

# A PLATAFORMIZAÇÃO DA MISOGINIA NA PUBLICIDADE DIGITAL EM REDES SOCIAIS: UM PRODUTO DA OPACIDADE ALGORÍTMICA E DA AUSÊNCIA DE MARCO REGULATÓRIO DA INTELIGÊNCIA ARTIFICIAL

## THE PLATFORMIZATION OF MISOGYNY IN DIGITAL ADVERTISING ON SOCIAL MEDIA: A PRODUCT OF ALGORITHMIC OPACITY AND THE ABSENCE OF REGULATORY FRAMEWORK FOR ARTIFICIAL INTELLIGENCE

Deise Brião Ferraz<sup>1</sup>  
Marli Marlene Moraes da Costa<sup>2</sup>

**RESUMO:** Esta pesquisa tem por objeto a “plataformização da misoginia” que é um termo livremente adaptado por estas autoras a partir do conceito desenvolvido por Khoo (2021) no âmbito do relatório “*Deplatforming Misogyny*”. O *Deplatforming* trabalha a ideia de remover ou banir, das plataformas digitais, indivíduos, grupos ou conteúdos capazes de promover atos misóginos ou discriminatórios contra mulheres. Já o que entendemos por plataforma da misoginia diz respeito ao contrário: o agrupamento, fortalecimento e disseminação da misoginia em plataformas de redes sociais a partir da *dataficação* da publicidade digital como meio capaz de promover e fortalecer este encontro. Tem-se por objetivo geral demonstrar que a plataforma da misoginia na publicidade digital está associada à opacidade algorítmica e à ausência de um marco regulatório para a Inteligência Artificial no Brasil. Para operacionalizar a proposta, tem-se, ainda, por objetivos: desenvolver o conceito de plataforma da misoginia e apontar sua ocorrência na publicidade digital em plataformas de redes sociais; relacionar a opacidade algorítmica e a ausência de um marco regulatório da IA como vetores que estimulam e potencializam esse cenário. Trata-se de pesquisa exploratória, com método de pesquisa bibliográfico e documental.

**Palavras-chave:** Gênero, inteligência artificial, discriminação algorítmica, misoginia, plataforma da misoginia.

**ABSTRACT:** This research aims to address the "platformization of misogyny," a term freely adapted by the authors from the concept developed by Khoo (2021) in the report "Deplatforming Misogyny." Deplatforming works on the idea of removing or banning individuals, groups, or content from digital platforms capable of promoting misogynistic or discriminatory acts against women. Conversely, what we understand by the platformization of misogyny concerns the aggregation, strengthening, and dissemination of misogyny on social media platforms through the datafication of digital advertising as a means to promote

---

<sup>1</sup> Doutoranda em Direito no PPGD/UNISC, com bolsa PROSUC/CAPES. Mestre em Direito e Justiça Social pelo PPGD/FURG, com bolsa CAPES/DS Líder de Pesquisa no grupo "Direito Sistemico e MASC" (CNPq/FURG) Membro do Grupo de Pesquisa Direito, Cidadania e Políticas Públicas (CNPq/UNISC) Jornalista, Advogada (OAB/RS 107.934) e Psicanalista

<sup>2</sup> Doutora em Direito pela Universidade Federal de Santa Catarina - UFSC, com pós-doutoramento em Direito pela Universidade de Burgos – Espanha.

and reinforce this encounter. The general objective is to demonstrate that the platformization of misogyny in digital advertising is one of the outcomes of algorithmic opacity and the absence of a regulatory framework for Artificial Intelligence in Brazil. To operationalize the proposal, the specific objectives are: to develop the concept of the platformization of misogyny and to indicate its occurrence in digital advertising on social media platforms; to relate algorithmic opacity and the absence of an AI regulatory framework as vectors that stimulate and potentiate this scenario. This is an exploratory research, utilizing bibliographic and documentary research methods.

**Keywords:** Algorithmic discrimination, artificial intelligence, gender, misogyny, platformization of misogyny.

## INTRODUÇÃO

A presente pesquisa tem por objeto a “plataformização da misoginia” - termo livremente adaptado por estas autoras a partir do conceito desenvolvido por Cynthia Khoo, no âmbito do *Women’s Legal Education and Action Fund* (LEAF), Canadá, publicado em 2021, cujo nome é “*Deplatforming Misogyny*”. O *Deplatforming* trabalha a ideia de remover ou banir, das plataformas digitais, indivíduos, grupos ou conteúdos capazes de promover atos misóginos ou discriminatórios contra mulheres. Já o que entendemos por plataformização da misoginia diz respeito ao contrário: o processo de agrupamento, fortalecimento e disseminação da misoginia em plataformas de redes sociais a partir da *dataficação*<sup>3</sup> da publicidade digital como meio capaz de promover e fortalecer este encontro.

Tem-se por objetivo geral demonstrar que a plataformização da misoginia na publicidade digital é um dos resultados da opacidade algorítmica e da ausência de um marco regulatório para a Inteligência Artificial (IA) no Brasil. Para operacionalizar a pesquisa, tem-se por objetivos específicos: desenvolver o conceito de plataformização da misoginia e apontar sua ocorrência na publicidade digital em plataformas de redes sociais a partir da compilação bibliográfica selecionada; relacionar a opacidade algorítmica e a ausência de um marco regulatório da IA como vetores que estimulam e potencializam esse cenário.

As plataformas de redes sociais não inventaram a misoginia, mas lucram com ela. Este é o achado reunido por esta pesquisa após a compilação de importantes estudos que demonstram o impacto algorítmico da publicidade digital no acirramento da desigualdade e violência de gênero, especialmente a partir dos recentes dados coletados no relatório “Observatório da indústria da desinformação e violência de gênero nas plataformas digitais”, conduzido pelo NetLab – o Laboratório de pesquisa em Internet e Redes Sociais da Universidade Federal do Rio de Janeiro (UFRJ), sob encomenda do Ministério das Mulheres do Governo Federal brasileiro, no âmbito da iniciativa Brasil Sem Misoginia.

---

<sup>3</sup> Pode-se dizer da tendência de transformar dados da vida de um usuário em informação que pode ser comercializada.

Diante da recenticidade dos dados e conclusões, bem como do alcance deste trabalho ao versar sobre tema cotidianamente presente na vida das pessoas e desacompanhado de uma reflexão correspondente ao seu impacto, justifica-se a relevância deste artigo. Trata-se de pesquisa exploratória, com método de pesquisa bibliográfico e documental e o método de procedimento é monográfico.

## 1 ENTENDENDO A PLATAFORMIZAÇÃO DA MISOGINIA

Compreenderemos a ideia de “plataformização da misoginia<sup>4</sup>” a partir do conceito original de onde partimos, que é o relatório escrito por Cynthia Khoo, no âmbito do *Women’s Legal Education and Action Fund (LEAF)*, Canadá, publicado em 2021, cujo nome é “*Deplatforming Misogyny*”. Trata-se do conceito de “desplataformar a misoginia” – em tradução livre -, que desenvolve a ideia de remover ou banir, das plataformas digitais, indivíduos, grupos ou conteúdos capazes de promover atos misóginos ou discriminatórios contra mulheres (Khoo, 2021).

Se “desplataformar” é a expressão que faz referência à remoção de atos, conteúdos e pessoas das plataformas de redes sociais, a tessitura que envolve o conceito livremente adaptado pelas autoras deste trabalho - de plataformização da misoginia - caminha no sentido contrário: busca entender o agrupamento, fortalecimento e disseminação da misoginia em plataformas de redes sociais a partir de interesses, conteúdos e grupos que se reúnem e encontram nesta esfera digital comungando do sentimento comum de desqualificar, categorizar, submeter ou estabelecer relações de dominação e poder em relação às mulheres e seus corpos. Especificamente neste trabalho, o recorte utilizado é a publicidade digital como meio capaz de promover e fortalecer este encontro através da *dataficação*, em um ambiente de opacidade algorítmica e ausência de regulação legal.

Khoo (2021) examinou o papel das plataformas digitais na proliferação da violência, abuso e assédio baseados em gênero facilitados pela tecnologia, ou em sua formulação original *Technology-facilitated gender-based violence*, adiante abreviado como TFGBV. Em seu referido relatório, examinou se, e como, as plataformas digitais *Facebook*, *YouTube* e *Twitter* devem ser responsabilizadas por isso através de regulamentação ou imposição legal – especificamente sob a lei canadense. Note-se que tais atividades e

---

<sup>4</sup> A opção pela não utilização do termo “redes sociais”, mas sim “plataformas”, se dá a partir do reconhecimento da distinção que isso implica. Há uma série de estudos organizados sobre o advento das plataformas digitais, são os *Platform Studies*, que tratam da peculiaridade de tais plataformas no modo como as pessoas se comunicam via redes sociais, já que elas exercem um tipo de mediação na comunicação muito mais abrangente do que o modelo baseado na *Web*. As plataformas coordenam mudanças e influências tecnológicas e políticas e estão envoltas em políticas de *Application Programming Interface (APIs)* (d’Andréa, 2018), que são mecanismos que permitem a comunicação entre aplicativos, sistemas e softwares em um mesmo lugar, aumentando a influência e o tempo empregado pelos usuários nessas plataformas. Mais detidamente, o termo “plataformização da misoginia” é uma adaptação das autoras deste artigo em relação ao termo *deplatforming misogyny*, desenvolvido por Cynthia Khoo, como explicado no corpo do artigo.

comportamentos englobados no TFGBV tem a tecnologia como aspecto central na perpetuação e difusão da violência de gênero contra meninas e mulheres – tanto cisgênero quanto transgênero.

As plataformas de redes sociais não inventaram a misoginia, mas lucram com ela, além de acelerar e ampliar o alcance da velha violência praticada na sociedade contra meninas e mulheres e que ganha mais um terreno para sua prática, exibição e coletivização de atos opressores que podem ganhar continuidade fora da plataforma, já que a membrana que separa o marco entre o mundo *online* e *offline* é totalmente permeável ou mesmo inexistente. Afinal, aquilo que ocorre e é dito, pensado, organizado de forma *online* implica em uma correlação com o cotidiano e pode produzir seus efeitos nele. O *online* levanta a assustadora ideia do “Quem é você quando ninguém está vendo?”, embora seja um ambiente com algumas legislações específicas em termos de responsabilidade civil e penal.

As atividades que se enquadram no âmbito da TFGBV, segundo Khoo (2021), incluem discurso de ódio, ameaças e intimidação, espionagem e monitoramento de contas ou interceptação de comunicações privadas, assédio online em grupo, campanhas coordenadas de denúncia, exploração sexual resultante de aliciamento online, difamação, distribuição não consensual de imagens íntimas, abuso baseado em imagens (incluindo *deepfakes*). O vasto e apurado rol de práticas tem gênero e tecnologia como ponto de convergência, atuando na perpetuação da misoginia.

Não se pretende olhar com purismo para as plataformas questionando a sua importância, nem de retornarmos a tempos anteriores onde a tecnologia era embrionária. A questão é que agora há, sem dúvidas, um agrupamento muito mais potente, sem barreiras de tempo e localização, com senso de comunidade, que conta com o refinamento dos algoritmos e utilização de dados pessoais capazes de conectar pessoas com todo tipo de interesse afim, e que tudo isso fomentou uma escalada na misoginia – anteriormente existente, mas analógica.

Este universo em discussão não trata apenas de uma violência de gênero ou práticas misóginas explícitas, mas também da formulação de toda uma racionalidade de controle e dominação que organiza uma masculinidade que tenta suprimir, apagar e esvaziar tudo o que foge do “feminino clássico” e seus papéis de gênero patriarcalmente estabelecidos. Tratam-se de verdadeiras comunidades antifeministas. Lilly (2016) nomeou a política representacional da comunidade antifeminista online como “manosfera”. Segundo ela:

A manosfera é a expressão online do movimento dos direitos dos homens offline, que surgiu na década de 1980. Diversas análises sugeriram que o surgimento de vários fenômenos políticos sob o nome 'movimento dos direitos dos homens' (MRM) são essencialmente tentativas de reação por homens para reafirmar ou redefinir uma visão retrógrada particular de masculinidade que está em conflito com os ideais feministas e os papéis em mudança das mulheres (Lilly, 2016 p. 1)

Ao longo de sua pesquisa, a autora examinou as representações da masculinidade e descobriu que o ideal masculino na “manosfera” enfatiza o controle e dominação, honra e “valores de guerreiro”, ao qual ela nomeou como “masculinidade clássica” relacionada a um passado mítico. Todo o discurso se

organiza em torno de uma masculinidade em suposta crise e ameaçada pelas supostas tentativas de “feminização do homem” e da sociedade. Logo, para defender e reestruturar esta masculinidade mítica, é propagada uma ideia de união em torno de uma causa comum que pretende colocar o feminino “no seu lugar” – lugar esse que é sempre subalterno, submisso e estereotipado.

Toda essa violência de gênero baseada em tecnologia na “manosfera” está sendo aqui esquadrihada para que fique claro que não se trata de uma organização social emergente, organizada em torno de uma pauta masculinista que já existia socialmente em larga escala e que tinha orgulho de sua existência. Não. Há uma tônica um pouco mais complexa: é verdade que a misoginia sempre existiu, mas sempre houve um movimento empenhado em justificar sua existência biológica e historicamente, ou através da religião, dos papéis de gênero, romantizando e conferindo ares de liberdade e igualdade formal.

O que muda agora é que a tecnologia manifesta nas plataformas, gera o senso de comunidade reunindo pessoas com interesses escusos, inclusive a partir da categorização da publicidade digital, somada à ausência de regulamentação legal tanto no que diz respeito à responsabilização das plataformas e usuários pela misoginia que promove quanto a ausência de regulamentação da Inteligência Artificial (IA) que permite a opacidade algorítmica e o emprego de sistemas de classificação de interesses e categorias pelas plataformas para a distribuição de anúncios e conteúdo. Toda essa engrenagem gendrada torna o terreno fértil para o desenvolvimento da violência baseada em gênero facilitada pela tecnologia.

A plataformização da misoginia conta com a aceleração tecnológica para sua difusão, além de um outro novo ponto que merece melhor elucidação: o uso de algoritmos sofisticados de aprendizado de máquina capazes de direcionar e entregar anúncios e conteúdos de forma não aleatória, mas baseada em interesses catalogados a partir de dados coletados pelas plataformas de redes sociais. Logo, o direcionamento do *Facebook* não torna todos os usuários disponíveis para os anunciantes, mas torna os indivíduos com maior propensão de se interessarem pelo assunto impulsionado, disponíveis. Segundo Cotter, Medeiros, Pak e Thorson (2021), o processo de classificação de usuários é um processo de mercantilização, que envolve a construção (algorítmica) de categorias de interesse para direcionamento de anúncios e a classificação de usuários nelas.

Note-se que a segmentação de usuários depende de uma percepção de valor daquele público para o anúncio ou conteúdo impulsionado. De forma bastante atuante as plataformas distribuem conteúdo que explora e difunde a misoginia, colaborando com a construção de uma comunidade orientada em torno de uma mesma pauta. Assim como os sistemas de classificação tradicionais, a classificação algorítmica envolve escolhas éticas, mediando relações sociais, políticas e econômicas, frequentemente beneficiando grupos dominantes em categorias como gênero, raça, etnia e sexualidade, classe e outros marcadores.

O fenômeno de segmentação e categorização de usuários é o que torna a mercantilização da misoginia tão atraente e rentável para a publicidade digital nas plataformas:

Os desenvolvedores tomam decisões para conceituar e operacionalizar modelos de previsão e, ao fazer isso, "colocam uma

moldura filosófica particular no mundo que o torna propício ao trabalho de código e algoritmos" (Kitchin; Dodge, 2011, p. 247). Os desenvolvedores determinam quem é representado nos conjuntos de dados e como os dados são "preparados para o algoritmo" (Gillespie, 2014). Os desenvolvedores também tomam decisões sobre quais categorias identificar, o que elas significam, quem e o que deve ser classificado nelas e o que fazer com elas (Gillespie, 2014). Em um nível mais elevado, os valores e interesses dos proprietários de plataformas guiam o desenvolvimento de algoritmos e intervenções posteriores em seus resultados e resultados. Assim, as classificações algorítmicas representam "leituras interessadas da realidade dataficação" (Rieder, 2017: 110) (Cotter *et al.* 2021, n.p., tradução nossa).

É assim que a plataformização da misoginia não apenas cria o ambiente de encontro de pessoas e grupos organizadas em torno de um pseudo ideal de subalternização de mulheres e dominação, poder e controle de seus corpos, papéis e atitudes, como também é capaz de mercantilizar a misoginia a partir de classificações algorítmicas opacas e do próprio desenvolvimento dos algoritmos de caixa preta que orientam o funcionamento da plataforma.

"Desplataformar" a misoginia passa, ao nosso ver, por ações que transcendem a própria plataforma, pois banimento, remoção e bloqueio de usuários ainda dependem de políticas internas e orientações éticas das plataformas – que sim, são importantes, mas não podem ser o alicerce principal quando o que está em jogo é a violência de gênero. Nos cabe compreender, agora, de forma mais aprofundada a ocorrência da misoginia na publicidade antes de falarmos mais detidamente de caminhos de solução.

## **2 ACHADOS DE PESQUISA SOBRE A MISOGINIA NA PUBLICIDADE DIGITAL EM PLATAFORMAS SOCIAIS**

Os achados de pesquisa utilizados neste trabalho dizem respeito à publicação dos resultados da primeira etapa do relatório "Observatório da indústria da desinformação e violência de gênero nas plataformas digitais", conduzida pelo NetLab – o Laboratório de pesquisa em Internet e Redes Sociais da Universidade Federal do Rio de Janeiro (UFRJ), sob encomenda do Ministério das Mulheres do Governo Federal brasileiro, no âmbito da iniciativa Brasil Sem Misoginia. A fim de uma melhor contextualização, cabe destacar que o objetivo principal da pesquisa é identificar, arquivar e analisar anúncios tóxicos que impactaram mulheres e foram impulsionados em uma ou mais plataformas da *Meta* — o que compreende as principais redes sociais utilizadas no Brasil, quais sejam *Facebook*, *Instagram*, *Messenger* e *Audience Network*.

Conforme resultados publicizados na primeira fase do relatório, em vinte e oito dias de coleta, a partir de peças publicitárias que estavam ativas entre 01 e 09 de fevereiro de 2024, foram identificados 1.565 anúncios tóxicos que impactaram mulheres e apresentaram indícios de comportamento misógino, de golpes ou fraudes direcionados a mulheres ou de irregularidades na oferta de produtos ou serviços para o público feminino (Santini *et al.*, 2024). Para melhor compreensão de como isso se deu, o caminho metodológico percorrido buscou

por anúncios ativos e não categorizados como sensíveis<sup>5</sup> a partir de 165 termos ou palavras-chave relacionadas à saúde, misoginia, autonomia financeira e espiritualidade femininas, para realizar uma análise exploratória sobre as principais narrativas e tipos de anúncios, por meio da interface da Biblioteca de Anúncios da *Meta*. Destes, 56 termos de busca foram mantidos para a coleta final. Foram analisados *posts* impulsionados que citavam ou apresentavam imagens (foto ou vídeo) de mulheres ou, ainda, que se referiam a mulheres como interlocutoras, além de outros indicativos de que os anúncios buscavam interagir com o público feminino (Santini *et al*, 2024).

Após a coleta, houve classificação das categorias mais comuns de ameaças às mulheres, partindo inicialmente de três níveis de toxicidade na publicidade online: problemática, irregular ou ilegal/fraudulenta. Foi considerada problemática a publicidade “[...]que reforça estereótipos de gênero, cultura masculinista e/ou explora vulnerabilidades psicossociais e econômicas, sem necessariamente incidir em crimes ou irregularidades.”; irregular a publicidade que “[...] viola normas de publicidade e de Conselhos Profissionais (como CFM, CFO, OAB, etc.) e/ou outras normas setoriais.”; ilegal ou fraudulenta a publicidade “[...] que viola a norma jurídica brasileira, em especial a Constituição Federal, o Código Penal, o Código de Defesa do Consumidor e a resolução 96/2008 da Anvisa, que regula a propaganda e a publicidade de medicamentos.” (Santini *et al*, 2024, p. 12)

Não se trata de fazer uma revisão sistemática da pesquisa elaborada pelo NetLab, mas de trabalhar com dados tão recentes e que foram amplamente publicizados nos meios de comunicação brasileiros, especialmente por sua ligação institucional com o Ministério das Mulheres. Portanto, algumas informações serão, inevitavelmente, replicadas com a devida referência da origem, para oportunizar a contribuição deste artigo em suas demais seções a partir da hipótese de que tais acontecimentos não são fenômenos aleatórios, mas que, pelo contrário, contam com a opacidade algorítmica na distribuição publicitária, aliada à ausência de regulamentação das plataformas e da própria Inteligência Artificial no Brasil.

Retomando-se a coleta sistematizada na pesquisa em análise, dado interessante foi que 61% dos anúncios redirecionam usuárias para o *WhatsApp* – principal aplicativo de mensageria em utilização no país. Destaca-se que nesse aplicativo as conversas são criptografadas de ponta a ponta e contam com um ambiente privilegiado para a prática de qualquer prática ilícita ou fraudulenta já que não passa por nenhum tipo de interferência, com revisão de sua atuação apenas mediante determinação judicial.

Outro achado da pesquisa é que as postagens publicitárias parecem recorrer a argumentos de autoridade e credibilidade para conquistar a confiança das mulheres destinatárias de seus produtos e serviços. É o caso dos anúncios identificados como sensíveis relacionados à saúde que fazem uso indevido de imagem, como no vídeo replicado pelo menos 21 vezes na amostra — e nem sempre identificado como sensível, que apresenta imagens manipuladas do

---

<sup>5</sup> Segundo a *Meta*, são consideradas categorias especiais de anúncios aquelas que versam sobre ofertas de crédito, moradia, temas sociais, eleições ou políticas. A escolha da categoria faz parte dos padrões de publicidade estabelecidos pela plataforma.

presidente da Anvisa, Antônio Barra Torres, na CPI da Covid e cuja transação da fraude é concluída no redirecionamento da conversa para o *WhatsApp*:

A sobreposição de áudio falso é facilitada pela presença de uma máscara cobrindo a boca de Barra Torres. No vídeo editado, ele estaria dando a chancela da Anvisa a um produto “milagroso” para emagrecimento, que seria comercializado apenas pelo “WhatsApp de um especialista credenciado”, cujo contato estaria disponível na publicação, que traz link para chat no aplicativo de mensagens. Replicado em diferentes anúncios, o vídeo também traz a logomarca da Globo e/ou da Anvisa. Em todos os anúncios, o texto do post diz que “Nós, da ANVISA, encorajamos a todos que buscam a diminuição do peso de forma saudável a adquirir o produto exclusivamente através dos canais regulamentados. (Santini *et al*, 2024, p. 15).

Como no vídeo *fake* acima mencionado, o corpo feminino foi pauta de 80% dos anúncios tóxicos analisados, seja através de tópicos relacionados à estética, emagrecimento, contendo fraude ou golpe ou reforçando estereótipos de gênero. Desde promessas de resultados, beleza sem esforço, soluções milagrosas para emagrecimento, exploração de imagem de mulheres, oferta de tratamento sem comprovação científica (Santini *et al*, 2024, p. 17).

Os achados da pesquisa também apontam que 79% dos anúncios representam risco à saúde pública ou individual (promessas de resultados e criação de expectativas em relação a resultados de beleza, saúde, bem-estar e emagrecimento, sem comprovação científica; 66% dos anúncios traz desinformação, publicidade abusiva ou enganosa (promessas de mudança de vida, resultados rápidos em aspectos financeiros, perda de peso, desrespeitando conselhos profissionais, e sem comprovação científica, mas utilizando-se de informações supostamente científicas).

Quanto ao reforço de estereótipos de gênero, machismo, e objetificação da mulher, 44% dos conteúdos foi catalogado assim, por determinarem como mulheres devem ser e qual aparência devem ter, colocando-as como subordinadas ou como objetos sexuais. Foram encontrados, também, fraudes ou golpes em 35% dos anúncios com abordagem sensacionalista que se utilizavam de gatilhos que davam a ideia de grandes benefícios e preços fora do comum e direcionam consumidores para sites falsos; 34% dos anúncios utiliza identidade falsa, utilizando pessoas amplamente conhecidas, como celebridades, para dar credibilidade ao golpe.

Tais dados lançam luz sobre os estereótipos de gênero, a intensa pressão da indústria da beleza, a busca por pertencimento e aceitação e a forma como mulheres são colocadas em posição de objeto de desejo ou de tornarem-se um, deixando de ocupar sua posição de sujeitos nas relações, a partir de papéis desenhados e lugares previamente separados para serem ocupados: “[...] é mãe na casa, cidadã na *polis* e trabalhadora no mercado. E existe ainda uma outra grande função que é exercida – ou demandada a ser exercida -, que é a de ser uma mulher desejada.” (Homem; Calligaris, 2019, p. 80). Sugerem os autores sobre a necessidade de se reconfigurar o pacto social a respeito da posição das mulheres como sujeitos com desejos e não como objetos do desejo de outros. Só um sujeito é capaz de desejar para si e existir não como a negação ou contraposição do Outro, mas como fim em si mesmo:

Dentro desse espectro foram popularizadas e inscritas na ordem do simbólico as ideias de que pessoas que nascem marcadas pelo sexo feminino são seres com constituição psíquica naturalmente empática, sensível, dócil e com habilidades manuais mais desenvolvidas, além de dotadas de uma certa dificuldade na tomada de decisão, raciocínio lógico e trato estatístico. Esse viés falacioso é erroneamente reforçado e confirmado através da linguagem e dos estereótipos de gênero que falam sobre “coisa de mulher” e “lugar de mulher”. Utiliza-se para esse fim, comumente, apenas a expressão “feminino”, como um guarda-chuva semântico que reafirma um comportamento social condizente com o sexo biológico.

No mesmo modo, o masculino também é tomado como adjetivo universal que nomeia e orienta a linguagem, suas significações e certas características e papéis gendrados, sendo adotado como o parâmetro a partir do qual se baseiam os Outros. A mulher é tida como o Outro, o que surge como o oposto do homem: o que não é homem, o que não é masculino, feminino é. Mulher é (Ferraz; Costa, 2023a, p. 121)

O uso indevido de marcas renomadas, de programas e jornais conhecidos para dar credibilidade aos produtos e serviços fraudulentos totalizou 19% dos anúncios; 13% dos anúncios promovia conteúdo masculinista, conhecido como sendo aquele destinado ao desenvolvimento masculino com o objetivo de tornar mulheres submissas e reforçar uma lógica descabida de que homens têm sido vítimas de uma suposta ditadura feminista; por fim, 9% dos anúncios descumpriu normas de conselhos profissionais (mostrando rosto de pacientes em procedimentos, sem utilização de número de registro profissional).

Deste apanhado, é possível perceber que, em um período inferior a um mês, foi possível identificar 1.565 anúncios problemáticos e que representam um risco para mulheres. Elas não estão seguras em um ambiente que deveria ser destinado à sua socialização e ócio, seja como alvo de anúncios fraudulentos ou ilícitos – muitos dos quais as conduzem para o ambiente nebuloso de uma conversa criptografada no *WhatsApp* -, ou estão sendo alvo e objeto, simultaneamente de estereótipos de gênero reforçados em campanhas publicitárias. O que todo o contexto sinaliza é que o ócio das mulheres tem representado um risco para elas por toda uma racionalidade que as objetifica sexualmente e as categoriza como coisas que podem ser livremente dispostas e apropriadas (Ferraz; Costa, 2023c, p. 388).

Se a publicidade é uma das principais fontes de receita da plataforma, o ônus da moderação desta atividade naturalmente deve recair sobre ela. Se isso não está acontecendo, também podemos compreender que todo um sistema social gendrado e construído sobre engrenagens patriarcais, lucra com a desinformação e a violência digital de gênero. É importante entender agora a importância da opacidade algorítmica em todo esse processo e a ausência de regulamentação sobre a responsabilização das plataformas e da própria IA no Brasil.

### **3 OPACIDADE ALGORÍTMICA E A AUSÊNCIA DE UM MARCO REGULATÓRIO PARA A INTELIGÊNCIA ARTIFICIAL NO BRASIL**

Para melhor compreender a opacidade algorítmica e a mercantilização das segmentações de interesses nas plataformas de redes sociais, imagine-se que, há alguns bons anos atrás, boas possibilidades para veiculação de material publicitário com objetivo de despertar interesse em consumidores estavam ligadas à veiculação de criativos na televisão aberta, em horário nobre; à veiculação nas páginas de um jornal de grande circulação ou mesmo na rádio; ou até mesmo expondo uma propaganda em um grande e bem localizado *outdoor*. A partir dessas possibilidades, espera-se um amplo alcance do material publicitário. Mas um grande alcance não significa, necessariamente, grandes vendas já que essa atuação vai se estender a um público nem sempre qualificado como comprador ideal.

Agora imagine este mesmo produto sendo veiculado com uma segmentação de público específico que atenda ao ideal de comprador da empresa/marca/anunciante, utilizando-se da *dataficação* das plataformas para uma entrega de criativos publicitários para pessoas que já manifestaram interesses por determinada categoria de produtos/mensagens, classificadas por gênero e até mesmo classe social (dependendo do país, esta é uma informação elegível na categorização de anúncios da *Meta*), idade, e com possibilidades, inclusive, de excluir determinados grupos do recebimento da publicidade. Tudo isso a partir de um sistema de distribuição baseado em um leilão capaz de determinar o alcance e número de impressões dos criativos do anúncio de acordo com o investimento.

Considerando-se os dois cenários, é notável a vantagem existente no caso das plataformas. O que consumidores entregam para ter acesso às redes sociais são os seus dados, gostos, escolhas, interações. Como tudo isso é feito a partir de uma política de privacidade, através da autorização por termos e condições de uso que a grande maioria dos usuários jamais leu ou lerá – apesar de aceitar -, a noção que se instala é que as pessoas têm acesso às redes sem qualquer contrapartida. Justamente por não compreender, a maioria das pessoas não dá o devido valor aos seus dados.

De fato, pouco se sabe sobre o desenvolvimento do algoritmo que está associado às redes sociais da *Meta* e de outras tantas empresas. Ele enquadra-se dentro dos sistemas de caixa preta – *black box* – que são aqueles em que a referida compreensão e explicabilidade das decisões e resultados automatizados não são claros nem mesmo para especialistas. Nesse caso, sistema de caixa preta é “*noción que se emplea para describir la dificultad que sobreviene para explicar el resultado obtenido por la IA en términos comprensibles por el común de las personas*” (Caparrós, 2022, p. 6)

Pesquisas realizadas por Ali *et al* (2019) observaram uma distorção significativa na entrega dos anúncios ao longo de linhas de gênero e raça para "anúncios reais" de oportunidades de emprego e moradia, apesar dos parâmetros de segmentação neutros. /os resultados demonstram mecanismos previamente desconhecidos que podem levar a uma entrega de anúncios potencialmente discriminatória, mesmo quando os anunciantes definem seus parâmetros de segmentação para serem inclusivos. E também é por isso que não se pode esperar uma regulamentação ética vinda por parte das empresas – apenas -, se desacompanhada de uma regulamentação legal.

As possibilidades de segmentação são imensas, inclusive com utilização de recursos de segmentação sensíveis - dados demográficos e interesses do usuário -, grupos de usuários podem ser excluídos de receber determinados anúncios com base nas escolhas de segmentação dos anunciantes. E, considerando-se que estes anúncios não dizem respeito apenas à publicidade de consumo, mas também a possibilidades de crédito, moradia e emprego, política, a discriminação algorítmica pode simplesmente excluir e invisibilizar determinados públicos de fatos e oportunidades da vida.

Segundo, descobrimos que a entrega distorcida pode ocorrer devido ao conteúdo do próprio anúncio (ou seja, o título do anúncio, texto e imagem, coletivamente chamados de criativo do anúncio). Por exemplo, anúncios direcionados ao mesmo público-alvo, mas que incluem um criativo que estereotipicamente seria de maior interesse para homens (por exemplo, musculação), podem ser entregues para mais de 80% de homens, e aqueles que incluem um criativo que estereotipicamente seria de maior interesse para mulheres (por exemplo, cosméticos) podem ser entregues para mais de 90% de mulheres. Da mesma forma, anúncios que se referem a conteúdo cultural estereotipicamente de maior interesse para usuários negros (por exemplo, hip-hop) podem ser entregues para mais de 85% de usuários negros, e aqueles que se referem a conteúdo estereotipicamente de interesse para usuários brancos (por exemplo, música country) podem ser entregues para mais de 80% de usuários brancos, mesmo quando direcionados de forma idêntica pelo anunciante. Assim, apesar de fazer o mesmo lance para o mesmo público-alvo, a entrega do anúncio pelo anunciante pode ser fortemente distorcida com base apenas no criativo do anúncio (ALI, *et al.*, 2019, n.p.)

Em tratando-se de *black box* e sem a cooperação das plataformas, torna-se praticamente impossível compreender os caminhos de discriminação algorítmicos percorridos no processo de distribuição publicitário e a extensão de seus efeitos discriminatórios. Tem-se a noção, a partir de pesquisas como a de Ali *et al* (2019) que as imagens utilizadas nos criativos dos anúncios são capazes de afetar seu direcionamento e distribuição, baseando-se essencialmente em estereótipos de gênero.

Nesse sentido, há Projeto de Lei nº 2.630/2020 que Institui a Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet, com o objetivo “[...] de proteger a liberdade de expressão e o acesso à informação e fomentar o livre fluxo de ideias na internet [...]”, estabelecendo que os provedores de redes sociais e de serviços de mensageria privada, no âmbito e nos limites técnicos de seu serviço, devem adotar medidas para identificar todos os conteúdos impulsionados e publicitários cuja distribuição tenha sido realizada mediante pagamento ao provedor de redes sociais.

O projeto que ficou conhecido como “PL das *Fake News*” despertou incômodo em muitos políticos e também em plataformas como o *Google* e o próprio *Facebook*, de forma que hoje está tramitando na Câmara dos Deputados sem qualquer garantia de aprovação. Dentre os seus objetivos estão o fortalecimento do processo democrático por meio do combate às *fake news*, a defesa da liberdade de expressão, a busca por maior transparência das práticas de moderação de conteúdos postados em redes sociais e, de acordo com o

interesse deste trabalho, o inciso IV revela o objetivo de adoção de mecanismos e ferramentas de informação sobre conteúdos impulsionados e publicitários disponibilizados para o usuário.

Embora tenha previsto, no artigo 3º, referente aos princípios da lei, o inciso que prevê “IX – transparência nas regras para veiculação de anúncios e conteúdos pagos” e mesmo em sua seção V, que versa acerca da Transparência e obriga os provedores de redes sociais a produzirem relatórios trimestrais de transparência para informar os procedimentos e as decisões de tratamento de conteúdos gerados por terceiros no Brasil, não há nenhum tipo de arquitetura algorítmica prevista. Os relatórios são baseados predominantemente por números de contas, de medidas, de moderação de contas. Pouco se sabe sobre a discriminação algorítmica produzida. Em que pese a criação de um Conselho de Transparência e Responsabilidade na Internet, o alcance de toda essa regulamentação não resolve o cerne da questão da plataforma da misoginia.

São necessárias, outrossim, transparência e governança algorítmica dessa caixa preta, com processos e decisões tomadas por sistemas de IA transparentes, auditáveis e explicáveis, permitindo a responsabilização dos desenvolvedores e operadores. Estas lacunas somente poderão ser supridas a partir da instituição de um Marco Regulatório para a Inteligência Artificial no Brasil – o que está em andamento a partir do PL 2.338/2023.

O PL que regulamenta a IA, está sob exame e discussão na Comissão Temporária Interna sobre Inteligência Artificial no Brasil (CTIA) no Senado, cujo Colegiado deve examinar os projetos concernentes ao relatório final aprovado pela Comissão de Juristas responsável por subsidiar a elaboração de substitutivo sobre Inteligência Artificial no Brasil (CJSUBIA). A CJSUBIA também foi uma Comissão Temporária Interna do Senado instituída em 2022 com a finalidade de elaborar minuta de substitutivo para instruir a apreciação dos PL em trâmite que buscam a regulamentação da IA no Brasil.

As regulamentações éticas não têm capacidade de lidar com todos os desdobramentos provenientes da IA em suas diferentes fases. Diante do alcance e expressividade da IA com suas manifestações diárias na vida das pessoas, desde o direcionamento publicitário ao qual elas estão expostas até a análise de elegibilidade de políticas públicas e acesso à linhas de crédito, é imprescindível que se discuta instrumentos de governança, garantia de direitos fundamentais e responsabilidade civil, o que requer uma regularização jurídica e não apenas ética, além do desenvolvimento de políticas públicas específicas.

A IA não deve criar ou reforçar preconceitos ou alimentar vieses cognitivos capazes de impactar de maneira injusta ou desproporcional determinados indivíduos, promovendo discriminação de qualquer natureza, sobretudo em relação a marcadores sensíveis como gênero, raça, etnia ou classe social. Considerando-se que os vieses discriminatórios estão presentes na sociedade, deve-se admitir, por óbvio, que também estejam presentes nos sistemas de Inteligência Artificial. Não se tratam de sistemas neutros que se comportam de forma igual com todas as pessoas, pelo contrário: a construção desse raciocínio acerca da suposta neutralidade e objetividade da IA se desenvolve a partir da falta de conhecimento sobre seu funcionamento que é alimentada pela

opacidade em sua aplicação, formulação e ausência de explicabilidade das decisões automatizadas.

Ausentes qualquer previsibilidade, justificção, responsabilidade e conhecimento sobre essa verdadeira caixa preta que tem sido a IA no Brasil, a luta encampada pelos Direitos Humanos em todos os seus processos de internalização, ratificação, discussão e constitucionalização é ameaçada em alguma medida diante desse novo terreno tecnológico que ainda carece de regulamentação quanto ao funcionamento, segurança, transparência e governança. O impacto da discriminação algorítmica ataca frontalmente os direitos fundamentais, sobretudo o direito à igualdade e a não discriminação. E um marco legal tem que prever todas essas questões, tomando como parâmetros os direitos fundamentais e uma perspectiva de gênero.

## CONCLUSÃO

Diante do alcance e expressividade da IA, com suas manifestações diárias na vida das pessoas, desde o direcionamento publicitário ao qual elas estão expostas até a análise de elegibilidade de políticas públicas e acesso à linhas de crédito, é imprescindível que se discuta instrumentos de governança, garantia de direitos fundamentais e responsabilidade civil, o que requer uma regularização jurídica e não apenas ética, além do desenvolvimento de políticas públicas específicas.

Neste sentido, responsáveis por projetar e implantar sistemas de IA devem também ser responsáveis pelo funcionamento de seus sistemas. Para isso são necessárias normas gerais para o seu uso que garantam fiscalização, transparência, previsibilidade de resultados, com o objetivo de proteger os direitos fundamentais, conferir segurança jurídica e garantir a dignidade da pessoa humana, do regime democrático e do desenvolvimento científico e tecnológico. Mas não apenas isso. Leis desacompanhadas de políticas públicas não possuem efeito transformativo.

Não obstante, mesmo antes de uma regulamentação promulgada no país, há uma série de outros passos capazes de mitigar os efeitos da discriminação algorítmica na publicidade das plataformas: sem dúvida alguma, a promoção de relações mais equânimes na sociedade em relação à economia do cuidado, que compreende a horizontalidade dos papéis nas relações humanas como algo inerente ao próprio trato humano e não a gênero; a reorganização da divisão sexual do trabalho com relações paritárias no mercado de trabalho e possibilidade de acesso a ele; a naturalização de outras possibilidades de corpos para além daqueles instituídos e perseguidos, aumentando a representatividade e reconhecimento de mulheres trans, gordas, pessoas com deficiências, padrões contrassenso, especialmente no mercado publicitário digital; a ocupação e reafirmação de espaços públicos por mulheres, o que compreende espaços digitais – tudo isso que reflete a macroestrutura que ambienta a misoginia. Logo, o *offline* reflete-se no *online*, se é que é possível traçar essa dicotomia no atual *framework*.

Em uma escala mais localizada, soluções mais pontuais em relação à publicidade digital podem envolver a prática de acionar conselhos profissionais

diante de criativos publicitários que extrapolem limites, que exponham pacientes e resultados, que façam promessas milagrosas, que não identifiquem o número de registro no conselho profissional. Tais medidas podem ser efetivas já que estes conselhos já têm seus direitos constituídos, bastando ser reivindicados.

Ademais, as plataformas poderiam/deveriam investir em um trabalho mais criterioso de equipes de moderação dos conteúdos e naquilo que diz respeito aos sistemas automatizados das plataformas, que os algoritmos sejam treinados para identificar e banir conteúdo discriminatório muito antes de sua veiculação. Tais treinamentos só funcionam lucidamente em um contexto de diversidade de equipes, com formação focada na detecção e correção de vieses discriminatórios.

Como tão bem se sabe, problemas estruturais como a misoginia não tem um caminho único de solução e nem dependem meramente de uma legislação para que o problema se resolva. Este é um daqueles problemas organizados por camadas – algumas mais profundas, outras mais visíveis e superficiais -, que dependem de uma solução estrategicamente articulada e multisetorial, que deve incluir, naturalmente, um marco regulatório capaz de dar conta dos avanços tecnológicos. Não é só a tecnologia que avança, a misoginia também.

## REFERÊNCIAS

ALI, Muhammad; SAPIEZYNSKI, Piotr; BOGEN, Miranda; KOROLOVA, Aleksandra; MISLOVE, Alan; RIEKE, Aaron. Discrimination through optimization: How Facebook's ad delivery can lead to skewed outcomes, **Proceedings of the ACM on Human-Computer Interaction**, v. 3, n. CSCW, p. 1-30, nov. 2019. Disponível em: <https://arxiv.org/abs/1904.02095>. Acesso em: 19 mar. 2024.

BRASIL. PL 2.630/2020. **Institui a Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet**. Brasília, DF: Câmara dos Deputados, 2020. Disponível em: [https://www.camara.leg.br/proposicoesWeb/prop\\_mostrarintegra?codteor=1909983](https://www.camara.leg.br/proposicoesWeb/prop_mostrarintegra?codteor=1909983). Acesso em: 12 mar. 2024.

\_\_\_\_\_. PL 2.338/2023. **Dispõe sobre o uso da Inteligência Artificial**. Brasília, DF: Senado Federal. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/157233>. Acesso em: 07 out. 2023.

CAPARRÓS, Mariana Sánchez. Los riesgos de la inteligencia artificial para el principio de igualdad y no discriminación. Planteo de la problemática y algunas aclaraciones conceptuales necesarias bajo el prisma del Sistema Interamericano de Derechos Humanos. **elDial.com Contenidos Jurídicos**, 2022, p. 1-22. Disponível em: [https://www.researchgate.net/publication/361510404\\_Los\\_riesgos\\_de\\_la\\_inteligencia\\_artificial\\_para\\_el\\_principio\\_del\\_igualdad\\_y\\_no\\_discriminacion\\_planteo\\_d](https://www.researchgate.net/publication/361510404_Los_riesgos_de_la_inteligencia_artificial_para_el_principio_del_igualdad_y_no_discriminacion_planteo_d)

e\_la\_problematika\_y\_algunas\_aclaraciones\_conceptuales\_necesarias\_bajo\_el\_prisma\_del\_Sistema\_Intera. Acesso em: 07 out. 2023.

COTTER, Kevin; MEDEIROS, Meghaan; PAK, Chong; THORSON, Kjerstin. “Reach the right people”: The politics of “interests” in Facebook’s classification system for ad targeting, **Big Data & Society**, v. 8, n. 1, p. 1–16, 2021. Disponível em: <https://doi.org/10.1177/2053951721996046>. Acesso em: 19 mar. 2024.

d’ANDRÉA, Carlos Frederico de Brito. Cartografando controvérsias com as plataformas digitais: apontamentos teórico-metodológicos. **Revista Galaxia**, n. 38, p. 28-39, maio-ago. 2018. Disponível em: <https://revistas.pucsp.br/index.php/galaxia/article/view/34208/25695>. Acesso em: 01 mar. 2024.

ETICAS CONSULTING. **Inteligencia Artificial y mujeres, una historia de discriminación**. Disponível em: <https://www.eticasconsulting.com/inteligencia-artificial-y-mujeres-una-historia-de-discriminacion/>. Acesso em: 7 out. 2023.

FERRAZ, Deise Brião. COSTA, Marli Marlene Moraes da. O Protocolo de Julgamento com Perspectiva de Gênero como respostas institucional à pretensa universalização do feminino, amparada nos esforços internacionais de eliminação de todas as formas de discriminação contra as mulheres, **Revista de Direito Internacional**, v. 20, n. 1, p. 115-127, 2023a. Disponível em: <https://uniceub.emnuvens.com.br/rdi/article/view/9070/pdf>. Acesso em: 01 abr. 2024.

FERRAZ, Deise Brião; COSTA, Marli Marlene Moraes da. Violência de gênero, subjetivação e homosociabilidade masculina: Nem sempre os processos de socialização acompanham a evolução legislativa. In: Clovis Gorczewski e André Viana Custódio (org.). **Direito & políticas públicas**, v. III, 1 ed. Porto Alegre: Free Press, 2023b, p. 309-328.

FERRAZ, Deise Brião; COSTA, Marli Marlene Moraes da. Políticas Públicas contra o assédio e violência sexual: estudo comparado entre o Protocolo de Barcelona “No Callamos” e o projeto de lei brasileiro que cria o Protocolo “Não é Não”. **Revista Vertentes do Direito**, [S. l.], v. 10, n. 2, p. 379–410, 2023c. Disponível em: <https://sistemas.uft.edu.br/periodicos/index.php/direito/article/view/16566>. Acesso em: 2 abr. 2024.

HOMEM, Maria; CALLIGARIS, Contardo. **Coisa de menina? Uma conversa sobre gênero, sexualidade, maternidade e feminismo**. Campinas: Papyrus, 2019, p. 80.

KHOO, Cynthia. **Deplatforming Misogyny. Report on platform liability for technology-facilitated gender-based violence**. Women’s Legal Education

and Action Fund, 2021. Disponível em: <https://www.leaf.ca/publication/deplatforming-misogyny/>. Acesso em: 18 mar. 2024.

LILLY, Mary. **‘The World Is Not a Safe Place for Men’: The Representational Politics of the Manosphere.**”. Ottawa: University of Ottawa, 2016. Disponível em: <https://ruor.uottawa.ca/server/api/core/bitstreams/1eee5112-7f22-4ffc-a49d-a978a56bed05/content>. Acesso em: 17 mar. 2024.

MARTÍN, Nuria Belloso. La problemática de los sesgos algorítmicos (con especial referencia a los de género). ¿Hacia un derecho a la protección contra los sesgos? *In*: MARTÍN, Joaquín Garrido; JIMÉNEZ, Ramón Valdivia (Coord.) **Inteligencia artificial y filosofía del derecho**. Murcia: Ediciones Laborum, 2022, p. 45-78.

\_\_\_\_\_. Sobre fairness y machine learning: El algoritmo ¿puede (y debe) ser justo? **Anales de la Cátedra Francisco Suárez**, n. 57, p. 7-38, 2023.

ANTINI, R. Marie *et al.* **Golpes, fraudes e desinformação na publicidade digital abusiva contra mulheres**. Rio de Janeiro: NetLab – Laboratório de Estudos de Internet e Redes Sociais, Universidade Federal do Rio de Janeiro (UFRJ). Publicado em 08 de março de 2024. Disponível em: <https://netlab.eco.ufrj.br/post/golpes-fraudes-e-desinformac-a-o-na-publicidade-digital-abusiva-contra-mulheres>. Acesso em: 19 mar. 2024.

ZANELLO, Valeska. **Saúde mental, gênero e dispositivos: cultura e processos de subjetivação**. 1. ed. Curitiba: Appris, 2018.