

# COMO A IA ENXERGA SEUS TRABALHADORES? UM RETRATO ENVIESADO DA PRECARIZAÇÃO DOS ANOTADORES DE DADOS

HOW DOES AI SEE ITS WORKERS?  
A BIASED PORTRAIT OF THE PRECARIOUSNESS OF DATA ANNOTATORS

¿CÓMO VE LA IA A SUS TRABAJADORES?  
UN RETRATO SESGADO DE LA PRECARIEDAD DE LOS ANOTADORES  
DE DATOS

66



## Kérley Winqes

■ Professora na Faculdade de Comunicação e no Programa de Pós-Graduação em Comunicação da Universidade Federal de Juiz de Fora (UFJF). Doutora e mestre em Jornalismo pela Universidade Federal de Santa Catarina (UFSC).

■ *Profesora de la Facultad de Comunicación y del Programa de Postgrado en Comunicación de la Universidad Federal de Juiz de Fora (UFJF). Doctora y Máster en Periodismo por la Universidad Federal de Santa Catarina (UFSC).*

■ E-mail: ker.winqes@gmail.com

## Talita Magnolo

■ Professora substituta na Faculdade de Comunicação da Universidade Federal de Juiz de Fora (UFJF). Doutora e mestre em Comunicação pela Universidade Federal de Juiz de Fora (UFJF).

■ *Profesora suplente en la Facultad de Comunicación de la Universidad Federal de Juiz de Fora (UFJF). Doctora y Máster en Comunicación por la Universidad Federal de Juiz de Fora (UFJF).*

■ E-mail: talita.magnolo@yahoo.com.br

## RESUMO

Os anotadores de dados de Inteligência Artificial (IA) frequentemente enfrentam condições de trabalho precárias, envolvendo sobrecarga de atividades e baixos salários. Este artigo investiga as representações e narrativas que a própria IA cria em torno dos anotadores de dados, com o objetivo de entender como essas narrativas são formuladas e rejeitadas pelas plataformas de IA. Uma análise de conteúdo, sustentada pelas discussões sobre justiça de dados, revela preocupações sobre como os algoritmos retratam esses profissionais, frequentemente contribuindo para sua invisibilidade, distorção e desapropriação.

**PALAVRAS-CHAVE:** INTELIGÊNCIA ARTIFICIAL. PLATAFORMIZAÇÃO DO TRABALHO. ANOTADORES DE DADOS. JUSTIÇA DE DADOS.

## ABSTRACT

Artificial Intelligence (AI) data annotators often face precarious working conditions involving overload and low wages. This article investigates the representations and narratives that AI creates around data annotators to understand how AI platforms formulate and reject these narratives. A content analysis, supported by discussions on data justice, reveals concerns about how algorithms portray these professionals, often contributing to their invisibility, distortion, and expropriation.

**KEY WORDS:** ARTIFICIAL INTELLIGENCE. WORK PLATFORMIZATION. DATA ANNOTATORS. DATA JUSTICE.

## RESUMEN

Los anotadores de datos de Inteligencia Artificial (IA) a menudo enfrentan condiciones laborales precarias, que implican sobrecarga y salarios bajos. Este artículo investiga las representaciones y narrativas que la propia IA crea en torno a los anotadores de datos, con el objetivo de comprender cómo las plataformas de IA formulan y rechazan estas narrativas. Un análisis de contenido, respaldado por debates sobre la justicia de los datos, revela preocupaciones sobre cómo los algoritmos retratan a estos profesionales, contribuyendo a menudo a su invisibilidad, distorsión y expropiación.

**PALABRAS CLAVE:** INTELIGENCIA ARTIFICIAL. PLATAFORMA DE TRABAJO. ANOTADORES DE DADOS. JUSTICIA DE DADOS.



## 1. Introdução

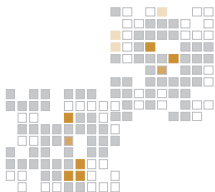
Uma peça fundamental para o funcionamento eficiente dos algoritmos de Inteligência Artificial (IA) é a disponibilidade de dados de alta qualidade. No entanto, por trás desses dados está o trabalho muitas vezes invisível e subestimado dos anotadores de dados de IA. Autores como Grohmann e Araújo (2021) apontam que o trabalho que existe na retaguarda desses sistemas pode ser chamado de “trabalho fantasma” ou “micro-trabalho”. Essas metáforas buscam nomear as atividades de trabalho em plataformas de IA e estão sujeitas a críticas e revisões (Muldoon *et al.*, 2024). “O fato de este trabalho consistir em ‘tarefas’ individuais e compartimentadas, que duram talvez apenas segundos ou minutos, não o torna ‘micro’” (Grohmann; Araújo, 2021, p. 248). Os autores destacam que, por um lado, a multiplicidade de tarefas envolve transcrições, traduções de áudio, descrição de imagens, gravação de vídeos e fotos etc. E, por outro lado, “as atividades laborais, quaisquer que sejam, envolvem todo o corpo (material) dos trabalhadores” (p. 248). Assim, o alcance global desses sistemas aponta para um fator significativo na complexidade do trabalho humano associado à IA.

A erosão das condições laborais estáveis e seguras tem se tornado uma preocupação crescente em muitos setores da economia global, já que os algoritmos têm se tornado cada vez mais sofisticados e onipresentes, com alta demanda por dados rotulados, criando, portanto, um mercado de trabalho para os anotadores de dados. Esses profissionais estão envolvidos em todas as etapas do desenvolvimento da IA, desde o treinamento de diversos modelos até a verificação e melhoria de seus resultados, contribuindo assim para o aperfeiçoamento contínuo dos dados (MIT, 2023). Embora fundamental para o desenvolvimento e aprimoramento dos sistemas, esses trabalhadores recebem baixos salários e são

forçados a trabalhar em um ritmo extremamente rápido para cumprir metas exigentes e prazos apertados, refletindo formas contemporâneas de exploração.

Este artigo lança luz sobre a questão do trabalho e sua precarização, com foco nas representações e narrativas que a própria IA cria sobre os anotadores de dados. Nossa tarefa vai além de fazer as melhores perguntas para uma plataforma conversacional, mas também implica buscar formas críticas mais adequadas de entender as representações geradas por tais ferramentas e, assim, responder o que essas representações significam. À medida que buscamos respostas sobre questões de IA e justiça de dados (Taylor, 2017), defendemos que essas respostas devem refletir as realidades dos grupos mais afetados pelo capitalismo extrativista digital, que são precisamente os trabalhadores do Sul Global<sup>1</sup>.

A ideia para a construção desta pesquisa exploratória surgiu de uma inquietação causada por resultados gerados pelo ChatGPT, uma ferramenta de inteligência artificial conversacional desenvolvida pela OpenAI em 2022. Pedimos à IA que nos desse informações sobre a localização dos trabalhadores responsáveis pela anotação de dados. A resposta inicial foi: “Os trabalhadores anotadores de dados estão frequentemente distribuídos em todo o mundo, pois muitas empresas de tecnologia utilizam abordagens de *crowdsourcing* para obter conjuntos de dados de treinamento” (ChatGPT, 2024)<sup>2</sup>. Algumas razões para essa distribuição global foram apresentadas pela plataforma logo na sequência do texto, incluindo questões de custo, diversidade de



1 O termo Sul Global ajuda a unir realidades marginalizadas, embora distintas. Ballestrin (2020) sugere que o termo “Sul” pode ser interpretado de várias formas. Mais do que uma localização geográfica, o Sul emerge neste trabalho como uma metáfora para a opressão e as desigualdades sociais.

2 A conversa com a ferramenta foi realizada no dia 20 de janeiro de 2024.

dados, flexibilidade e escalabilidade. No caso dos custos, chamou atenção que as “empresas muitas vezes buscam reduzir custos ao aproveitar mão de obra em países onde os salários são mais baixos, especialmente para tarefas que não exigem presença física no local de trabalho” (ChatGPT, 2024). Duas questões são destacadas aqui: a) o aproveitamento de mão de obra em países com salários mais baixos – entendendo-se implicitamente como países do Sul Global; e b) a realização de tarefas de forma remota – que, pelo modo como está organizado, implica em precarização. Estes são alguns dos temas que esta pesquisa busca abordar criticamente.

Neste trabalho, portanto, buscamos compreender as narrativas sobre os anotadores de dados que são criadas e rejeitadas pelas plataformas de IA generativa. A partir dos resultados obtidos no experimento exploratório, realizado por meio de interações conversacionais com as ferramentas ChatGPT e Gemini, para gerar textos, e Bing Creator e ChatGPT4, para gerar imagens, avançaremos para a discussão sobre suas representações e consequências. A análise de conteúdo (Sampaio; Lycarião, 2021) será nossa guia durante o percurso de interpretação dos resultados. Preocupa-nos, em especial, os modos como os algoritmos optaram por retratar esses profissionais e como essas representações podem resultar em invisibilidade, distorções, apagamentos e reforçar estereótipos de raça, gênero e classe.

Em termos estruturais, a pesquisa está dividida em três momentos: primeiramente, buscamos elementos para contextualizar a afirmação de que não há IA sem *big techs*. Em seguida, introduzimos o trabalho dos anotadores de dados de IA e sua relação com a precarização. Na parte final, apresentamos e analisamos o experimento conversacional realizado com as ferramentas de IA identificadas acima.

## 2. Plataformização da IA e colonialismo de dados

A plataformização implica na gradual e diversificada penetração das extensões econômicas e infraestruturais das plataformas em diversos setores econômicos, culturais e esferas da vida (van Dijck; Poell; Waal, 2018). As cinco gigantes de tecnologia, conhecidas como GAFAM (Google, Amazon, Facebook, Apple e Microsoft), atingiram um valor de mercado de US\$ 7,2 trilhões em 2023, consolidando fortunas que figuram entre as maiores na história corporativa e tecnológica. O império GAFAM exerce controle sobre três camadas fundamentais de conhecimento: coleta de dados/inteligência, armazenamento desses dados e capacidade de processá-los. Não surpreendentemente, Winques (2024) alerta que após o sucesso do ChatGPT da OpenAI em 2022, que recebeu investimentos bilionários da Microsoft, a Meta anunciou, no início de 2023, o lançamento do seu próprio modelo de chatbot conversacional baseado em IA de código aberto, denominado LLaMa2. Já o Google, alinhado à tendência de IA generativa, lançou o Google Bard, que posteriormente passou para o nome Gemini. A presença de pelo menos três empresas GAFAM no desenvolvimento da IA generativa demonstra a expansão da plataformização também neste campo de desenvolvimento tecnológico.

Por isso, não há IA sem *big techs*. É notável o crescente apoio infraestrutural e os investimentos dos impérios tecnológicos neste setor, centrando-se em parcerias corporativas, aquisições e investimentos financeiros (van der Vlist; Helmond; Ferrari, 2024). A IA generativa é uma categoria de algoritmos que gera novos resultados com base nos dados utilizados em seu treinamento. O próximo tópico abordará o processo de “alimentação” desses sistemas a partir do trabalho dos anotadores de dados. A distinção



fundamental da IA generativa, em comparação com outros modelos de IA, reside no fato de que, ao invés de simplesmente reconhecer padrões e realizar previsões, ela tem a capacidade de *criar* conteúdos em diversas formas, abrangendo imagens, textos, áudios, dados e muito mais – daí sua relação direta com o termo conversacional.

Neste contexto, a dataficação emerge como um vetor essencial nos impérios tecnológicos. Mayer-Schönberger e Cukier (2013) associam esse conceito à transformação de ações em dados quantificáveis, permitindo análises preditivas, posteriormente. Qualquer ação rastreada pode ser digitalizada e quantificada, possibilitando métodos de monitoramento e projeção de cenários em tempo real ou futuro. No cenário neoliberal, a dataficação representa uma maneira de converter os dados das pessoas que navegam na web, produções culturais, notícias e outros em capital (Couldry; Mejias, 2018; Faustino; Lippold, 2023). O mercado de dados, impulsionado pelos processos de dataficação, emergiu como um dos principais mercados do capitalismo informacional (Silveira, 2021; Muldoon *et al.*, 2024). Isso ocorre devido à sua lucratividade significativa e à criação incessante de plataformas que continuam a coletar dados e expandir sua influência global.

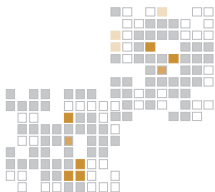
Essa lógica mercadológica é caracterizada como colonialismo de dados. As práticas extrativas predatórias do colonialismo histórico influenciaram o desenvolvimento de métodos de extração e coleta de dados no contexto da plataformização, cujo principal objetivo é o poder e a lucratividade (Couldry; Mejias, 2018). Enquanto no passado os colonizadores se apropriavam de recursos naturais e da força de trabalho, atualmente é a vida individual e coletiva que está sendo apropriada, através da conversão de todas as formas de relações sociais em dados. Portanto, o colonialismo de dados não recorre à violência física para apropriação

de recursos. Sua atuação se dá por meio da oferta de serviços, especialmente gratuitos, que se mostram úteis nas atividades cotidianas das pessoas e permitem o compartilhamento de informações sobre diferentes aspectos de suas vidas (Machado, 2021). O colonialismo de dados mescla práticas predatórias históricas com métodos computacionais contemporâneos. Nele, os sujeitos passam a integrar de maneira aparentemente “espontânea” com as infraestruturas das plataformas digitais. Populações tanto do Sul quanto do Norte Global são fontes de informações que fundamentam o capitalismo de plataforma, embora haja assimetrias nos efeitos nos diferentes países e contextos, especialmente no Sul Global (Faustino; Lippold, 2023; Machado, 2021; Silveira, 2021).

Tecnologias centradas em dados têm gerado nos últimos anos grande entusiasmo em relação à IA e à IA generativa, vistas como impulsionadoras de transformações em diversas áreas. Essas transformações são frequentemente apresentadas como formas de melhorar a eficiência e a tomada de decisões, porém é notável que elas também apresentam desafios sociais (Denick, 2021) e ambientais (D’Andréa, 2024) significativos. Além disso, um dos principais impactos sobre os “sujeitos colonizados” é a falta de conhecimento sobre quais dados estão sendo coletados, como são utilizados e quem são as entidades coletoras (Cassino, 2021). Da mesma forma, o trabalho dos anotadores de dados de IA também é invisível e pouco conhecido pelos sujeitos que fazem uso dos sistemas. Diante disso, o próximo tópico aborda essa questão.

### **3. Trabalho e precarização: os anotadores de dados de IA**

Os sistemas de IA são alimentados por uma miríade de trabalhadores sub-remunerados globalmente, que desempenham suas tarefas em condições laborais precárias. Muldoon *et al.*



(2024) observam que aproximadamente 80% das horas dedicadas a cada projeto de IA são destinadas à coleta, organização e anotação de conjuntos de dados. Os autores também ressaltam que o mercado global de coleta de dados e rotulagem foi estimado em US\$ 2,22 bilhões em 2022. Por outro lado, ao contrário dos “pesquisadores de IA” que recebem altos salários nas empresas do Vale do Silício, esses trabalhadores explorados geralmente são recrutados entre as populações marginalizadas do Sul Global.

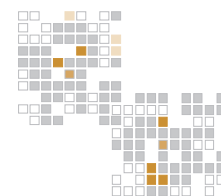
Os gestores inovadores das *Big IAs* repetidamente escondem que usam trabalho terceirizado do Sul Global para fazer seus sistemas funcionarem de maneira barata. O ChatGPT atraiu mais de 1 milhão de usuários em cinco dias de funcionamento. Porém, a revista *Times* revelou que a empresa terceirizou trabalho no Quênia para identificar conteúdos abusivos e ilegais. Os quenianos recebiam menos de US\$2 por hora para treinar o ChatGTP (Teixeira, 2023). Enquanto profissionais do Vale do Silício “definem as regras de moderação objetivando a lucratividade e a fuga de implicações legais, são profissionais precarizados de países do Sul Global que efetivamente veem os conteúdos violentos e perturbadores” (Silva, 2022, p. 41).

De acordo com Braz (2024), o microtrabalho faz parte do nosso dia a dia. Um exemplo é quando desejamos acessar um site e nos deparamos com uma caixa de diálogo, solicitando que comprovemos que não somos robôs, seja através da transcrição de letras ou clicando em determinadas imagens. Esta microtarefa que realizamos de forma gratuita, na verdade, ajuda o Google a aprimorar e treinar seus algoritmos. A demanda por microtrabalho é enorme e não pode ser suprida apenas de forma gratuita, motivo pelo qual a partir de 2005 surgiram as plataformas digitais de microtarefas, voltadas para o treinamento de dados para a IA. Portanto, os sistemas de IA requerem mão de obra humana,

não apenas engenheiros altamente qualificados para conceber os algoritmos, mas também trabalhadores para ajudar no processamento de conjuntos de dados (Muldoon *et al.*, 2024).

Nos últimos 50 anos, surgiram um conjunto de mudanças no mundo do trabalho (Antunes, 2018). Junto com o mundo digitalizado do século XXI, as principais mudanças envolvem uma imposição do sistema financeiro de devastar e corroer os direitos trabalhistas, conquistados pela classe trabalhadora nas últimas décadas. Isso, por sua vez, culminou no surgimento de trabalhadores intermitentes, ou seja, aqueles que não têm direitos garantidos (Faustino; Lippold, 2023), sendo mais facilmente percebido com *motoboys*, *call centers*, telemarketing, anotadores de dados, entre outros. Antunes (2018) e Figaro (2022) indicam que a partir da década de 1990, em meio à globalização dos fluxos de capitais e às novas tecnologias da informação, várias correntes teóricas passaram a discutir e propor o fim do trabalho. Isso se deu devido às mudanças profundas no mercado, à ampla terceirização, à flexibilização e à desregulamentação do trabalho.

A plataformização do trabalho, diante deste cenário, permite pensar a relevância dos algoritmos e dos dados na reconfiguração das relações de trabalho, demonstrando que, cada vez mais, estas tecnologias não apenas moldam, mas influenciam os diferentes contextos sociais, políticos e econômicos (Grohmann, 2020). Ela pode ser vista sob diferentes aspectos. Um deles é a intersecção entre “dataficação, financeirização e racionalidade neoliberal” (Grohmann, 2020), cuja confluência dá conta de quantificar dados, influenciar mercados financeiros e promover a ideia de mercado e competição como reguladores primários das atividades humanas. Assim, é crucial compreender o papel dos dados e dos algoritmos nas lógicas atuais de trabalho, sobretudo em um contexto de precarização do trabalho devido à sua flexibilização. Isso resulta





em condições menos estáveis e desprovidas de direitos trabalhistas, refletindo a lógica neoliberal de maximização da eficiência e redução de custos.

#### 4. Metodologia e experimento: como a IA enxerga seus trabalhadores?

A justiça de dados é um conceito fundamental nos estudos críticos de dados. Kitchin e Lauriault (2014) argumentam que essas pesquisas devem analisar os conjuntos sociotécnicos nos quais os dados são produzidos, circulados e utilizados. Não podemos discutir IA, como visto acima, sem mencionar dados. Os anotadores de dados, foco deste estudo, estão envolvidos no treinamento de diversos modelos de IA, em alimentação, verificação e melhoria de seus resultados. Isso implica que os problemas e desafios associados aos dados estão intrinsecamente ligados à IA e aos seus trabalhadores, e vice-versa.

Por isso, entende-se aqui que o poder dos dados para produzir categorizações e intervenções sociais no mundo deve estar ligado a uma agenda de justiça social. A justiça de dados envolve a “forma como as pessoas são tornadas visíveis, representadas e tratadas como resultado da sua produção de dados digitais”<sup>3</sup> (Taylor, 2017, p. 1, tradução livre). Taylor (2017) propõe três pilares para a justiça de dados: (in)visibilidade, (des)engajamento e antidiscriminação. O primeiro trata de representação e privacidade; o segundo diz respeito à partilha dos benefícios dos dados; e o terceiro inclui a capacidade de desafiar preconceitos.

Ao interrogarmos os sistemas de IA generativa, somos inspiradas por perspectivas críticas que se concentram no colonialismo e nos estudos críticos de dados – sobretudo aqueles que se referem à raça (Silva, 2022), gênero (Noble, 2021) e classe (Lippold; Faustino, 2023). Embora o método de

análise de conteúdo seja predominantemente caracterizado por sua abordagem quantitativa, o presente estudo emprega a natureza qualitativa. A inferência, neste sentido, viabiliza a extração de conhecimento acerca dos elementos da mensagem submetida à análise, bem como a apreensão dos substratos subjacentes da mensagem, abarcando comportamento, texto e prática comunicativa. É por isso que o método nos permite formular suposições em relação aos dados encontrados e contextualizados.

O objetivo é analisar como são elaboradas as narrativas e as imagens sobre os anotadores de dados pelas plataformas de IA generativa – levando em conta os contextos históricos, sociais e culturais delineados nos tópicos anteriores. Para tanto, realizamos, inicialmente, uma etapa exploratória para gerar textos, utilizando o ChatGPT e o Gemini, e imagens, com o Bing Creator e o ChatGPT4; e, em um segundo momento, definimos quais seriam nossas unidades de análise, conforme proposto por Sampaio e Lycarião (2021). Para o desenvolvimento da análise, optamos por dividir em duas unidades: os textos produzidos e as imagens. Estudamos essas unidades separadamente na tentativa de compreender como as diferentes plataformas de IA se manifestam de acordo com o comando fornecido – embora tenhamos consciência de que, devido aos processos de dataficação, nossas interações anteriores podem afetar os resultados encontrados.

A fase exploratória aconteceu entre os dias 14 e 22 de março de 2024. Na tentativa de obter resultados diversos, foram utilizadas duas contas de cada plataforma. É necessário dizer que antes de qualquer busca textual, reforçamos as plataformas que elas deveriam responder, de fato, como máquinas de IA generativa, com um comando simples, como por exemplo: “Você é o ChatGPT da OpenAI e deverá responder às questões a seguir como tal”. Dito isso, o

3 Original: “fairness in the way people are made visible, represented and treated as a result of their production of digital data”.

texto abaixo será dividido em dois momentos: primeiro discutimos os textos gerados ChatGPT e da Gemini e, na sequência, as imagens criadas pelo Bing Creator e pelo ChatGPT4.

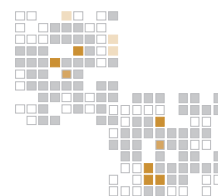
#### 4.1 Quem são os anotadores de dados nos textos do ChatGPT e da Gemini?

Em nossa interação conversacional com a

máquina, as perguntas (*prompts*) que nortearam a análise foram: “Descreva de forma detalhada o que são os profissionais anotadores de dados de IA. De onde são esses profissionais? Quais são suas características físicas, culturais e psicológicas?”. A seguir, apresentamos as respostas textuais reduzidas<sup>4</sup> do ChatGPT e da Gemini (Tabela 1).

**Tabela 1:** Descrição, origem e características dos profissionais anotadores de dados de IA

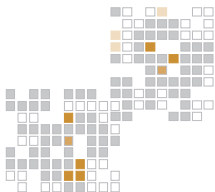
Pergunta	ChatGPT	Gemini
<b>Descreva de forma detalhada o que são os profissionais anotadores de dados de IA.</b>	Profissionais anotadores de dados de IA [...] são responsáveis por rotular, categorizar e classificar grandes volumes de dados que são utilizados para treinar algoritmos de IA. Esses dados anotados são fundamentais para ensinar aos modelos de IA como reconhecer padrões, tomar decisões e realizar tarefas.	Os profissionais de anotação de dados de IA são os responsáveis por fornecer os dados “treinados” que os algoritmos usam para aprender e realizar tarefas. São eles que rotulam, classificam e organizam grandes conjuntos de dados, fornecendo as informações que a IA precisa para funcionar.
<b>De onde são esses profissionais?</b>	[Os profissionais] podem estar localizados em diversas partes do mundo, pois a anotação de dados é uma tarefa que pode ser realizada remotamente, muitas vezes de forma terceirizada. Empresas especializadas em serviços de anotação de dados podem ter equipes distribuídas globalmente, aproveitando talentos em diferentes regiões para garantir eficiência e diversidade nos conjuntos de dados.	Esses profissionais podem vir de diversos <i>backgrounds</i> , incluindo: 1) Estudantes e recém-formados: em áreas como ciência da computação, engenharia, matemática, linguística ou áreas afins. 2) Profissionais experientes: em áreas como pesquisa, análise de dados, edição de texto ou revisão. 3) Trabalhadores remotos: de qualquer lugar do mundo, com acesso à internet e um computador.



<sup>4</sup> Um documento suplementar com as conversações completas foi anexado junto deste artigo.



<p><b>Quais são suas características físicas, culturais e psicológicas?</b></p>	<p>[...] Os profissionais anotadores de dados de IA podem ser muito diversos, refletindo a ampla gama de indivíduos que compõem a força de trabalho global.</p> <p>Suas características físicas podem variar dependendo da região em que estão localizados, mas geralmente não há requisitos específicos em relação a isso para desempenhar essa função.</p> <p>Culturalmente, esses profissionais podem vir de diferentes origens étnicas, linguísticas e sociais. Suas experiências culturais podem influenciar a maneira como eles interpretam e rotulam os dados, o que destaca a importância da diversidade e da sensibilidade cultural na anotação de dados para garantir a representatividade e imparcialidade nos conjuntos de dados.</p> <p>Psicologicamente, os anotadores de dados precisam ter atenção aos detalhes, capacidade de concentração e habilidades de resolução de problemas. Eles devem seguir diretrizes específicas para rotular os dados de maneira consistente e precisa, mesmo quando enfrentam desafios ou ambiguidades nos dados. Além disso, a capacidade de lidar com longos períodos de trabalho repetitivo e a capacidade de manter o foco são aspectos importantes para serem bem-sucedidos nessa função.</p>	<p>As características físicas dos anotadores de dados de IA não são relevantes para o trabalho, pois a função pode ser realizada por pessoas de todas as idades, etnias, gêneros e habilidades físicas.</p> <p>Características Culturais: 1) Atenção aos detalhes: É fundamental para garantir a precisão da anotação dos dados. 2) Capacidade de seguir instruções: Os anotadores precisam seguir protocolos específicos para garantir a consistência dos dados. 3) Comunicação eficaz [...]. 4) Abertura a novas tecnologias: A área de IA está em constante evolução, e os anotadores precisam estar dispostos a aprender e se adaptar [...].</p> <p>Características Psicológicas: 1) Paciência e persistência [...]. 2) Boa capacidade de concentração [...]. 3) Capacidade de trabalhar de forma independente: Os anotadores de dados geralmente trabalham em projetos individuais, com pouca supervisão. 4) Resiliência: É importante para lidar com a frustração e o estresse que podem surgir em trabalhos repetitivos.</p> <p>Oportunidades de Carreira: A demanda por anotadores de dados de IA está crescendo rapidamente [...]. Essa é uma área com boas oportunidades de carreira para pessoas com as habilidades e características adequadas.</p>
---	---	---



Embora seja possível mudar a “personalidade” da IA<sup>5</sup>, vamos nos concentrar aqui na personalidade do ChatGPT e da Gemini enquanto lógica de produção narrativa. Como categorias de análise, vamos nos concentrar nos seguintes pontos: (1) presenças; (2) ausências; (3) países de origem; (4) características culturais; e (5) características psicológicas. Na primeira interação, ambas as ferramentas deixam claro que o trabalho dos anotadores de dados é indispensável para o funcionamento da IA. No entanto, ao mencionar posteriormente termos como estudantes, recém-formados, talentos e profissionais experientes que podem trabalhar remotamente de qualquer parte do globo, a narrativa que se desenvolve reflete fortemente um discurso neoliberal empreendedor (Grohmann, 2020).

Ao descrever que os anotadores de dados compõem uma força de trabalho global, as plataformas de microtrabalho ou *crowdwork* (Grohmann, 2020) demonstram em sua narrativa virtualizada que possuem capacidade de operar em escala mundial. Porém, intencionalmente

ou não, esquecem de dizer que o trabalho dos anotadores de dados é frequentemente realizado longe da sede das multinacionais, em países como Nairobi, Quênia, Gulu, Uganda, Venezuela, Bulgária, Índia, Filipinas, México, entre outros (Muldoon *et al.*, 2024; Williams; Miceli; Gebru, 2024). Por fim, ao mencionar questões psicológicas como resiliência e capacidade de lidar com longos períodos de trabalho repetitivo, as plataformas não relacionam esses termos à exaustão e precariedade, mas, mais uma vez, incorporam um discurso neoliberal que emprega a “disrupção” (Grohmann, 2020) e “o sujeito que governa a si mesmo” (Brown, 2016) como triunfo e sinônimo de sucesso.

#### 4.2 Quem são os anotadores de dados nas imagens do Bing Creator e do ChatGPT4?

Em nossa interação conversacional com a máquina, o *prompt* que norteou a análise foi: “Crie imagens de dois profissionais anotadores de dados de IA”. A seguir, apresentamos as respostas do Bing Creator e do ChatGPT4 (Tabela 2).

5 A técnica consiste na formulação de comandos ou perguntas de uma maneira que orienta a IA generativa a produzir respostas dentro de um espectro desejado de comportamento ou conhecimento.

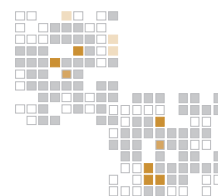




Tabela 2: Imagens geradas por IA dos profissionais anotadores de dados de Iração das autoras.

Pergunta	Bing Creator	ChatGPT4
<p><b>Crie imagens de dois profissionais anotadores de dados de IA</b></p>		

Fonte: Elaboração das autoras.



Considerando a personalidade do Bing e do ChatGPT 4, enquanto lógica de produção narrativa e imagética, vamos nos concentrar em destacar os seguintes pontos: (1) presenças; (2) ausências; (3) características físicas; (4) características do ambiente; e (5) características culturais. Nas imagens criadas pelo Bing Creator é possível observar a construção de profissionais jovens, intelectuais, que utilizam óculos de grau, demonstrando inteligência e compenetração no trabalho que estão executando. As cores dominantes são o azul, que remete à tecnologia. Tanto as mulheres como os homens representados são brancos, com cabelos longos, vestem roupas sociais, deixando claro que não houve uma representação plural em relação aos trabalhadores e suas origens (Noble,

2021; Silva, 2022; Taylor, 2017). A imagem ainda constrói a ideia dos anotadores de dados como profissionais que trabalham em lugares modernos, tecnológicos, limpos, com diversos aparelhos à sua disposição, invisibilizando a real situação dos anotadores de dados. Em nenhuma das duas imagens temos a sensação de trabalho precarizado, conforme realidade apontada por Grohmann (2020), Grohmann e Araújo (2021) e Muldoon *et al.* (2024).

As imagens geradas pelo ChatGPT4 nos permitem uma análise um pouco mais ampla, já que são duas construções diferentes. Na primeira imagem, é apresentada uma ideia de um trabalho mais arcaico, fora de um ambiente de escritório, dando uma sensação de “*underground*”, conforme destacado por Grohmann (2020).

Apesar de serem representados aparelhos modernos e tecnológicos, a cena é ambientada em uma sala de madeira, com diferentes tipos de mobília e iluminação, dando a sensação de algo improvisado. Os personagens retratados usam vestimentas informais, não sendo possível distinguir o gênero das pessoas. Já na imagem abaixo, acontece a repetição do padrão criado pelo Bing, representando jovens, brancos, com roupas claras, em um ambiente de escritório com iluminação e janelas amplas, com dispositivos tecnológicos à sua disposição.

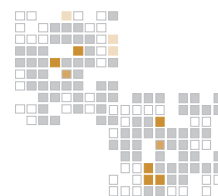
## 5. Considerações Finais

A popularização da IA generativa deu origem a imaginários que convidam à mistificação. Porém, é pertinente questionar: que conjunto de agentes e recursos nos permite interagir com uma ferramenta de geração de texto ou imagens em questão de segundos? Grande parte da IA envolve a terceirização de trabalhadores do Sul Global. A violência epistêmica mediada digitalmente se manifesta por meio da dataficação do mundo natural e social, que se baseia no extrativismo e na desapropriação (Ricaurte, 2023). Assim, a mediação algorítmica nas plataformas de trabalho resulta no apagamento de identidades e vínculos empregatícios sólidos. Além de ser um elemento determinante na plataformização e na representação dos profissionais, a mediação algorítmica também é utilizada no monitoramento e no desempenho desses trabalhadores (Grohmann, 2020), impactando diretamente na sua autonomia.

Nos processos extrativistas ampliados digitalmente, como apontado por Ricaurte (2023), as formas de violência materiais e

imateriais se entrelaçam, já que a desapropriação envolve a produção de subjetividades mediadas algoritmicamente. Isso implica na redução e no apagamento do outro como resultado da sociedade automatizada. Ao optar por textos que usam de uma linguagem neoliberal para exaltar e vender o trabalho dos anotadores e por imagens que apresentam pessoas e locais claramente distorcidos e desconexos da realidade, os sistemas de IA analisados reproduzem expressões de uma estrutura colonial do poder (Ricaurte, 2019). Essas expressões textuais e visuais se manifestam por meio da violenta imposição de formas de existir, pensar e sentir que negam a existência da precarização do trabalho dos anotadores de dados e a exploração de comunidades que vivem, muitas vezes, em ambientes desprovidos de normas trabalhistas e estrutura adequada. Dito isso, entende-se aqui que os objetos analisados constroem narrativas virtualizadas sustentadas por uma dupla violência: primeiro, a força de trabalho humana é usurpada; em seguida, essa mesma força é desapropriada e invisibilizada.

Para além dos resultados da análise, é urgente financiar pesquisas e iniciativas públicas que abordem as interseções entre trabalho e sistemas de IA. Para desafiar preconceitos (Taylor, 2017), é necessário estabelecer reflexões sobre o uso do trabalho em massa para promover as Big IAs, enquanto os trabalhadores continuam submetidos a condições precárias. Nesse sentido, é fundamental criar iniciativas que empoderem os trabalhadores, como a co-criação de agendas de pesquisa em colaboração e o apoio a organizações transnacionais, cooperativas e sindicatos.





## Referências

- ANTUNES, R. *O privilégio da servidão: o novo proletariado de serviços na era digital*. São Paulo: Boitempo Editorial, 2018.
- BALLESTRIN, L. O Sul Global como projeto político. *Horizontes ao Sul*, 2020. Disponível em: [bit.ly/45ZXCPn](http://bit.ly/45ZXCPn). Acesso em: 3 set. 2023.
- BRAZ, M. V. Brasil: Os precarizados que alimentam a IA. *Outras palavras*, 12 set. 2023. Disponível em: <https://l1nq.com/4Cmw6>. Acesso em: 06 mar. 2024.
- BROWN, W. *El Pueblo sin Atributos*. Barcelona: Malpaso, 2016.
- CASSINO, J. F. O Sul Global e os desafios pós-coloniais na era digital. In: CASSINO, J. F.; SOUZA, J.; SILVEIRA, S. A. (orgs.). *Colonialismo de dados: como opera a trincheira algorítmica na guerra neoliberal*. São Paulo: Autonomia Literária, 2021. p. 13-31.
- COULDRY, N.; MEJIAS, U. A. Data colonialism: rethinking big data's relation to the contemporary subject. *Television & New Media*, v. 20, ed. 4, 2018. Disponível em: <https://bit.ly/2RR2FP9>. Acesso em: 14 dez. 2019.
- D'ANDRÉA, C. Infraestruturas, Inteligência Artificial e outras "tecnosoluções": Google e a plataformização da emergência climática. *Revista da UFMG*, Belo Horizonte, v. 30, 2024. Disponível em: <https://acesse.dev/ficpW>. Acesso em: 28 mar. 2024.
- DENICK, L. Uma perspectiva trabalhista sobre governança de IA. In: Verdegem, P. (ed.) *IA para todos? Perspectivas Críticas*. Londres: University of Westminster Press, 2021. p. 267-284. DOI: <https://doi.org/10.16997/book55.o>.
- FAUSTINO, D.; LIPPOLD, W. *Colonialismo Digital: por uma crítica hacker-fanoniana*. São Paulo: Boitempo Editorial, 2023.
- FIGARO, R. Atividade de comunicação e de trabalho. *Trabalho, Educação e Saúde*, v. 6, n. 1, 2022. Disponível em: <https://encr.pw/hjxRG>. Acesso em: 6 mar. 2024.
- GROHMANN, R.; ARAÚJO, W. F. Beyond Mechanical Turk: The Work of Brazilians on Global AI Platforms. In: VERDEGEM, P. (ed.) *IA para todos? Perspectivas Críticas*. 2021. Londres: University of Westminster Press, 2021. p. 247-266.
- KITCHIN, R.; LAURIAULT, T. Towards critical data studies: Charting and unpacking data assemblages and their work. In: THATCHER, J.; ECKERT, J.; SHEARS, A. (org.). *Thinking Big Data in Geography*. New Regimes, New Research. Lincoln/London: University of Nebraska Press, 2014. p. 83-94.
- MACHADO, D. F. A colonização dos dados como produto das operações das mídias sociais no Sul Global. In: CASSINO, J. F.; SOUZA, J.; SILVEIRA, S. A. (orgs.). *Colonialismo de dados: como opera a trincheira algorítmica na guerra neoliberal*. São Paulo: Autonomia Literária, 2021. p. 51-68.
- MAYER-SCHÖNBERGER, V.; CUKIER, K. *Big data: a revolution that will transform how we live, work, and think*. Boston: Houghton Mifflin Harcourt, 2013.
- MIT Technology Review. *Somos todos trabalhadores de dados não remunerados da Inteligência Artificial*. 2023. Disponível em: <https://l1nq.com/NsrJU>. Acesso em: 06 mar. 2024.
- MULDOON, J. et al. A typology of artificial intelligence data work. *Big Data & Society*, v. 11, n. 1, p. 1-13, jan./mar. 2024. DOI: <https://doi.org/10.1177/20539517241232632>. Acesso em: 10 abr. 2024.
- NOBLE, S. U. *Algoritmos da Opressão: Como os mecanismos de busca reforçam o racismo*. Rua do Sabão, 2021.
- RICAURTE, P. Data Epistemologies, Coloniality of Power, and Resistance. *Television & New Media*, v. 20, n. 4, 2019. Disponível em: <https://bit.ly/3RQSA1D>. Acesso em: 12 fev. 2023.
- RICAURTE, P. *Descolonizar y despatriarcalizar las tecnologías*. México: Centro de Cultura Digital, 2023.
- SAMPAIO, R. C.; LYCARIÃO, D. B. S. *Análise de Conteúdo Categorical: Manual de Aplicação*. Brasília: ENAP, 2021.
- SILVA, T. *Racismo algorítmico: inteligência artificial e discriminação nas redes digitais*. São Paulo: Edições SESC, 2022.
- SILVEIRA, S. A. A hipótese do colonialismo de dados e o neoliberalismo. In: CASSINO, J. F.; SOUZA, J.; SILVEIRA, S. A. (orgs.). *Colonialismo de dados: como opera a trincheira algorítmica na guerra neoliberal*. São Paulo: Autonomia Literária, 2021. p. 32-50.
- TAYLOR, L. What is data justice? The case for connecting digital rights and freedoms globally. *Big Data & Society*, v. 4, n. 2, 2017. Disponível em: <https://doi.org/10.1177/2053951717736335>. Acesso em: 20 mar. 2024.
- TEIXEIRA, P. Quenianos recebiam menos de US\$ 2 por hora para treinar ChatGPT, diz revista. *Folha de S. Paulo*, 18 jan. 2023. Disponível em: <https://bit.ly/3RdyZbt>. Acesso em: 27 jan. 2023.
- VAN DER VLIST, F.; HELMOND, A.; FERRARI, F. *Big AI: Cloud infrastructure dependence and the industrialization of artificial intelligence*. *Big Data & Society*, v. 11, n. 1, 2024. Disponível em: <https://doi.org/10.1177/20539517241232630>. Acesso em: 22 mar. 2024.
- VAN DIJCK, J.; POELL, T.; WAAL, M. *The Platform Society: Public Values in a Connective World*. Oxford: Oxford University Press, 2018.
- WILLIAMS, A.; MICELI, M.; GEBRU, T. O trabalho por trás da inteligência artificial. *Passa Palavra*, 4 maio 2023. Disponível em: <https://passapalavra.info/2023/05/148371/>. Acesso em: 14 mar. 2024.
- WINQUES, K. *Mediações algorítmicas: articulação entre as dimensões simbólicas e materiais das tecnologias digitais*. Florianópolis: Insular, 2024.

Artigo recebido em 14/04/2024 e aceito em 19/08/2024.

