação (humana — ao menos na maior parte das vezes) para que seu significado seja apreendido. Isso significa, para recuperar o aspecto discursivo da representação, que o significado da obra não é dado imediatamente à percepção visual, mas precisa ser discursivamente construído e, portanto,

mediado pela palavra e pela ordem do simbólico. (Coli & Sousa, 2025, p. 14)

A *irrepresentação* está associada à sonoridade musical e à capacidade de interiorização, além das relações espaço-temporais. A dimensão sonora dos vídeos gerados por IA — no *Google Veo 3*, os vídeos já vêm com áudios e falas dos personagens gerados por IA — também contemplam este modo de presença. Os efeitos sonoros podem, em sua qualidade representativa, afetar o ânimo e a percepção dos espectadores, de maneira que o ultrarrealismo visual talvez não alcance.

Já a *perpresentação* é a junção ou sobreposição dos media apreativos, representativos e irrepresentativos. “Nesse modo de presença, comparecem todos os outros três, independentemente de um juízo de valor sobre a qualidade estética ou o comprometimento ético-político” (Duarte, 2023b, p. 168). Os vídeos ultrarrealistas gerados pelo *Veo 3* são, por natureza, polimedias e, portanto, perpresentativos, pois combinam imagem, som, narrativa e elementos discursivos para criar uma experiência complexa.

Cumprе salientar que o *Google Veo 3* tem capacidade generativa de vídeos de até 8 segundos por vez, quando do momento desta pesquisa, o que estabelece também uma nova relação de produção audiovisual, com vídeos extremamente curtos e narrativas fragmentadas.

É oportuno frisar ainda que os traços imagéticos dos personagens costumam ser altamente estereotipados e até se repetem em diferentes vídeos, bem como o tom da voz, estetizando modelos de reprodução.

Identificamos ainda perfis que produzem vídeos feitos por humanos com atores e atrizes imitando, reproduzindo, os modelos estéticos

gerados a partir dos vídeos ultrarrealistas, o que aponta para a materialização desta nova estética audiovisual gerada pela IA que argumentamos.

Defendemos, desse modo, uma nova visada para os vídeos ultrarrealistas, um novo modo de presença que se caracterizaria como *ultra-apresentativo*, ao sintetizar com muita intensidade elementos que estão em todos os modos descritos anteriormente e que extrapolam o campo simbólico representativo, a imediaticidade visual apresentativa, a sonoridade irrepresentativa e a sobreposição perrepresentativa. Na mesma caixa de diálogo com a IA, vídeos diferentes podem ser criados com objetos sensíveis diferentes utilizando o mesmo comando, bastando ao usuário solicitar novas versões. Além disso, componentes se repetem em vídeos diferentes, apontando para uma criação automatizada da estética a partir das estatísticas, com vozes iguais, perfis de fenótipos que se reproduzem em vídeos diferentes a partir dos contextos.

O topos nostálgico e idealizado dos vídeos ultrarrealistas

Georges Didi-Huberman (2020) argumenta que a imaginação é a nossa primeira potência livre de reorganizar o mundo de outra forma, de uma maneira mais justa. Os vídeos ultrarrealistas podem ser vistos como uma manifestação dessa potência da imaginação, uma vez que reorganizam a realidade criando novas narrativas e tonalidades visuais que antes não eram possíveis ou eram limitadas pela produção audiovisual tradicional.

Entre diversos aspectos da contemporaneidade, da compreensão do humano e de sua imaginação como propulsora do desejo de estabelecer ordenamento social, vale destacar o que George Steiner (1974/2003) conceitua como *mitologias seculares*. O autor destaca tais mitologias como o produto do colapso das grandes narrativas religiosas que foram

responsáveis por estruturar o Ocidente durante séculos. Este moderno sistema de crenças que elegeram, nos últimos séculos, personagens e ideologias messiânicas, não possuem necessariamente caráter religioso, mas atuam psicologicamente como religiões, o que colabora com a nossa capacidade de criar novas imagens.

Seja por meio da filosofia política de Marx, da psicologia freudiana, nos estudos antropológicos de Lévi-Strauss ou dos pressupostos da astrologia, Steiner observa que as tentativas de preencher o vazio com ilusões — como ocultismo, drogas psicodélicas ou religiões orientais — são insuficientes e muitas vezes enganosas na busca desesperada pelo sentido absoluto. No entanto, é na identificação com o sentido existencial, na construção de estruturas morais, na ideia de pertencimento comunitário e na esperança pela redenção que se ancora a nostalgia pelo absoluto.

Desse modo, a busca pelo preenchimento de um vazio de sentido em realidades alternativas visa preencher um vazio emocional e existencial da sociedade contemporânea. De certo, pode-se indicar entre as mitologias contemporâneas aquela que emana da era digital, na qual as produções ultrarrealistas de vídeos construídas por IA representam o culto à tecnologia e a ideia de inovação, inclusive com o que podemos chamar de templo: o Vale do Silício.

Um reflexo possível desta nostalgia pode ser encontrado em perfis de redes sociais que reconstroem narrativas bíblicas contadas a partir dos vídeos ultrarrealistas produzidos por IA. É possível também encontrar vídeos com personagens bíblicos em contextos e modos de aparição novos ou diferentes do imaginário simbólico lançado por meio da própria narrativa bíblica, a exemplo do *vlog* de Jesus ressuscitando

Lázaro (Figura 2), ou de Maomé abrindo o Mar Vermelho (Figura 3). No mesmo sentido, também é possível encontrar vídeos que reconstroem cenários remotos de cidades a partir de *prompts* de comando com descrições sobre o espaço urbano. Outros vídeos recuperam discursos de um tempo antigo e redesenham paisagens do passado vivido pelos usuários, em um completo movimento nostálgico.

Figura 2

Imagem de Jesus gerada pelo Google VEO 3



Nota. Reprodução de vídeo publicado no perfil
E se a IA pudesse contar a Bíblia? (2025).

Exemplos populares incluem canais no *TikTok* e no *YouTube*, e perfis no *Instagram* que utilizam IA para criar *deepfakes* históricos ou cenas de filmes “perdidos”, gerando milhões de visualizações e forte adesão. A capacidade dessas produções de evocar uma sensação de nostalgia por um passado que nunca existiu da forma apresentada — ou que é uma versão idealizada dele — corrobora a análise de Zygmunt Bauman (2017) sobre a *retrotopia*. Em sua obra póstuma, Bauman aponta a retrotopia como a nostalgia por um passado social e político idealizado, que serve como refúgio de um presente incerto e um futuro

desencantador que se alinhando perfeitamente à busca por referenciais sólidos na crise contemporânea.

Figura 3

Imagem de Maomé gerada pelo Google VEO 3



Nota. Reprodução de vídeo publicado no canal do YouTube BR Metaverso (2025)

Ao abordar a noção de uma “epidemia global de nostalgia”, conforme desenvolvida por Svetlana Boym, professora de Literatura Eslava Comparada em Harvard, a nostalgia é compreendida como um sentimento marcado por perda, deslocamento e pela idealização subjetiva de um passado imaginado. A partir desse referencial, Bauman argumenta que, no percurso histórico moderno, a até então dominante “epidemia frenética de progresso” foi gradualmente substituída por uma epidemia de nostalgia, que passou a ocupar o centro das dinâmicas sociais e culturais contemporâneas (p. 9).

Esta é uma atividade do campo da imaginação, lançada do agente humano ao agente maquínico, que abre utopias concretas, ou

heterotopias (Didi-Huberman, 2020). A IA, ao gerar esses vídeos a partir de *prompts* feitos por humanos, demonstra um poder quase ilimitado de criar e explorar as possibilidades visuais e narrativas. Tecnologias como o *Google Veo 3* estão expandindo o campo do audiovisual para além das representações convencionais, empregando outra forma de imaginar o audiovisual. No contexto dos vídeos hiperrealistas, as heterotopias de Michel Foucault (2009/2013) se manifestam nos “outros espaços” que essas criações digitais constroem. Esses vídeos criam realidades alternativas, simulacros que, embora pareçam reais, não o são, atuando como o exemplo que Foucault cita dos espelhos que refletem e distorcem a realidade percebida (p. 37). Assim como um cemitério ou um teatro são heterotopias (pp. 37-40), esses vídeos gerados por IA são *topos* virtuais onde o tempo e o espaço são recombinações e a realidade é subvertida, oferecendo uma contra-localização aos espaços cotidianos.

A relação entre o sensível e o inteligível na criação de imagens utiliza a imaginação como a faculdade que conecta esses dois aspectos (Didi-Huberman, 2020). Os vídeos criados com IA manipulam elementos sensíveis (sons, imagens, movimentos) e constroem narrativas e discursos que buscam ser inteligíveis e, em alguns casos, até mesmo persuasivos, como nos exemplos em que simulam a linguagem do telejornalismo ou apresentam conteúdo ideológico ou político. O autor afirma que a densidade do mundo tem de ser imaginada e que a imaginação redistribui o mundo, o *reatmosferiza* de maneira igualitária (p. 11). Os vídeos gerados por IA, ao criarem realidades alternativas ou simulações, estão, de certa forma, *reatmosferizando* o mundo real, criando novas densidades e percepções, mesmo que sejam artificiais, e, portanto, heterotopias.

Neste cenário emergente, a imaginação, antes um domínio exclusivamente humano, torna-se um terreno compartilhado com a inteligência artificial. Isso nos coloca diante de questões cruciais sobre autoria, propriedade e a própria construção da realidade. À medida que as heterotopias digitais se proliferam e a retrotopia ganha novas formas através da IA, é fundamental que continuemos a explorar as implicações éticas, sociais e culturais dessa reconfiguração da nossa experiência.

Uma nova gramática para compreensão da antropotécnica

Ao realizar sua argumentação acerca do contemporâneo, Giorgio Agamben (2008/2009) se apoia na ideia de que o contemporâneo não aborda somente o que se pode chamar de presente cronológico, mas uma relação particular com o tempo. Para o autor, a ideia de contemporâneo está intimamente ligada a capacidade de compreender as sombras do seu tempo, as inatualidades, os anacronismos, aquilo que ainda não se manifestou ou que já passou mas ressoa no presente, as heterotopias, retrotopias e utopias. Se faz necessário recorrer a Agamben para poder compreender, em alguma medida, o pensamento do filósofo e pensador alemão pós-frankfurtiano Peter Sloterdijk, que nos aproxima das discussões da genesis contemporânea que projeta luz sobre nosso objeto de estudo, o uso de IA para produção de vídeos ultrarrealistas, e do conceito de antropotécnica.

Ao introduzir o conceito de antropotécnica na obra *Regras para o parque humano: uma resposta à carta de Heidegger sobre o humanismo*, Sloterdijk (1999/2000) aborda que a cultura humanista, baseada na leitura e na escrita, aparece como um modelo histórico de antropotécnica. O autor afirma que “os homens são os animais que

devem ser domesticados por outros meios que não a violência bruta” (p. 16) ao mesmo tempo que, na contemporaneidade, se trava uma luta entre “impulsos domesticadores e os bestializadores e os respectivos meios de comunicação” (p. 46). Sloterdijk destaca que a capacidade dos meios letrados de produzir sujeitos e ordenamentos sociais foi substituída por uma cultura audiovisual de massas, que reconfigura os mecanismos de formação.

Em *O desprezo das massas*, ele observa que “o espaço público foi invadido por uma *hiperpresença* do trivial, e a alta cultura perdeu sua função reguladora (Sloterdijk, 2000/2016, p. 38). Desse modo, convém indagar o real papel do uso de recursos tecnológicos, como as IAs produtoras de vídeos, na reestruturação dos modos de pensar, ser e estar. Certamente Sloterdijk questionaria: a IA é a nova ferramenta de domesticação ou de bestialização? Ela nos *amansa* ao terceirizar tarefas complexas ou nos *bestializa* ao atrofiar nossas capacidades cognitivas inatas?

Em *Has de cambiar tu vida*, Sloterdijk (2009/2012) enfatiza que o ser humano é um ser exercitante e que a história da humanidade pode ser lida como uma sequência de empreendimentos de autocultivo (pp. 16-19). Essa ideia de autocultivo se estende à forma como as sociedades se organizam e se representam. É nesse ponto que a emergência dos vídeos hiperrealistas gerados por IA se insere de forma provocadora na teoria da antropotécnica.

As tecnologias de geração de vídeos ultrarrealistas por IA podem ser lidas como uma forma de antropotécnica de segunda ordem, isto é, práticas que moldam a percepção do mundo e ao mesmo tempo também produzem regimes de realidade. Trata-se de uma “nova escrita”

não mais sobre papel, mas sobre os sensores e afetos dos indivíduos. A verossimilhança dessas imagens simula a presença e compromete a distinção entre o real e o produzido, produzindo uma forma de verdade performática e afetiva.

Na obra *Filosofia da caixa preta*, Flusser (1983/2011) destaca que o aparelho fotográfico não reproduz o mundo, mas o reprograma segundo códigos operacionais. Em um estágio mais avançado, as inteligências artificiais operam como sistemas simbólicos autônomos, capazes de estabelecer métricas algorítmicas da realidade. A antropotécnica, aqui, se faz pela repetição e intensificação sensorial, produzindo efeitos de autoridade simbólica sem a mediação da crítica ou do *logos*.

Em um contexto no qual “o virtual e o real se entrelaçam de modos antes inimagináveis”, como sugeriria Sloterdijk (1999/2000) ao analisar a nossa *esfera* midiaticizada, a IA atua como uma antropotécnica que reconfigura não apenas como nos comunicamos, mas também como percebemos a verdade e como construímos nossas próprias realidades.

Zygmunt Bauman define a cultura moderna como uma “fábrica de ordens” (Bauman, 2011/2013), um mecanismo destinado a produzir distinções, hierarquias e classificações que orientem o comportamento. No entanto, na modernidade líquida, essa fábrica deixa de operar por meio de grandes narrativas e passa a se articular via fluxos informacionais e sistemas de consumo.

Em 2025, o Governo Federal brasileiro decidiu adotar o uso dos vídeos ultrarrealistas para fazer o enfrentamento do debate político a respeito das pautas que defendia. Foram milhares de vídeos sobre Imposto de Renda, Pix, taxaço de super ricos, soberania nacional, entre outros temas, postados nas redes sociais oficiais do governo.

No contexto digital contemporâneo, a superficialidade do consumo não apenas reflete um afastamento das narrativas profundas e complexas, mas também atua como um mecanismo de controle social, criando uma aparência de ordem e previsibilidade. Nesse cenário, os usuários são constantemente bombardeados por informações simplificadas, *clickbait*s e conteúdos que reforçam suas crenças, transformando o consumo em um processo de alfabetização superficial, no qual o conhecimento é frequentemente reduzido à adesão a tendências momentâneas e efêmeras.

Assim, é pertinente refletir sobre as ideias apresentadas por Zygmunt Bauman (1997/1998) em *O mal-estar da pós-modernidade*. Bauman discute como a noção de cultura, originada da comparação feita por Cícero entre o cultivo da mente (*cultura animi*) e a agricultura, evoluiu para uma metáfora da ordem social, em que a estrutura cultural divide a sociedade entre aqueles que moldam e aqueles que são moldados. Essa divisão, que distingue os *escultores* dos *esculpidos* e os *professores* dos *alunos* (p. 162), sugere um modelo de sociedade ordenada, onde as instituições operam como *fábricas de ordem*, projetando previsibilidade e controle sobre os comportamentos individuais.

Ao aplicar essa ideia ao fenômeno da IA audiovisual, percebemos que os vídeos gerados por inteligência artificial constituem novas formas de fabricação de ordens simbólicas, que não são explicitamente coercitivas, mas operam pela sedução, pela imersão e pela simulação do real. Em outras palavras, são formas de governo do sensível (Rancière, 2000/2009), que reorganizam o campo do que pode ser visto, sentido e acreditado.

Assim, a produção massiva de conteúdo sintético criado por IA impõe um desafio à formação do sujeito contemporâneo. Como adverte

Sloterdijk, “num mundo em que todos os discursos são tecnicamente simuláveis, a diferença entre formação e adestramento tende a desaparecer” (1999/2000, p. 24). Surge, então, a necessidade de uma nova paideia tecnológica, uma forma de alfabetização simbólica, a construção de uma nova gramática capaz de resistir à captura total pela mídia sintética.

A distopia de um parque humano automatizado por algoritmos se realiza na medida em que as antropotécnicas não são reconhecidas como tais, e passam a operar de modo invisível. Como adverte Yuval Harari (2018), as máquinas de IA conhecem melhor os nossos desejos do que nós mesmos, e esse tipo de saber pode ser usado para fins de controle afetivo e político.

Diante deste cenário, é necessária uma preocupação com o futuro desta tecnologia e como outros vídeos deste tipo serão utilizados — se é que já não estão sendo utilizados — como objetos políticos e de controle social, para manipulação ou reforço de vieses. Afinal, como aponta Muniz Sodré (2023), “as rédeas do poder social se confundem cada vez mais com os dispositivos de controle da informação”, e “o relacionamento do sujeito humano com a realidade hoje passa necessariamente pela tecnologia, em especial as tecnologias da informação, em todos os seus modos de realização” (pp. 24-25).

Conclusão

A partir da análise desenvolvida neste artigo, compreendemos que a produção de vídeos ultrarrealistas com inteligência artificial, como os criados pelo *Google Veo 3*, representa uma virada profunda na estética audiovisual contemporânea. Esses conteúdos, gerados a partir de comandos simples, carregam uma complexidade simbólica que recria

narrativas, imagens e sons com precisão suficiente para borrar as fronteiras entre o real e o artificial. Percebemos que a máquina reproduz a realidade e também constrói novos sentidos, revelando uma autonomia estética que tensiona as relações clássicas entre autor, obra e público.

Além disso, observamos que essa nova estética, marcada pela polimedialidade e pela lógica da ultrarrepresentação, configura uma espécie de sequestro da representação e estatização dos modos de presença. A repetição de estereótipos visuais, sonoros e discursivos por parte da IA revela uma padronização que atua como filtro cultural, promovendo uma estética algorítmica. Essa constatação nos leva a refletir sobre os riscos de homogeneização simbólica, apagamento da diversidade cultural e substituição progressiva da criação humana por simulações maquínicas de subjetividade.

Identificamos também que os vídeos ultrarrealistas frequentemente constroem uma retrotopia, ou seja, uma nostalgia idealizada de um passado que, muitas vezes, nunca existiu da forma como é apresentado. Essa operação responde à necessidade contemporânea de preenchimento simbólico e afetivo diante das incertezas do presente. Ao analisar essas produções, percebemos como a IA se consolida como uma nova força simbólica e produtiva, capaz de moldar afetos, crenças e estruturas de sentido coletivas por meio da recombinação estética de memórias e desejos.

Por fim, propomos uma leitura da IA generativa como uma nova forma de antropotécnica. Inspirados pelos autores citados no trabalho, argumentamos que essas tecnologias não apenas geram conteúdos, mas atuam como dispositivos de formação subjetiva e de produção de regimes de verdade. Diante disso, consideramos urgente desenvolver

uma nova alfabetização simbólica, que nos permita reconhecer os mecanismos de poder e mediação estética operados pelas máquinas. Afinal, a inteligência artificial, ao reprogramar a imaginação humana, também nos convoca a reimaginar criticamente a nossa própria condição cultural, política e sensível.

Referências

- Agamben, G. (2009). *O que é o contemporâneo? e outros ensaios* (Vinícius Nicastro Honesko, Trad.). Argos. (Obra original publicada em 2008)
- Bauman, Z. (2013). *A cultura no mundo líquido moderno* (Carlos Alberto Medeiros, Trad.). Zahar. (Obra original publicada em 2011)
- Bauman, Z. (1998). *O mal-estar da pós-modernidade* (M. Gama & C. M. Gama, Trad.). Jorge Zahar. (Obra original publicada em 1997)
- Bauman, Z. (2017). *Retrotopia* (R. Aguiar, Trad.). Zahar.
- Bostrom, N. (2014). *Superintelligence: paths, dangers, strategies*. Oxford University Press.
- BR Metaverso (2025). *Moisés meteu o louco e os caras foram de arrasta!* [Vídeo]. YouTube. <https://www.youtube.com/shorts/61pyEtr6UF8>
- Coli, A. L. & Sousa, T. G. L. C. (2025). Inteligência artificial generativa e o mundo da arte: novos modos de presença ou meros sistemas algorítmicos?. *Artefilosofia*, 19(38), 1-25. <https://doi.org/10.69640/raf.v19i38.7849>

- Didi-Huberman, G. (2020). La imaginación, nuestra comuna. *Theory Now: Journal of Literature, Critique and Thought*, 3(2), 5-21. <https://doi.org/10.30827/tmj.v3i2.13931>
- Duarte, R. (2023a). Modos de presença nas manifestações estéticas contemporâneas. In R. Duarte & R. Rocha (Org.), *Modos de presença nos fenômenos estéticos* (pp. 9-22). Relicário.
- Duarte, R. (2023b). Modos de presença revisitados (à guisa de um posfácio). In R. Duarte & R. Rocha (Org.), *Modos de presença nos fenômenos estéticos* (pp. 151-168). Relicário.
- E se a IA pudesse contar a Bíblia? [@ia.bibl.ia]. (2025). *POV: Lázaro ressuscitou*. [Vídeo]. Instagram. https://www.instagram.com/p/DKSpGE_xUhT/
- Flusser, V. (2011). *Filosofia da caixa preta: ensaios para uma futura filosofia da fotografia*. Annablume. (Obra original publicada em 1983)
- Foucault, M. (2013). *O corpo utópico, as heterotopias* (S. T. Muchail, Trad.) (2a ed.). N-1 Edições. (Obra original publicada em 2009)
- Harari, Y. N. (2018) *21 lições para o século 21* (P. Geiger, Trad.). Companhia das Letras.
- Manovich, L. (2018). Automatizando a estética: inteligência artificial e cultura das imagens (D. Hora, Trad.). *Esferas*, 1(11), 119-126. <https://doi.org/10.31501/esf.v1i11.9586>. (Obra original publicada em 2017)

- Marx, K. (2013). *O capital: crítica da economia política - livro I: o processo de produção do capital* (R. Enderle, Trad.). Boitempo. (Obra original publicada em 1867)
- Pedreño Muñoz, A., González Gosálbez, R., Mora Illán, T., Pérez Fernández, E. M., Ruiz Sierra, J. & Torres Penalva, A. (Eds.). (2024). *La inteligencia artificial en las universidades: retos y oportunidades*. (Informe anual sobre IA y educación superior). Grupo 1 million bot.
- Rancière, J. (2009). *A partilha do sensível: estética e política* (M. C. Netto, Trad.). (2a ed.). EXO experimental org.; Editora 34. (Obra original publicada em 2000)
- Silva, T. (2022). *Racismo algorítmico: inteligência artificial e discriminação nas redes digitais*. Edições Sesc São Paulo.
- Sloterdijk, P. (2012). *Has de cambiar tu vida. Sobre antropológica* (P. M. de Veza, Trad.). Pre-Textos. (Obra original publicada em 2009)
- Sloterdijk, P. (2016). *O desprezo das massas: ensaio sobre as lutas culturais na sociedade moderna*. (C. Cavalcanti, Trad.). Estação Liberdade. (Obra original publicada em 2000)
- Sloterdijk, P. (2000). *Regras para o parque humano: uma resposta à carta de Heidegger sobre o humanismo* (J. O. de A. Marques, Trad.). (4a ed.). Estação Liberdade. (Obra original publicada em 1999)
- Sodré, M. (2023). A ruptura paradigmática da comunicação. *MATRIZES*, 17(3), 19-27. <https://doi.org/10.11606/issn.1982-8160.v17i3p19-27>
- Steiner, G. (2003) *Nostalgia do absoluto* (J. G. Flores, Trad.). Relógio D'Água. (Obra original publicada em 1974)

ÍNDICE REMISSIVO

Símbolos

- 3D 24, 182, 261, 309
- 360° 305, 308, 315, 316, 317, 318, 321, 322, 328, 330, 331, 333, 335
- 360 graus 298, 300, 301, 305, 306, 307, 308, 309, 310, 315, 318, 320, 322, 324, 325, 326, 328, 330, 335

A

- algoritmo 34, 40, 44, 45, 115, 156, 236, 252, 255
- algoritmos 18, 20, 22, 23, 24, 29, 35, 37, 38, 40, 47, 48, 49, 51, 53, 67, 69, 74, 88, 107, 109, 110, 117, 119, 140, 149, 154, 165, 169, 178, 180, 189, 198, 200, 205, 206, 248, 249, 250, 255, 256, 257, 258, 264, 299, 304, 375, 378, 460, 473
- audiovisual 62, 306, 431, 432, 441, 450, 453, 455, 456, 460, 463, 464, 468, 470, 472, 473

B

- big techs 19, 20, 22, 29, 456

C

- Cartas Telepáticas 431, 436, 437, 438, 439, 440, 441, 442, 443, 447
- ChatGPT 18, 19, 29, 32, 72, 75, 89, 90, 92, 214, 216, 219, 223, 276, 280, 281, 286, 287, 294, 295, 308, 319, 320, 321, 322, 323, 324, 329, 332, 333, 342, 355, 357, 358, 359, 362, 365, 368, 370, 371, 373, 391, 392, 393, 401, 406, 407, 408, 409, 410, 412, 413, 414, 415, 421, 422, 425, 426, 428, 429
- cidadania 247, 248, 250, 254, 255, 256, 259, 264, 366
- composição 14
- comunicação 31, 35, 36, 41, 42, 43, 44, 45, 53, 54, 58, 59, 60, 61, 64, 66, 67, 70,

- 74, 155, 177, 178, 179, 181, 185, 186, 189, 191, 194, 195, 196, 198, 208, 209, 215, 266, 270, 337, 338, 339, 340, 341, 342, 345, 346, 348, 349, 351, 352, 353, 356, 357, 360, 361, 364, 366, 368, 373, 375, 377, 432, 442, 446, 470, 477

- comunicação 79, 80, 82, 85, 86, 87, 90, 91, 93, 100, 102, 103, 104, 107, 108, 109, 111, 114, 116, 117, 119, 121, 123, 124, 126, 127, 132, 133, 136, 137, 141, 391, 392, 411, 412, 421, 428, 430

D

- desaceleração 24
- Design 4
- desinformação 18, 22, 26, 27, 34, 35, 36, 39, 45, 46, 47, 48, 49, 50, 56, 58, 59, 60, 61, 63, 64, 65, 66, 67, 68, 70, 71, 72, 73, 74, 77, 150, 218, 279, 281, 282, 283, 379
- desinformación 27, 33, 79, 80, 83, 85, 86, 91, 98, 99, 100, 102, 103, 104, 108, 111, 126, 130, 131, 132, 133, 136, 137, 138, 139, 140, 141, 143, 144, 145, 232, 236
- diseño 84, 114, 224, 226, 239, 391, 392, 393, 396, 399, 405, 410, 412, 413, 414, 415, 422, 426, 427, 428, 429
- docente 223, 224, 225, 226, 228, 230, 231, 232, 233, 234, 239, 242, 243, 244, 267, 268, 269, 270, 281, 282, 284, 286, 287, 291, 292

E

- ecologia dos meios 13, 14, 247, 249, 348
- Edgar Pêra 433, 434, 435, 436, 437, 438, 443, 444, 448
- educação 59, 68, 73, 148, 158, 208, 210, 247, 248, 249, 250, 251, 252, 253, 254, 255,

256, 259, 264, 265, 268, 271, 280, 288,
289, 292, 293, 294, 296, 324

educación 95, 96, 97, 98, 99, 100, 103, 106,
138, 223, 229, 234, 235, 242, 477

ética 13, 69, 108, 112, 118, 125, 139, 148,
149, 150, 151, 152, 153, 154, 157, 159,
160, 161, 163, 166, 167, 168, 169, 170,
208, 218, 227, 241, 242, 248, 250, 252,
256, 258, 264, 265, 281, 282, 283, 287,
288, 335, 376, 379, 386, 388

F

Facebook 22, 67, 84, 137, 307

fake news 27, 51, 53, 54, 64, 75, 76, 82, 83,
104, 105, 106, 107, 108, 109, 111, 117,
118, 119, 120, 122, 123, 124, 127, 129,
144, 145, 211

feminino 174, 175, 178, 179, 185, 186, 187,
189, 192, 194, 196, 197, 198, 199, 200

fotografía 14

I

IA 2, 4, 12, 16, 24, 28, 29, 30, 31, 58, 60, 61,
69, 70, 71, 72, 73, 74, 86, 88, 89, 90,
91, 92, 107, 108, 109, 110, 111, 116,
119, 123, 125, 127, 133, 135, 143, 148,
149, 150, 151, 152, 153, 154, 155, 156,
157, 158, 159, 160, 161, 162, 164, 165,
166, 167, 169, 170, 206, 207, 210, 213,
214, 218, 219, 225, 226, 227, 230, 231,
232, 233, 234, 235, 237, 239, 240, 241,
242, 243, 247, 248, 249, 250, 251, 252,
253, 254, 256, 258, 261, 262, 263, 264,
265, 268, 270, 274, 275, 276, 277, 278,
279, 280, 281, 282, 283, 286, 288, 289,
290, 291, 301, 320, 321, 323, 324, 325,
338, 339, 342, 343, 344, 345, 354, 355,
356, 357, 360, 364, 367, 368, 372, 373,
375, 376, 377, 378, 379, 380, 382, 385,
386, 403, 408, 413, 422, 428, 439, 441,
450, 453, 454, 455, 456, 457, 458, 459,
460, 461, 462, 463, 464, 465, 466, 468,
469, 470, 471, 472, 473, 474, 476, 477

IA generativa 108, 149, 242, 320, 339, 342,
357, 373, 453, 454, 460, 474

IA generativas 92, 249

IAs 21, 234, 274, 275, 276, 280, 341, 343,
344, 345, 355, 356, 365, 454, 456,
462, 470

imagens 25, 26, 37, 59, 69, 70, 71, 202, 205,
206, 207, 209, 216, 219, 247, 249, 252,
298, 299, 300, 302, 303, 304, 305, 307,
309, 310, 315, 317, 318, 320, 321, 322,
323, 324, 326, 328, 335, 336, 379, 431,
438, 439, 440, 441, 442, 443, 444, 445,
447, 448, 451, 452, 457, 460, 462, 465,
468, 471, 474, 476

Instagram 23, 25, 67, 186, 199, 200, 209,
217, 219, 220, 221, 300, 304, 306, 307,
308, 315, 316, 318, 319, 321, 322, 330,
332, 335, 466, 476

inteligencia artificial 79, 81, 85, 86, 88, 90,
91, 98, 99, 104, 105, 106, 107, 108,
110, 111, 112, 115, 117, 118, 119, 122,
123, 124, 125, 126, 127, 130, 132, 133,
134, 135, 141, 143, 145, 146, 223, 225,
228, 229, 230, 235, 237, 238, 239, 240,
241, 242, 243, 391, 393, 477

inteligência artificial 18, 24, 28, 33, 53, 58,
69, 70, 72, 74, 148, 152, 154, 160, 161,
163, 164, 167, 168, 169, 182, 203, 204,
205, 206, 208, 213, 215, 216, 219, 252,
253, 255, 256, 261, 262, 267, 268, 279,
281, 301, 308, 309, 311, 320, 324, 337,
342, 348, 354, 355, 368, 369, 371, 372,
374, 375, 386, 388, 434, 436, 438, 441,
442, 443, 444, 445, 446, 447, 450,
451, 452, 454, 460, 469, 472, 473,
475, 476, 477

inteligencia artificial generativa 107, 130,
223, 225, 228, 230, 235, 238, 239,
242, 391

inteligência artificial generativa 58, 203, 204,
215, 216, 262, 337, 348, 355, 368, 454

interativo 374

interativos 263, 279, 298, 377

L

linguagem 24, 64, 69, 157, 168, 206, 214, 217, 256, 257, 274, 299, 301, 322, 330, 337, 338, 339, 340, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 360, 361, 362, 363, 364, 365, 366, 367, 368, 371, 372, 373, 375, 378, 380, 452, 468

linguagem simples 337, 338, 339, 340, 342, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 360, 361, 362, 363, 364, 365, 366, 367, 368, 371

M

mídia 19, 22, 41, 52, 71, 75, 176, 177, 178, 196, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 215, 217, 218, 219, 272, 274, 278, 293, 294, 308, 315, 374, 435, 461, 473

mediática 27, 39, 61, 73, 191, 297

mediáticas 184, 186, 193, 210, 271, 319, 448

mediático 14, 63, 175, 177, 190, 196, 199, 218, 255, 267, 292, 299, 301, 306, 338, 347, 348

mediáticos 18, 193

modulação 35, 36, 37, 38, 39, 44, 45, 46, 47, 48, 49, 50, 55, 56

O

OpenAI 18, 28, 29, 32, 72, 73, 76, 216, 220, 274, 277, 295, 308, 319, 320, 321, 333, 343, 344, 370, 371, 373, 375, 380, 393, 429, 445

P

plataforma 29, 44, 84, 115, 116, 123, 137, 138, 186, 218, 226, 255, 277, 294, 303, 308, 309, 311, 312, 314, 320, 323, 325, 326, 329, 432, 437, 452, 456, 457

plataformas 19, 20, 21, 26, 35, 36, 37, 38, 44, 45, 46, 47, 50, 51, 54, 63, 64, 65, 66, 67, 68, 83, 93, 100, 102, 109, 112, 113, 114, 118, 120, 131, 133, 137, 138, 140, 165, 174, 175, 176, 178, 179, 180, 181, 182, 183, 184, 185, 187, 188, 189, 192, 194, 195, 196, 197, 198, 199, 200, 214, 235, 236, 247, 248, 250, 257, 277, 280, 288, 299, 300, 303, 304, 306, 309, 315, 319, 328, 329, 330, 334, 338, 433, 440, 443, 456

prompts 24, 231, 234, 236, 281, 288, 308, 320, 356, 366, 380, 381, 428, 447, 451, 452, 458, 460, 466, 468

R

regulación 130, 131, 134, 137, 139, 143, 242

S

Satiko 174, 175, 176, 178, 179, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 209, 221

Slow Living 13

T

TDIC 267, 280, 284

textos 5, 14, 24, 69, 92, 153, 205, 206, 225, 231, 238, 242, 247, 249, 254, 276, 277, 300, 337, 338, 341, 342, 344, 345, 351, 355, 356, 366, 367, 373, 374, 375, 377, 379, 383, 385, 386, 435, 447

textos jornalísticos 374, 379

textos jurídicos 337, 344, 356

TikTok 21, 109, 115, 116, 124, 126, 127, 213, 304, 306, 308, 315, 317, 318, 319, 321, 322, 323, 336, 431, 432, 437, 441, 442, 443, 445, 466

Trump 107, 108, 109, 111, 113, 115, 120, 122, 123, 124, 125, 127

Twitter 84, 128, 145, 201

U

ultrarrealista 455

ultrarrealistas 451, 452, 453, 455, 458, 461,
462, 463, 464, 465, 469, 470, 471,
473, 474

V

video 77, 81, 83, 109, 121, 125, 136, 137,
457, 462

video 215, 324, 330, 335, 451, 452, 455,
456, 459, 466, 467

videos 81, 83, 109, 115, 124, 136, 331

videos 20, 21, 23, 24, 30, 59, 69, 70, 133,
192, 205, 206, 216, 254, 307, 315, 317,
331, 333, 352, 379, 440, 441, 442, 443,
445, 450, 451, 452, 453, 455, 456, 457,
458, 460, 461, 462, 463, 464, 465,
466, 468, 469, 470, 471, 472, 473, 474

X

X 32, 67, 104, 129, 137, 172, 294, 358, 361

Y

YouTube 83, 137, 304, 307, 329, 331, 466,
467, 475

RIA

Editorial