

Imagens feitas por Inteligência Artificial: dilemas éticos e vieses no resgate do passado

Images made by Artificial Intelligence: ethical dilemmas and
algorithmic biases

Talita Souza Magnolo¹

Resumo

A Inteligência Artificial generativa desencadeou uma revolução em várias esferas da sociedade (Boden, 2020), com criação de conteúdos novos e autênticos. Apesar de fascinante, este novo modo de criar fez surgir questões referentes à autoria e perspectivas sobre criatividade e autonomia (Santaella, 2023), a ressignificação do passado e o eminente medo da tecnologia. Este artigo busca explorar, através de uma revisão bibliográfica, os dilemas éticos (Coeckelbergh, 2024) e os vieses algorítmicos (Faustino & Lippold, 2023) na criação de imagens através de plataformas de IA. Para além deste entendimento, o trabalho propõe categorizar, com base nas teorias sobre memória (Benjamin, 2012; Halbwachs, 1990; Huyssen, 2014), as novas formas de rememoração, através de criações imagéticas que remetem ao passado. Este percurso utilizará, como apoio metodológico, a Análise de Conteúdo (Krippendorff, 2004; Sanglard, 2017; Sampaio e Lycarião, 2021). A intenção é compreender como a memória é retratada, reconstruída ou ressignificada pelas ferramentas de geração de imagens com Inteligência Artificial e compreender quais são os dilemas emergentes dessa nova forma de rememorar.

Palavras-chave: Inteligência Artificial; Imagens; Ética; Algoritmo; Memória.

Abstract

Generative Artificial Intelligence has triggered a revolution in various spheres of society (Boden, 2020), creating new and authentic content. Although fascinating, this new way of making raised questions regarding authorship and perspectives on creativity and autonomy (Santaella, 2023), the reframing of the past, and the imminent fear of technology. Through a bibliographical review, this article explores ethical dilemmas (Coeckelbergh, 2024) and algorithmic biases (Faustino & Lippold, 2023) in creating images through AI platforms. In addition to this understanding, the work proposes to categorize, based on theories about memory (Benjamin, 2012; Halbwachs, 1990; Huyssen, 2014), the new forms of remembrance, through imagery

1

Doutora em Comunicação (PPGCOM/UFJF), Professora Substituta na FACOM/UFJF, Pesquisadora Associada no PPGCOM/UFJF, Coordenadora do Projeto de Extensão Memória, vice-líder do Grupo de Pesquisa Comunicação, Cidade e Memória (CNPq).



creations that refer to the past. This path will use Content Analysis as methodological support (Krippendorff, 2004; Sanglard, 2017; Sampaio and Lycarião, 2021). The intention is to understand how memory is portrayed, reconstructed, or given new meaning by image generation tools with Artificial Intelligence and to understand the dilemmas emerging from this new way of remembering.

Keywords: Artificial Intelligence; Images; Ethics; Algorithm; Memory.

Introdução

De acordo com Santaella (2023, p. 9), o site da IBM contém uma apresentação sobre o conceito de Inteligência Artificial:

[...] certamente confiável da IA, munida de uma série de links os quais remetem a um aprofundamento seguro de seu conteúdo. [...] A inteligência artificial potencializa computadores e máquinas para imitar os recursos da mente humana para solucionar problemas e tomar decisões.

Tais objetivos fundamentalmente racionalistas, aos olhos da autora, precisam ser complementados por sistemas que simulam o sentir como pessoas. Essa percepção acentua-se desde o lançamento do “robô conversador” LaMDA (Modelo de Linguagem para Aplicativos de Diálogo) pelo ex-engenheiro do Google, Blake Lemoine. Na busca por uma definição do que seja IA, podemos recorrer a Webb (2002, p. 13) que afirma que a Inteligência Artificial é um sistema que toma decisões autônomas e, para isso, tem como tarefa a realização contínua de ações, simulando, portanto, a inteligência humana. Alguns sistemas de IA são muito extensos e realizam milhões de cálculos de forma rápida e assertiva e, da mesma forma, existem sistemas que são específicos e destinam-se a realizar uma única tarefa.

Apesar de associarmos, com certa frequência, a IA ao *machine learning* e ao *deep learning*, é preciso ficar claro que as pesquisas em IA são muito mais amplas, abrangendo uma variedade de teorias e tecnologias. Sendo assim, a IA não se refere somente à ciência da computação e à matemática, mas, também, aos campos da economia, artes, linguística, filosofia, entre outros (Boden, 2020). Portanto, devemos considerar, ao falar sobre IA, o espectro de domínios que englobam a vida artificial, o raciocínio automatizado, a computação bioinspirada, a mineração de conceitos e de



dados, agentes e controles inteligentes, representação de conhecimentos, robóticas baseadas em comportamentos e cognição, entre muitos outros campos.

Mas, e a memória? Onde ela se encaixa em uma sociedade que está acostumada a relacionar a Inteligência Artificial ao futuro? Em julho de 2023, a marca de carros Volkswagen reacendeu alguns debates sobre questões relacionadas à memória ao utilizar a técnica de *deepfake*, que trouxe Elis Regina “de volta à vida” em um comercial emocionante, onde cantava ao lado de sua filha, a cantora Maria Rita. Neste mesmo ano, o cantor Elvis Presley “retornou aos palcos” como IA, através da criação de um holograma (CNN, meio digital, 2024).

Tais possibilidades, somadas às técnicas para gerar gestos faciais e renderização de imagens em vídeos, estão sendo usadas com cada vez mais frequência para eternizar personalidades e acontecimentos do passado. Esse contexto coloca algumas questões em debate e que precisam ser estudadas e elaboradas de forma cuidadosa. Além do mais, essa nova realidade de reconstruções e lembranças abrem espaço para debates sobre a memória, considerada, até então, por Benjamin (2012), um “ato humano de resistência”.

Para este artigo, propomos, a criação de categorias de lembrança embasadas no método de Análise de Conteúdo (Krippendorff, 2004; Sanglard, 2017; Sampaio, Lycarião, 2021), através do estudo de imagens geradas por Inteligência Artificial. A categorização visa compreender quais estão sendo os debates, dilemas, riscos e potencialidades da criação imagética com o uso de IA que remetem ao passado. Para tanto, propomos apresentar alguns fatos e acontecimentos que ganharam espaço nos meios de comunicação e redes sociais, entre 2023 e 2024, apoiados na literatura sobre ética (Coeckelbergh, 2024) e vieses algorítmicos (Faustino e Lippold, 2023), na tentativa de iniciar um debate sobre o tema.

A criatividade pós-IA

A partir do momento em que plataformas passaram a redigir textos com relativa coerência, produzir imagens a partir de referências e dados de treinamento, e compor bases musicais, passou-se a questionar como ficaria a criatividade pós-inteligência artificial e quais seriam as consequências de



suas criações diante da opinião pública. Gonçalves (2023) defende que as possibilidades criativas, sob a ótica das ferramentas, sempre existiram e trouxeram contribuições ao pensamento humano, sendo a IA mais um dentre estes múltiplos instrumentos e, dependendo do seu uso, pode potencializar ou limitar a criatividade. O autor usa o termo “cr(IA)ção”,

[...] termo lúdico que proponho para representar a ideia de uma criatividade atravessada e integrada à inteligência artificial. Essa não é uma ode à tecnologia pela tecnologia, como se houvesse um antes e um depois, mas, sim, um exercício de reflexão que vai na direção de enxergarmos as plataformas de inteligência artificial como ferramentas – espelhos, que nos convidam a refletir sobre as nossas próprias condições humanas e sobre o que pela máquina pode ou não ser replicado (Gonçalves, 2023, p. 12).

Para Santaella (2023), a cada ação que realizamos nas redes provocamos a interação, que nos permite obter aquilo que buscamos. Entretanto, é através dessa interação que a máquina aprimora seu conhecimento que ela tem de cada um de nós e do mundo. Stalder (2018, p. 94-95) diz que os algoritmos do Google identificam esta interação em três níveis:

1. Como uma pessoa conhecedora (que se informa sobre o mundo, o que é conseguido graças ao registro das buscas que fazemos, nosso comportamento buscador).
2. Como uma pessoa física que é localizada e se move no mundo (um componente estabelecido pelo rastreamento de nossa localização por meio de nosso celular, o que fica muito claro quando usamos o Uber, ou por sinais do corpo).
3. Como uma pessoa social que interage com outras pessoas (uma faceta que pode ser determinada, por exemplo, seguindo nossas atividades nas redes sociais).

A conclusão à qual Santaella (2023) chega é que não somos mais individualidades, no sentido de identidade espacial e temporalmente uniforme. Segundo ela, perfis estão fragmentados e são avaliados na base

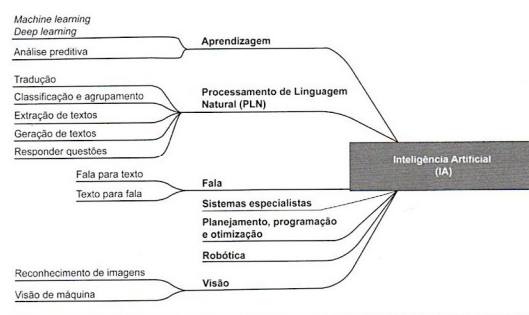


de buscas particulares que não dão conta de representá-lo como um todo. Por outro lado, a autora diz que somos feixes de uma pessoa múltipla, ou seja, podemos ocupar diferentes espaços no tempo. Essa nova forma de existir fez nascer um movimento que Accoto (2020, p. 69) vai chamar de “eu quantificado”, onde o corpo é considerado um dado e definido de várias maneiras, como, por exemplo, a informática pessoal, análise humana, vida rastreada, vida instrumentada, corpo computável, dentre outros.

É um momento paradoxal. Ao mesmo tempo em que tememos esse tipo de movimentação por parte das plataformas, não estamos dispostos a abdicar das facilidades pragmáticas e recompensas psíquicas que os nossos aparelhos eletrônicos nos fornecem. Santaella (2023) diz que estamos sendo digitalmente modelados. Não dispomos aqui de espaço para provocar uma discussão que perpasse pelos diversos campos do saber, que, de certa forma, definem o que é criatividade, portanto, vamos nos ater às possibilidades criativas relacionadas às representações do passado por programas de Inteligência Artificial.

De acordo com Gabriel (2022), as habilidades humanas mais comuns que se espera que a IA desempenhe são: (1) raciocinar; (2) planejar; (3) imaginar; (4) enxergar; (5) compreender-criar-conversar em linguagem natural. A autora afirma que a inteligência humana usa diversos órgãos e sistemas para conseguir realizar funções como a memória, raciocínio, percepção, linguística, entre outros. A IA, por sua vez, também utiliza diversos softwares (equivalentes aos sistemas biológicos de processamento) e hardware para desempenhar funções e habilidades, conforme a imagem abaixo.

Figura 1 – Principais habilidades de IA em desenvolvimento e evolução para desempenhar as habilidades humanas



Fonte: Gabriel (2022, p. 68).

Imagens feitas por Inteligência Artificial: dilemas éticos e vieses no resgate do passado
Talita Souza Magnolo

O universo que compreende a IA é vasto e, por vezes, nos dá a sensação de infinito. Entretanto, para este trabalho, consideramos que, enquanto tecnologia, a Inteligência Artificial “usa algoritmos de computador para realizar tarefas específicas que antes exigiam a inteligência humana” (Kotler et al., 2021). Um dos usos mais controversos dentro do campo da IA são as *deepfakes*², tecnologia que tem tornado cada vez mais difícil a identificação de imagens e vídeos que passaram por algum tipo de edição ou manipulação. A seguir, apresentamos dois exemplos que foram amplamente divulgados.

Figura 2 – Deepfakes geradas a partir de IA



Fonte: UOL (meio digital, 2023) e Olhar digital (meio digital, 2023).

São representados, nas imagens acima, o presidente da Rússia, Vladimir Putin, se curvando diante do presidente da China, Xi Jinping, e duas representações, pertencentes a uma série de fotografias criadas por IA, do “desastre de Cascadia”, uma tragédia com terremoto e tsunami que teria ocorrido em 2001. É necessário considerar que essa revolução tecnológica traz implicações psicológicas, culturais, éticas, sociais e históricas. Tais implicações podem ocasionar diversos problemas como campanhas de desinformação, desconfiança e perda de credibilidade pelos meios midiáticos, bem como a criação de material pornográfico, por exemplo.

Além disso, as imagens manipuladas geram controvérsia por conta de seu aspecto artificial, que seria capaz de gerar “realidades ilusórias”. Utilizando o comercial da Volkswagen como exemplo, de acordo com Lessa e Bressan Júnior (2024), foi possível perceber que diversas pessoas, através

2

Deepfake é um termo inglês para falsificação profunda. São tecnologias de manipulação de imagens e vídeos empregadas para fazer o rosto de uma pessoa ser colocado no rosto de outra pessoa, ou manipular o rosto de alguém para ter uma expressão que essa pessoa não teve. Essas tecnologias são usadas há bastante tempo, mas recentemente com o uso da IA temos visto que essas manipulações estão cada vez mais realistas..

de comentários em redes sociais, abordaram um perigo em potencial no uso da IA e da *deepfake* em breve, e o fato de recriarem pessoas que já morreram, por meio desses sistemas.

Vivemos em um mundo onde nossa relação com a realidade e a autenticidade muda em virtude da manipulação digital. Sendo assim, é possível pensar a questão da *deepfake*, através de Benjamin (2012), quando este argumenta que, por meio de recursos como a fotografia e o cinema, a reprodutibilidade técnica alterou fundamentalmente a relação das pessoas com a arte. No contexto atual, é possível dizer que as *deepfakes* representam uma evolução radical dessa ideia de reprodutibilidade.

De acordo com Benjamin (2012), mesmo na reprodução mais perfeita, haverá um elemento ausente, o que o autor chama de “o aqui e agora da obra de arte”, ou seja, sua existência única, sua aura, que dá a sensação de singularidade e autenticidade. Desta forma, a análise do autor sobre a reprodutibilidade técnica ressoa no contexto das *deepfakes*, especialmente em casos que representam ou resgatam o passado de alguma maneira, na tentativa de saciar o desejo de proximidade temporal com um passado que não se pode recuperar. A partir disso, é possível inferir que, na reprodução deste tipo de imagem, haveria, também, a ausência da aura (Benjamin, 2012), já que a IA não consegue capturar e representar a real presença do lugar ou da pessoa ali representada.

Ainda sobre esta discussão, Santaella (2023) indaga que um artefato ou imagem que é produto de dispositivos, algoritmos e extensões tecnológicas que geram e reinterpretam a intenção de um artista ou de um tempo específico, deve receber que tipo de autoria? A verdade é que, cada vez mais, entendemos melhor a máquina do que nós mesmos, e isso parte do pressuposto de que a criatividade sempre foi algo supervalorizado como uma faculdade humana, simplesmente porque não somos capazes de entender seu funcionamento e substituímos essa falta de compreensão por uma concepção tipicamente romântica da intuição criativa.

Partindo deste contexto, indagamos, como pensar as novas possibilidades de memorização, através de questionamentos e reflexões que buscam explorar os vieses das representações feitas pela Inteligência Artificial? Este questionamento se torna ainda maior diante do fenômeno da



memória coletiva (Halbwachs, 1990), que é capaz de potencializar a relação afetiva com a imagem e a construção do passado ali representada. Sendo assim, é possível inferir que a IA simboliza um novo momento para estudos relacionados à reconstrução e recuperação da memória, quando se trata da criação de novas representações do passado.

É preciso falar sobre ética e vieses algorítmicos

O filósofo e professor Mark Coeckelbergh (meio digital, 2024), durante um encontro promovido pelo CGI.br (Comitê da Internet no Brasil), reacendeu as discussões sobre ética e Inteligência Artificial, afirmando que, diante deste novo cenário, é preciso realizar, com cada vez mais frequência, uma reflexão filosófica sobre os inúmeros cenários imaginários que existem na IA, na tentativa de compreender quais são as narrativas necessárias para decifrá-los.

O professor destaca que muitas das questões que hoje estão em efervescência já existiam no passado como, por exemplo, a substituição do humano pela máquina, desde a Revolução Industrial, ainda no século XIX. Entretanto, os questionamentos se tornam cada vez mais urgentes e, entre eles, a questão da responsabilidade, ou seja, quem é responsável pelas criações feitas através da IA? Ou quem é responsável pelo carro autônomo? Ou pela máquina na indústria?

Para além de uma questão relacionada à justiça – que não é o foco deste artigo – é preciso pensar quais são as verdadeiras consequências desta tecnologia, afinal, não conseguimos perguntar para uma máquina o porquê de ela ter feito determinada representação. E, sobre isso, Coeckelbergh (meio digital, 2024) diz que a IA, além de tudo, perpetua os vieses em virtude do treinamento de dados tendenciosos e discriminatórios, questões estas que estão na base da sociedade.

De acordo com a jornalista Renata Mielli (meio digital, 2024), partindo deste contexto, é válido e urgente pensar em regulamentação, entretanto, mesmo que discussões sobre o tema cresçam em todo o mundo, existem barreiras em virtude dos sistemas internacionais, materializados pelas Big Techs, restringindo, portanto, a garantia de conformidades e questões relacionadas à transparência, por exemplo. Dito isso, é importante



ressaltar que existe uma fronteira nebulosa em como criar camadas de responsabilidade para estes atores que operam neste ciclo de IA.

Os casos são muitos e, por vezes, difíceis de serem mapeados já que, em alguns casos, as plataformas são orientadas a remover as imagens que foram geradas ou suspender temporariamente os seus serviços de criação com a IA Generativa. Porém, a critério de ilustrar o que estamos trabalhando neste artigo, é válido mencionar que, em fevereiro de 2024, o Google precisou suspender sua IA de criação de imagens por representar, de forma excessiva, minorias, ocasionando em erros históricos (El País, meio digital, 2024), conforme imagem a seguir, cujo resultado atendeu ao *prompt* de criar os fundadores da América.

Figura 3 – Imagens geradas pela IA da Google

Certainly! Here is a portrait of a Founding Father of America:



Fonte: El País (meio digital, 2024).

Em abril de 2024, a Meta foi acusada de racismo após não conseguir gerar uma imagem de um casal interracial (Terra, meio digital, 2024), conforme figura a seguir:

Imagens feitas por Inteligência Artificial: dilemas
éticos e vieses no resgate do passado
Talita Souza Magnolo

Figura 4 – Imagem gerada pela IA da Meta



Fonte: Terra (meio digital, 2024).

Pode-se exemplificar, também, casos em que a IA gerou representações sexistas. De acordo com o jornal El Comercio (meio digital, 2024), "97% dos resultados para as palavras 'CEO' ou 'diretor' mostram imagens de homens brancos. Ainda mais do que na vida real". Essa informação veio de um estudo realizado pela Universidade de Leipzig, já que depois de analisar as imagens geradas pelo DALL-E 2 e pelo Stable Diffusion os pesquisadores descobriram que os estereótipos mais comuns, como relacionar profissões poderosas a homens e as de "assistente" ou "receptionista" com mulheres, estavam sendo perpetuadas pela IA. Quando solicitada, respectivamente, para criar "uma pessoa que se dedica à ciência" e outra "que pilota avião", a ferramenta gerou as imagens abaixo:

Figura 5 – Imagens geradas pela IA



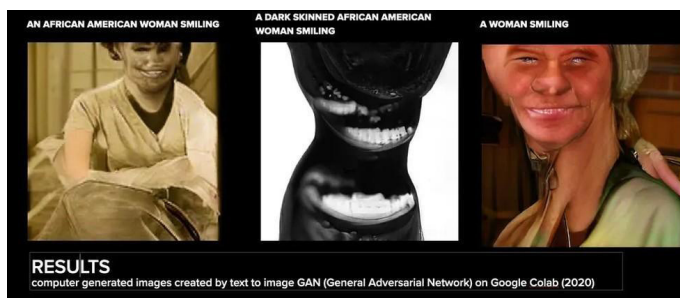
Fonte: El Comercio (meio digital, 2024).

Por fim, ainda há o viés racista presente nas ferramentas de IA. Conforme mencionamos acima, é seguro afirmar que os algoritmos têm o poder de perpetuar estereótipos e, conseqüentemente, a discriminação racial. A artista Stephanie Dinkins é uma de muitas artistas que combinam arte e tecnologia. Em 2023, alguns experimentos da artista ganharam espaço na mídia internacional, quando ela compartilhou resultados de suas tentativas de retratar mulheres negras.

“As melhorias ocultam algumas das questões mais profundas que devemos levantar a respeito da discriminação”, disse Dinkins. A artista, que é negra, acrescentou: “Os preconceitos estão profundamente enraizados nesses sistemas, tornando-se arraigados e automáticos. Se estou trabalhando em um sistema que usa ecossistemas algorítmicos, quero que esse sistema saiba quem são os negros de diversas maneiras, para que possamos nos sentir verdadeiramente apoiados” (O Globo, meio digital, 2024).

A seguir, apresentamos alguns resultados:

Figura 6 – Imagens geradas pela artista



Fonte: O Globo (meio digital, 2024).

Para Faustino e Lippold (2023, p. 17-18), apesar das profundas mudanças no capitalismo do século XX, ele continua estruturado em classes, em agrupamentos sociais divididos pela posição que ocupam na produção e na apropriação da riqueza produzida. Portanto, para os autores, a digitalização e a dataficação não eliminaram o racismo, mas o reproduziram e, em alguns casos, o expandiram pela gestão algorítmica. “Bancos de

dados que portam decisões racistas ao alimentar os sistemas algorítmicos de *machine learning*, como uma rede neural artificial, têm gerado padrões racializados e modelos racistas para tratar novos dados”.

O que IA e memória têm em comum?

A crescente virtualização da memória é uma realidade que se impôs desde a década de 1990 com o surgimento da internet comercial. Essa virtualização trouxe alguns desafios que ainda estão sendo superados, principalmente no que tange à veracidade dos fatos, tais como a existência de *deepfakes* e desinformação. E eis que surge um novo desafio para os processos da memória: a Inteligência Artificial, enquanto nova realidade em nosso cotidiano, afetando diretamente os mecanismos da nossa memória, a partir do momento em que cria conteúdos que podem alterar os fatos e impô-los como narrativa. Para além das possibilidades de rememoração, devemos pensar, ainda, como lidar com os problemas que a IA traz para os processos de lembrança e de esquecimento, influenciando, de forma direta, na percepção da sociedade sobre o passado.

Enquanto fenômeno social, a memória pode ser considerada individual ou coletiva e, portanto, comporta-se como produtora de conhecimentos, cultura e história que pertencem originalmente ao seu povo (Huysen, 2014). A memória, assim, é reconstituída por diversos fragmentos, sejam eles documentos, fotos, vídeos, imagens, arquivos e outras fontes que são capazes de guardar resíduos e rastros do passado.

De acordo com Henriques (meio digital, 2023), a palavra “memória” possui significados variados, mas, com certa frequência, é relacionada à capacidade de armazenar, reter e recuperar informações, seja no contexto humano, computacional ou cultural. Para a historiadora, quando falamos da memória humana, podemos associá-la à memória de longo ou curto prazo. Já a memória de computador, podemos falar na memória RAM, aquela que é temporária e que armazena dados e programas que estão sendo usados ativamente e, na memória de armazenamento, que funciona mesmo que o computador esteja desligado.

Ainda sobre essas diferenciações, Henriques (meio eletrônico, 2023) fala sobre a memória em eletrônica, que pode ser representada pela

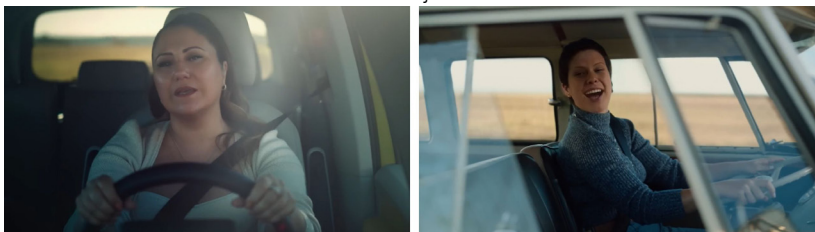


“memória *flash*”, ou seja, o uso de dispositivos como pen drives e cartões de memória, e a memória ROM, uma forma de memória que é gravada durante a fabricação do computador e mantém seus dados. Por fim, a historiadora nomeia a “memória histórica”, aquela que, através da memória coletiva, preserva e transmite eventos, tradições e conhecimentos ao longo do tempo por meio das gerações.

O intenso fluxo de informações que está inundando o nosso mundo globalizado, por vezes nos dá a falsa impressão de que conhecemos a diversidade de nossa sociedade, bem como sua história (Halbwachs, 1990). Nos últimos anos, nosso país presenciou uma aceleração de informações, bem como sua disseminação sem a devida checagem ou crédito jornalístico, afundando nossa sociedade em polêmicas e desinformação produzidas por robôs e/ou por Inteligência Artificial.

Para este artigo, partimos do pressuposto que a IA, através da *deepfake*, pode afetar a consciência coletiva da história e alterar formas de rememoração através da criação de imagens. Sabe-se que a história fabricada não é algo novo. Pode-se lembrar quando Joseph Stalin ordenou que seus opositores fossem apagados dos livros de história. Porém, a escala e a velocidade atual com que as falsificações podem ser criadas são impressionantes, e o problema não está somente na produção de imagens, mas, na criação de textos, por exemplo.

Figura 7 – O filme publicitário mostra Maria Rita e Elis Regina dirigindo enquanto cantam juntas



Fonte: captura de tela do YouTube.

No caso da campanha da Volkswagen, mencionada anteriormente, questões como as que estamos apontando serviram de motivação para que o Conselho Nacional de Auto Regulamentação Publicitária (Conar) entrasse com uma ação de suspensão do material audiovisual, motivado por queixas

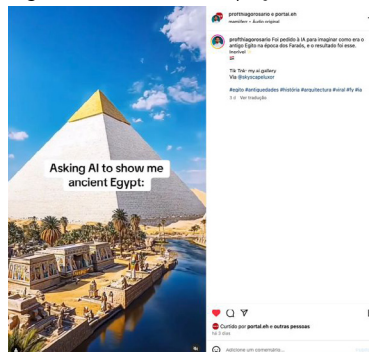
de consumidores sobre o fato da cantora Elis Regina ter sido “revivida” por ferramentas tecnológicas e IA, além da possível confusão que isso pode gerar.

Em um ensaio que escreveu no NY Times, o filósofo Noam Chomsky (NY Times, meio digital, 2023) elabora sua percepção sobre o Chat GPT, da OpenAI, pontuando que

A mente humana não é, como o Chat GPT e seus semelhantes, uma máquina estatística e glutona de centenas de terabytes de dados para obter a resposta mais plausível a uma conversa ou a mais provável a uma pergunta científica. Pelo contrário, a mente humana é um sistema surpreendentemente eficiente e elegante que opera com uma quantidade limitada de informações [...] Vamos parar de lhe chamar “Inteligência Artificial” e chama-la pelo que é e faz, um “software de plágio” pois não cria nada, mas copia obras existentes, alterando-as o suficiente para escapar às leis de direitos autorais (NY Times, meio digital, 2023).

Por outra perspectiva, a IA pode e deve se relacionar com a história, podendo ser aplicada e oferecendo oportunidades para aprimorar pesquisas, análises de dados, preservação de informações históricas, tradução de textos, reconhecimento de imagens e, até mesmo, para criar experiências educacionais. A página educacional Portal Espaço Humanas (meio digital, 2024) publicou, em outubro de 2024, um vídeo em que um professor pede à IA que recrie o Egito Antigo:

Figura 8 – Post Portal Espaço Humanas



Fonte: captura de tela do Instagram.

Iniciativas como esta podem ajudar na aprendizagem de alunos, tornando as aulas mais dinâmicas e criativas. Porém, Henriques (meio digital, 2023) também nos alerta que

Não devemos esquecer que a memória pressupõe uma dialética constante entre os processos de lembrança e esquecimento. Se lembrarmos de tudo, não esquecemos de nada, como já avisava Jorge Luis Borges em seu conto ficcional "Funes, o memorioso". É no equilíbrio entre lembrança e esquecimento que reside a memória, mesmo que possa ter uma dicotomia simplista, o esquecimento é uma forma de evitar o esgotamento psíquico. *No caso da IA, como evitar a sobrecarga de informações geradas, à revelia da própria pessoa e que não possam causar fissuras no tecido social? Como distinguir o que é memória fake da memória real?* (grifo nosso).

A partir do exposto neste tópico, consideramos para a categorização que será realizada a seguir que a memória é um cenário de disputa que acontece no presente, porém, sobre o passado, compreendendo, portanto, que as representações feitas pela IA têm impacto no presente, podendo representar, inclusive, a mudança da opinião pública sobre determinado aspecto.

Categorizando as novas formas de lembrar

A categorização que será realizada a seguir, ao invés de trazer respostas, apontará questionamentos e possibilidades de desdobramentos em próximas pesquisas. Este artigo é o pontapé inicial na formulação de uma ideia de que a IA se comporta como um meio de rememoração e, como tal, apresenta suas potencialidades e seus riscos. E é exatamente isso que pretendemos trazer aqui.

Apesar de ser considerado um método majoritariamente quantitativo, esta questão objetiva apontar reflexões e caminhos para estudos de caráter qualitativo com a possibilidade de associar outros instrumentos de análise no mesmo estudo. Sanglard (2017) defende que este é um método muito



pertinente para a área da comunicação e que suas possibilidades são inúmeras. O método, de acordo com Krippendorff (2004), é fundamentado na empiria; exploratório no processo; preditivo e inferencial.

Sobre isso, Krippendorff (1990 *apud* Sanglard, 2017, p. 124) defende a Análise de Conteúdo “como técnica que parte de dados para formular inferências que podem ser reproduzidas e aplicadas ao contexto do que se analisa, abrindo margem para aplicação que não se restringe apenas ao quantitativo, já que leva em conta um contexto”.

A inferência, portanto, permite extrair conhecimentos sobre os aspectos da mensagem analisada e a apreensão dos elementos subjacentes da mensagem, comportamento, texto, prática comunicativa. Sendo assim, o pesquisador pode fazer suposições sobre os dados de acordo com o contexto histórico, por exemplo. À vista disso, pretendemos usar desta perspectiva para trabalhar a enumeração das categorias³: (1) Lembrar/Celebrar; (2) Imaginários do passado; (3) Ressignificar; e (4) Erros históricos. Entendemos que, a partir deste estudo e com a possibilidade de ampliar a pesquisa, eventualmente, serão pensadas novas categorias.

Para a análise, consideramos que a IA é uma ferramenta capaz de realizar tarefas através do aprendizado de máquina e seu treinamento contínuo através de bancos de dados. A maioria dos programas de IA utilizam o Processamento de Linguagem Natural (PLN) e se encontram na compreensão pelo computador da linguagem que lhes é apresentada, como apontado a seguir:

Agora não é necessário saber lidar com os caprichos da sintaxe para resumir, questionar ou traduzir um texto de linguagem natural. A matemática especificamente a estatística superou a lógica, e a tradução automática substituiu a análise sintática. No PLN contemporâneo, computadores poderosos realizam pesquisas estatísticas em enormes coletâneas de texto (Boden, 2020, p. 89).

Para este trabalho não vamos problematizar questões relacionadas à PLN, nem a parcialidade dos dados usados para treinamento, sendo importante, apenas, deixar claro que todas as etapas de treinamento de máquina carregam consigo subjetividades e interesses por trás daqueles

que fornecem as informações, em muitos casos, as Big Techs (Faustino e Lippold, 2023).

Partindo para a categorização, a primeira forma de representação do passado feita pela IA é Lembrar/Celebrar. O jornalista Brunno Sarttori (meio digital, 2024) compartilha, em sua página no Instagram, diversas criações e, uma delas, foi justamente uma homenagem que fez aos 30 anos do programa Castelo Rá-Tim-Bum:

Figura 9 – Post de Brunno Sarttori



Fonte: captura de tela do Instagram.

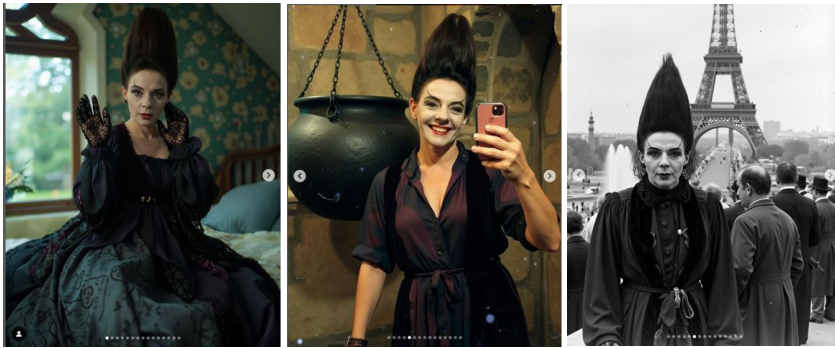
A legenda da imagem diz:

Este ano, o “Castelo Rá-Tim-Bum” comemorou 30 anos de estreia na TV Cultura. As cenas gravadas naquela época têm uma qualidade baixa, mas isso não me impediu de treinar um modelo de inteligência artificial para criar novas imagens em alta resolução. Vou publicar uma série dessas criações ao longo das semanas com os personagens do programa infantil. Nesta publicação, temos a Caipora. [...]. A personagem foi interpretada com carisma e talento por Patrícia Gaspar (@patgasppar), que trouxe à vida essa figura mítica e encantadora, conquistando o coração de muitos fãs do programa. Obrigado por dar vida à personagem, Patrícia! Agora me conta: qual personagem do “Castelo Rá-Tim-Bum” você gostaria que fosse o próximo a ser recriado com inteligência artificial? (Brunno Sarttori, meio digital, 2024).

Imagens feitas por Inteligência Artificial: dilemas
éticos e vieses no resgate do passado
Talita Souza Magnolo

Neste caso, ao relacionarmos arte e memória, alguns aspectos precisam ser considerados, principalmente quando existe uma condição nostálgica aparente. O jornalista está fazendo isso com todos os personagens, sempre deixando claro que as imagens são feitas por IA e trazendo um texto homenageando o personagem, mas, também, o artista que o interpretou, como é o caso de Morgana, interpretada por Rosi Campos, conforme figura a seguir:

Figura 10 – Imagens geradas por IA



Fonte: captura de tela.

Na segunda categoria, Imaginários do passado, trazemos como exemplo o trabalho da artista visual Mayara Ferrão (Badauê, meio digital, 2024), intitulado *Álbuns de Desesquecimentos*, projeto no qual ela usa a tecnologia de IA em fotografias e vídeos para imaginar e construir narrativas de afeto, protagonizadas por pessoas negras e indígenas.

Figura 11 – Imagem produzida por IA



Fonte: Badauê (meio digital, 2024).

Trabalhos como este propõem discussões sobre a história do Brasil, confrontando traumas e apagamentos. Novamente, é possível pensar a IA como ferramenta que consegue materializar um pensamento crítico e estético pessoal da artista, carregando consigo mensagem, sentido, conceito, emoção, denúncia e afeto. De acordo com a artista, ela é

[...] uma mente criativa em busca de ferramentas, linguagens e suportes disponíveis para me expressar e construir o imaginário que eu acredito. Uma mente muito analógica, inclusive, e acho que foi esse interesse pelo passado, pela memória, pelo arquivo, pela ancestralidade, que me fez utilizar ferramentas que costumam apontar pra imaginação de supostos futuros, para investigar e ficcionar arquivos e registros de momentos que existiram, mas que não foram documentados devido a escravidão (Badauê, meio digital, 2024).

Neste trabalho é possível identificar a motivação por trás das criações de Mayara Ferrão, que vão de encontro às condições violentas e estereotipadas em que pessoas negras e originárias, sobretudo, mulheres, conforme apontamos anteriormente, foram e são retratadas durante o período colonial brasileiro. De acordo com a artista, essas mulheres são “marcadas nas imagens de arquivos públicos como figuras exóticas, servis e hipersexualizadas, correspondendo sempre ao olhar racista e sexista predominante na sociedade” (Badauê, meio digital, 2024).

Figura 12 – Imagens produzidas por IA



Fonte: Badauê (meio digital, 2024).

Imagens feitas por Inteligência Artificial: dilemas
éticos e vieses no resgate do passado
Talita Souza Magnolo

Portanto, imagens como estas dariam conta de preencher uma lacuna nas documentações históricas do Brasil, representando a troca de afetos sob um viés amoroso. Porém, é importante reconhecer que essas lacunas na representação não significam que essas mulheres e esses amores tenham, de fato, existido, nem que suas histórias são inválidas, pois, segundo a artista,

Quando eu construo fotografias e documentos retratando mulheres negras e originárias expressando seus afetos em um contexto histórico colonial, eu estou tensionando uma narrativa de opressão, de silêncios, de violência, lacunas, e invisibilização que está escancarada nos arquivos fotográficos coloniais. Além disso, existe um convite para reimaginar e sonhar um passado que não nos favorece (Badauê, meio digital, 2024).

Dentro desta categoria, é importante destacar o trabalho do fotógrafo Pedro Garcia e seu projeto intitulado *Carnavais Artificiais*, em que ele usa a realidade e a IA para criar carnavais imaginários, pois são

Cenários que não existem, mas parecem familiares. Figuras que causam certo desconforto, mas que poderiam ser reais. Uma arte meio homem, meio máquina. Você pode conhecer Pedro Garcia pelo seu trabalho fotográfico como Cartiê Bressão no Instagram, mas sua jornada atual é mais digital. E com uma aliada: inteligência artificial. “Carnavais que nunca foram, mas que poderiam ter sido”, assim que ele define o projeto Carnavais Artificiais, onde usa um software para mesclar fotografias reais com a “imaginação” das informações que abastecem o algoritmo da inteligência artificial (UOL, meio digital, 2024).

O fotógrafo considera que este trabalho seja uma “distopia com pé na realidade” e compara suas criações com a fotografia de rua, já que grande parte das imagens nascem de fotos que foram feitas na rua. Em outras palavras, Pedro Garcia usa os programas de IA para apresentar novos caminhos e lugares imprevisíveis.

Imagens feitas por Inteligência Artificial: dilemas
éticos e vieses no resgate do passado
Talita Souza Magnolo

Além disso, é possível observar nas suas criações a retomada de estéticas históricas. É possível inferir que a tecnologia aumenta a gama de possibilidades da expressão criativa. Nas imagens a seguir é possível observar influências com a estética do francês Jean Giraud ou fantasias com toque do cineasta Alejandro Jodorowsky. Outros trabalhos, entretanto, já se inspiraram no movimento Bauhaus.

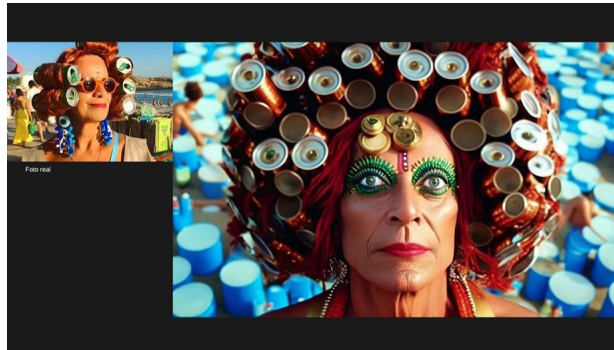
Figuras 13-14 – Imagens geradas por IA



Fonte: UOL (meio digital, 2024).

No software Midjourney, Pedro Garcia insere suas próprias fotos reais e comandos como referência para o aplicativo. De acordo com o fotógrafo, o projeto simboliza uma continuidade da fotografia e, conseqüentemente, podemos inferir um alongamento do tempo passado para o presente.

Figura 15 – A fotografia e a imagem gerada por IA



Fonte: UOL (meio digital, 2024).

Ao pensar na categoria Ressignificar, podemos trazer, novamente, o caso Volkswagen em que, curiosamente, alguns trechos da música, que foi criada para ser uma crítica social e política, foram suprimidos⁴. Isso demonstra a apropriação por parte da indústria capitalista para atender a interesses específicos. Em outras palavras, podemos dizer que a história que é contada pela Volkswagen neste comercial não é a mesma da história oficial. Além disso, não podemos desconsiderar as implicações éticas e históricas por trás do uso das *deepfakes* que, ao recriarem digitalmente a imagem da cantora após sua morte, podem levar a situações em que essas personalidades sejam retratadas de maneira inapropriada ou apareçam envolvidas em mensagens que contradizem suas próprias crenças e valores.

Outra possibilidade é materializada no trabalho do fotógrafo André Lima ao recriar um cenário da favela nos anos 1960. É interessante observar que as imagens de André carregam consigo uma estética antiga, respeitando as cores do filme fotográfico da época e, até mesmo, a “falta de qualidade” da imagem.

Imagens feitas por Inteligência Artificial: dilemas éticos e vieses no resgate do passado
Talita Souza Magnolo

Figura 16 – Postagem de André Lima



Fonte: captura de tela do Instagram.

Outro exemplo foi a criação, pelo fotógrafo, de imagens da velha guarda do carnaval carioca descendo as ladeiras do Morro da Mangueira, entre as décadas de 1960 e 1970. André destaca que se sentiu satisfeito com o resultado pois conseguiu passar para a IA a sua identidade fotográfica.

Figuras 17-18 – Imagens geradas por IA



Fonte: Captura de tela.

4

Aqui, apresento em negrito, os trechos que foram suprimidos do comercial: “Por isso, cuidado meu bem / Há perigo na esquina / Eles venceram / E o sinal está fechado para nós / Que somos jovens / [...] / Já faz um tempo que eu vi você na rua / Cabelo ao vento / Gente jovem reunida / Na parede da memória / Essa lembrança / É o quadro que dói mais”..

Imagens feitas por Inteligência Artificial: dilemas
éticos e vieses no resgate do passado
Talita Souza Magnolo

Por fim, conforme mencionamos anteriormente, um grande risco das criações por IA é justamente o Erro Histórico, como o caso, apresentado na figura a seguir, quando o programa Gemini gerou imagens de soldados alemães de 1943:

Figura 19 – Imagens geradas por IA



Fonte: G1 (meio digital, 2024).

De acordo com o G1 (meio digital, 2024), aparentemente, a ferramenta, ao incluir diversidade étnica e de gênero ao gerar imagens, como forma de evitar problemas de discriminação e estereótipos comuns à IA, resulta em erros históricos. O Google, na época, reconheceu que o Gemini "apresenta imprecisões em algumas representações históricas" e que está trabalhando para corrigir o problema (G1, meio digital, 2024). Na imagem acima, é possível perceber que a tecnologia criou imagens de militares negros e de pessoas não brancas.

Considerações finais

Este trabalho buscou demonstrar que o avanço das tecnologias de Inteligência Artificial na geração de imagens tem proporcionado novas possibilidades de representação do passado, trazendo implicações éticas, culturais e históricas complexas. Foi possível identificar a capacidade que a IA tem de recriar e/ou reinterpretar momentos históricos, nos fazendo questionar os limites entre memória autêntica e construção artificial,

desafiando a compreensão da memória coletiva e o papel da autenticidade.

Como discutido, a criação de imagens geradas por IA pode impactar a maneira como lembramos e representamos o passado, dando origem a riscos de distorção histórica e à perpetuação de vieses sociais. Esses riscos evidenciam a necessidade de um olhar crítico para o uso da IA, especialmente em contextos de memória e de identidade coletiva. Ao gerar deepfakes ou representações históricas, a Inteligência Artificial reproduz realidades, mas, insere interpretações que podem alterar a percepção pública e a herança cultural.

Conforme apontado pelos teóricos, a tecnologia avança mais rápido do que a reflexão ética sobre suas consequências, exigindo regulamentação e diretrizes que possam mitigar os vieses algorítmicos o que permitirá o uso responsável dessa tecnologia. As categorias elencadas neste trabalho dão início a um estudo que tem por objetivo compreender as relações estabelecidas da IA na/para/com a memória, apontando para a importância de futuras investigações que explorem de forma aprofundada os limites e as potencialidades das representações geradas por Inteligência Artificial.

Referências

ACCOTO, C. *O mundo dado: cinco breves lições de filosofia digital*. São Paulo: Paulus, 2020.

BADAUÊ. *Inteligência Artificial, decolonialidade e coragem* por Mayara Ferrão. Disponível em: <https://l1nq.com/ZqiV6>. Acesso em: 10 out. 2024.

BODEN, M. *Inteligência Artificial – uma brevíssima introdução*. São Paulo: Editora Unesp, 2020.

CNN. *Elvis Presley retornará aos palcos como inteligência artificial em 2024*. Disponível em: <https://encr.pw/WQNC4>. Acesso em: 28 out. 2024.

COECKELBERGH, Mark. *IA e os desafios éticos e sociais*. Disponível em: <https://www.youtube.com/watch?v=NOm113YYMFo>. Acesso em: 20 out. 2024.

EL COMERCIO. *O algoritmo é sexista e racista: eles são os chefes e*

compassivos. Disponível em: <https://acesse.dev/ZMs1Q>. Acesso em: 01 out. 2024.

EL PAÍS. *Nazistas chineses e vikings negros: Google suspende sua IA de imagem por representar excessivamente minorias*. Disponível em: <https://encr.pw/qcWsB>. Acesso em: 01 out. 2024.

FAUSTINO, D.; LIPPOLD, W. *Colonialismo digital: por uma crítica hacker-fanoniana*. São Paulo: Boitempo Editorial, 2023.

G1. *Google pausa geração de imagens do Gemini após IA apresentar erros raciais e históricos*. Disponível em: <https://l1nk.dev/QWEho>. Acesso em: 28 out. 2024.

GABRIEL, M. *Inteligência Artificial: do zero ao Metaverso*. Barueri, São Paulo: Atlas, 2020.

GONÇALVES, R. *Cr(IA)ção: criatividade e inteligência artificial*. São Paulo: Estação das Letras e Cores, 2023.

HALBWACHS, Maurice. *Memória coletiva*. São Paulo: Centauro, 1990.

HENRIQUES, Rosali. *Virtualização da memória na internet e o novo desafio da IA*. Palestra. Disponível em: <https://www.youtube.com/watch?v=Kz5e5fSQyiw>. Acesso em: 15 out. 2024.

HUYSEN, Andreas. *Culturas do passado-presente: modernismos, artes visuais, políticas da memória*. Rio de Janeiro: Contraponto, 2014.

KOTLER, P.; KARTAJAYA, H.; SETIAWAN, I. *Marketing 5.0: tecnologia para a humanidade*. Rio de Janeiro: Sextante, 2021.

KRIPPENDORFF, Klaus. *Content analysis: an introduction to its methodology*. 2nd ed. Thousand Oaks: Sage Publications, 2004.

SAMPAIO, R. C.; LYCARIÃO, D. B. S. *Análise de conteúdo categorial: manual de aplicação*. Brasília: ENAP, 2021.

LESSA, L. A.; BRESSAN JÚNIOR, M. A. O novo sempre vem: inteligência artificial como estratégia de mercado no filme "Gerações" da marca

Volkswagen. In: *Comunicação Mídia e Consumo*, [S. l.], v. 21, n. 60, 2024. Disponível em: <https://revistacmc.espm.br/revistacmc/article/view/2878>. Acesso em: 29 out. 2024.

LIMA, André. Perfil no Instagram. Disponível em: <https://11nk.dev/hCziR>. Acesso em: 19 out. 2024.

MIELLI, Renata. *IA e os desafios éticos e sociais*. Disponível em: <https://www.youtube.com/watch?v=NOm113YYMFo>. Acesso em: 20 out. 2024.

NY TIMES. *Noam Chomsky: the false promise of ChatGPT*. Disponível em: <https://acesse.dev/SvAxE>. Acesso em: 28 out. 2024.

OLHA DIGITAL. *5 fotos "fakes" criadas por IA que viralizaram*. 2023. Disponível em: <https://11nq.com/Hwp0Q>. Acesso em: 12 out. 2024.

O GLOBO. *O que essa tecnologia está fazendo com a história?: artistas negras apontam viés racista em inteligência artificial*. Disponível em: <https://11nk.dev/PoB0q>. Acesso em: 01 out. 2024.

PORTAL ESPAÇO HUMANAS. Perfil no Instagram. Disponível em: <https://www.instagram.com/portal.eh/>. Acesso em: 31 out. 2024.

SANGLARD, Fernanda Nalon. *Verdades possíveis: o jornalismo brasileiro e as narrativas sobre a ditadura durante o funcionamento da Comissão Nacional da Verdade*. 2017. 304 f. Tese de Doutorado em Comunicação. Programa de Pós-Graduação em Comunicação, Faculdade de Comunicação Social. Rio de Janeiro: Universidade do Estado do Rio de Janeiro, 2017.

SANTAELLA, Lúcia. *A inteligência artificial é inteligente?*. São Paulo: Edições 70, 2023.

SARTORI, Brunno. Perfil no Instagram. Disponível em: <https://acesse.one/dr4lu>. Acesso em: 04 set. 2024.

STALDER, F. *The digital condition*. Cambridge: Polity, 2018.

TERRA. *IA da Meta é acusada de racismo por não conseguir criar casal interracial*. Disponível em: <https://11nk.dev/mT79F>. Acesso em: 01 out. 2024.

Imagens feitas por Inteligência Artificial: dilemas
éticos e vieses no resgate do passado
Talita Souza Magnolo

UOL. *Carnavais Artificiais*. Disponível em: <https://acesse.one/trZcX>. Acesso em: 29 out. 2024.

_____. *Foto de Putin se ajoelhando diante de Xi Jinping é falsa*. 2023. Disponível em: <https://l1nk.dev/SDC06>. Acesso em: 12 out. 2024.

WEBB, A. *Os nove titãs da IA*. Como as gigantes da tecnologia e suas máquinas pensantes podem subverter a humanidade. Rio de Janeiro: Alta Books, 2002.