

# Imagem gerada por Inteligência Artificial na escrita antropológica<sup>1</sup>

Imagen generada por la inteligencia artificial en la escritura antropológica

Image generated by artificial intelligence in anthropological writing

Juliane Cristina Helanski Cardoso<sup>2</sup>

Universidade Estadual de Campinas. São Paulo, Brasil

julianhelanski@gmail.com

ORCID: 0000-0001-8649-8986

**Citar como:** Helanski Cardoso, J. C. (2024). Imagem gerada por Inteligência Artificial na escrita antropológica. *Desde el Sur*, 16(2), e0023.

## RESUMO

O artigo discute o papel das imagens produzidas por inteligência artificial (IA) na antropologia visual, investigando sua capacidade de representar identidade e experiência. Aborda-se também os desafios técnicos e éticos da classificação de dados de IA e sua interação com redes sociotécnicas, questionando a neutralidade tecnológica. As implicações técnicas incluem a categorização de dados que pode perpetuar preconceitos e relações de poder. Destaca-se a simplificação e distorção das representações pela IA exigindo uma análise crítica das histórias incorporadas nas categorizações. Propõe-se que os antropólogos examinem a relação entre imagem, rótulo e referente, reconhecendo diferenças e semelhanças em seus papéis e dos rotuladores de IA na produção de conhecimento. Além disso, discute-se como as imagens de IA podem influenciar a interpretação e análise antropológicas ao misturar realidade e emoção.

1 Trabalho financiado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes).

2 Doutoranda em Ciências Sociais pela Universidade Estadual de Campinas (UNICAMP). Seu interesse de pesquisa é a ciência e tecnologia, especialmente a Inteligência Artificial. Participe também como pesquisadora colaboradora em dois centros de excelência em Inteligência Artificial: o Centro de Inteligência Artificial da Universidade de São Paulo (C4AI/IBM/FAPESP) e o Understanding Artificial Intelligence (UAI/IEA/USP); e em antropologia visual: o Laboratório Antropológico de Grafia e Imagem (La'grima).

Argumenta-se que é necessário um engajamento crítico com as implicações éticas e técnicas da geração e uso dessas imagens.

#### **PALAVRAS-CHAVE**

Antropologia, Etnografia, Inteligência Artificial, Imagem

#### **RESUMEN**

El artículo discute el papel de las imágenes producidas por la inteligencia artificial (IA) en la antropología visual, investigando su capacidad para representar la identidad y la experiencia. También aborda los desafíos técnicos y éticos de la clasificación de datos de IA y su interacción con redes sociotécnicas, cuestionando la neutralidad tecnológica. Las implicaciones técnicas incluyen la categorización de datos que puede perpetuar sesgos y relaciones de poder. Se destaca la simplificación y distorsión de las representaciones por parte de la IA, lo que requiere un análisis crítico de las historias incrustadas en las categorizaciones. Se propone que los antropólogos examinen la relación entre imagen, etiqueta y referente, reconociendo diferencias y similitudes en sus roles y los de los etiquetadores de IA en la producción de conocimiento. Además, analiza cómo las imágenes de IA pueden influir en la interpretación y el análisis antropológicos al mezclar realidad y emoción. Se argumenta que es necesario un compromiso crítico con las implicaciones éticas y técnicas de la generación y el uso de estas imágenes.

#### **PALABRAS CLAVE**

Antropología, etnografía, inteligencia artificial, imagen

#### **ABSTRACT**

The article discusses the role of images produced by artificial intelligence (AI) in visual anthropology, highlighting their ability to represent identity and experience. It also addresses the technical and ethical challenges of AI data classification and its interaction with socio-technical networks, questioning technological neutrality. Technical implications include data categorization that can perpetuate biases and power relations. The simplification and distortion of

representations by AI is highlighted, requiring a critical analysis of the stories embedded in the categorizations. It is proposed that anthropologists examine the relationship between image, label, and referent, recognizing differences and similarities in their roles and those of AI labelers in knowledge production. In addition, it discusses how AI images can influence anthropological interpretation and analysis by blending reality and emotion. It is argued that a critical engagement with the ethical and technical implications of the generation and use of these images is necessary.

#### KEYWORDS

Anthropology, ethnography, artificial intelligence, image

### Introdução

Neste artigo, analiso como as imagens produzidas por IA podem contribuir para a antropologia visual, investigando como elas mostram a identidade e a experiência na escrita antropológica. Também abordo os problemas técnicos e éticos da classificação de dados de IA verificando como a IA interage com as redes sociotécnicas. Além disso, discuto os desafios e oportunidades das imagens digitais criadas pela IA e reflito sobre suas implicações para a antropologia visual, baseado em críticas às narrativas hegemônicas sobre IA. Isso envolve uma reflexão sobre a materialidade dos algoritmos e como eles são formados pelas práticas materiais, questionando a ideia de que a tecnologia é neutra e separada da sociedade e da história. As implicações técnicas incluem os problemas da categorização de dados em sistemas de IA, que podem repetir e reforçar preconceitos e relações de poder existentes. O estudo destaca os rótulos usados por plataformas como a *ImageNet*, que podem simplificar, distorcer e gerar rótulos incoerentes que não representam a complexidade da identidade e da experiência. Esses aspectos técnicos levantam questões sobre a objetividade e a justiça das imagens geradas pela IA e seu impacto nas narrativas e entendimentos antropológicos.

Argumento que, embora as imagens de IA possam ser úteis para explorar conceitos como identidade e experiência, é preciso se engajar criticamente com as implicações técnicas e éticas de sua geração e uso. Coloco que as práticas de categorização da IA podem simplificar e distorcer as representações, refletindo visões tendenciosas dos dados de treinamento usados. Isso exige uma análise crítica das histórias sociais e políticas incorporadas nessas categorizações. Além disso, sugiro que os

antropólogos devem examinar a relação entre imagem, rótulo e referente, reconhecendo tanto as semelhanças quanto as diferenças nos papéis dos antropólogos e dos rotuladores de IA na produção de conhecimento. O argumento se estende ao conceito de equivalência entre texto e imagem em conteúdo gerado por IA, que pode misturar realidade e emoção, afetando assim a interpretação e análise antropológicas.

### **Produção de imagens por antropólogo e Inteligência Artificial na pesquisa etnográfica**

Nesta seção, investigo como as imagens geradas pela IA podem enriquecer a escrita antropológica, olhando para sua representação da identidade e experiência. A análise se baseia em uma visita ao Centro de Inteligência Artificial da Universidade de São Paulo (C4AI/USP), onde faço uma pesquisa etnográfica. Sustento que a origem e o contexto dos dados da internet são importantes, pois eles afetam o treinamento da IA, que depende de descrições humanas e programadas. Assim, as imagens levantam questões sobre idade, gênero, etnia e padrões de beleza, refletindo ideais midiáticos e preconceitos. Isso é essencial para compreender como o DALL-E 3 cria as imagens.

Uma das vezes que visitei o Centro de IA, acompanhada pelo meu principal interlocutor, chamei a atenção para uma sala fechada à chave. Na porta de madeira havia uma janela de vidro pequena, que me permitiu ver um pouco do interior da sala. Notei vários computadores e muitos cabos e fios, que me pareciam iguais. Interessada, perguntei se aquela era a sala do laboratório do Centro. Meu anfitrião disse que não, mas que era uma espécie de «TI» do Centro. Disse um «ah» simples, mais constrangida do que impressionada.

Esse episódio teve um impacto duradouro na minha pesquisa. Por que eu assumi que a sala era o laboratório do Centro de IA? Eu imaginava encontrar um laboratório de supercomputadores e cientistas envolvidos em seus projetos, mas, na realidade, eram somente computadores —um cluster, como eu descobriria, grupo de computadores que aumentam o processamento de dados e são essenciais para o avanço das pesquisas. Havia, implicitamente nessa conversa, preconceitos e imaginários que moldaram minha expectativa. No entanto, o episódio gerou certos afetos que foram relevantes para a pesquisa e, de alguma forma, influenciaram esta experimentação.

Decidi expressar os sentimentos que essa conversa despertou em mim, além de escrever, por meio de imagens geradas por IA. O motivo da minha escolha foi a minha maior familiaridade com as ferramentas tecnológicas.

Recorri a uma descrição baseada na conversa e nas emoções que senti naquele momento para criar a descrição que produziria as imagens.

Escolhi especificar o objeto a ser gerado na imagem: «eu, uma mulher branca».³ Usei o seguinte texto de entrada (*input/prompt*) no DALL-E 2: *uma mulher branca perdida em uma sala cheia de computadores, cabos e fios*. O DALL-E 2 produziu a seguinte saída (*output*) com a sequência de imagens abaixo (Figura 1). Testei outra IA similar, o *Midjourney* na sua versão conectada ao *Discord*, usando o mesmo texto em português (Figura 2) e depois em inglês: *a white woman lost in a room full of computers, cables and wires* (Figura 3). Depois usei a mesma descrição em inglês (Figura 4) e depois em português (Figura 5), no DALL-E 3 em março de 2024. Em seguida, repeti a descrição sem especificar o adjetivo (branca) do sujeito (mulher) em inglês (Figura 6) e português (Figura 7). Depois, testei sem especificar o gênero, usando apenas a categoria «pessoa» em inglês (Figura 8) e português (Figura 9). Observe que eu não citei a «cor» da minha pele como branca e nem falei sobre o estilo de imagem, a minha idade, a cor dos meus olhos, o meu corte ou a cor do cabelo, nem a minha nacionalidade ou região.



FIGURA 1. DALL-E 2: uma mulher perdida em uma sala cheia de computadores, cabos e fios

3 A literatura sobre as implicações de gênero e etnia em ambientes tecnocientíficos, predominantemente ocupados por homens brancos, é vasta, embora haja mulheres cientistas em geral. Há também uma ampla discussão sobre a situacionalidade do conhecimento produzido por mulheres (por exemplo, ver Haraway, 1995). Haraway (2016), por exemplo, orientada teórica e metodologicamente pela ficção científica, pelo feminismo, pela especulação e pela fabulação, imagina uma reconfiguração das relações humanas com a terra e todos os seus habitantes em meio a uma devastação ecológica. Para Haraway, o atual estado do planeta requer *simpoiesis*, *fazer-com*, em vez de *autopoiesis* ou *auto-fabricação*, para aprender a viver com o problema de viver e morrer juntos em uma terra danificada.



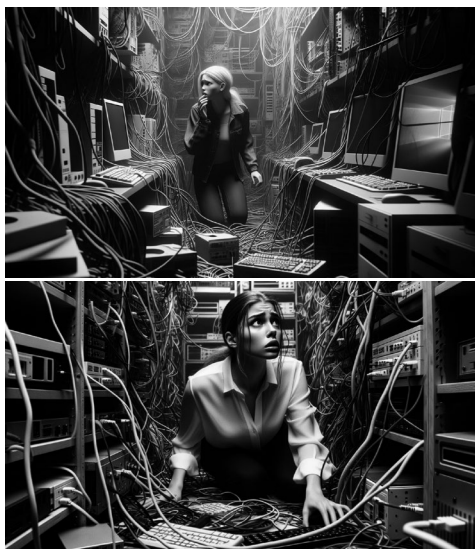
**FIGURA 2.** Midjourney: uma mulher perdida em uma sala cheia de computadores, cabos e fios



**FIGURA 3.** Midjourney: a white woman lost in a room full of computers, cables and wires



**FIGURA 4.** DALL-E 3: a white woman lost in a room full of computers, cables and wires



**FIGURA 5.** DALL-E 3: uma mulher branca perdida em uma sala cheia de computadores, cabos e fios



**FIGURA 6.** DALL-E 3: a woman lost in a room full of computers, cables and wires



**FIGURA 7.** DALL-E 3: uma mulher perdida em uma sala cheia de computadores, cabos e fios





**FIGURA 8.** DALL-E 3: uma pessoa perdida em uma sala cheia de computadores, cabos e fios



**FIGURA 9.** DALL-E: a person lost in a room full of computers, cables and wires

Uma questão é se precisava usar todas essas características descritivas - *uma mulher branca, perdida em uma sala cheia de computadores* - e qual seria o efeito do resultado sem elas. Em especial como o termo «perdida» fez a IA gerar uma atmosfera angustiante - o que corresponde ao sentimento que tive em campo naquela ocasião. A falha era nas legendas do usuário ou nas imagens e legendas da IA? Também há que considerar que este teste foi feito no começo de 2023, na versão grátis do DALL-E 2 que a OpenAI tirou e substituiu pelo DALL-E 3 pago no GPT 4. Depois disso, houve muitos progressos na criação de imagens, principalmente na diversidade e na representação de pessoas nos dados, mas com limitações. A discussão sobre esses limites parte do princípio de que a IA pode ser treinada com dados sintéticos, além dos dados reais obtidos da sociedade, que seriam feitos para refletir o que é moral e politicamente correto e socialmente justo.

Um argumento oposto a essa ideia é que usar filtros para mudar o texto de entrada seria alterar a realidade, que é moral e politicamente problemática. Esse argumento oposto assume que não importa o quão ruim seja a realidade, o mais ético é não automatizar, e caso se opte por simular a realidade, com dados sintéticos, estaria a modificando. Esse



argumento oposto não considera a origem do problema se aceitarmos que os seres (humanos ou não) podem coexistir relativamente. Se o relativismo for impossível e o conflito entre as diferenças e as visões de mundo for inescapável, esse argumento não vale. De qualquer forma, seria melhor não digitalizar a realidade ou não usar a IA para esses propósitos, como neste experimento. Isso parece um dilema político e epistemológico, mas o relativismo ontológico não implica necessariamente um método analítico diferente.

Algumas diferenças técnicas e humanas nas imagens chamam a atenção. A qualidade das imagens do DALL-E 2 é menor que as do *Midjourney* (que não eram pagas quando foram feitas as Figuras 1 e 2). As imagens do DALL-3 (Figuras 3 a 8) no GPT-4, feitas em 2024, são mais nítidas que as do DALL-E 2. Outro ponto técnico é que os programas aprenderam em inglês, uma língua com sua própria cultura e valores. A origem dos dados da Internet também importa, pois as fontes desses dados são influenciadas pelo contexto e pela cultura, e quem publicou imagens na rede pode ter ajudado a treinar a IA, o que inclui parte da população mundial. Da mesma forma, não podemos esquecer que as pessoas fazem as descrições das imagens, tanto as que publicam as imagens quanto as que as rotulam. No DALL-E 3 há um processo que mistura as descrições dos humanos com as do programa. Observe que o problema técnico tem uma dimensão social inseparável, e essa separação é só um recurso analítico.

As imagens trazem questões sobre idade, gênero, etnia e padrões de beleza. Até a descrição que não fala da cor ou do gênero da pessoa tem um viés social, como a idealização e a generalização de como uma pessoa deve ser para ser bonita pela mídia e a discriminação e o preconceito de gênero, idade e etnia nas tecnologias como algoritmos e em contextos tecnológicos - temas discutidos por autoras como Noble (2018) e Benjamin (2019). Vamos analisar essas questões, explicando como o DALL-E 3 produz as imagens.

A pesquisa sobre imagens geradas por IA revela complexidades técnicas e sociais que impactam diretamente a representação da identidade e experiência. A análise realizada a partir da visita ao Centro de IA enfatiza a importância do contexto e da origem dos dados da internet na formação dessas imagens. Questões de idade, gênero, etnia e padrões de beleza são levantadas, evidenciando preconceitos embutidos nas descrições e no próprio processo de geração de imagens. O experimento com diferentes programas e textos de entrada ressalta a influência da língua e cultura na aprendizagem da IA. Além disso, a integração de descrições humanas e programadas destaca a interação entre o humano

e o algoritmo. Nesse sentido, a discussão sobre os limites éticos e técnicos da IA na produção de imagens é essencial para uma compreensão mais ampla de seu impacto nas narrativas sociais e antropológicas.

## **A lógica da descrição «sintética» da imagem gerada por Inteligência Artificial**

Nesta seção, analisamos como a IA DALL-E 3 é treinada e as questões éticas e sociais de representar identidade e experiência nas imagens. O DALL-E 3 melhora a geração de imagens a partir de descrições complexas e legendas sintéticas (escritas pelo programa) ou humanas, mas não resolve o problema da categorização e representação de pessoas nas imagens.

Um artigo da *OpenAI* e da *Microsoft* (Betker *et al.*, s. d.) mostrou as novidades do DALL-E 3 para gerar imagens a partir de descrições mais complexas. Eles usaram um modelo de legenda de imagens melhorado e treinado em legendas sintéticas. Segundo eles, os modelos de imagem atuais não seguem bem as descrições complexas e ignoram ou mudam o sentido da legenda. Eles acham que isso é por causa da baixa qualidade das legendas nos dados de treinamento. O DALL-E 3 foi treinado com um modelo de legenda de imagens que faz descrições detalhadas e corretas das imagens. Depois, o modelo foi usado nos dados de treinamento para fazer legendas mais sofisticadas. Assim, eles treinaram vários modelos de texto para imagem nessas legendas sintéticas e testaram sua capacidade de seguir legendas. O resultado foi um novo sistema de geração de texto para imagem que tem mais coerência e estética do que os outros (Betker *et al.*, s. d.).

Betker *et al.* (s. d.) utilizam imagens e textos da web que as descrevem. O texto simples foca no elemento principal da imagem, mas deixa de lado outros detalhes (objetos comuns, localização, quantidade, cores, tamanhos etc.). As vezes os textos são errados ou não têm a ver com a imagem. Os autores querem melhorar as descrições geradas pelo computador. Para isso, usam um legendador de imagens, que cria as descrições como um modelo de linguagem como o GPT-4. O modelo de linguagem prevê a próxima palavra na frase usando a matemática das palavras anteriores. O DALL-E conecta o legendador a uma imagem com uma ferramenta chamada CLIP. Essa ferramenta ajuda o legendador, pois ele não precisa olhar todos os elementos da imagem. Assim, o legendador fica mais inteligente para descrever imagens (Betker *et al.*, s. d.).

Betker *et al.* (s. d.) descrevem o programa de computador que gera legendas para fotos em duas fases: primeiro, usa legendas curtas sobre o objeto principal da imagem para reconhecê-lo melhor. Depois, usa legendas maiores sobre todos os elementos da imagem para fazer

legendas mais completas. Eles dizem que as legendas sintéticas ajudaram o modelo a aprender melhor e propõem que as legendas reais e artificiais sejam combinadas para aumentar o desempenho do modelo. Eles defendem que essa combinação é importante porque os modelos podem se enganar se só virem um tipo de legenda. Por exemplo, se um modelo só vê legendas com um espaço no início, ele pode ficar confuso com as legendas sem um espaço. Assim, a combinação de legendas evita que o modelo aprenda coisas erradas. Por outro lado, eles argumentam, as legendas reais são como as pessoas escrevem, então misturar as legendas aleatoriamente ajuda o modelo a aprender de forma «mais humana».

O modelo foi treinado com as mesmas imagens (500.000 vezes) em várias versões de um programa de computador. Eles usaram 2.048 imagens por lote, totalizando 1 bilhão de imagens. Depois, cada versão gerou 50.000 imagens novas a partir de frases especiais. A qualidade das imagens novas dependia de quão bem elas combinavam com as frases. A qualidade é medida pelos códigos da imagem e da frase. Eles compararam os códigos para dar uma pontuação. Isso foi repetido para todas as 50.000 imagens e frases, depois multiplicado por 100.

Betker *et al.* (s. d.) testaram três programas para criar imagens a partir de textos (legendas). Um usou legendas reais (*ground truth*) para treinar. Os outros usaram principalmente legendas sintéticas, curtas ou longas (95% sintéticas). Cada programa foi avaliado com legendas reais e sintéticas longas. Os programas com legendas sintéticas foram um pouco melhores que os com legendas reais, e muito melhores com as legendas sintéticas longas. Os autores dizem que as legendas sintéticas ajudam os programas, não os atrapalham, e os testes foram mais coerentes com elas.

Betker *et al.* (s. d.) usaram amostras de treinamento com objetos, cenas e paisagens assim: um diário antigo cheio de desenhos e histórias de criaturas e aventuras fantásticas; um cartaz dos anos 1960 incentivando a migração interplanetária, com um foguete retrô e um planeta colorido; um furão fofo em um pote de doces, em uma cozinha aconchegante; um gnomo de jardim guerreiro em armadura de folha e casca, protegendo seu reino de jardim; uma paisagem gelada com uma cachoeira congelada e um fogo iluminando a neve; um portal giratório e colorido em um mar de café, para uma dimensão fantástica.

Um homem e uma mulher em uma *graphic novel* são as únicas pessoas usadas como exemplos de treinamento. Eles são descritos assim: «Uma figura de uma *graphic novel*. Uma rua da cidade lotada à luz da lua. As calçadas estão cheias de gente aproveitando a noite. Na esquina, uma ruiva com um manto de veludo está comprando algo do vendedor velho

e carrancudo. O vendedor elegante, um homem alto com um bigode notável, está falando animadamente em seu telefone *steampunk*» (Betker *et al.*, s. d., p. 3).

A UNESCO (2024) encontrou viés e discriminação de gênero, homofobia e racismo nas ferramentas de IA que geram texto, chamadas LLM (*Large Language Model*, mais conhecido como Inteligência Artificial Generativa). O estudo disse que os LLMs atribuem papéis desiguais e estereotipados a homens e mulheres, e geram conteúdo negativo sobre gays e certos grupos étnicos. Os modelos gratuitos e de código aberto, como Llama 2 e GPT-2, foram os mais enviesados por gênero. O estudo propôs medidas práticas para promover a igualdade de gênero na criação das ferramentas de IA. Há duas formas de enfrentar o desafio de combinar texto e imagem na geração de imagem por IA: uma técnica e outra política. Betker *et al.* (s. d.) sugeriram que o problema está na descrição da imagem, e que bastaria ajustar as legendas com legendas sintéticas para resolvê-lo. Note que o problema está ligado à descrição do mundo real, que pode ter dados errados que não refletem o entendimento das pessoas. Por outro lado, o relatório da UNESCO (2024) afirmou e recomendou que o problema está na baixa representação de gênero no desenvolvimento da IA.

A análise da IA DALL-E 3 e seu treinamento revela avanços significativos na geração de imagens a partir de descrições complexas. Embora o modelo tenha sido aprimorado com o uso de legendas sintéticas e técnicas avançadas de treinamento, ainda persistem desafios éticos e sociais na representação de identidade e experiência nas imagens geradas.

O estudo liderado por Betker *et al.* (s. d.) destaca a importância de uma abordagem combinada de legendas reais e artificiais para otimizar o desempenho dos programas de geração de texto para imagem. No entanto, questões como viés de gênero, homofobia e racismo identificadas pela UNESCO em ferramentas de IA destacam a necessidade de considerações políticas mais amplas no desenvolvimento dessas tecnologias. A combinação de abordagens técnicas e antropológicas pode ajudar a enfrentar os desafios relacionados à representação justa e equitativa na geração de imagens por IA promovendo uma narrativa mais inclusiva e sensível às diversas identidades e experiências humanas.

Um aspecto crítico é a forma como as imagens são rotuladas nos conjuntos de dados usados para treinar os sistemas de IA. O ImageNet, por exemplo, é um conjunto de dados controverso e influente que levanta preocupações sobre preconceitos e estereótipos nas categorizações de imagens. O ImageNet tem milhões de imagens rotuladas em mais de 20 mil categorias, algumas estranhas ou ofensivas. O experimento ImageNet

Roulette usou um sistema de IA treinado apenas na categoria de pessoa do ImageNet para rotular as fotos dos usuários, mostrando assim os preconceitos e suposições do conjunto de dados. O experimento mostrou que os conjuntos de dados que classificam as pessoas por características físicas, emocionais, comportamentais ou morais são perigosos e injustos, e que a estrutura e o conteúdo desses conjuntos de dados têm grandes efeitos (Crawford e Paglen, 2019).

As imagens com milhões de categorias que a IA usa são etiquetadas por pessoas. Algumas imagens são de coisas, outras são de gente e seus trabalhos, como soldado, mergulhador, professor e médico. Mas muitas vezes as pessoas recebem etiquetas que mostram preconceito e estereótipo, como uma mulher de biquíni ser vulgar ou um homem com cerveja ser alcoólatra. Essas etiquetas variam conforme as normas sociais, os contextos históricos e geográficos, e o modo de fazer os conjuntos de dados. A categoria pessoa simplifica as identidades humanas complexas. Isso é político porque envolve poder que afeta como a IA classifica as pessoas, e pode gerar problemas, como reconhecimento facial ou algoritmos de recrutamento, que usam dados tendenciosos e causam rotulagem errada, exclusão, identificação falsa e uso indevido de dados (Crawford e Paglen, 2019).

As pessoas anotam imagens para treinar sistemas de aprendizado de máquina que as reconhecem e interpretam. Esses dados moldam o conhecimento da IA e afetam a sociedade. Mas as imagens de treino usadas em sistemas de IA são baseadas em pressupostos instáveis e tendenciosos que o campo raramente questiona. As imagens não têm um único sentido, mas vários possíveis, dependendo de como são vistas e rotuladas, como mostram os estudos sobre a relação inconstante entre imagens e significados (Crawford e Paglen, 2019). Essas imagens podem servir para diferentes propósitos e mudar de sentido conforme a cultura, o espectador e o contexto. A IA e seus dados não são imparciais na forma de categorizar o mundo, pois a rotulação revela que eles refletem a política, os vieses e as subjetividades da história (Crawford e Paglen, 2019).

Kate Crawford e Trevor Paglen (2019) afirmam que muitas empresas de IA ocultam como usam e classificam as imagens. Elas também podem usar fotos não autorizadas do Google para rotular e treinar seus sistemas. O ImageNet fez isso. Os rótulos do ImageNet muitas vezes simplificam e distorcem imagens, gerando rótulos sem sentido e incoerentes, e refletindo visões enviesadas sobre a relação entre imagens e o texto. Como a rotulação de imagens com base em suposições sobre a aparência física e os papéis sociais, que têm histórias e implicações políticas obscuras (Crawford e Paglen, 2019).

Crawford e Paglen (2019) citam a empresa IBM como exemplo de uma tentativa de aprimorar o reconhecimento facial para variar tons de pele. A empresa criou o conjunto de dados *Diversity in Faces* (DiF) com quase um milhão de imagens do Flickr, buscando equilíbrio em tons de pele, traços faciais, idade e gênero. Eles também adicionaram categorias novas, como simetria facial e forma da cabeça, dizendo que elas dão mais informações sobre os rostos do que só idade, gênero e cor da pele. As empresas querem ter dados variados para melhorar suas ferramentas para mercados maiores, mas classificar as pessoas é complicado e político. A IBM acabou usando matemática para definir diversidade e justiça nos dados se baseando na opinião de alguns rotuladores para estimar a idade e o gênero de alguém por uma foto, o que não é muito científico (Crawford e Paglen, 2019).

Os sistemas de IA visual baseiam-se em conceitos como comida ou emoções que não são objetivos nem universais, mas dependem de quem os rotula e do seu contexto histórico. Eles também presumem que há uma conexão clara e mensurável entre o aspecto de algo e sua identidade, e que isso pode ser quantificado. Os dados usados na IA se sustentam em crenças duvidosas sobre imagens e rótulos. Tirar da Internet dados de IA problemáticos pode parecer bom para a privacidade, mas não evita que eles sejam usados em sistemas e pesquisas de IA. Além disso, isso pode esconder como os vieses e erros na IA foram gerados, o que é importante para entender e consertar (Crawford e Paglen, 2019).

A IA suscita questões sobre como classificar e representar imagens. As críticas ao ImageNet e sua história mostram os problemas e vieses nos sistemas de IA. O DALL-E, que gera imagens a partir de textos, mostra como a relação entre imagem, rótulo e referente depende de contextos culturais. Mais transparência das empresas de IA e mais diversidade nos dados são geralmente recomendados pela crítica social da tecnologia como passos importantes para reduzir preconceitos e injustiças, mas Amba Kak e Sarah West (2024) dizem que não basta diversificar atores no desenvolvimento da IA se as entidades comerciais seguirem determinando os rumos da pesquisa e houver pouca resistência ao seu domínio.

O trabalho antropológico e a IA podem ser entendidos pela relação entre imagem, rótulo e referente. O antropólogo pode analisar essa relação e seu papel como usuário da IA para produzir conhecimento antropológico, que às vezes se assemelha e às vezes se diferencia do papel do rotulador. Se assemelha porque o antropólogo também rotula uma experiência de campo com base em sua percepção pessoal, que vira texto, que vira legenda, que vira imagem, que é gerada por metadados usados para treinar a IA. Se diferencia porque o rotulador não tem propósito

antropológico, seja científico ou especulativo. O rotulador segue os rótulos e categorias preestabelecidos por outros rotuladores, destacando diferenças e semelhanças entre os propósitos e processos envolvidos.

Os dados vêm de fontes diferentes, que podem ser objetivas ou subjetivas e influenciar de várias maneiras, desde a construção de um conhecimento, uma realidade até a valorização de uma individualidade, emoções e percepções na produção de imagem de pessoas na IA, nem sempre coerentes ou justas.

Portanto, a IA e a antropologia podem colaborar na análise de como a IA transforma elementos culturais em imagens que codificam aspectos das experiências humanas, enfatizando a importância de uma abordagem crítica e reflexiva na produção e interpretação de imagens geradas por IA.

Em síntese, a evolução da IA, exemplificada pelo DALL-E 3, têm revelado avanços notáveis na geração de imagens a partir de descrições complexas. No entanto, este progresso é acompanhado por desafios éticos e sociais, particularmente na representação precisa e justa da identidade e experiência humanas. A combinação de legendas reais e artificiais é essencial para a otimização do desempenho dos sistemas de geração de texto para imagem, mas questões como viés de gênero, homofobia e racismo em ferramentas de IA enfatizam a necessidade de considerações políticas mais amplas no desenvolvimento dessas tecnologias.

A análise crítica dos conjuntos de dados utilizados para treinar os sistemas de IA, como o ImageNet, ressalta a importância de uma abordagem reflexiva e transparente na categorização de imagens. A interseção entre a IA e a antropologia destaca como a relação entre imagem, rótulo e referente pode ser compreendida e analisada para produzir conhecimento significativo, enfatizando a importância de uma abordagem crítica e inclusiva na produção e interpretação de imagens geradas por IA.

### **Equivalências entre texto e imagem na escrita antropológica**

A análise da IA DALL-E 3 e seu treinamento revela avanços significativos na geração de imagens a partir de descrições complexas, destacando a crescente intersecção entre tecnologia e representação visual. No entanto, a questão da representação justa e equitativa ainda persiste, especialmente quando se trata da representação de identidade e experiência nas imagens geradas. Este dilema ecoa as preocupações levantadas por Etienne Samain (1995) sobre o poder das imagens na antropologia e a necessidade de uma análise mais aprofundada das práticas verbais e visuais na disciplina. Assim como Betker *et al.* (s. d.) destacaram a importância de uma abordagem combinada de legendas



reais e artificiais, os debates antropológicos sobre a relação entre texto e imagem ressaltam a necessidade de considerar tanto os aspectos técnicos quanto antropológicos na interpretação e produção de imagens. A interseção entre antropologia e tecnologia, exemplificada pelo uso da IA na geração de imagens, oferece novas possibilidades para a prática antropológica, mas também levanta questões éticas e epistemológicas sobre a representação e a agência das imagens na sociedade contemporânea. Ao confrontar esses desafios, os antropólogos são confrontados com a tarefa de reconciliar a ciência e a imaginação, buscando uma antropologia mais crítica e sensível.

A relação entre a antropologia e a fotografia surge com a vontade de retratar pessoas e culturas. O desafio da fotografia na antropologia é como ela mostra, revela ou esconde alguém ou algo. Podemos ver as imagens feitas por IA como um campo de disputas políticas, em que novas éticas podem emergir? Uma coisa é clara, o poder dessas imagens é frágil. E como se usa o poder das imagens? Ou podemos pensar nessas imagens sem ligá-las a significados e instrumentos de poder? A imagem tem uma força performativa de criar sentido?

Samain (1995) explica como a fotografia pode ser útil para a pesquisa antropológica quando usada junto com a etnografia. Ele nos convida a explorar formas de olhar para as imagens na antropologia. O autor defende uma reflexão sobre temas antigos com novos olhares, mas, para este debate, interessa como as tecnologias da verbo-visualidade, como a fotografia, o som, o cinema, o vídeo e a computação, especialmente, e neste caso a IA, fundamentam a prática da antropologia visual. A questão é como compreender antropológicamente as pessoas e sociedades em suas experiências pelos novos modos de verbo-visualidade.

Samain (1995) argumenta que muitos pesquisadores não se capacitam para a antropologia visual e erram por não entenderem o seu objeto de estudo ou não saberem como comunicar as suas ideias com imagens. Por isso, é importante pensar em como aliar as ciências humanas à comunicação. Essa aliança é mais necessária hoje, pois as TICs (Tecnologias de Informação e Comunicação) permitem novas formas de compreender, interpretar e organizar as sociedades, nas quais a imagem pode incentivar e suscitar debates. O que a fotografia e a escrita têm em comum é que ambas são modos de representar outras representações. O ato de representar envolve geralmente três passos: observar, explicar e significar. Assim, há uma interpretação das representações de representações que pode ser feita com diferentes meios de comunicação. Isso significa questionar a fotografia como um registro fiel da realidade e, talvez, isso também se aplique à imagem produzida por IA.

Quando examinamos a imagem produzida pela IA e os processos de mediação que acontecem até o produto, percebemos que a imagem não pode ser vista apenas como a representação do real ou expressão do olhar subjetivo do seu autor. Isso significa que a imagem é, primordialmente, um processo. Neste processo, propomos o conceito de *equivalência* (Bisilliat, 1974) entre texto e imagem para analisar as imagens produzidas por IA na descrição etnográfica. A equivalência é uma conexão ou relação de coisas diferentes por meio de um elemento comum ou de uma chave. Essa relação pode ser real ou fictícia, concreta ou abstrata, individual ou coletiva; uma recordação, um saber ou uma experiência, e pode mudar ou colocar uma coisa em outro lugar, e até mesmo fazer algo desaparecer ou ocultar-se. A equivalência entre texto e imagem indica que a imagem representa um sistema codificado pelo qual a imagem produzida com a IA pode ser interpretada como algo que eu vi ou quis ver. Essa equivalência produziu imagens cujos referentes mesclou realidade e emoção.

Outra conexão possível entre antropologia, ficção e literatura é feita por James Clifford (2011; 2016) e Michael Taussig (2011), que utilizaram a escrita e o desenho para narrar histórias sobre outras culturas. Essas histórias são influenciadas pelas perspectivas e experiências do antropólogo. Clifford argumenta que a antropologia como literatura e ficção pode desafiar as noções convencionais de verdade e objetividade, e que a antropologia pode ser vista como uma forma de arte que usa habilidades literárias para criar narrativas mais cativantes e emotivas do que as descrições objetivas e técnicas da antropologia tradicional.

Stuart McLean (2017) é um autor que aborda de outra forma esses temas, propondo que se faça antropologia como uma fabulação: habilidade criativa de inventar personagens e histórias; criação que participa dos processos de construção do mundo material, nos quais os atos de criatividade humana já estão sempre presentes. McLean (2017) afirma que a antropologia tem, desde sempre, a capacidade de subverter as fronteiras convencionais entre documentário e ficção, ao diminuir a distância que mantém essas fronteiras. Isso abre a realidade à dúvida e, possivelmente, à transformação.

As imagens que ultrapassam esses limites desafiam a representação da identidade, da experiência, da situação e da realidade. A fabulação não é só criar imagens de uma realidade, que pode ser virtual ou real, verdadeira ou falsa. É também envolver-se, pela percepção e pela ação, na própria formação das coisas. A imaginação, assim, é um modo de nos relacionarmos com o mundo à nossa volta de forma profunda e transformadora (Ingold, 2012).

A criatividade por trás dessas imagens gerou uma ficção, por assim dizer, que de várias formas afeta a realidade e talvez mude algo ou alguém ao estimular debates que transcendem a própria imagem e minha pesquisa. Como Bruno Latour (1994) mostrou, os textos ligam humanos e não humanos com retórica, estratégia, escrita, contextualização e semiótica, revelando como eles distribuem seus papéis, ações e capacidades, e como se definem o que é animal, material, humano ou não humano. Assim, para interagir com a IA, é preciso levar em conta esses aspectos nas redes sociotécnicas em que ela está envolvida, pois estas influenciam as imagens produzidas.

Tim Ingold (2015) questiona a separação entre pessoas e objetos. Ele os chama de coisas e diz que se influenciam. Até uma pedra tem agência no seu meio. Mas Ingold (2022) diz que, com a tela em vez da página, deixamos de tocar as superfícies do mundo. A tecnologia afasta o olho da mão, a superfície da linha. O teclado e a tela sensível ao toque são lisos e impedem o rabisco. Ingold acha que a revolução digital vai se acabar, pois o mundo não suporta mais a energia e os metais tóxicos que ela usa. Ingold é realista, embora pessimista. Ele sabe que a degradação e o uso insustentável do meio ambiente são reais, como provam vários estudos da infraestrutura da IA<sup>4</sup>.

Segundo Vilém Flusser (2009), as imagens técnicas são uma forma de simplificação que reduz as dimensões dos objetos, consistindo em uma abstração de pixels que subtrai. Tais imagens sugerem uma sociedade totalitária programada centralmente, na qual os receptores e produtores de imagens estão sob controle rigoroso. Essas imagens indicam uma sociedade telemática, na qual os criadores e colecionadores de imagens estão mais conectados. Neste caso, a imagem gerada por IA pode ser entendida como uma forma de simplificação da realidade.

Neste cenário de incerteza, Flusser (2009) disse algo importante: as imagens técnicas são relevantes para os interesses existenciais futuros, pois uma catástrofe é imprevisível. O desafio é, aparentemente, conciliar o conhecimento científico, ou pelo menos o apreço pela ciência, com uma imaginação realmente flexível (Clarke, 1984).

A IA produz imagens que questionam as concepções tradicionais de representação e verdade na antropologia, assim como o desenho tira o peso da verdade, nos termos de Taussig (2011). A relação entre texto

---

4 Ver Pasquinelli e Joler (2020), Crawford e Joler (2018) e Crawford (2021). Essas referências são importantes para desmistificar a IA e compreender seus efeitos reais no mundo. Elas nos convidam a questionar as narrativas dominantes sobre a IA e a buscar formas mais justas e sustentáveis de desenvolver e usar essa tecnologia.

e imagem, mediada por tecnologias como a IA, permite-nos explorar novos modos de narrar e interpretar a realidade. A antropologia visual, combinada com a ficção e a literatura, pode superar a dicotomia entre objetividade e subjetividade. A fabulação, a criação de personagens e histórias, possibilita problematizar a identidade, a experiência e a própria realidade.

A produção de imagem com a IA nos faz reconsiderar a relação entre humano e objeto, levando em conta as redes sociotécnicas em que estão envolvidos. A agência das coisas e o papel da tecnologia na forma como nos conectamos com o mundo são aspectos relevantes a serem observados. O desafio é, então, conciliar o saber científico com a imaginação, usando as ferramentas da IA para criar uma antropologia mais ativa e transformadora. Portanto, as imagens geradas por IA abrem um novo horizonte de possibilidades para a antropologia demandando uma reflexão crítica sobre seus métodos e instrumentos. Ao confrontar esses desafios, os antropólogos são convocados a reconciliar a ciência e a imaginação, buscando uma antropologia mais crítica e sensível. Em última análise, a incorporação da tecnologia na prática antropológica pode ampliar nosso entendimento da realidade, desde que seja acompanhada por uma reflexão cuidadosa sobre suas implicações éticas e epistemológicas.

### **A política da imagem produzida por Inteligência Artificial**

Enfatize a necessidade de integrar ciência e imaginação na antropologia, olhando para a forma como a tecnologia amplia a compreensão da realidade, e nesta seção destaque a importância de examinar as complexas relações entre humanos e não humanos na produção de imagens, especialmente como a IA é influenciada por contextos políticos, interesses diversos e relações sociotécnicas complexas.

Portanto, a relação com a IA não apenas instiga os antropólogos a repensar a relação entre humanos e objetos na prática antropológica, especialmente em um contexto em que a tecnologia desempenha um papel cada vez mais significativo na produção e interpretação de imagens.

A IA é uma tecnologia que envolve relações entre humanos e não humanos, que rompe com as separações tradicionais entre natureza e cultura, entre ciência e política (Latour, 1994). Logo, as possibilidades técnicas e tecnológicas dependem da combinação do contexto político com suas metas, projetos e expectativas, com os recursos e o acesso materiais (Mura, 2011). A IA mostra, entre outras coisas, uma ligação causal entre diferentes atores que colaboram, formando sistemas sociotécnicos, refletindo forças e interesses variados que podem ter origens, grupos

sociais e étnicos diferentes, apresentando visões de mundo, competências e objetivos diversos e, por vezes, opostos. Assim, a imagem pode ser vista como uma co-criação entre antropólogo e IA (e todas as outras pessoas envolvidas neste processo). Essa co-criação pode ser entendida como uma relação entre texto escrito, imagem e IA, estratégias e experiências do pesquisador no campo. Podemos pensar a imagem produzida por IA como resultado da interação entre prática, percepção, ação e ambiente (nos termos de Ingold, 2015), e em vários níveis de sua expressão técnica e social, e que gera um conhecimento sobre algo ou alguém.

Quando criei a imagem presumi uma identidade fixa. Essa identidade e experiência me fizeram compreender algo específico? Não foi um aprendizado sobre mim, mas sobre o meu lugar no campo de pesquisa, pensando em como se deu o meu contato com o Centro de IA, e os dados de treinamento da IA. De acordo com Roy Wagner (1991), o que chamamos de identidade pode ser visto como uma pessoa fractal, ou seja, uma entidade que cria versões de si mesma em níveis diferentes de relação, sem se limitar a uma unidade ou um conjunto. Assim, a pessoa fractal não é singular nem plural, mas uma entidade que se reproduz em escalas e dimensões diferentes. Do mesmo modo, o sentido é indivisível, não podendo ser dividido ou separado em partes, ele contraria a ideia de que o significado pode ser obtido por meio de métodos gramaticais ou semióticos, e mostra que o significado é um fenômeno fractal que se reproduz em escalas diferentes. Suely Kofes (2015) diz que o conceito de alteridade tem mais a ver com uma forma de manter a especificidade antropológica e, assim, potencialmente criar conceitos, ou inventá-los, como Roy Wagner defende.

O conceito de experiência usado nessa identidade enfrenta muitas críticas. Alguns apontam que ele pressupõe visibilidade, transparência, individualidade, autenticidade, anteriores ao discurso, à narrativa, aos conceitos. Outros questionam sua ligação com a subjetividade e a consciência individual. Kofes (2015) fala de Jason Throop (2003) e diz que o termo «experiência» tem várias interpretações e diferenças teóricas. Ela nota que alguns conceitos na antropologia são fundamentais, mas usados sem reflexão. Um deles é a experiência. Ela mostra sua presença em vários campos do saber, como um elemento-chave na antropologia, no feminismo, na psicologia e na teoria crítica, inclusive na etnografia crítica. Em todos esses campos, a experiência é importante e central, mas há uma ilusão sobre sua objetividade ou propriedades, pois falta definição clara.

Segundo Victor Turner (1980; 1982), a experiência é uma estrutura que conecta percepções, memórias, eventos e emoções vividos, com os sentidos e valores que se expressam (Kofes, 2015). Para Turner, essa

expressão da experiência é a unidade estrutural dela. Isso quer dizer que a experiência, ao virar narrativa, relaciona fatos e sentimentos, dando-lhes sentidos e valores. A narrativa cria, assim, uma estrutura de relação entre vida, pesquisa e escrita antropológica, sem esconder seu aspecto problemático formado por linhas bem diferentes (Kofes, 2015).

Kofes (2015) cita Lévy-Bruhl (1938), que fala da experiência como uma forma de entender os aspectos afetivos e conceituais que uma ideia limitada de realidade deixaria de lado como irrealis e irracionais. A experiência se confunde com o que é observado e sentido diretamente, e Kofes (2015) afirma que a solução para o conflito entre experiência e método está em Turner. Ou seja, não é preciso contrapor os dois, pois as narrativas dramáticas acabam com essa e outras dicotomias, segundo Turner (1980; 1982). Ao juntar narrativas diferentes e espalhadas, o resultado é uma dramatização e um envolvimento do antropólogo na história, que ele conta a partir de um nós que supera a análise e só resta a experiência. Assim, Kofes (2015) diz que a experiência é criada na pesquisa ou como um «momento etnográfico» que vem das conexões entre epistemologias, as observações do campo e a análise posterior, como sugere Strathern (1999).

Wagner (1991) sugeriu uma forma clara de encarar o problema: os antropólogos sociais podem usar termos estruturalistas para decompor os fenômenos sociais e culturais em vários eixos e obter entendimentos que mantêm sua escala com elegância e surpresa às formas generalizantes de conceito e pessoa que não são singulares nem plurais. A fractalidade, assim, se relaciona ao todo, se transforma nele e o replica, e é algo diferente de uma soma ou de uma parte isolada. Como uma forma que é holográfica ou que define sua própria escala, não é como uma organização social ou uma ideologia cultural que ordena e organiza, explica ou interpreta, um conjunto de elementos diversos. Ela é uma expressão dos próprios elementos.

A IA, que usa a tecnologia para relacionar humanos e não humanos, combina contexto político, recursos e acesso a materiais para criar possibilidades técnicas e tecnológicas. A imagem feita pela IA é co-criada pelo antropólogo, a IA e os outros participantes, a partir da interação entre prática, percepção, ação e ambiente, que produz conhecimento (Ingold, 2015). Ao fazer a imagem com base em uma identidade fixa, a pergunta é se essa identidade e experiência ensinaram algo específico. A resposta é não. O conceito de identidade fractal, de Wagner, diz que a identidade se repete em diferentes níveis de relação, sem se limitar a uma unidade ou conjunto. Isso desafia a ideia de um «eu» único e permite a criação de um «nós» na pesquisa. No meu caso, esse «nós» inclui a antropóloga, meus

contatos, o Centro de IA, os dados de treinamento da IA e as pessoas e processos envolvidos.

A IA aborda a materialidade como um aspecto que se relaciona, entre outras perspectivas, à ligação entre os algoritmos como uma realização ou imposição de ideias matemáticas abstratas sobre dados concretos (ou como uma suposta representação objetiva e completa da realidade). No entanto, a partir de uma história social da IA, Pasquinelli (2019) argumenta que a forma do algoritmo emergiu de práticas materiais, da divisão cotidiana de espaço, tempo, trabalho e relações sociais. Essas práticas deram origem aos números e às tecnologias numéricas, e não o contrário. Ao examinar a genealogia industrial das oficinas, observa-se que a computação se desenvolveu gradualmente de formas concretas para formas cada vez mais abstratas.

Um exemplo recente dessa abstração é nos veículos autônomos que usam IA para estimar o espaço 3-D de uma estrada a partir de vídeos. A IA não copia um motorista individual, mas sim uma força de trabalho coletiva que usa dados que refletem uma produção coletiva de espaço, tempo, trabalho e relações. Isso pode ser visto como uma industrialização da visão e automação da percepção, gerando debates sobre a confiança em uma máquina para interpretar uma realidade objetiva, a natureza da imagem virtual e a computação da visão.

A IA usa o trabalho e o conhecimento coletivos, gerando *big data* - dados da Internet armazenados em grandes *data centers* sem regulamentação. A IA também precisa de processadores potentes e algoritmos de compressão de dados, que a nutrem e a treinam. Essa infraestrutura requer vários minerais, como as terras raras usadas em *laptops* e *smartphones* para reduzir seu tamanho e peso, e o silício usado para fazer os chips que são essenciais para a computação avançada, especialmente para a IA. O silício passa por vários processos físicos e químicos com o ambiente e os humanos. Elementos diferentes, como substâncias químicas (silício, quartzo, ácidos, cobre, ouro, prata) e fatores físicos (fundição, óptica, temperatura, umidade, tempo, espaço, partículas) se combinam, se organizam, se controlam e se alteram com o trabalho e o conhecimento coletivo, incluindo leis, instituições, empresas, campos de pesquisa, agentes de pesquisa e financiamentos.

Para compreender como a ciência e a tecnologia são feitas, é preciso levar em conta os vários elementos que as compõem. Esses elementos revelam a diversidade e a complexidade das práticas científicas diárias, como já mostraram diversos estudos (Latour, 1994; 2000; 2001). Nesse sentido, as imagens e os textos dos sistemas de IA têm aspectos físicos,



lógicos, econômicos e ideológicos que influenciam sua produção, circulação e uso na sociedade (Pasquinelli, 2023). Uma análise crítica desses aspectos implica questionar os interesses e os efeitos dos algoritmos que atuam sobre as imagens e os textos, bem como as formas de resistir e desafiar suas lógicas de dominação.

Em síntese, a relação entre ciência e imaginação na antropologia se torna mais relevante frente à crescente influência da tecnologia, em particular da IA, na produção e interpretação de imagens. A complexidade das relações entre humanos e não humanos, especialmente no contexto sociotécnico da IA, demanda uma abordagem crítica e reflexiva dos antropólogos. Ao analisar as implicações sociais e políticas da interação com a IA, fica claro que a produção de imagens não é apenas um processo técnico, mas uma co-criação entre humanos, tecnologia e contextos sociopolíticos diversos. Essa co-criação expressa visões de mundo, interesses e valores variados, evidenciando a necessidade de uma análise cuidadosa das relações de poder subjacentes.

A noção de identidade fractal sugerida por Wagner desafia a ideia de singularidade e enfatiza a natureza relacional e coletiva da pesquisa antropológica. Da mesma forma, a compreensão da experiência como uma estrutura narrativa complexa, que integra percepções, memórias e valores, ressalta a importância de considerar múltiplas perspectivas na construção do conhecimento.

A história social da IA mostra como a tecnologia surge de práticas materiais e relações sociais complexas, enfatizando a necessidade de uma análise crítica das lógicas de dominação presentes nos algoritmos e nas infraestruturas tecnológicas. A compreensão da ciência e tecnologia como processos sociais e políticos revela a importância de questionar os interesses subjacentes à produção e circulação de imagens.

A interseção entre ciência, imaginação e tecnologia na antropologia requer uma abordagem interdisciplinar e reflexiva, que considere as complexas relações entre humanos, não humanos e contextos sociopolíticos. Somente por meio de uma análise crítica das práticas e narrativas envolvidas na produção de imagens é possível compreender as formas de poder e dominação presentes na era da IA.

## **Conclusão**

Argumentei que a IA produz imagens como um modo de escrita antropológica por meio da revisão de estudos sobre antropologia visual e escrita. Para tanto, abordei os desafios, limites e possibilidades das imagens da IA para a antropologia visual. Descrevi minha experiência em um Centro de IA e como usei as imagens da IA para transmiti-la, o que me

fez refletir sobre o processo de escrita antropológica. Além disso, analisei imagens feitas por programas como o DALL-E e seu potencial para explorar conceitos como identidade e experiência na escrita antropológica. Também considerei visões críticas sobre a classificação dos dados da IA, como as de Crawford e Paglen sobre o ImageNet, para questionar os impactos técnicos e éticos das imagens da IA. A pesquisa também situou a discussão nas redes sociotécnicas mais amplas, seguindo a análise social da IA de Pasquinelli.

Analisei também como as imagens produzidas por IA representam pessoas ou coisas, e o que isso significa para a antropologia visual. Discuti as questões políticas da representação visual na IA e o papel dos humanos nesses sistemas. Mostrei que a IA pode produzir imagens que contribuem para a escrita antropológica, explorando novas maneiras de usar as imagens na descrição etnográfica e de suscitar debates, mas também alertei para as complexidades políticas da representação visual em IA destacando a importância de um olhar crítico sobre como as imagens são feitas, compartilhadas e usadas pela sociedade. Essas limitações indicam a necessidade de uma análise cuidadosa das consequências do uso de imagens de IA na prática antropológica, pois elas não são neutras, mas são influenciadas pelo conhecimento, trabalho e subjetividade humana. Da mesma forma que os algoritmos e os dados humanos e sintéticos que os moldam, reforçando e favorecendo perspectivas e situações sociais que podem causar danos ou prejuízos para determinadas pessoas, grupos ou sociedades.

### **Contribución de autoría**

Juliane Cristina Helanski Cardoso cumplió con todas las funciones de CRediT.

### **Fuente de financiamiento**

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

### **Potenciales conflictos de interés**

Ninguno.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Benjamin, R. (2019). *Race after technology. Abolitionist tools for the new Jim code*. Polity Press.
- Betker, J., Goh, G., Jing, L., Brooks, T., Wang, J.-f., Li, L., Ouyang, L., Zhuang, J., Lee, J. e Guo, Y. (2023). Improving image generation with better captions. Computer Science. <https://cdn.openai.com/papers/dall-e-3.pdf>
- Clarke, A. C. (1984). *Profiles of the future. An inquiry into the limits of the possible*. Holt, Rinehart, and Winston.
- Clifford, J. (2011). *A experiência etnográfica. Antropologia e literatura no século XX*. Editora UFRJ.
- Clifford, J. e Marcus, G. (Eds.). (2016). Sobre a alegoria etnográfica. In *A escrita da cultura: poética e política da etnografia*. Eduerj e Papéis Selvagens.
- Crawford, K. e Paglen T. (September 19, 2019). *Excavating AI: The politics of training sets for machine learning*. <https://excavating.ai>
- Crawford, K. (2021). *The Atlas of AI. Power, politics, and the planetary costs of artificial intelligence*. Yale University Press. <https://doi.org/10.2307/j.ctv1ghv45t>
- Crawford, K. e Joler, V. (2018). Anatomia de um sistema de IA: o Echo da Amazon como um mapa anatômico do trabalho humano, dados e recursos planetários. *AI Now Institute e Share Lab*. <https://anatomyof.ai>
- Flusser, V. (2009). *Universo das imagens técnicas. Elogio da superficialidade*. Annablume.
- Fotografia e literatura nos livros de Maureen Bisilliat. (2024). *Instituto Moreira Salles*. <https://ims.com.br/fotografia-e-literatura-nos-livros-de-maureen-bisilliat/>
- Haraway, D. J. (1995). Saberes localizados: a questão da ciência para o feminismo e o privilégio da perspectiva parcial. *Cadernos Pagu*, 5, 7-41. <https://ieg.ufsc.br/storage/articles/October2020/31102009-083336haraway.pdf>
- Haraway, D. J. (2016). *Staying with the trouble. Making Kin in the Chthulucene*. Duke University Press. <https://doi.org/10.2307/j.ctv11cw25q>
- Ingold, T. (2012). *Introduction*. In *Imagining landscapes: past, present and future*. Routledge.
- Ingold, T. (2015). *Estar vivo. Ensaios sobre movimento, conhecimento e descrição*. (1st ed.). Editora Vozes.
- Ingold, T. (2022, 22 de agosto). Beyond writing and drawing: In praise of scribble [Vídeo de YouTube]. <https://www.youtube.com/watch?v=iVUSbKAEGEO>

- Kak, A. e West, S. (2024) A modern industrial strategy for AI?: Interrogating the US Approach. In *AI Nationalism(s): Global Industrial Policy Approaches to AI*. AI Now Institute. <https://ainowinstitute.org/publication/a-modern-industrial-strategy-for-aiinterrogating-the-us-approach>
- Kofes, S. (2020). As grafias – traços, linhas, escrita, gráficos, desenhos - como perturbação no conhecimento antropológico. *Revista de Antropologia da UFSCar*, 12(2), 12-26. <https://doi.org/10.52426/rau.v12i2.345>
- Latour, B. (1994). *Jamais fomos modernos. Ensaio de Antropologia Simétrica*. Editora 34.
- Latour, B. (2000). *Ciência em Ação. Como seguir cientistas e engenheiros sociedade afora*. Unesp.
- Latour, B. (2001). *A esperança de pandora. Ensaio sobre a realidade dos estudos científicos*. Edusc.
- Lévy-Bruhl, L. (2018). *A experiência mística e os símbolos entre os primitivos*. Teodoro Editor.
- McLean, S. (2017). *Fictionalizing anthropology. Encounters and fabulations at the edges of the human*. University of Minnesota Press.
- Mura, F. (2011). De sujeitos e objetos: um ensaio crítico da antropologia da técnica e da tecnologia. *Horizontes Antropológicos*, 17(36), 95-125. <https://doi.org/10.1590/S0104-71832011000200005>
- Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York University Press.
- Pasquinelli, M. e Joler, V. (2020). O manifesto Nooscópio: Inteligência Artificial como Instrumento de Extrativismo do Conhecimento. KIM Research Group (Karlsruhe University of Arts and Design) e Share Lab (Novi Sad). <https://nooscope.ai>
- Pasquinelli, M. (2019). 3000 years of algorithmic rituals: The emergence of ai from the computation of space. *e-flux*, 101, 1-12. <https://hdl.handle.net/10278/5043608>
- Pasquinelli, M. (2023). *The eye of the master. A social history of artificial intelligence*. Verso.
- Pasquinelli, M. (Ed.). (2015). *Alleys of your mind. Augmented intelligence and its traumas*. Meson Press. <https://doi.org/10.14619/014>
- Samain, E. (1995). «Ver» e «dizer» na tradição etnográfica: Bronislaw Malinowski e a fotografia. *Revista Horizontes Antropológicos*, 1(2), 23-60. [http://www.cchla.ufpb.br/etienne\\_samain\\_unicamp/wp-content/uploads/2018/01/Samain-1995-Ver-e-dizer-Malinowski.pdf](http://www.cchla.ufpb.br/etienne_samain_unicamp/wp-content/uploads/2018/01/Samain-1995-Ver-e-dizer-Malinowski.pdf)
- Strathern, M. (2023). *Property, substance, and effect. Anthropological essays on persons and things*. HAU Books.

Strathern, S. (2014). *O efeito etnográfico e outros ensaios*. (1st ed.). Cosac & Naify.

Taussig, M. (2011). *I swear I saw this. Drawings in fieldwork notebooks, namely my own*. The University of Chicago Press.

Throop, C. J. (2003). Articulating experience. *Anthropological Theory*, 3(2), 219-241. <https://doi.org/10.1177/1463499603003002006>

Turner, V. (1980). Social dramas and storys about them. *Critical Inquiry*, 7(1), 141-168, <https://doi.org/10.1086/448092>

Turner, V. (1982). *From ritual to theater. The human seriousness of play*. Performing Arts Journal Publications.

UNESCO e International Research Centre on Artificial Intelligence. (2024). Challenging systematic prejudices: an investigation into bias against women and girls in large language models. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000388971>

Wagner, R. (1991). The Fractal Person. In Marilyn Strathern e Maurice Godelier (orgs.), *Big men and great men. Personifications of power in Melanesia*. Cambridge University Press.

Recepción: 12/12/2023

Aceptación: 4/3/2024