



Carolina Bouchardet Dias

**Moderação Algorítmica e Autodefesa Digital:
A autoexposição de usuárias do Instagram como resistência ao
governo algorítmico das condutas**

Dissertação de Mestrado

Dissertação apresentada ao Programa de Pós
Graduação em Direito da PUC-Rio como requisito
parcial de obtenção de título de Mestre em Direito

Orientador: Prof. Antonio Pele

Rio de Janeiro
Junho 2022



Carolina Bouchardet Dias

**Moderação Algorítmica e Autodefesa Digital:
A autoexposição de usuárias do Instagram como resistência ao
governo algorítmico das condutas**

Dissertação apresentada ao Programa de Pós
Graduação em Direito da PUC-Rio como requisito
parcial de obtenção de título de Mestre em Direito.

Prof. Antonio Pele

Orientador

Departamento de Direito PUC-Rio

Prof. Caitlin Sampaio Mulholland

Departamento de Direito PUC-Rio

Prof. Bethânia de Albuquerque Assy

Departamento de Direito PUC-Rio

Rio de Janeiro, 27 de Junho de 2022

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Carolina Bouchardet Dias

Bacharel em Direito pela Pontifícia Universidade Católica do Rio de Janeiro (2019). Advogada.

Dias, Carolina Bouchardet

Moderação Algorítima e Autodefesa Digital: A autoexposição de usuárias do Instagram como resistência ao governo algorítmico das condutas / Carolina Bouchardet Dias; orientador: Ilié Antonio Pele. – 2022.
129 f.; 30 cm

Dissertação (mestrado)–Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Direito, 2022.
Inclui bibliografia

1. Direito – Teses. 2. Resistência. 3. Discriminação. 5. Governamentalidade algorítmica. 6. Moderação. 7. Instragram. I. Pele, Ilié Antonio. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Direito. III. Título.

CDD: 340

À minha mãe, Christiane,
ainda bem que você tá aqui.

Agradecimentos

Uma semana de aulas presenciais e dois anos de mestrado *online* e pandemia. Assim foi preciso: levar a experiência do mestrado concomitante com o início da vida profissional na advocacia e com uma crise global sanitária. Girar os pratos, se exaurir no trabalho e abrir um livro, se exaurir nos livros e voltar pro trabalho. Tudo isso em isolamento social e, é verdade, em meio a muitos privilégios sociais que precisam ser reconhecidos. Adicionalmente, o desafio da saúde mental durante dois anos de expectativas frustradas e testemunhando a pandemia de Covid-19 e o des-governo do Brasil. No meio tempo, felizmente, o amor dentro de casa, no Zoom, no WhatsApp, e em cada uma das pessoas tão especiais que me acompanharam nos anos solitários de mestrado e pandemia, a quem agradeço aqui.

À minha mãe, Christiane, por ter me segurado, me curado e nunca saído do meu lado. Por ter me dito de madrugada que eu deixasse pra decidir amanhã se eu realmente queria desistir, “ainda bem que você tá aqui”. Ao meu pai, Alberto, pelos abraços entre as sessões de estudo, as tentativas de apaziguar minhas angústias e por ceder seu espacinho na Rua Otávio Carneiro onde eu podia ir me isolar para escrever nos fins de semana. Ao meu irmão, Dudu, pelo carinho diário, cumplicidade e ser um exemplo de disciplina e gentileza. Aos meus avós, Arlete, Alberto e Alcy, pelo apoio de sempre, mesmo por ligações de voz e vídeo.

Ao meu orientador, Antônio Pele, pelas infinitas recomendações de leitura, música e filmes. Também agradeço por me mostrar a poesia, mas também a realidade da pesquisa acadêmica, pela disponibilidade e interesse de sempre, por ter criado pensamento juntos, aberto portas, pelo ensinamento técnico e estratégico dentro da academia, por ter me chamado a atenção para a resistência feminina, e, enfim, ter se tornado, pra mim, um exemplo de pessoa e de carreira acadêmica.

Às minhas colegas de turma, Natália Penque e Vitória Westin, pelas discussões virtuais noturnas, risadas e apoio emocional, tendo me ensinado que os anos de mestrado podem ser ligeiramente menos solitários.

Ao Eduardo Sartori, Edu, por ter aparecido num turbilhão de emoções, me mostrado paz quando eu precisava, me ensinado a resgatar interesses artísticos que contribuíram para um pensar mais livre, a desligar o celular, esquecer minhas relações, viver o presente, focar na dissertação, e, no tempo livre, fazer o nada com todo o meu propósito.

Ao Gabriel Lando, por ter me acompanhado no início do mestrado e da pandemia, em Copacabana, com discussões desafiantes sobre as minhas aulas no Zoom, e por, mesmo depois de ir embora pra Alemanha, sustentar nossa amizade e reconfortar as minhas angústias acadêmicas com o budismo, lembrando que tudo e todo mundo tem o mesmo fim.

À Adriana Vidal, por ter continuado me inspirando na pós-graduação e me mostrado que o pensamento crítico serve para expandir a si mesmo, não para cercar. À Viviana Ribeiro, Vivi, por ter reacendido meu interesse pelo feminismo e, com as aulas do IPIA, me fazer gostar novamente do estudo à noite, pelas discussões por videochamada, recomendações de leitura, por se empolgar honestamente com a minha pesquisa e se disponibilizar para pensarmos juntas.

Ao Luis André Azevedo, pela gentileza cotidiana, por valorizar a vida acadêmica, seja relacionada ou não ao Direito Societário, por me indicar livros de budismo, acreditar em mim, me chamar pra mesa dos advogados adultos e entender e incentivar meus compromissos do mestrado. Ao João Pedro Hennings, por ter se juntado ao time e segurado a barra quando eu precisava de mais tempo para estudar. Ao Peter Sester, pelas oportunidades de trabalharmos juntos, lembrar a gramática normativa do português, pelas discussões sobre direito e pelas dicas acadêmicas.

Ao pessoal do escritório, em especial Helena Teich, Mayara Campelo, Harife Eugenio e João Macarini por debaterem comigo os textos do processo seletivo de mestrado e pelas conversas por videochamada que tornavam a semana mais leve e me enchiam de saudades das nossas discussões jurídicas e da cumplicidade no aquário do Eizirik Advogados. Ao Nelson Eizirik, por ser sempre referência de humildade e gentileza e, mesmo à distância nos anos de pandemia, demonstrar preocupação. À Raquel Gomes, pela parceria de sempre, pela compreensão, conselhos, motivação e carinho da grande amiga que se tornou. Também agradeço à Jussara Duarte, Erika Mendes, Pablo Arraes, Ruan dos Santos e Fernanda Rodrigues pelo trabalho absolutamente fundamental que fazem no escritório. À Marlene, que me acompanhou desde o início do mestrado com carinho de vó e foi embora no finalzinho, agradeço pelo cuidado e amor de sempre (e pela plantinha na minha sala).

Aos “Trançados”, Gabriel Torres, Eduardo Vieira (Eddie), Roberta Dias (Robertinha), Mayra Dias, Gabriel Costa (Goda), Maria Luiza Fernandes (Malu), Fernanda Lopes, Caio Bonancin, Paulo Henrique Dias e Maitê Paes, por preencherem os últimos dois anos de isolamento social com amor, risadas e cerveja. Ao Lucas

Rodine, pela parceria e por me lembrar a importância de viver além dos livros e do trabalho. À Larissa Campos, Luiza Ribeiro, Leonardo Palermo, Bernardo Kauffman, Guilherme Abrantes, Mariana Gianotti e Beatriz Fleming, por também terem me acompanhado até aqui. A todos, por entenderem minhas ausências e fazerem do mestrado durante a pandemia uma experiência leve de alguma forma.

Ao Lucas Souza e Mauro Pedra, que me acolheram no Muay Thai, onde eu encontrei o equilíbrio necessário para escrever cada dia mais uma página. Agradeço por além de professores de luta, terem me ensinado sobre disciplina e outras formas de pensar a realidade.

À Lurdinha, pelos abraços e conversas nas manhãs em que nos encontrávamos, e pelos trabalhos de reprodução da vida humana ainda economicamente desvalorizados no Brasil, que me permitiram dedicar mais tempo aos estudos.

À Denise Barros, grande professora da Unirio e amiga mais divertida da faculdade de comunicação que eu não fiz. Agradeço pelos ensinamentos de metodologia da pesquisa e pela amizade que emprestei da minha mãe.

À Lucimara Rase, por ter me ajudado a transformar no melhor que eu pude a experiência de um mestrado à distância.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Resumo

Dias, Carolina Bouchardet; Pele, Ilié Antonio. **Moderação Algorítmica e Autodefesa Digital: A autoexposição de usuárias do Instagram como resistência ao governo algorítmico das condutas**. Rio de Janeiro, 2022. 129p. Dissertação de Mestrado – Departamento de Direito, Pontifícia Universidade Católica do Rio de Janeiro.

Com o crescente emprego de inteligência artificial na moderação de conteúdo do Instagram, a empresa é cada vez mais criticada pelos efeitos discriminatórios de seus sistemas algorítmicos, que parecem revelar uma hierarquia de humanidades dentro da rede social. Ainda que a autoexposição dos usuários seja a fonte primária de dados usados pelo Instagram em sua geração de lucros, alguns corpos detêm plena e irrestrita visibilidade enquanto outros são desproporcionalmente ocultados pela moderação automatizada. A moderação algorítmica na rede social integra um novo tipo de governo das condutas, baseado em dados e correlações estatísticas, que assume traços específicos com relação a usuárias mulheres. Trata-se de um poder que, ocultado sob discursos tecnoutópicos e corporativistas do Instagram, não se reconhece como poder e se torna mais dificilmente detectável. Existe, contudo, a possibilidade de resistência a tal governo das condutas. Enquanto grande parte dos trabalhos acadêmicos sobre resistência online é focada na resistência ora a violências interpessoais entre usuários ora a práticas e discursos sexistas, é escassa a literatura sobre a resistência ao poder dos sistemas algorítmicos de moderação de conteúdo. Esta pesquisa pretende contribuir com o debate sobre o tema, explorando como técnica de resistência ou “autodefesa digital”, a autoexposição das usuárias do Instagram voltada à denúncia do poder algorítmico da rede social, em especial aos silenciamentos e discriminações por ele perpetuados.

Palavras-chave

Resistência; discriminação; governamentalidade algorítmica; moderação; Instagram.

Abstract

Dias, Carolina Bouchardet; Pele, Ilié Antonio(Advisor). **Algorithmic Moderation and Digital Self-Defense: The self-exposure of Instagram female users as resistance to the algorithmic government of conduct.** Rio de Janeiro, 2022. 129p. Dissertação de Mestrado – Departamento de Direito, Pontifícia Universidade Católica do Rio de Janeiro.

With the increasing use of artificial intelligence in Instagram's content moderation, the company is increasingly criticized for the discriminatory effects of its algorithmic systems, which seem to reveal a hierarchy of humanities within the social network. Even though users' self-exposure is the primary source of data used by Instagram in its profit generation, some bodies hold full and unrestricted visibility while others are disproportionately hidden by automated moderation. Algorithmic moderation on the social network integrates a new kind of government of conducts, based on data and statistical correlations, that takes on specific traits with regard to female users. This is a power that, hidden under Instagram's technoutopian and corporatist discourses, does not recognize itself as power and becomes more difficult to detect. There is, however, the possibility of resistance to such government of conduct. While much of the academic work on online resistance is focused on resistance either to interpersonal violence between users or to sexist practices and discourses, the literature on resistance to the power of algorithmic content moderation systems is scarce. This research intends to contribute to the debate on the theme, exploring as a resistance technique or “digital self-defense”, the self-exposure of female Instagram users aimed at denouncing the algorithmic power of the social network, especially the silences and discriminations perpetuated by it.

Keywords

Resistance; discrimination; algorithmic governmentality; moderation; Instagram.

Sumário

Lista de Figuras	11
Introdução	12
1. O Regime de visibilidades do Instagram	17
1.1 Diretrizes da Comunidade	19
1.2 Exclusão automatizada de conteúdo	30
1.3 <i>Shadowban</i>	40
2. Os valores subjacentes à moderação algorítmica.....	48
2.1 Capitalismo de vigilância e a lógica do lucro	49
2.2 Colonialidade automatizada e a hierarquização de humanidades	56
3. O governo algorítmico das condutas	65
3.1 A governamentalidade algorítmica.....	66
3.2 A subjetividade feminina produtiva no Instagram: Quem é a “boa sujeita de dados”?.....	72
3.3 Complicação dos processos de subjetivação e possibilidades de resistência	78
4. Autoexposição e resistência das usuárias do Instagram	85
4.1. Polly Oliveira: Performances humorísticas.....	88
4.2 Morgan Barbour: Digilantismo feminista	95
4.3 Autodefesa Digital: Autoexposição e contraexposição.....	101
5. Considerações Finais	109
Referências Bibliográficas	112
Anexo	127

Lista de Figuras

Figura 1 – Postagem de @pollyoliveirareal em 20.02.2021	127
Figura 2 – Postagem de @pollyoliveirareal em 04.03.2021	127
Figura 3 – Postagem de @cummunitystandards em 20.12.2020	128
Figura 4 – Postagem de @cummunitystandards em 20.12.2020	128
Figura 5 – Postagem de @cummunitystandards em 26.03.2020	129

Introdução

De plataforma digital para o compartilhamento de fotografias a uma das redes sociais mais utilizadas no mundo, o Instagram assistiu ao rápido crescimento de sua base de usuários em pouco tempo, expandindo sua proposta inicial e provocando mudanças até mesmo nos padrões de estética e na cultura como um todo. Concebida inicialmente como uma plataforma para o compartilhamento de fotografias, com o tempo ela passou a articular recursos similares aqueles oferecidos por outras redes sociais e recepcionar postagens de vídeos curtos temporários, transmissões ao vivo e até mesmo textos. Mesmo com as mudanças pelas quais passou a empresa, seu foco sempre foi a *imagem*, o que se reflete nas constantes atualizações de recursos oferecidos aos usuários: a possibilidade de edição de fotos com diferentes filtros; a expansão das capacidades das câmeras de smartphones, com a opção de incluir efeitos, caixas de perguntas e enquetes nas imagens publicadas; os Stories, que permitem a postagem de fotos e vídeos visíveis aos demais usuários por apenas 24h; os Arquivos, que possibilitam o salvamento de imagens de outros perfis em pasta pessoal; dentre outros artifícios cujo foco é primordialmente a edição e exposição de mídias visuais¹.

Apesar de a empresa ter sido adquirida pelo Facebook, Inc. em 2012 – que, desde outubro de 2021, chama-se Meta Platforms, Inc.² –, o foco em conteúdo visual permaneceu, aliado ao compartilhamento da intimidade, de onde advêm seu potencial de lucro. O aplicativo gira em torno de um tipo de autoexposição espetacularizada, que chamo de “autoexposição espetacular”³: a publicação, pelos usuários, de fotos e vídeos pensados para suscitar desejos nos demais indivíduos que

¹ MANOVICH, Lev. *Instagram and Contemporary Image*. Disponível em: <<http://manovich.net/index.php/projects/instagram-and-contemporary-image>>. Acesso em: 10 Out. 2021; e CALDEIRA, Sofia P.; RIDDER, Sander De; BAUWEL, Sofie Van, “Between the Mundane and the Political: Women’s Self-Representations on Instagram”. *Social Media + Society*, julho-setembro 2020, p. 7.

² UNITED STATES, SECURITIES AND EXCHANGE COMMISSION. Form 10-K for the fiscal year ended December 31, 2021, Meta Platforms, Inc. Disponível em: <<https://d18rn0p25nwr6d.cloudfront.net/CIK-0001326801/14039b47-2e2f-4054-9dc5-71bcc7cf01ce.pdf>>. Acesso em: 3 abr. 2022.

³ SIBILIA, Paula. *O show do eu: a intimidade como espetáculo*. Rio de Janeiro: Nova Fronteira, 2008.

utilizam a plataforma. A autoexposição espetacular comumente diz respeito à exposição de uma intimidade performada para a audiência da rede social, estética, apelativa, e que tenha o potencial de suscitar o engajamento dos seguidores daquele perfil – isto é, suscitar reações como curtidas, compartilhamento e comentários. É sobre a autoexposição espetacular que se baseia o modelo de negócios da plataforma, ou melhor, sobre a geração de dados decorrente dela.

A composição do ambiente imagético onde os usuários do Instagram interagem advém principalmente das *técnicas de curadoria e moderação* de conteúdo, as quais passaram a envolver cada vez mais inteligência artificial para lidar com a crescente produção de conteúdo na rede social. Enquanto as técnicas de curadoria do Instagram objetivam exibir a cada usuário publicações selecionadas para atenderem a desejos estatisticamente prováveis, as técnicas de moderação têm o objetivo de excluir ou reduzir a visibilidade do que é julgado pela rede social como violador das políticas internas ou inapropriado, capaz de suscitar nos usuários reações e comportamentos indesejados.

Neste trabalho, pretendo focar nas técnicas automatizadas de *moderação* de conteúdo, as quais servem a um propósito punitivo dos usuários que tenham descumprido as regras internas ou publicado conteúdo inapropriado. Sua finalidade é suprimir ou reduzir o alcance de conteúdo que não possa permanecer visível aos usuários, a fim de atingir os objetivos da empresa – em geral, reter/aumentar audiência e engajamento dos usuários. Uma vez que tanto a curadoria quanto a moderação de conteúdo são centradas na visibilidade dos usuários dentro da rede social e na distribuição de oportunidades de autoexposição, acabam tendo implicações mútuas, por vezes confundindo-se. Em razão disso, ocasionalmente se fará menção também à curadoria, como técnica que é integrante de um sistema maior de controle do ambiente do Instagram.

O emprego de sistemas algorítmicos na moderação de conteúdo do Instagram foi catalisado pela pandemia de Sars-Cov 2, uma vez que os empregados antes responsáveis por essa função tiveram que ser mantidos em casa para respeitar as ordens de isolamento social⁴. Esse aumento foi acompanhado de críticas à inconsistência

⁴ GILLESPIE, Tarleton, “Content moderation, AI, and the question of scale”. *Big Data & Society*, julho-dezembro 2020, p. 1.

da aplicação das Diretrizes da Comunidade por tais sistemas. Em resposta, a empresa se defende dizendo que tais efeitos consistem em erros dos sistemas algorítmicos ou impressões equivocadas dos usuários sobre seu funcionamento.

Tem crescido o interesse na academia pelos estudos sobre o emprego de inteligência artificial na moderação de conteúdo em redes sociais e as implicações políticas e éticas decorrentes disso⁵⁻⁶. Em especial, tem-se estudado como as técnicas de moderação de conteúdo, suprimindo mais conteúdo do que o “necessário”, atentam contra a liberdade de expressão dos usuários⁷, silenciando-os com técnicas de moderação automatizadas que produzem resultados inconsistentes⁸ e discriminatórios em razão do gênero do usuário⁹.

Por outro lado, ainda é incipiente a literatura que problematiza a moderação discriminatória de conteúdo a partir da resistência dos usuários. Enquanto grande parte dos trabalhos acadêmicos sobre resistência *online*, especialmente resistência feminista, é focada na resistência a violências interpessoais entre usuários¹⁰ ou na resistência ao racismo e ao sexismo como um todo¹¹, é escassa a literatura sobre a

⁵ GORWA, Robert; BINNS, Reuben; KATZENBACH, Christian, “Algorithmic content moderation: Technical and political challenges in the automation of platform governance”. *Big Data & Society*, janeiro-junho 2020, pp. 1-15; GILLESPIE, Tarleton, “Content moderation, AI, and the question of scale”. Op. Cit., pp. 1-5.

⁶ SPISÁK, Sanna; PIRJATANNIEMI, Elina; PAALANEN, Tommi; PAASONEN, Susanna; VIHLMAN, Maria, “Social Networking Sites’ Gag Order: Commercial Content Moderation’s Adverse Implications for Fundamental Sexual Rights and Wellbeing”. *Social Media + Society*, abril-junho 2021, pp. 1-9.

⁷ LLANSÓ, Emma J., “No amount of ‘AI’ in content moderation will solve filtering’s prior-restraint problem”, *Big Data & Society*, Janeiro-junho 2020, pp. 1-16.; GILLESPIE, Tarleton, “Governance of and by platforms”. In: BURGESS, Jean; POELL, Thomas; MARWICK, Alice (Eds.). *The SAGE Handbook of Social Media*. London: SAGE Publications Ltd., 2018, pp. 254-278; ELKIN-KOREN, Niva, “Contesting algorithms: Restoring the public interest in content filtering by artificial intelligence”. *Big Data & Society*, julho-dezembro 2020, pp. 1-13.

⁸ MYERS-WEST, Sarah, “Censored, suspended, shadowbanned: User interpretations of content moderation on social media platforms”. *New Media & Society*, v. 20., n. 11, 2018, pp. 4366-4383.

⁹ ZOLIDES, Andrew, “Gender moderation and moderating gender: Sexual content policies in Twitch’s community guidelines”. *New Media & Society*, v. 23, n. 10, 2021, pp. 2999-3015; DUGUAY, Stefanie; BURGESS, Jean; SUZOR, Nicolas, “Queer women’s experiences of patchwork platform governance on Tinder, Instagram, and Vine”. *Convergence: The International Journal of Research into New Media Technologies*, v. 26, n. 2, 2020, pp. 237-252; GERRARD, Ysabel; THORNHAM, Helen, “Content Moderation: Social Media’s Sexist Assemblages”. *New Media & Society*, v. 22, n. 7, 2020, pp. 1266-1286.

¹⁰ KURASAWA, Fuyuki; RONDINELLI, Elisabeth; KILICASLAN, Gulay, “Evidentiary activism in the digital age: on the rise of feminist struggles against gender-based online violence”. *Information, Communication & Society*, v. 24, n. 14, 2021, pp. 2174-2194; RODIS, Paulina d. C. Inara, “Let’s (re)tweet about racism and sexism: responses to cyber aggression toward Black and Asian women”. *Information, Communication & Society*, v. 24, n. 14, 2021, pp. 2153-2173.

¹¹ SARMENTO, Rayza, “Ativismo Feminista Online: mapeando eixos de atuação”. *Revista Sul-Americana de Ciência Política*, v. 7, n. 1, 2021, pp. 19-37; SOUZA, Emanuella de; LEÃO, Odilair; BOAKARI, Francis, “Nós por nós: protagonismo e resistência de mulheres afrodescendentes no Instagram”. *Educação, Cultura e Comunicação*, v. 12, n. 24, 2021, pp. 78-94. Disponível em:

resistência ao poder dos sistemas algorítmicos de moderação de conteúdo. Nesse âmbito, Myers-West estudou de que modo usuários de redes sociais silenciados por sistemas automatizados de moderação de conteúdo lidavam com o silenciamento por meio das ferramentas disponibilizadas pela plataforma¹². Velkova e Kaun¹³, por sua vez, exploraram a resistência de usuários a partir das próprias estruturas algorítmicas que os oprimem para corrigir seus resultados, e Bridges¹⁴ analisou a estratégia de provocar falhas algorítmicas como resistência às suas categorizações normativas.

A maioria dos trabalhos sobre o tema, no entanto, assume o ponto de vista da empresa de tecnologia ou dos sistemas algorítmicos – por exemplo, tendo como pressuposto a explicação da rede social sobre seus sistemas de moderação e buscando soluções a serem implementadas na forma de moderação de conteúdo. Além disso, são poucos os trabalhos que adotam uma perspectiva feminista nesse tema, ou que abordam os vieses de raça e gênero dos sistemas algorítmicos de moderação do Instagram.

Esta pesquisa pretende contribuir para o debate sobre a discriminação da moderação de conteúdo nas redes sociais ao fazer uma inversão: partir da experiência das usuárias no Instagram para pensar o poder algorítmico da rede social e explorar suas respostas como estratégias de autodefesa a esse poder. Em vez de estudar simplesmente a não possibilidade de autoexposição no Instagram (ou as possibilidades de autoexposição diferenciadas) devido à moderação de conteúdo, busco estudar a autoexposição das usuárias como autodefesa ao poder algorítmico do Instagram.

Pretendo, em primeiro lugar, realizar um estudo do caráter discriminatório de técnicas de moderação de conteúdo do Instagram, com ênfase na moderação algorítmica. Por moderação algorítmica, quero me referir às técnicas de moderação

<<http://publicacoes.unifatea.edu.br/index.php/ECCOM/article/view/1358>>. Acesso em: 9 nov. 2021; GOLDBECK, Jennifer; ASH, Summer; CABRERA, Nicole, “Hashtags as online communities with social support: A study of anti-sexism-in-science hashtag movement”. *First Monday*, v. 22, n. 9, 2017. Disponível em: <<https://firstmonday.org/ojs/index.php/fm/article/view/7572/6529>>. Acesso em: 2 fev. 2022; JANE, Emma A, “‘Dude...stop the spread’: Antagonism, agonism and #manspreading on social media”. *International Journal of Cultural Studies*, v. 20, n. 5, 2016, pp. 459-475.

¹² MYERS-WEST, Sarah, “Censored, suspended, shadowbanned: User interpretations of content moderation on social media platforms”. *New Media & Society*, v. 20., n. 11, 2018, pp. 4366-4383.

¹³ VELKOVA, Julia; KAUN, Anne, “Algorithmic resistance: media practices and the politics of repair”. *Information, Communication & Society*, v. 24, n. 4, 2021, pp. 523-540.

¹⁴ BRIDGES, Lauren E., “Digital Failure: Unbecoming the ‘good’ data subject through entropic, fugitive, and queer data”. *Big Data & Society*, janeiro-junho 2021, pp. 1-17.

que empregam modelos algorítmicos para a identificação de conteúdo supostamente ofensivo ou violador das políticas internas do aplicativo, e subsequente exclusão do conteúdo ou restrição de visibilidade do perfil responsável por seu compartilhamento¹⁵. Neste trabalho utilizarei, para fins de fluidez da leitura, as expressões “algorítmica” e “automatizada” intercaladamente, como sinônimos (ainda que não o sejam), para me referir às técnicas de moderação de conteúdo que se valem de sistemas algorítmicos.

Uma das maiores limitações desse trabalho diz respeito à restrição de informações disponíveis sobre os processos de moderação automatizada do Instagram. O principal material de apoio para as análises são artigos acadêmicos, postagens da própria empresa, documentos oficiais do Instagram e reportagens jornalísticas. Além disso, as técnicas automatizadas de moderação de conteúdo estão em constante mutação, sendo provável que tenham mudado substancialmente entre a escrita deste trabalho e sua apresentação. Apesar disso, as discussões levantadas pela análise dos efeitos dos sistemas atuais é que importa, sobretudo na análise da micropolítica operada por usuárias e usuários que resistem.

Para cumprir com o objetivo da pesquisa, pretendo explorar brevemente, no Capítulo 1, os efeitos de algumas das técnicas algorítmicas de moderação de conteúdo, a fim de identificar eventuais discriminações, bem como traços do funcionamento dessas técnicas que revelam como é a relação travada com os usuários. No Capítulo 2, analiso os valores subjacentes à moderação algorítmica do Instagram – especificamente, a busca incessante pelo lucro, a hierarquização de humanidades herdada da colonialidade e a epistemologia datacêntrica, que deposita uma fé acrítica no potencial de dados descreverem objetivamente a realidade e possibilitarem a otimização contínua da rede social. No Capítulo 3, analiso o governo algorítmico das condutas no Instagram do qual participa a moderação automatizada, bem como o governo especificamente da conduta de usuárias identificadas como mulheres, e, por fim, onde se encontram as possibilidades de resistência a esse governo. Por fim, no Capítulo 4, por meio de dois estudos de caso, exploro a possibilidade de uso da

¹⁵ GORWA et. al. usam uma definição de moderação algorítmica semelhante à utilizada neste trabalho: “We define algorithmic commercial content moderation (referred to as algorithmic moderation for brevity in the following sections) as systems that classify user generated content based on either matching or prediction, leading to a decision and governance outcome (e.g. removal, geoblocking, account takedown).” (GORWA, Robert; BINNS, Reuben; KATZENBACH, Christian, “Algorithmic content moderation: Technical and political challenges in the automation of platform governance”, pp. 1-15)

autoexposição por usuárias do Instagram como autodefesa ao governo algorítmico das condutas.

O Regime de visibilidades do Instagram

Em 2016, o Instagram anunciou que o *feed* da plataforma – isto é, a seção que exhibe as publicações do próprio usuário e as dos perfis por ele seguidos¹⁶ – deixaria de ser ordenado com base em um critério puramente cronológico reverso para passar a ser um *feed* “curado”, isto é, que exibisse postagens mais recentes relacionadas aos interesses do usuário¹⁷. Com essa alteração, os modelos algorítmicos usados na curadoria e moderação do conteúdo a ser exibido a cada usuário passaram a gerar controvérsias não só sobre vieses humanos embutidos em sua construção como também compartilhados pelos usuários, cujo comportamento treina tais modelos.

A moderação automatizada de conteúdo do Instagram é recorrentemente associada à discriminação e censura, sobretudo de corpos femininos, transexuais e racializados, dentre outros grupos sociais discriminados com base em características físicas, visíveis no conteúdo publicado¹⁸. Isso porque, apesar de todos os usuários e perfis estarem formalmente submetidos às mesmas regras escritas de visibilidade (as Diretrizes da Comunidade), a moderação algorítmica é acusada de ter efeitos consistentemente inconsistentes, distribuindo oportunidades de autoexposição, representação e discurso conforme a adequação dos corpos expostos a certas matrizes de gênero, estéticas e raciais dominantes.

¹⁶ INSTAGRAM, “Feed”. Disponível em: <<https://business.instagram.com/instagram-feed>>. Acesso em: 10 fev. 2022.

¹⁷ MOSSERI, Adam. “Shedding More Light on How Instagram Works”. *Instagram*, 08.06.2021. Disponível em: <<https://about.instagram.com/blog/announcements/shedding-more-light-on-how-instagram-works>>. Acesso em: 30 jul. 2021.

¹⁸ MORALES, Eduardo, “How Instagram Upholds the Censoring of Black Bodies and the Suppression of POC-Run Accounts”. *An Injustice*, 03.08.2020. Disponível em: <<https://aninjustice-mag.com/how-instagram-upholds-the-censoring-of-black-bodies-and-the-suppression-of-poc-run-accounts-cd1d33ff4dc5>>. Acesso em: 17 nov. 2021; CHRISTIE, Lacey-Jade, “Instagram censored one of these photos but not the other. We must ask why”. *The Guardian*, 19.10.2020. Disponível em: <<https://www.theguardian.com/technology/2020/oct/20/instagram-censored-one-of-these-photos-but-not-the-other-we-must-ask-why>>. Acesso em: 17 nov. 2021; ARE, Carolina, “How Instagram’s algorithm is censoring women and vulnerable users but helping online abusers”. *Feminist Media Studies*, v. 20, n. 5, 2020, pp. 741-744; DICKSON, E.J., “Why Did Instagram Confuse These Ads Featuring LGBTQ People for Escort Ads?”. *Rolling Stone*, 11.07.2019. Disponível em: <<https://www.rollingstone.com/culture/culture-features/instagram-transgender-sex-workers-857667/>>. Acesso em: 20 nov. 2021.

Portanto, é como se existisse um regime *ocultado* de visibilidade dentro do Instagram. Ocultado, e não oculto, porque, apesar de suas regras serem percebidas a partir dos efeitos da moderação algorítmica, elas não constam expressamente nas políticas internas da rede social e são recorrentemente *negadas* pela empresa quando alguma contestação ou escândalo midiático vêm à tona – normalmente sob argumentos técnicos sobre o funcionamento e a complexidade algorítmicos.

Assim, para além das Diretrizes da Comunidade – as regras efetivamente escritas que determinam o que não pode ser visível dentro da plataforma –, técnicas como a exclusão automatizada de conteúdo e a restrição de visibilidade (*shadowbanning*) reforçariam e constituiriam regras *ocultadas* de visibilidade. Em um paralelo com o *racismo por denegação*¹⁹ localizado por Lelia Gonzalez na sociedade brasileira, que, racista, nega o próprio racismo sob discursos de democracia racial, é como se a atividade algorítmica praticasse discriminações raciais e de gênero [além de outras] por denegação. A autora se baseia na categoria freudiana de *denegação* como “[p]rocesso pelo qual o indivíduo, embora formulando um de seus desejos, pensamentos ou sentimentos, até aí recalcado, continua a defender-se dele, negando que lhe pertença”²⁰. Embora os sistemas automatizados de moderação de conteúdo do Instagram deixem transparecer o racismo e o sexismo que integram a estrutura dos produtos da empresa, a Meta Platforms Inc. continua a negá-los, rechaçando os adjetivos “racista” e “sexista” atribuídos aos seus sistemas sob argumentos tecnocêntricos – isto é, focados em minúcias técnicas e exclusivamente no potencial benéfico da tecnologia para a vida social.

Tendo isso em vista, neste Capítulo 1, pretendo explorar o regime ocultado de visibilidades do Instagram a partir dos efeitos de técnicas de moderação automatizada de conteúdo. Para tanto, analisarei três elementos que integram a moderação algorítmica da plataforma: (i) as Diretrizes da Comunidade, isto é, as políticas internas formalmente usadas como guia para determinar o tipo de conteúdo que não pode ser publicado no Instagram e que, em tese, *fundamentam as técnicas automatizadas de exclusão e omissão de conteúdo*, (ii) a exclusão automatizada de conte-

¹⁹ GONZALEZ, Lélia. *Por um Feminismo afro-latino-americano: ensaios, intervenções e diálogos* (Organização: Flávia Rios e Márcia Lima). Rio de Janeiro: Zahar, 2020, p. 115.

²⁰ LAPLANCHE, Jean; e PONTALIS, Jean-Bertrand. *Vocabulário da psicanálise*. Santos: Martins Fontes, 1970 *apud* Idem.

údo, cuja proposta é identificar e excluir postagens que violem as Diretrizes da Comunidade, e (iii) o *shadowban*, uma técnica algorítmica de restrição de visibilidade de determinado perfil após a postagem de conteúdo que, apesar de não violar as regras da plataforma, é considerado ofensivo ou limítrofe às proibições da rede social.

Com essa exploração inicial do regime ocultado de visibilidades no Instagram, pretendo investigar como os *efeitos* das técnicas algorítmicas de moderação selecionadas refletem e atualizam discriminações de raça e de gênero e revelam uma nova relação de poder, travada entre usuários e plataforma, por meio de sistemas algorítmicos.

Na Seção 1.1, a partir da análise principalmente da proibição de imagens de mamilos femininos pelas Diretrizes da Comunidade, pretendo explorar o tratamento diferenciado dispensado a corpos femininos e masculinos, bem como a não compreensão de corpos alheios à divisão binária dos gêneros nessas regras escritas. Na Seção 1.2., no estudo da exclusão automatizada de conteúdo, examino efeitos relatados por pesquisa da Salty Algorithmic Bias Collective e reportagens jornalísticas a fim de investigar quais são outras eventuais discriminações sentidas por usuários do Instagram nessa técnica de moderação de conteúdo. Por fim, no estudo do *shadowban*, busco identificar nessa prática alguns dos traços presentes em uma nova forma de poder exercida pelos sistemas algorítmicos do Instagram como um todo: o caráter punitivo dos sistemas de moderação, a hierarquização de humanidades com base em características físicas e a negação do poder como tal, de onde advém a dificuldade de localizá-lo, entendê-lo como poder e responder a ele. Essas ideias serão desenvolvidas com mais detalhes no Capítulo 3.

1.1.

Diretrizes da Comunidade

Nesta subseção, a partir da análise da proibição de mamilos femininos no Instagram argumentarei que as Diretrizes da Comunidade refletem concepções de gênero e feminilidade que se atém à sexualização do corpo feminino e à binariedade de gênero. Buscarei demonstrar que a redação das regras acaba cristalizando o que é considerado “normal”, padrões de normalidade não ofensiva à comunidade do

Instagram, o que importa na *imposição de significados* e no *policimento desproporcional* do que é identificado pelos sistemas automatizados de moderação como “anormal”.

As Diretrizes da Comunidade consistem em um conjunto de regras que regulam o tipo de conteúdo proibido no Instagram cuja implementação é garantida pelas técnicas de moderação da plataforma – sejam elas humanas ou algorítmicas²¹. As Diretrizes da Comunidade são apresentadas como regras objetivas e claras sobre o tipo de publicação que não é tolerado dentro da rede social, com o objetivo declarado de garantir a segurança e o respeito nas interações travadas em seu espaço e fomentar uma comunidade positiva e diversificada²².

Tendo esses objetivos em mira, as regras da plataforma vedam, por exemplo, publicações que contenham: (i) o uso de conteúdo alheio; (ii) imagens de nudez ou mamilos femininos; (iii) ameaças reais ou discurso de ódio; (iv) discursos direcionados a humilhar ou degradar pessoas específicas; e (v) temas sociopolíticos entendidos como polêmicos que possam incentivar condutas tipificadas como crimes (lesões a outrem ou a si mesmo (automutilação), suicídios, assassinatos, crimes de ódio, interrupção voluntária da gravidez, etc.)²³:

- **Compartilhe somente as fotos que você tirou e os vídeos que gravou ou conteúdo que você tem direito de compartilhar.**
Como sempre, você é dono do conteúdo que publica no Instagram. Publique conteúdo autêntico e não publique nada copiado ou obtido da internet sem ter o direito. Saiba mais sobre os direitos de propriedade intelectual.
- **Publique fotos e vídeos apropriados para um público diversificado.**
Sabemos que algumas pessoas podem querer compartilhar imagens com nudez de natureza artística ou criativa. No entanto, por vários motivos, não permitimos nudez no Instagram. Isso inclui fotos, vídeos e alguns conteúdos criados digitalmente que mostram relações sexuais, genitais e destaque em nádegas totalmente expostas. Algumas fotos de mamilos femininos também estão proibidas, mas permitimos fotos no contexto de amamentação, parto e pós-parto, situações relacionadas à saúde (por exemplo, pós-mastectomia, conscientização sobre câncer de mama ou cirurgia

²¹ ARE, Carolina, “The Shadowban Cycle: an autoethnography of pole dancing, nudity and censorship on Instagram”. *Feminist Media Studies*, maio 2021, pp. 4-5. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/14680777.2021.1928259>>. Acesso em: 20 nov. 2021.

²² Nos termos da versão em português das Diretrizes da Comunidade: “O Instagram é um reflexo da nossa comunidade feita de diferentes culturas, idades e crenças. Passamos muito tempo pensando sobre os diferentes pontos de vista que criariam um ambiente aberto e seguro para todos. Elaboramos as Diretrizes da Comunidade para que você nos ajude a promover e a proteger essa comunidade maravilhosa.” (INSTAGRAM, “Community Guidelines”. Disponível em: <https://help.instagram.com/477434105621119/?helpref=hc_fnav>. Acesso em: 2 fev. 2022).

²³ INSTAGRAM, “Diretrizes da Comunidade”. Disponível em: <<https://www.facebook.com/help/instagram/477434105621119>>. Acesso em: 9 abr. 2022.

de confirmação de gênero) ou ato de protesto. Nudez em imagens de pinturas e esculturas também é permitida.

[...]

- **Respeite os outros membros da comunidade do Instagram.**

Queremos promover uma comunidade diversificada e positiva. Removemos conteúdo que apresente ameaças reais ou discurso de ódio e conteúdo direcionado a pessoas físicas cuja a intenção seja degradá-las ou constrangê-las.

[...]

- **Mantenha o nosso ambiente acolhedor. Não romantize a automutilação.**

Todos na comunidade do Instagram se preocupam uns com os outros. Ela é um espaço que muitas vezes é procurado por pessoas que enfrentam problemas difíceis (como distúrbios alimentares, autoflagelação ou outros tipos de automutilação) para gerar conscientização ou encontrar apoio. [...]

Incentivar ou solicitar que as pessoas pratiquem automutilação vai contra esse ambiente de apoio. Se recebermos denúncias sobre isso, removeremos o conteúdo ou desativaremos as contas que o publicarem. Também podemos remover conteúdo que identifique vítimas ou sobreviventes de automutilação com intuito de ataque ou humor.

A violação das Diretrizes da Comunidade pode importar na exclusão de conteúdo e/ou de perfis; bloqueio temporário (por período determinado ou não) para postagens e uso dos recursos da rede social²⁴; ou em “outras restrições” – expressão cujo significado não é definido pelo Instagram. Para fins desta pesquisa, serão consideradas abrangidas as restrições de visibilidade de perfis por meio do “*shadowban*”, abordado mais adiante.

Apesar do objetivo declarado de zelar pela pluralidade, como qualquer conjunto de normas as Diretrizes da Comunidade são redigidas de forma enviesada. A seguir, tratarei de uma das regras das Diretrizes da Comunidade que mais explicitamente manifesta discriminação no tratamento de corpos expostos na rede social conforme o seu gênero: a vedação a imagens exclusivamente de “mamilos femininos”.

Pelo menos nos últimos 10 anos, algumas controvérsias envolvendo a redação das Diretrizes da Comunidade do Instagram foram alvo de atenção midiática. Uma das mais antigas, que acompanhou a rede social desde os seus anos iniciais e

²⁴ Dentre alguns dos recursos a que se faz referência, estão: (i) o *Feed*, em que os usuários encontram as publicações do próprio perfil e daqueles que seguem; (ii) os *Stories*, recurso que permite a publicação de imagens, textos e vídeos disponíveis para a visualização dos demais usuários por apenas 24h; (iii) o *Reels*, uma ferramenta que permite que os usuários publiquem vídeos com a duração máxima de 60 segundos e o editem com filtros de imagem e áudios do aplicativo; e (iv) as *Lives*, que permitem a gravação e transmissão de vídeos ao vivo e a interação em tempo real com a audiência. Esses são alguns dos recursos que podem ser bloqueados para uso por determinado perfil que tenha violado as Diretrizes da Comunidade.

recebeu intensa atenção da mídia e de personalidades famosas, foi aquela relacionada à proibição de imagens de mamilos [exclusivamente] femininos (regra até hoje existente na plataforma, com algumas flexibilizações)²⁵. Enquanto fotografias em que se viam mamilos masculinos eram permitidas sem quaisquer restrições, aquelas que expunham mamilos femininos eram absolutamente vedadas sob os imperativos das Diretrizes da Comunidade de manter a segurança e o respeito, ainda que se tratassem de fotos não sexuais ou eróticas, como imagens de resultados de mastectomias²⁶, de amamentação²⁷ e artísticas²⁸.

Tal censura, no entanto, não foi facilmente acatada. Em resposta, usuárias do Instagram iniciaram, por exemplo, a campanha de *hashtag* #freethenipple, em 2012, contestando o olhar masculino e patriarcal sobre o corpo das mulheres que é reproduzido pelas Diretrizes da Comunidade quando sexualiza incondicionalmente os mamilos femininos e determina que sejam escondidos. Também houve a criação do perfil @genderless_nipples, em 2016, voltado a publicar imagens de mamilos em que não é possível identificar o gênero da pessoa retratada²⁹. A maioria das fotos permanecia no ar e não era identificada como violação às Diretrizes da Comunidade, tendo apenas algumas sido excluídas pela moderação. Outra manifestação de resistência crítica às Diretrizes da Comunidade do Instagram foi a de Micol Hebron, que, inconformada após ter derrubada pelo Facebook uma foto sua sem camisa em uma exibição de arte sobre câncer de mama, criou adesivos de mamilos masculinos

²⁵ GRETCHEN, Faust, “Hair, Blood and the Nipple: Instagram Censorship and the Female Body”. In: KÖHN, Steffen; FOX, Samantha; TERRY, Mike; FRÖMMING, Urte Undine (Coord.). *Digital Environments. Ethnographic Perspectives across Global Online and Offline Spaces*. Bielefeld: Verlag, 2017, pp 159-170; e SURINGAR, G.A. Fighting Instagram Censorship of the Female Nipple: @exotic.cancer and @genderless_nipples Seeking the Boundaries of the Instagram Guidelines to Critique the Sexualisation of the Female Body. Dissertação de Mestrado. Universidade de Utrecht. 29f. TCS Postcolonial and Gender Studies, 2020.

²⁶ MAJSIAK, Brianna, “Instagram and Facebook, Please Stop Sexualizing Our Mastectomy Photos”. *Everyday Health*, 22.10.2020. Disponível em: <<https://www.everydayhealth.com/breast-cancer/instagram-and-facebook-please-stop-sexualizing-our-mastectomy-photos/>>. Acesso em: 2 fev. 2022.

²⁷ ANDERSSON, Erin, “Breastfeeding on social media: How one picture led to a Toronto mother’s ban from Instagram”. *The Globe and Mail*, 13.05.2014. Disponível em: <<https://www.theglobeandmail.com/life/parenting/a-breastfeeding-mom-on-instagram-call-in-the-censors/article18644323/>>. Acesso em: 2 fev. 2022.

²⁸ VOLKERS, R., “Perverse Media: How Instagram limits the potential of feminist art”. *EUC Student Academic Journal – ESAJ*, n. 2. Disponível em: <<https://esaj.nl/Roos-Volkers-2/>>. Acesso em: 2 fev. 2022.

²⁹ HUNT, Elle, “Male or female? Genderless Nipples account challenges Instagram’s sexist standards”. *The Guardian*, 19.01.2017. Disponível em: <<https://www.theguardian.com/world/2017/jan/19/genderless-nipples-generates-50000-followers-on-instagram>>. Acesso em: 2 fev. 2022.

que mulheres poderiam usar para tirar fotos caso quisessem expor seus seios “ofensivos”³⁰ no Instagram³¹.

Essas são algumas das manifestações de ocupação, disputa e resistência na plataforma que denunciavam a sexualização incondicional do corpo feminino – isto é, o corpo socialmente determinado como de “mulher” devido à sua composição biológica – pelas Diretrizes da Comunidade. Apesar de a regra que veda a exposição de mamilos femininos não ser a única dentro das Diretrizes da Comunidade digna de críticas por permitir que usuários sejam tratados de forma diferenciada com base em seu gênero, é aquela que mais *explicitamente* manifesta o olhar patriarcal das políticas internas da plataforma sobre o corpo feminino.

Após inúmeros protestos das usuárias e atenção midiática, a redação das regras sobre nudez feminina foi alterada, e, em 2022, permite a exibição de mamilos femininos “no contexto de amamentação, parto e pós-parto, em situações relacionadas à saúde (por exemplo, pós-mastectomia, conscientização sobre câncer de mama ou cirurgia de confirmação de gênero) ou como ato de protesto”³². No entanto, as Diretrizes da Comunidade ainda estipulam como regra geral a vedação aos “mamilos femininos”³³, sem qualquer tentativa de entender a diversidade de corpos existentes ou de regulamentar de forma semelhante o corpo socialmente lido como masculino.

É possível apreender a vedação à exposição dos mamilos femininos como uma continuidade do tipo de sensibilidade coletiva que não admite a exposição do

³⁰ HEBRON, Micol, “Nipples! Articles about my digital pasty designed to fight censorship of female nipples”. Disponível em: <<http://micolhebron.artcodeinc.com/pages/nipples-articles-about-my-digital-pasty-designed-to-fight-censorship-of-female-nipples/>>. Acesso em: 2 jan. 2022.

³¹ FERRIER, Morwenna, “How to beat Instagram and Facebook’s female nipple ban? Use male nipples instead”. *The Guardian*, 07.07.2015. Disponível em: <<https://www.theguardian.com/life-andstyle/shortcuts/2015/jul/07/instagram-facebook-female-nipple-ban-use-male-nipples-instead>>. Acesso em: 3 fev. 2022.

³² “Sabemos que há momentos em que as pessoas podem querer compartilhar imagens de nudez de natureza artística ou criativa, mas, por vários motivos, não permitimos nudez no Instagram. Isso inclui fotos, vídeos e alguns conteúdos criados digitalmente que mostram relações sexuais, genitais e foco em nádegas totalmente expostas. Algumas fotos de mamilos femininos também estão incluídas, mas fotos no contexto de amamentação, parto e pós-parto, em situações relacionadas à saúde (por exemplo, pós-mastectomia, conscientização sobre câncer de mama ou cirurgia de confirmação de gênero) ou como ato de protesto são permitidas. Nudez em imagens de pinturas e esculturas também é permitida.”. (INSTAGRAM, “Diretrizes da Comunidade”). Disponível em: <<https://www.facebook.com/help/instagram/477434105621119>>. Acesso em: 9 abr. 2022).

³³ Também são proibidas imagens de seios femininos sendo apertados, “definido como o movimento de pegar neles com os dedos curvados, deixando marcas e uma mudança nítida no formato dos seios” (Idem).

corpo da mulher em espaços “públicos” – no caso de Instagram, em espaços acessados pela audiência da plataforma³⁴. No entanto, em vez de buscar uma genealogia da sexualização dos mamilos femininos para compreender sua vedação pelo Instagram, importa considerar essa proibição das Diretrizes da Comunidade dentro do contexto cultural e jurídico em que está situada a empresa para entender sua lógica.

O Instagram não veda a exposição de mamilos femininos pura e simplesmente porque adere de forma consciente a um olhar patriarcal. Um dos predicados suscitado pela empresa é o de ser uma rede social que preza pela sua “comunidade diversa”, constituída por “diferentes culturas, idades e crenças”. A vedação a mamilos femininos supostamente tem como objetivo assegurar um lugar “seguro”, evitar que usuários se ofendam e manter a audiência de diversas faixas etárias da plataforma – de onde provém seu lucro, como será visto nos Capítulos 2 e 3.

Nesse sentido, duas das razões que podem explicar o combate ferrenho a representações da sexualidade feminina no Instagram para manter a moral e a segurança dizem respeito: (i) à tradição puritana dos Estados Unidos, país em que foi criada e está sediada a empresa e (ii) à legislação local. A primeira das razões busca evitar eventuais ofensas a outros usuários e a segunda das razões visa a evitar riscos legais à própria empresa Instagram – ou sua controladora, Meta Platforms, Inc. De todo modo, existe um esforço grande em dessexualizar a rede social – seja para reter a audiência “plural” que alega proteger, seja para evitar responsabilidade legal.

O Instagram é permeado por valores culturais específicos do país em que localizado – os Estados Unidos, onde existe uma tradição cultural puritana segundo a qual o sexo e a nudez são associados ao risco e culpa. A tese de Paasonen et. al é que tal tradição acaba sendo refletida na moderação de redes sociais como o Instagram e aplicada indistintamente ao conteúdo e usuários de outras nacionalidades. Isso explicaria em alguma medida o tratamento gravoso dispensado à nudez e se-

³⁴ Com base no estudo de Silvia Federici, sobre a caça às bruxas na Europa como processo político [de acumulação primitiva do capital] que disseminou socialmente um imaginário atribuído à figura feminina como sexualmente descontrolada, traiçoeira e obscena, além de ter relegado o corpo feminino ao espaço privado, pode-se encontrar similaridades entre essa sensibilidade que legitimava a caça às bruxas e o disciplinamento do corpo feminino e a sensibilidade ainda circulante na sociedade ocidental que vê como legítimo o ocultamento dos mamilos femininos (FEDERICI, Silvia. *Calibã e a Bruxa: mulheres, corpos e acumulação primitiva*. São Paulo: Elefante, 2017).

xualidade pela moderação de conteúdo, atrelando-as a risco e insegurança abstratamente concebidos, sem considerar-se o contexto de produção e compartilhamento do conteúdo³⁵.

Além disso, a aprovação da legislação FOSTA/SESTA nos Estados Unidos em 2018 teve impacto sobre o combate à nudez feminina e sexualidade na moderação de conteúdo de redes sociais como o Instagram³⁶. Com o objetivo de dificultar o tráfico ilegal de sexo *online*, a legislação criou uma exceção ao *Communications Decency Act* (Seção 230), permitindo a responsabilização de sites pelo conteúdo de terceiros quando se trate de publicidade de prostituição (inclusive trabalhos sexuais consensuais). Apesar do objetivo nobre, a legislação contribuiu para que redes sociais passassem a adotar regras mais rígidas sobre a exposição do corpo feminino na moderação de seus espaços para evitarem ser responsabilizadas pela facilitação de trabalhos sexuais³⁷.

Além da sexualização do corpo feminino, a vedação a “mamilos femininos” contida nas Diretrizes da Comunidade deixa clara sua adesão à binariedade de gênero. Não há, nessas regras internas, qualquer ponderação sobre a identidade com a qual a pessoa retratada que expõe seus mamilos se identifica, ou se ela, anteriormente definida como mulher por sua constituição biológica, passou por uma transição médica de gênero vindo a ser socialmente enxergada como homem – e, portanto, com mamilos não sexualizados³⁸.

Não importa o contexto, o caráter da imagem publicada (se sexual ou não) nem o gênero com o qual a pessoa retratada se identifica³⁹, se mamilos identificados

³⁵ PAASONEN, Susanna; JARRETT, Kylie; LIGHT, Ben. *NSFW: Sex, Humor and Risk in Social Media*. Cambridge, MA: The MIT Press, 2019.

³⁶ *Ibid.*, p. 62.

³⁷ PAASONEN, Susanna; ARE, Carolina, “Sex in the shadows of celebrity”. *Porn Studies*, v. 8, n. 4, setembro 2021, pp. 414.

³⁸ JACOBS, Julia, “Will Instagram Ever ‘Free the Nipple’?”. *The New York Times*, 22.11.2019. Disponível em: <<https://www.nytimes.com/2019/11/22/arts/design/instagram-free-the-nipple.html>>. Acesso em: 22 fev. 2022.

³⁹ A binariedade de gênero refletida nas Diretrizes da Comunidade, isto é, a divisão dos seres humanos simplesmente em homens e mulheres, faz surgir indagações sobre a determinação do gênero de um mamilo independentemente da identidade com a qual a pessoa retratada se identifique, ou se ela, anteriormente definida como mulher por sua constituição biológica, passou por uma transição médica de gênero. Rain Dove (@raindovemodel), modelo não-binária, após publicar imagens suas sem camisa expondo seus mamilos no Instagram teve uma delas apagada pela plataforma, sob o pretexto de serem mamilos femininos – uma vez que seu corpo é socialmente lido como o de uma mulher. Sua identidade é prontamente descartada pelas Diretrizes da Comunidade e seu corpo recebe a imposição de uma matriz de inteligibilidade que ignora sua existência. (DOVE, Rain. *I am NOT “Female”*. 27 jun. 2018. Instagram: @raindovemodel. Disponível em: <<https://www.instagram.com/p/BkibsngHp73/>>. Acesso em: 2 fev. 2022)

pela moderação como de “mulheres” forem visíveis, tais imagens violam as Diretrizes da Comunidade, devendo ser ocultadas para manter a moral e a segurança dentro da plataforma e preservar a “comunidade diversa” que as regras internas alegam proteger. Com essa proibição, no entanto, as Diretrizes da Comunidade fazem o oposto.

Ao reduzirem a nudez feminina à obscenidade e ofensividade, sempre punível, e sequer reconhecerem a existência de corpos que não sejam simplesmente masculinos ou femininos, as Diretrizes da Comunidade suprimem em alguma medida a possibilidade de diversidade da comunidade, além de instituírem um padrão de normalidade, de normatização dos corpos, reforçando conjuntos de significados sociais atribuídos a eles em função de sua aparência física, sua constituição biológica imediatamente apreensível nas imagens postadas.

O reducionismo com o qual as Diretrizes da Comunidade tratam a nudez feminina importa, afinal, na *imposição de significados* aos corpos identificados como de mulheres. A aplicação global de uma política contida dentro da binariedade de gênero e culturalmente situada, sem qualquer adaptação, acaba resultando na punição de usuários cujas imagens não se amoldem aos significados corporais consolidados nas Diretrizes da Comunidade. Seja a imposição de significados sexuais a corpos femininos que sequer pretendiam ser sexuais ou a imposição de significados de gênero a corpos que não se identificam com a categorização binária dos corpos, identidades alheias estão constantemente recebendo atribuições de significados hegemônicos que não respeitam sua humanidade e autodeterminação. Formulando essa conclusão de outra forma, significados *hegemônicos* de gênero tentam *homogeneizar* os corpos circulantes na rede social. O objetivo das Diretrizes da Comunidade de preservar uma comunidade diversa é, afinal, tornado inócuo, uma vez que as ideias e imagens de corpos dissidentes da norma devem lutar para existir no Instagram, exceto se se conformarem aos padrões escritos – e discriminatórios – de visibilidade.

A fixação de um padrão de normalidade, que supostamente não ofenda nem afete negativamente a audiência do Instagram, é necessariamente acompanhado de um policiamento dos limites desse padrão, para identificar e punir os desvios às Diretrizes da Comunidade que coloquem em risco os objetivos da rede social. Além de instituírem um regime de visibilidades escrito, que cristaliza o que é normal e inofensivo para os usuários do Instagram, as Diretrizes da Comunidade estabelecem

as sanções aplicáveis aos desvios. Isso importa em que o não cumprimento com os significados sociais de gênero consolidados nas regras internas leva a uma punição pela moderação da plataforma.

Existe, portanto, um *policimento desproporcional*, pela moderação de conteúdo da rede social, do corpo e sexualidade identificados como femininos (ainda que não o sejam) e das regras de inteligibilidade dos gêneros. O foco acentuado das Diretrizes da Comunidade em proibir exclusivamente mamilos femininos e combater a sexualidade – melhor dizendo, o corpo marcado pelo seu sexo, o feminino – resulta no ininterrupto escrutínio dos corpos de mulheres, a fim de verificar se estão vestidos ou posados de forma potencialmente “ofensiva”, se sexuais, o quanto de pele é mostrada, como estão interagindo com a composição da cena, etc. É preciso assegurar que o corpo feminino não ofenda por sua sexualidade ou cause nojo, arriscando a perda de amplitude de audiência para a rede social. As imagens de corpos trans acabam recebendo o mesmo tratamento por não serem identificados simplesmente como homens ou mulheres – ou, ao contrário, por serem identificados como mulheres quando não o são.

É expressivo desse policiamento o fato de que, pelo menos até julho de 2019, a redação das regras de exposição de corpos femininos do Instagram tenha se baseado nas diretrizes de publicidade da *Victoria's Secret*⁴⁰, marca de lingerie cujo marketing se tornou famoso pela promoção de corpos inatingíveis e de um padrão de beleza centrado exclusivamente na juventude, magreza, cisgenereidade, branquitude e traços europeus. Tais regras, que serviram de base para o Instagram, traziam tópicos descrevendo *minuciosamente* como corpos femininos *seminus* deveriam estar dispostos para serem considerados “aceitáveis”, não havendo, por outro lado, qualquer norma dirigida a corpos masculinos em roupas íntimas ou de banho. Afinal, o corpo não marcado pelo seu sexo não precisa ser exaustivamente regulamentado e fiscalizado para que atenda ao olhar do outro; basta-lhe existir dentro de seus próprios termos. Além disso, as mesmas regras utilizavam a todo tempo a expressão “meninas”, o que, além de apontar para um padrão etarista de atração física, limita-se ao binarismo de gênero e à existência de dois tipos de corpos. Quaisquer

⁴⁰ SALTY, “Exclusive: Victoria’s Secret Influence on Instagram’s Censorship Policies”. *Salty*, 22.11.2019. Disponível em: <<https://saltyworld.net/exclusive-victorias-secret-influence-on-instagram-censorship-policies/>>. Acesso em: 28 jan. 2022.

outros, alheios à norma (de inteligibilidade dos gêneros⁴¹), são tidos como violações. Esses, enfim, são os valores imediatamente refletidos na redação das Diretrizes da Comunidade, o regime escrito de visibilidade da plataforma.

A título de exemplo do policiamento desproporcional de corpos não alinhados à divisão binária dos gêneros⁴², em 2019 gerou repercussão na mídia⁴³ o caso do boletim de notícias Salty (@salty.world), que tinha suas publicações no Instagram recorrentemente excluídas sob a alegação de estarem “promovendo serviços de acompanhantes”, o que não seria tolerado pela plataforma. As postagens, no entanto, consistiam em capas de sua revista mensal, retratando pessoas trans racializadas, portadoras de deficiências físicas ou gordas, totalmente vestidas ao lado dos títulos das reportagens trazidas pela revista – geralmente envolvendo experiências vividas por esses grupos socialmente marginalizados⁴⁴.

O policiamento desproporcional não tem como alvo apenas o visual do material publicado, mas também as palavras que compõem o conteúdo. Alguns termos são incluídos em listas de bloqueio (*blocklists*), principalmente por serem indicativas de conteúdo sexual, um dos tipos de material mais combatidos pela rede social, como já foi o caso de palavras que se referem à sexualidade de minorias⁴⁵⁻⁴⁶. Desse modo, esse público frequentemente encontra mais dificuldades para promover seus anúncios, manifestar ideias ou expor seus corpos (não nus).

⁴¹ BUTLER, Judith. *Problemas de Gênero: Feminismo e subversão da identidade* (tradução de Renato Aguiar), 18ª edição. Rio de Janeiro: Civilização Brasileira, 2019.

⁴² A propósito do policiamento de anúncios voltados ao público queer: O’HARA, Mary Emily, “Queer and Feminist Brands say they are Being Blocked from Running Ads on Instagram and Facebook”. *MTV News*, 19.07.2019. Disponível em: <<https://www.mtv.com/news/3131929/queer-and-feminist-brands-say-they-are-being-blocked-from-running-ads-on-instagram-and-facebook/>>. Acesso em: 2 fev. 2022; BECK, Willa, “Why are Certain Bodies being Censored on IG? : Uye Surana”. *Office*, 18.11.2021. Disponível em: <<http://officemagazine.net/why-are-certain-bodies-being-censored-ig-uye-surana>>. Acesso em: 22 abr. 2022.

⁴³ O’HARA, Mary Emily, Op. Cit.; PARSONS, Vic, “Instagram accused of ‘sexualising, policing and censoring trans bodies’”. Disponível em: <<https://www.pinknews.co.uk/2021/04/22/instagram-trans-bodies-censorship-we-deserve-to-be-here/>>. Acesso em: 20 abr. 2022.

⁴⁴ SALTLY. *Unlike other big brands and media outlets, Salty is unable to advertise on @Instagram or @Facebook because the algorithms assume that Salty is “promoting escort services”*. 8 jul. 2019. Instagram: @salty.world. Disponível em: <<https://www.instagram.com/p/BzrYz9jF4Qt/>>. Acesso em: 23 fev. 2022.

⁴⁵ WATSON, Imogen, “Tackling LGBT+ blocklists: publishers and agencies rally against digital censorship”. *The Drum*, 05.03.2020. Disponível em: <<https://www.the-drum.com/news/2020/03/05/tackling-lgbt-blocklists-publishers-and-agencies-rally-against-digital-censorship>>. Acesso em: 20 abr. 2022.

⁴⁶ Anúncios LGBTQIA+ são alvo desproporcional de policiamento por parte de ferramentas automatizadas de moderação de conteúdo e acabam sendo injustamente silenciados com maior recorrência. A propósito, ver EUROPEAN DIGITAL RIGHTS, “How online ads discriminate: Unequal harms of online advertising in Europe”, 2021. Disponível em: <https://edri.org/wp-content/uploads/2021/06/EDRi_Discrimination_Online.pdf>. Acesso em: 3 fev. 2022.

Esse é apenas um dos casos [recorrentes] de impedimento à circulação ou redução de visibilidade de conteúdo que exponha corpos atravessados por múltiplos eixos de opressão social, ideias relacionadas a eles ou acompanhadas dessas imagens⁴⁷. Em uma rede social imagética como o Instagram, a crueldade desse silenciamento fica mais patente: afinal, a categorização e hierarquização social com base em *características físicas* é continuamente atualizada, a cada supressão de imagens que retratem corpos historicamente marginalizados. O estereótipo de que pessoas trans negras trabalham se prostituindo, a sexualização incondicional do corpo feminino, o descaso com as identidades de gênero em prol da categorização biológica dos corpos e a atribuição de caráter inapropriado aos corpos de mulheres gordas ou portadoras de deficiências físicas são revividos a cada censura de suas imagens que são classificadas como violação às Diretrizes da Comunidade, quando não passam de fotografias toleradas quando postadas por indivíduos socialmente mais privilegiados.

As Diretrizes da Comunidade, por serem a política que determina o que é tolerado ou não dentro do Instagram, guiam os mecanismos de *enforcement*, como a exclusão de conteúdo e o *shadowban*. Por esse motivo, sua redação carregada de valores hegemônicos tem peso sobre os resultados da moderação, contribuindo para que corpos e discursos dissidentes, não compreendidos em seus termos escritos, não sejam identificados pelas técnicas de moderação de conteúdo – ou melhor, sejam identificados como violações. Assim, grupos já marginalizados fora de ambientes virtuais, precisam enfrentar barreiras, também no virtual, que grupos mais privilegiados não enfrentam para conseguir expor seus corpos e ideias com a mesma liberdade e alcance.

A redação das Diretrizes da Comunidade importa porque tal conjunto de regras baseia a moderação humana e algorítmica de conteúdo. Por mais que normas estejam sempre sujeitas a interpretação humana, sua redação pode tender a uma maior ou menor segurança de significados. A consideração de realidades e corpos diversos na elaboração das Diretrizes da Comunidade poderia reduzir as situações

⁴⁷ O'HARA, Mary Emily, "Queer and Feminist Brands say they are Being Blocked from Running Ads on Instagram and Facebook". *MTV News*, 19.07.2019. Disponível em: <<https://www.mtv.com/news/3131929/queer-and-feminist-brands-say-they-are-being-blocked-from-running-ads-on-instagram-and-facebook/>>. Acesso em: 2 fev. 2022.

nas quais um corpo trans não é reconhecido como humano pela moderação algorítmica da rede social, ou em que são censuradas mulheres gordas que expõem seus corpos para criticar padrões de beleza hegemônicos e mamilos femininos expostos sem pretensão sexual. Isso, contudo, não é o que acontece.

1.2.

Exclusão automatizada de conteúdo

Vistas as discriminações imediatamente apreensíveis a partir da proibição aos mamilos femininos das Diretrizes da Comunidade, nesta seção me voltarei à exclusão automatizada de conteúdo para investigar quais discriminações são reveladas por seus efeitos. Na análise da exclusão automatizada de conteúdo do Instagram, pretendo evidenciar duas contradições percebidas a partir dos efeitos dos sistemas algorítmicos de moderação como um todo. Em primeiro lugar, o binômio hipervisibilidade-invisibilidade que permeia a experiência de determinados usuários no Instagram: com base em sua aparência física, eles são invisíveis em sua humanidade aos sistemas algorítmicos de moderação e, ao mesmo tempo, hipervisíveis como violações às Diretrizes da Comunidade, apesar da inocuidade de seu conteúdo. Em segundo lugar, argumento que o Instagram combate pesadamente a nudez – sobretudo a classificada como feminina – ao mesmo tempo em que não demonstra o mesmo empenho em coibir assédio sexual e violências entre usuários com base em gênero e raça. A partir dessas contradições e dos exemplos abordados, será possível investigar quais discriminações estão presentes nos efeitos da moderação automatizada da rede social, ainda que ausentes das Diretrizes da Comunidade.

A exclusão de conteúdo consiste em uma técnica de moderação aplicada a publicações que o Instagram considere diretamente violadoras das Diretrizes da Comunidade ou outras políticas internas, como os Termos de Uso. Prevista expressamente neste segundo documento, a exclusão de conteúdo em tese pode ser aplicada a postagens que, por exemplo, violem direitos autorais, exaltem o terrorismo, grupos de propagação de ódio, ou que apresentem “ameaças reais ou discurso de ódio e conteúdo direcionado a pessoas físicas cuja intenção seja degradá-las ou constrangê-las”⁴⁸.

⁴⁸ INSTAGRAM, “Community Guidelines”. Disponível em: <https://help.instagram.com/477434105621119/?helpref=hc_fnav>. Acesso em: 2 fev. 2022. Nos termos da versão

Antes de alterações promovidas na rede social em 2020, a decisão de exclusão de conteúdo era irreversível, e ao usuário só restava acatar a punição do Instagram. Atualmente, no entanto, se o usuário que tiver sua postagem removida entende que não violou as políticas da rede social, pode recorrer da decisão, solicitando que seja reanalisada juntamente com o conteúdo. Caso o Instagram julgue ter cometido um erro, a postagem excluída será publicada novamente; caso mantenha a sua decisão, o usuário poderá novamente apelar para o Comitê de Supervisão⁴⁹, o qual, criado em maio de 2020, analisa um número limitado de apelações.

A operacionalização da remoção de conteúdo no Instagram é feita tanto pela atuação humana quanto por inteligência artificial. Pessoas físicas, geralmente trabalhando em empresas terceirizadas localizadas em países periféricos⁵⁰ ou ganhando salários medíocres nos Estados Unidos⁵¹, são responsáveis por julgar, em segundos, as denúncias feitas por outros usuários a materiais considerados violadores das Diretrizes da Comunidade – o que, é claro, importa em muitos erros de classificação e inconsistências na exclusão de conteúdo⁵². Além disso, o Instagram também passou a empregar cada vez mais a inteligência artificial para ativamente identificar conteúdo considerado danoso. Dessa maneira, não seria necessário esperar um usuário se deparar com publicações chocantes ou ofensivas para somente

oficial em inglês do documento: “*Instagram is not a place to support or praise terrorism, organized crime, or hate groups. Offering sexual services, buying or selling firearms, alcohol, and tobacco products between private individuals, and buying or selling non-medical or pharmaceutical drugs are also not allowed. We also remove content that attempts to trade, co-ordinate the trade of, donate, gift, or ask for non-medical drugs, as well as content that either admits to personal use (unless in the recovery context) or coordinates or promotes the use of non-medical drugs. [...] We remove content that contains credible threats or hate speech, content that targets private individuals to degrade or shame them, personal information meant to blackmail or harass someone, and repeated unwanted messages.*”

⁴⁹ INSTAGRAM, “I don’t think Instagram should have taken down my post”. Disponível em: <https://www.facebook.com/help/instagram/280908123309761?helpref=uf_permalink>. Acesso em: 2 fev. 2022.

⁵⁰ PERRIGO, Billy, “Inside Facebook’s African Sweatshop”. *Time*, 17.02.2022. Disponível em: <<https://time.com/6147458/facebook-africa-content-moderation-employee-treatment/>>. Acesso em: 21 mai. 2022.

⁵¹ D’IGNAZIO, Catherine; KLEIN, Lauren F.. *Data feminism*. Cambridge, Massachusetts; London, England: The MIT Press, 2020, p. 182-183; JONES, Phil. *Work Without the Worker: Labour in the Age of Platform Capitalism*. Londres: Verso, 2021; ROBERTS, Sarah T. *Behind the Screen: Content Moderation in the Shadows of Social Media*. New Haven: Yale University Press, 2019.

⁵² As classificações iniciais desses moderadores humanos também podem ser usadas para alimentar os sistemas automatizados de exclusão de conteúdo. Esses sistemas comumente envolvem algoritmos de *hash*, que transformam um exemplo inicial (*input*) em uma sequência de dados (*output*) que servirá de base para identificar conteúdo semelhante a ser suprimido. Isso pode importar na replicação do mesmo “erro” de julgamento ou de vieses humanos em uma escala muito maior. Não se sabe se a moderação algorítmica do Instagram ainda se apoia sobre esse método e em qual proporção, comparativamente com as tecnologias de *machine learning*.

depois excluí-las; antes de alguém ter contato com o conteúdo indesejado, a inteligência de máquina o encontraria e removeria. Esse mecanismo automatizado, no entanto, vem recebendo críticas pelas censuras que opera.

A finalidade das remoções de conteúdo automatizadas é impedir a circulação de conteúdo danoso, para manter o respeito, a segurança e a boa convivência entre os usuários do Instagram. Assim, algumas das categorias priorizadas na remoção automática são as de discurso de ódio, bem como os conteúdos relacionados a autotutilação e nudez⁵³. Por mais que a empresa diga ter altas taxas de sucesso na identificação e combate a esse tipo de publicação, os relatos que vêm à tona na mídia⁵⁴ e em relatórios de pesquisa⁵⁵ revelam um outro lado da remoção de conteúdo: o dos erros e silenciamentos causados pelos sistemas algorítmicos empregados nessa técnica de moderação – ou melhor, pelos vieses humanos que neles estão integrados.

Em pesquisa conduzida pela *Salty Algorithmic Bias Collective* em colaboração com a Universidade de Michigan (2021)⁵⁶, cujo objetivo era entender o padrão de censura do Instagram às publicações de grupos marginalizados, as evidências coletadas indicam a existência de vieses, nos mecanismos automatizados de exclusão de conteúdo, contra pessoas racializadas, não binárias, transexuais, portadores de deficiências físicas, trabalhadoras e educadoras do sexo. A maioria desses grupos marginalizados havia experimentado censura pelo Instagram em percentuais maiores do que grupos socialmente mais privilegiados. Apesar de a rede social excluir as postagens de tais indivíduos sob a alegação de terem violado alguma regra das Diretrizes da Comunidade (geralmente aquelas relativas à nudez), em grande parte

⁵³ ZUCKERBERG, Mark, “A Blueprint for Content Governance and Enforcement”. *Facebook*, 2018 (última edição em 05.05.2021). Disponível em: <<https://www.facebook.com/notes/751449002072082/>>. Acesso em: 2 fev. 2022.

⁵⁴ LIM, Merlyna; ALRASHEED, Ghadah, “Beyond a technical bug: Biased algorithms and moderation are censoring activists on social media”. *The Conversation*. 16.05.2021. Disponível em: <<https://theconversation.com/beyond-a-technical-bug-biased-algorithms-and-moderation-are-censoring-activists-on-social-media-160669>>. Acesso em: 7 abr. 2022.

⁵⁵ SMITH, Shakira; HAIMSON, Oliver L; FITZSIMMONS, Claire; BROWN, Nikki Echarte (Salty Algorithmic Bias Collective in collaboration with Michigan University). “Censorship of Marginalized Communities on Instagram”, 27.09.2021. Disponível em: <<https://saltyworld.net/product/exclusive-report-censorship-of-marginalized-communities-on-instagram-2021-pdf-download/>>. Acesso em: 2 dez. 2021; EUROPEAN DIGITAL RIGHTS, “How online ads discriminate: Unequal harms of online advertising in Europe”, 2021. Disponível em: <https://edri.org/wp-content/uploads/2021/06/EDRi_Discrimination_Online.pdf>. Acesso em: 3 fev. 2022.

⁵⁶ Idem.

dos casos, as fotos e vídeos removidos sequer faziam jus ao motivo oferecido pela rede social⁵⁷.

Em linha com as reflexões levantadas na seção anterior, o combate ferrenho à nudez pelo Instagram, associado à atribuição de significados sexuais a corpos de mulheres, leva a que aqueles que divirjam da norma sejam os mais apagados pelos sistemas automatizados de moderação, que os enxergam como violações às Diretrizes da Comunidade. Além dessas discriminações, há ainda outras perpetuadas pelos sistemas algorítmicos da rede social.

Em um dos relatos coletados pela *Salty Algorithmic Bias Collective*, uma mulher hispânica e transexual dizia ter tido uma imagem sua removida apesar de ser do mesmo tipo publicado em páginas de anunciantes. Outra respondente identificada no grupo de portadores de deficiências físicas relatou que sua imagem de biquini, não diferente de qualquer outra postada na rede social por celebridades e influenciadoras de menor porte, havia sido removida por violar as diretrizes que vedam a “automutilação”, devido às suas cicatrizes. Na pesquisa, também era comum que transexuais e mulheres gordas relatassem terem tido fotografias suas removidas do Instagram por “nudez”, ainda que imagens semelhantes sejam corriqueiramente postadas por mulheres cis e magras e permitidas dentro da plataforma.

Se por um lado esses são corpos cuja humanidade é invisível à moderação – que os classifica erroneamente como violações às Diretrizes da Comunidade –, por outro, eles são corpos hipervisíveis, uma vez que são alvo das exclusões desproporcionais por conteúdo que, quando postado por outros usuários, é ignorado e deixado circular na plataforma: invisivelmente humanos e visivelmente violações.

Afinal esses corpos estão de fato transgredindo normas: as normas socialmente criadas que enrijecem as expressões sexuais reconhecidas como humanas, a constituição corporal considerada não patológica, as manifestações físicas de proporção hormonal tidas como normais e os traços fenotípicos vistos como superiores.

⁵⁷ Os resultados indicaram que as razões oferecidas pelo Instagram para a remoção do conteúdo variavam conforme o grupo marginalizado afetado. Em comparação com outros grupos marginalizados, pessoas portadoras de deficiências físicas eram mais propensas a serem censuradas por “auto-mutilação/auto-agressão” ou “violência”; pessoas trans e não binárias eram mais propensas a ter seu conteúdo removido em razão de “nudez”; usuários BIPOC (Black, Indigenous and People of Color) eram mais propensos a serem censurados em razão de “fake news/informação falsa”; e produtores de conteúdo acima do peso eram mais frequentemente acusados de violação às “diretrizes da comunidade”.

Essas normas estão por toda a parte na rede social: internalizadas pelos programadores dos sistemas algorítmicos de exclusão de conteúdo, espelhadas nas bases de dados, informações e classificações que treinaram os sistemas algorítmicos de reconhecimento de material proibido e no comportamento e mentalidade dos moderadores e dos usuários que classificam o conteúdo postado e produzem dados, alimentando os sistemas. O produto disso tudo é relatado pelos *efeitos discriminatórios da exclusão automatizada de conteúdo*: em sua invisibilidade humana, a hipervisibilidade da transgressão ganha contornos mais nítidos, chama a atenção dos sistemas algorítmicos e é, enfim, apagado.

É significativo que grande parte das postagens seja removida sob a alegação de terem infringido a proibição de nudez. Isso evidencia o caráter imagético do Instagram e a importância da corporalidade dentro da plataforma. Talvez por conta mesmo desse caráter o estudo da moderação discriminatória seja tão importante; afinal, a distribuição de possibilidades de discursos, comportamentos e visibilidade é diferenciada em razão da constituição corporal dos usuários que seja perceptível na rede social. Isto é, em razão de características físicas atreladas a determinada identidade social que sejam visíveis nas imagens publicadas. A visibilidade de um usuário é diretamente atrelada à sua aparência física, por mais que exista a tendência discursiva a separar a representação digital de um indivíduo de sua materialidade.

A existência digital dos usuários muitas vezes é tratada de forma desatrelada de suas identidades sociais no mundo *offline*, especialmente por discursos tecnoutópicos, que acreditam na neutralidade das novas tecnologias, em seu potencial de levar à sociedade a um novo patamar, mais elevado, de desenvolvimento e ignoram a história social da ciência da computação. Existe, contudo, uma corporalidade *online*. Como explicam Mariana Valente e Natália Neris, a Internet deixou de ser um espaço formado apenas por textos e passou a comportar a construção de “eus digitais”, atrelados a imagens, vídeos e informações que os tornam individualizáveis⁵⁸:

O avanço tecnológico fez com que em 20 anos a internet deixasse de significar, para a maior parte das pessoas, um conjunto de fóruns e chats em texto, para permitir a criação de verdadeiros *eus digitais*, individualizáveis e enriquecidos com abundância de imagens, vídeos e informações precisas sobre localização, gostos e atividades. A essa mudança podemos dar o nome de *corporificação*: se a experiência online, na década de 90, era em alguma medida dissociável da materialidade das pessoas que

⁵⁸ VALENTE, Mariana; NERIS, Natália, “Elas vão feminizar a Internet?”. *Revista SUR*, v. 15, n. 27, 2018, p. 108.

ali se comunicavam, o digital entrou posteriormente nos corpos e os corpos no digital, de forma a trazer questões sobre identidade e sexualidade para o centro da experiência nas redes, indissociável da experiência fora delas. E a discriminação e a violência cometidas contra os corpos das mulheres encontraram facilmente seus duplos no ambiente digital⁵⁹.

As violências *online* sofridas por mulheres são indicativas dessa corporalidade virtual atrelada a identidades sociais. Muitas jogadoras de vídeo game, por exemplo, passam a alterar o gênero de seus avatares em jogos *online* para evitar sofrer assédio de jogadores homens. Até mesmo em espaços cujo foco não é prioritariamente o compartilhamento de imagens dos usuários, como é o caso de fóruns – 4chan, Reddit etc. – a descoberta do gênero de uma usuária pode levar a assédio. É crescente a documentação sobre diferentes formas de violência praticada em ambientes virtuais com viés misógino⁶⁰. As identidades sociais são carregadas para a existência em plataformas digitais e quaisquer manifestações delas, sejam físicas (em imagens) ou verbais (por texto ou áudio) integram a corporalidade digital de um usuário.

Trata-se de uma corporalidade que se torna mais perceptível, no entanto, quando há imagens envolvidas, quando é possível saber rapidamente qual é a aparência física de um internauta e quais as identidades sociais a ela atreladas. O Instagram, por girar ao redor da autoexposição por imagens e vídeos evidencia como violências praticadas no âmbito digital muitas vezes estão atreladas a identidades sociais imediatamente apreensíveis a partir da *aparência física* do usuário. Mais do que isso, o caráter imagético do Instagram evidencia como o tratamento dispensado aos usuários pela rede social – por meio de técnicas de moderação de conteúdo, por exemplo – está atrelado a essas identidades sociais imediatamente apreensíveis.

O duplo padrão de avaliação do corpo identificado como feminino pela moderação da rede social, com alguns podendo ser exibidos e outros não, leva à conclusão de que o grau de escrutínio e risco de apagamento ao qual cada corpo está submetido varia conforme seu alinhamento maior ou menor às imagens e valores que informam os sistemas algorítmicos.

⁵⁹ Idem (itálicos no original).

⁶⁰ JANE, Emma. *Misogyny Online: A short (and brutish) history*. Los Angeles, London, New Delhi, Singapore, Washington DC, Melbourne: SAGE, 2017; THOMPSON, Laura, “I can be your tinder nightmare”: Harassment and misogyny in the online sexual marketplace”. *Feminism & Psychology*, v. 28, n. 1, 2018, pp. 69-89; COTE, Amanda, “I Can Defend Myself”: Women’s Strategies for Coping with Harassment While Gaming Online”. *Games and Culture*, v. 12, n. 2, 2017, pp. 136-155.

Além disso, importa que o conteúdo muitas vezes não é removido simplesmente pela parte visual. Apesar de o Instagram ser uma rede social imagética, extremamente focada na estética dos corpos, rostos e cenários que aparecem nas publicações, a exclusão automatizada de conteúdo também tem em mira publicações compostas apenas de discurso verbal. Na mesma pesquisa da *Salty Algorithmic Bias Collective*, um dos respondentes, por exemplo, relatou que suas publicações sobre direitos humanos recebiam significativamente menos visibilidade do que outros conteúdos que postasse no Instagram, vindo até mesmo a ser apagados. Em uma ocasião, após compartilhar na rede social infográficos e estatísticas sobre a população indígena nos Estados Unidos, reprovando a celebração do genocídio dos povos originários no Dia de Ação de Graças, o mesmo usuário teve todos os seus *Stories* automaticamente apagados sem qualquer explicação.

Como explicado por Zuckerberg, a inteligência artificial encarregada da exclusão de conteúdo como *enforcement* das Diretrizes da Comunidade é treinada para identificar postagens que contenham discurso de ódio, extremismo e terrorismo. Por mais que se argumente que situações como a descrita são erros do sistema, que identificou o ativismo indígena como discurso de ódio devido à escolha de palavras, fica a indagação sobre as bases de dados utilizadas para a alimentação desse sistema, a falta de consideração do contexto discursivo por sistemas automatizados⁶¹, os vieses carregados por seus criadores e os pontos cegos deixados por eles. A não consideração, na formulação dos sistemas automatizados, de discursos variados e voltados à afirmação da existência de povos culturalmente marginalizados leva a efeitos discriminatórios. Nenhuma explicação técnica defensava suscitando falhas dos modelos algorítmicos muda o fato de que seus *efeitos* são discriminatórios. Vale dizer, a “falha técnica” ou a incapacidade de sistemas algorítmicos preverem todas as complexidades da interação humana não torna os efeitos discriminatórios menos reprováveis nem explica seu direcionamento a determinados grupos.

Portanto, associar a moderação de conteúdo do Instagram à censura não é mera hipérbole retórica com a finalidade de construir um argumento político sobre

⁶¹ ELKIN-KOREN, Niva, “Contesting algorithms: Restoring the public interest in content filtering by artificial intelligence”. *Big Data & Society*, julho-dezembro 2020, pp. 1-13; GORWA, Robert; BINNS, Reuben; KATZENBACH, Christian, “Algorithmic content moderation: Technical and political challenges in the automation of platform governance”, p. 10.

certo grupo social. Existe uma razão material, verificável, para identificar determinadas exclusões de conteúdo do Instagram como censura: não só a redação das Diretrizes da Comunidade incorpora valores estéticos, raciais e de gênero específicos, mas os efeitos da aplicação de tais regras são consistentemente inconsistentes, afetando de forma diferenciada grupos sociais específicos. Ao mesmo tempo em que as exclusões de conteúdo deixam circular imagens de corpos femininos em roupas de banho e posições sensuais e excluem fotos semelhantes se o corpo for gordo ou transexual, repetem esse duplo padrão com recorrência.

Não há “falha” do sistema automatizado quando ele é construído, sob os imperativos das Diretrizes da Comunidade, para entregar esses resultados – isto é, para patrulhar minuciosamente os corpos não contidos dentro dos significados sociais que regem a moderação, corpos não contidos dentro da homogeneidade hegemônica da rede social. Afinal, tais sistemas são forjados por pessoas socialmente situadas – usualmente privilegiadas por sua raça –, usando bancos de dados que consideram uma noção estreita de “pessoa” ou ser humano, e treinados pelo comportamento e classificações dos moderadores ou dos usuários que, em massa, também privilegia determinado tipo de humanidade. Em razão disso, os grupos afetados pela exclusão não são aleatoriamente distribuídos⁶²; existe um padrão, indicativo dos vieses embutidos na plataforma.

Além disso, por mais que o Facebook/Instagram digam empregar, na moderação de seus espaços, inteligência artificial treinada para identificar e excluir postagens que contenham discurso de ódio e extremismo, o assédio sexual e ameaças de estupro e morte violentas recebidos por algumas usuárias, depois de denunciados, muitas vezes não resultam em qualquer medida protetiva ou punitiva. O discurso de ódio, o assédio e a misoginia *online* são tratados como problemas de menor potencial ofensivo e de caráter meramente privado e, após denúncias, tolerados pela moderação da plataforma⁶³. Se por um lado a nudez e a sexualidade femininas são

⁶² HAIMSON, Oliver, “Social Media Sites are Censoring Transgender, Black and Conservative Users – But for Very Different Reasons”. *Salty*. Disponível em: <<https://saltyworld.net/social-media-sites-are-censoring-transgender-black-and-conservative-users-but-for-very-different-reasons/>>. Acesso em: 7 mar. 2022.

⁶³ SWIFT, Skylar, “I Modeled Menstrual Products as a Trans Man and Instagram Let Neonazis Harass Me”. *Salty*. Disponível em: <<https://www.saltyworld.net/i-modeled-menstrual-products-as-a-trans-man-and-instagram-let-neonazis-harass-me/>>. Acesso em: 23 dez. 2021; CHRISTIE, Lacey-Jade, “Instagram censored one of these photos but not the other. We must ask why”. *The Guardian*, 19.10.2020. Disponível em: <<https://www.theguardian.com/technology/2020/oct/20/instagram-censored-one-of-these-photos-but-not-the-other-we-must-ask-why>>. Acesso em: 17 nov. 2021.

pesadamente combatidas e rapidamente censuradas para não “ofender” a comunidade “diversa” da rede social, as ameaças gráficas de violências extremas contra o corpo feminino são vistas como incapazes de gerar dano real, como meros insultos de ordem privada, e, portanto, toleradas.

A atitude da rede social revela um reflexo do patriarcalismo, que não vê a violência contra o corpo feminino como violência. Ou melhor, que enxerga essa violência como “*socialmente legitimada*”, nas palavras de Viviana Ribeiro⁶⁴. Um dos efeitos desse olhar é que o assédio e outras violências misóginas praticados *online* contra as mulheres – assim como as violências praticadas no *offline* – não chocam, não são vistos como violências⁶⁵, mas como riscos inerentes ao uso de redes sociais, que devem ser tolerados caso se opte por utilizar esses espaços. Espelho de uma dinâmica já conhecida: a uma mulher assediada ou violentada não se pergunta nada sobre a agressão e o agressor, mas sim sobre a agredida. Quais roupas estava usando, por onde estava andando, como se comportou? Em resumo, o que ela fez para justificar a violência sofrida? O risco e a responsabilidade pela agressão são da vítima, não de uma relação política que as coloca como inferiores e merecedoras de violência.

Por mais que a Meta Inc. tenha recentemente se proposto a desenvolver novos mecanismos de proteção aos usuários que sofrem assédio na rede social⁶⁶, muitos deles consistem em deslocar a responsabilidade pela experiência da violência *online* para as usuárias. Recursos como indexação de palavras indesejadas que, se identificadas, engatilham o ocultamento da mensagem ou pastas ocultas para as quais

⁶⁴ RIBEIRO, Viviana Mendes, “O controle da reprodução está no coração da organização sócio-política. A discussão sobre a descriminalização da prática da interrupção da gravidez voluntária no Brasil. *Revista da Defensoria Pública da União*, n. 15, janeiro-junho 2021, p. 234.

⁶⁵ Como chama atenção, Antonio Pele, o emprego da tecnologia faz surgir novas formas de circulação da violência – que não deveriam ser ignoradas como tais (PELE, Antonio, “Data Necropolitics”. *Abolition and Democracy*, v. 13, n. 13. Disponível em: <<https://blogs.law.columbia.edu/abolition1313/antonio-pele-data-necropolitics-2/>>. Acesso em: 20 mai. 2022).

⁶⁶ Em publicações compartilhadas em 2021, o Instagram anunciou que estavam sendo testadas novas ferramentas de proteção a vítimas de assédio e ameaças na rede social – como a possibilidade de os usuários habilitarem uma ferramenta, chamada “Limits”, que automaticamente esconde comentários e requisições de mensagens privadas de pessoas que eles não seguem ou que começaram a segui-los recentemente, e o recurso chamado “Hidden Words”, que permite que os usuários filtrem automaticamente palavras, frases e emojis ofensivos que são mantidos em uma pasta escondida que “você nunca precisa abrir se não quiser”. (INSTAGRAM, “An update on our work to tackle abuse on Instagram”. *Instagram*, 11.02.2021. Disponível em: <<https://about.instagram.com/blog/announcements/an-update-on-our-work-to-tackle-abuse-on-instagram>>. Acesso em: 3 fev. 2022; MOSSERI, Adam, “Introducing New Ways to Protect Our Community from Abuse”. *Instagram*, 10.08.2021. Disponível em: <<https://about.instagram.com/blog/announcements/introducing-new-ways-to-protect-our-community-from-abuse>>. Acesso em: 3 fev. 2022).

possam ser automaticamente encaminhados os recados abusivos ainda não reconhecem a gravidade do abuso cometido *online*, sobretudo de viés misógino, nem assumem responsabilidade por lidar com culturas sexistas, racistas e homofóbicas cujas manifestações são facilitadas pela arquitetura da rede social⁶⁷⁻⁶⁸. Tais agressões não são vistas como o discurso de ódio que a empresa diz combater nas Diretrizes da Comunidade.

Se por um lado, a Meta Inc. reconhece o interesse público em ocultar mamilos e nudez feminina, que poderiam ofender os usuários da plataforma, por outro trata violências interpessoais com base em gênero como agressões exclusivamente privadas. Se por um lado, trata a nudez com gravidade, reconhecendo seu potencial danoso de ofender e afastar os usuários da plataforma, por outro, não consegue entender o mesmo raciocínio ao discurso especificamente misógino, reconhecendo os danos reais que podem advir dessa prática, ainda que se proliferem os relatórios de pesquisa⁶⁹ descrevendo os impactos sobre a saúde física e mental, a sociabilidade e o exercício do trabalho produtivo pelas mulheres que são vítimas de ataques virtuais.

Em tese, as categorias priorizadas pela exclusão automática de conteúdo são a nudez e o discurso de ódio, mas, enquanto a moderação da rede social protege com garras as diretrizes que vedam a nudez e os mamilos femininos, falha em combater ou reconhecer sua responsabilidade no combate às violências virtuais sofridas

⁶⁷ DUGUAY, Stefanie; BURGESS, Jean; SUZOR, Nicolas, “Queer women’s experiences of patchwork platform governance on Tinder, Instagram, and Vine”. *Convergence: The International Journal of Research into New Media Technologies*, v. 26, n. 2, 2020, pp. 237-252.

⁶⁸ Essa forma de lidar com a misoginia e o assédio online reflete a crença em que violências contra a mulher são meramente interpessoais, e devem ser tratadas no âmbito privado. Internet Lab e IT for Change, em estudo sobre violências online contra as mulheres, alertam para a necessidade de tratar a misoginia em espaços públicos online como uma ofensa pública, não uma violação privada (IT FOR CHANGE, Índia; INTERNETLAB, Brazil, “General Assembly Report on Gender Justice: Submission to the Special Rapporteur on the right to freedom of opinion and expression”, junho 2021. Disponível em: <<https://itforchange.net/sites/default/files/1738/ITforChangeInternetLab-SubmissionToGARReportGenderJustice-June2021.pdf>>. Acesso em: 20 mar. 2022).

⁶⁹ CODING RIGHTS; INTERNETLAB. Violências de gênero na internet: diagnóstico, soluções e desafios. Contribuição conjunta do Brasil para a relatora especial da ONU sobre violência contra a mulher. São Paulo, 2017; PLAN INTERNATIONAL. “Free to be online? Girls’ and young women’s experiences of online harassment”, 2020. Disponível em: <<https://plan-international.org/publications/freetobeonline>>. Acesso em: 9 nov. 2020; UN WOMEN. “Online and ICT facilitated violence Against women and girls during COVID-19”, 2020. Disponível em: <<https://www.unwomen.org/en/digital-library/publications/2020/04/brief-online-and-ict-facilitated-violence-against-women-and-girls-during-covid-19>>. Acesso em: 26 set. 2021; WORLD WIDE WEB FOUNDATION. “Women’s Rights Online: closing the digital gender gap for a more equal world”, 2020. Disponível em: <<https://webfoundation.org/research/womens-rights-online-2020/>>. Acesso em: 24 set. 2021.

por usuárias mulheres. As decisões negociais do Instagram tomadas no âmbito da moderação automatizada de conteúdo têm efeitos sobre a conduta dos usuários. Afinal, eles estão em contínua relação com as técnicas de distribuição e supressão de conteúdo, adaptado suas condutas, ou tendo elas adaptadas conforme a dinâmica da rede social. Na seção a seguir, por meio do estudo do *shadowban*, serão levantadas algumas reflexões a propósito da relacionalidade entre técnicas de moderação automatizada de conteúdo e usuários, atentando para a relação de poder instaurada entre os dois lados.

1.3.

Shadowban

O terceiro e último elemento da moderação algorítmica do Instagram a ser analisado é a prática do *shadowbanning*, ou, em termos sucintos, a restrição de visibilidade de determinado usuário após ter postado conteúdo ofensivo ou delicado. Com a análise do seu funcionamento e efeitos, pretendo evidenciar como técnicas algorítmicas de moderação de conteúdo, constituindo punições a comportamentos indesejados dos usuários, podem ser vistas como uma nova forma de poder – de traços discriminatórios – que atua no entorno digital disciplinando condutas. Mais ainda, trata-se de um poder que, sob discursos corporativistas e tecnoutópicos⁷⁰, se nega como poder e diminui as críticas recebidas sob o argumento de serem atécnicos.

O *shadowban* consiste na redução substancial da visibilidade de determinado perfil para diminuir o alcance de conteúdo identificado como inapropriado, apesar de não violador das regras internas da rede social⁷¹. O *shadowban* teve sua existência negada múltiplas vezes pelo Instagram e pelo Facebook, mas os usuários da plataforma relatam eventos muito semelhantes que indicam o contrário. Devido tanto ao sigilo dos códigos e informações usados na moderação de conteúdo, quanto às diferenças dos relatos sobre os efeitos do *shadowban*, não existe um conceito

⁷⁰ Para a definição de discurso tecnoutópico, v. página 33 deste trabalho.

⁷¹ MYERS-WEST, Sarah, “Censored, suspended, shadowbanned: User interpretations of content moderation on social media platforms”. *New Media & Society*, v. 20., n. 11, 2018, pp. 4366-4383; PAASONEN, Susanna; ARE, Carolina, “Sex in the shadows of celebrity”. *Porn Studies*, v. 8, n. 4, setembro 2021, pp. 411-419; BRIDGES, Lauren, “Digital failure: Unbecoming the ‘good’ data subject through entropic, fugitive, and queer data”. *Big Data & Society*, janeiro-junho 2021, pp. 1-17.

unificado dessa prática, mas os trabalhos acadêmicos sobre o tema coincidem no efeito principal do *shadowban*: a restrição substancial da visibilidade de um perfil. A análise a seguir tem como base tanto esses trabalhos e publicações do Instagram em seu blog corporativo quanto uma manifestação de Mark Zuckerberg, postada no Facebook em 2018 e atualizada pela última vez em maio de 2021, em que ele explica a adoção de um novo sistema automatizado de restrição de visibilidade de perfis e os seus propósitos.

Com o *shadowban* de uma conta no Instagram, o perfil do usuário afetado se mantém plenamente operacional, mas não aparece nas recomendações algorítmicas, nos resultados de busca, nas sugestões “*autofill*” de pesquisa e seu conteúdo tem sua distribuição substancialmente reduzida aos demais participantes da rede social, até mesmo a seguidores do perfil. Tudo isso leva a que as métricas de engajamento do respectivo perfil caiam drástica e repentinamente, como visualizações, curtidas, compartilhamentos, reações e novos seguidores. Muitos relatos indicam que o usuário silenciado pelo *shadowban* não é informado sobre a aplicação dessa medida, e o percebe apenas quando o engajamento de outros usuários com o seu conteúdo cai sem qualquer razão aparente.

O motivo da aplicação do *shadowban* não é a violação frontal às Diretrizes da Comunidade ou aos Termos de Uso do Instagram por um usuário, mas sim a publicação de conteúdo que é considerado limítrofe ao que é proibido dentro da rede social. A título de ilustração, tornaram-se famosos casos de dançarinas de *pole dance* que, a partir de 2019, passaram a ter sua visibilidade substancialmente reduzida dentro do Instagram por exporem seus corpos em vídeos de performances de dança⁷². Não havia nudez no conteúdo – portanto, não havia a violação das regras proibitivas de nudez feminina –, mas as praticantes dessa modalidade de dança comumente usam poucas roupas, pois a pele é usada para ter firmeza no mastro. Por isso, a moderação considerava que seu conteúdo beirava à nudez e passou a restringir a visibilidade de contas de *pole dancers* e de *hashtags* relacionadas à dança.

Em postagem de 2018 no Facebook, Mark Zuckerberg explica as motivações por trás do mecanismo do *shadowban* – sem, contudo, utilizar essa nomenclatura,

⁷² ARE, Carolina. “The Shadowban Cycle: an autoethnography of pole dancing, nudity and censorship on Instagram”. *Feminist Media Studies*, maio 2021, pp. 9-10.

que é rechaçada pelo Facebook e pelo Instagram⁷³. Segundo o empresário, pesquisas realizadas dentro da companhia indicavam que os usuários tendem a engajar mais com conteúdo considerado limítrofe dentro das Diretrizes da Comunidade, isto é, conteúdo que, apesar de não violar frontalmente suas regras, é dotado de caráter “sensacionalista” ou “provocativo”, beirando ao tipo de material considerado proibido pelas políticas internas. Trata-se de um conteúdo inapropriado, mas não vedado. Isso ocorre com qualquer categoria, desde notícias até publicações de fotografias que beiram à nudez e postagens políticas⁷⁴.

Diante dessa tendência localizada pelo Facebook, tida como um problema a ser sanado para manter a ordem e evitar publicações ofensivas na rede social, uma das soluções encontradas foi a de reduzir a distribuição do conteúdo considerado “sensacionalista ou provocativo”, reduzindo o engajamento de outros usuários, para *desincentivar* a sua criação. Portanto, um tipo de penalidade àqueles que fazem publicações arriscadas. Zuckerberg exemplifica:

“fotos próximas ao limite de nudez, como [aquelas] com roupas reveladoras ou posições sexualmente sugestivas, obtiveram em média mais engajamento antes de mudarmos a curva de distribuição para desencorajar isto. O mesmo se aplica às postagens que não se enquadram na nossa definição de discurso de ódio, mas que continuam a ser ofensivos”.

Assim, além de utilizar a inteligência artificial para ativamente encontrar conteúdo ofensivo e violador das Diretrizes da Comunidade e removê-lo, a companhia passou a utilizá-la também para identificar conteúdo considerado inapropriado e reduzir sua distribuição e visibilidade dentro da rede social. Este, afinal, é o *shadowban*.

A partir do objetivo declarado por Zuckerberg, é possível perceber o *shadowban*, assim como a exclusão de conteúdo e outras técnicas de moderação de conteúdo, como *punições* aos usuários que violem as normas expressas ou tácitas

⁷³ O Instagram já mencionou o “*shadowban*” algumas vezes em pronunciamentos públicos, mas sempre dando a entender que é um fenômeno não existente, mera impressão dos usuários. Ao mesmo tempo, no entanto, a empresa já admitiu obliquamente haver a distribuição diferenciada de conteúdo dentro da plataforma. A propósito, ver ARE, Carolina, “Instagram Quietly Admitted Algorithm Bias, But How Will It Fight It”. *Blogger on Pole*, 25.06.2020. Disponível em: <<https://bloggeronpole.com/2020/06/instagram-quietly-admitted-algorithm-bias-but-how-will-it-fight-it/>>. Acesso em: 23 fev. 2022.

⁷⁴ ZUCKERBERG, Mark, “A Blueprint for Content Governance and Enforcement”. *Facebook*, 2018 (última edição em 05.05.2021). Disponível em: <<https://www.facebook.com/notes/751449002072082/>>. Acesso em: 2 fev. 2022.

da plataforma. Para além das outras técnicas de moderação, o diferencial do *shadowban* está em sua furtividade. Esse mecanismo evidenciou uma forma de disciplinamento furtiva, que se vale da voz – ou melhor, da produção de conteúdo, imágica ou verbal – do próprio usuário para agir sobre ele. Trata-se de um disciplinamento que não age diretamente sobre as condutas dos usuários, mas sobre o seu entorno, o ambiente digital em que ele atua. Em vez de castigo – como poderia ser uma exclusão de perfil – há a restrição de visibilidade, a manipulação ambiental para reduzir engajamento e interpelação dentro da plataforma.

A finalidade declarada do mecanismo automático de redução de visibilidade do conteúdo limítrofe é reduzir o sensacionalismo dentro da rede social para criar “um discurso mais saudável e menos polarizado do qual mais pessoas se sintam seguras participando”⁷⁵. Apesar do tom neutro, estratégico e corporativista da fala de Zuckerberg, os efeitos do *shadowbanning* revelam os vieses por trás da sua concepção.

Ao longo dos últimos anos, multiplicaram-se denúncias na mídia e em publicações em perfis do próprio Instagram relatando casos de influenciadores e influenciadoras negros⁷⁶, profissionais do sexo, mulheres gordas⁷⁷, e pessoas LGBTQIA+⁷⁸ que sofreram *shadowban* ou exclusão de postagens após terem publicado algum conteúdo que, apesar de interpretado pela moderação como impróprio, não destoava do que é aceito e amplamente veiculado quando postado por usuários cujo corpo é conformado a padrões estéticos, raciais e de gênero dominantes. Assim

⁷⁵ Idem.

⁷⁶ MORALES, Eduardo. “How Instagram Upholds the Censoring of Black Bodies and the Suppression of POC-Run Accounts”. *An Injustice*, 03.08.2020. Disponível em: <<https://aninjustice-mag.com/how-instagram-upholds-the-censoring-of-black-bodies-and-the-suppression-of-poc-run-accounts-cd1d33ff4dc5>>. Acesso em: 17 nov. 2021; SMITH, Adam. “Instagram Boss Says It Will Change Algorithm To Stop Mistreatment Of Black Users, Alongside Other Updates”. *Independent*, 16.06.2020. Disponível em: <<https://www.independent.co.uk/life-style/gadgets-and-tech/news/instagram-black-lives-matter-racism-harassment-bias-algorithm-a9567946.html>>. Acesso em: 17 nov. 2021; EL-WARDANY, Salma. “Like Our Society, Instagram In Biased Against Women of Colour”. *Refinery29*, 10.12.2020. Disponível em: <<https://www.refinery29.com/en-gb/2020/12/10150275/shadow-ban-instagram-censorship-women-of-colour>>. Acesso em: 17 nov. 2021.

⁷⁷ REID, Rebecca, “Plus-size model says she’s been shadow banned by Instagram because of her body”. *Metro*, 11.06.2018. Disponível em: <<https://metro.co.uk/2018/06/11/plus-size-model-says-shadow-banned-instagram-body-7621367/>>. Acesso em: 17 nov. 2021; GRADY, Lora, “Women are Calling Out Instagram for Censoring Photos of Fat Bodies”. *Flare*, 18.10.2018. Disponível em: <<https://www.flare.com/news/instagram-censorship-fat-plus-size/>>. Acesso em: 17 nov. 2021.

⁷⁸ JOSEPH, Chanté, “Instagram’s murky shadow bans just serve to censor marginalized communities”. *The Guardian*, 08.11.2019. Disponível em: <<https://www.theguardian.com/commentis-free/2019/nov/08/instagram-shadow-bans-marginalised-communities-queer-plus-sized-bodies-sexually-suggestive>>. Acesso em: 17 nov. 2021.

como a exclusão de conteúdo, o *shadowban* afeta desproporcionalmente grupos sociais historicamente marginalizados⁷⁹.

A inconsistência da moderação do Instagram não é aleatória, mas sim indicativa dos valores estéticos e comportamentais que determinam quais seres humanos são legíveis para a rede social. Os sistemas algorítmicos de moderação da plataforma reforçam uma hierarquia de humanidades já conhecida fora dos ambientes digitais: são humanos e plenamente livres para circular, falar e serem visíveis aqueles corpos que se enquadrem nos ideais hegemônicos de beleza feminina, na matriz binária de inteligibilidade de gêneros e na branquitude. Por outro lado, os corpos dissidentes dessas verdadeiras diretrizes da comunidade, gordos, não brancos, com pelos e estilizações corporais destoantes do estritamente feminino ou masculino, não conseguem ser lidos por esses sistemas algorítmicos, ou melhor, são lidos como desvios do padrão de humanidade da moderação, devendo ser ocultados para não ofenderem a comunidade nem perturbarem a ordem da rede social. Esse é o tipo de discurso “polarizado” e “inseguro” que a rede social pretende evitar.

A propósito dos vieses estéticos da moderação pelo *shadowban*, Carolina Are nota que há um tratamento diferenciado para a nudez e sexualidade femininas. Enquanto existe um amplo esforço para despir as interações sociais na rede social de sexo e nudez, o conteúdo sexual postado por celebridades e determinadas influenciadoras digitais é facilmente tolerado e, mais do que isso, promovido. Há determinados corpos cuja nudez é *valiosa* para o Instagram, porque atrai anunciantes, usuários e mais geração de dados. Se por um lado as Diretrizes da Comunidade dizem não permitir nudez feminina para evitar ofender os usuários da rede e constantemente derruba ou oculta imagens de mulheres gordas de biquini, por exemplo, sob a justificativa de violarem a vedação à nudez, por outro, o Instagram promove e

⁷⁹ A expressão “shadowban” ficou conhecida depois de passar a ser usada, em 2016, pelo site de extrema-direita “Breitbart” em diversas matérias que acusavam o Twitter e o Facebook de esconderem as postagens de conservadores, sem o seu conhecimento, para limitar sua influência. Apesar disso, o termo atualmente é majoritariamente associado ao silenciamento desproporcional de grupos sociais historicamente oprimidos, que têm sua visibilidade restringida pela moderação de determinadas redes sociais sob o argumento de terem violado as regras da plataforma. (RADSCH, Courtney, “Shadowban/Shadow Banning”. In: BELLI, Luca; ZINGALES, Nicolo; CURZI, Yasmine (Ed.). *Glossary of Platform Law and Policy Terms*. Rio de Janeiro: FGV Direito Rio, 2021, p. 296; ARE, Carolina. “The Shadowban Cycle: an autoethnography of pole dancing, nudity and censorship on Instagram”, p. 1-2)

recomenda largamente o conteúdo sexual postado por celebridades como Kim Kardashian e outras influenciadoras digitais de menor porte que se alinham a determinados valores estéticos⁸⁰.

Quanto aos vieses raciais da curadoria e moderação algorítmicas do Instagram como um todo (incluindo o *shadowban*), sua problematização não é inédita na prática de blogueiras brasileiras. No final de 2020, influenciadoras negras do Brasil realizaram experimentos com os algoritmos empregados na curadoria e recomendação de conteúdo no Instagram a fim de testar seus vieses de raça. Luana Carvalho⁸¹, Sá Ollebar e Triscilla Oliveira⁸² passaram a postar imagens de mulheres brancas para verificar eventuais alterações na amplitude de distribuição do conteúdo e de engajamento de outros usuários, verificando que tais publicações eram distribuídas a usuários da plataforma muito mais amplamente. A influenciadora digital Sá Ollebar, por exemplo, durante uma semana, publicou fotos de mulheres brancas promovendo o mesmo conteúdo de autocuidado que normalmente compartilha na plataforma. O alcance de suas publicações aumentou em 6.000% e os seguidores relataram que, apesar de não receberem suas postagens há meses, o Instagram havia circulado para eles todas as publicações da blogueira em que figuravam mulheres brancas.

Como destaca a cientista de dados Nina da Hora⁸³, dois problemas identificados em tais experimentos são que: (i) eles não são dotados de rigor científico, e, portanto, as conclusões extraídas de seus resultados não são suscetíveis de generalização, e (ii) os algoritmos testados pelas blogueiras são desconhecidos porque mantidos privados pelas companhias de tecnologia, e até que seu funcionamento seja entendido pelos usuários por meio de testes, já terão sido alterados e readaptados a novas complexidades das interações digitais; assim, o conhecimento sobre eles nunca é preciso.

Por mais que tais problemas sejam relevantes e dignos de consideração, os experimentos realizados por blogueiras servem como metodologia para chamar

⁸⁰ HORNE, Julian Van, “Shadowbanning is a Thing – and It’s Hurting Trans and Disabled Advocates”. *Salty*. Disponível em: <<https://saltyworld.net/shadowbanning-is-a-thing-and-its-hurting-trans-and-disabled-advocates/>>. Acesso em: 7 mar. 2022.

⁸¹ CARVALHO, Luana. *Último teste*. 11 out. 2020. Instagram: @lxccarvalho. Disponível em: <https://www.instagram.com/p/CGOB_2DHHMO/>. Acesso em: 2 fev. 2022.

⁸² OLIVEIRA, Triscilla. *Vamos ver...* Niterói. 12 out. 2020. Instagram: @soulanja. Disponível em: <<https://www.instagram.com/p/CGOk-xiBx1e/>>. Acesso em: 8 fev. 2022.

⁸³ HORA, Nina da. “Aula Aberta: Racismo e Algoritmos com Nina da Hora”. Disponível em: <<https://www.ninadahora.dev/>> (8:01 – 9:00). Acesso em: 3 fev. 2022.

atenção para determinado problema político decorrente de escolhas corporativas das empresas de tecnologia. A falta de rigor metodológico e a rápida obsolescência dos testes não invalida sua conclusão: a de que existem vieses raciais no funcionamento da curadoria e moderação de conteúdo do Instagram por aqueles que estão à frente da construção dessas tecnologias – em sua maioria homens brancos, cis e heterossexuais. O conhecimento técnico sobre os algoritmos empregados pela rede social nessas tarefas não anula o que seus *efeitos* revelam e o que é comunicado pelas blogueiras negras brasileiras: tais algoritmos foram pensados a partir da branquitude patriarcal⁸⁴. O componente humano da criação de tecnologias, invisibilizado por sua falsa neutralidade e pela adesão ao rigor científico, é suscitado e tornado visível por experimentos de blogueiras tais como os mencionados.

Mais do que isso, tais testes, apesar de não dotados de rigor científico, são uma forma de afirmar, dentro da rede social, que se *sabe* do poder exercido pelos sistemas algorítmicos de moderação sobre os indivíduos. Por mais que o Instagram use argumentos técnicos para se defender de críticas aos danos sociais causados por seus sistemas automatizados, seu poder é percebido por alguns indivíduos, principalmente aqueles que se deparam com as inconsistências da moderação de conteúdo. Trata-se de um poder silencioso, disfarçado sob a máscara de plataforma voltada ao entretenimento, sutil em sua forma de operação pelos dados e defendido pela empresa a cada manifestação pública que diminui as críticas recebidas. Um poder que se nega como poder.

Por meio da seleção de quais imagens devem ser ocultadas para manter um ambiente “saudável” e “seguro”, a rede social direciona comportamentos e molda sensibilidades. Por mais que diga que seus algoritmos não são racistas ou sexistas, os efeitos discriminatórios desses sistemas são sentidos e relatados pelas usuárias. Desse modo, elas afirmam sua capacidade de produzir conhecimento sobre a realidade e sobre o funcionamento da rede social, com base em suas experiências com

⁸⁴ Tarcizio Silva elaborou uma linha do tempo do racismo algorítmico em que se encontram exemplos do impacto que o racismo de desenvolvedores de tecnologia tem sobre os produtos (SILVA, Tarcizio, “Linha do Tempo do Racismo Algorítmico: casos, dados e reações”, 2019. Disponível em: <<https://tarciziosilva.com.br/blog/destaques/posts/racismo-algoritmico-linha-do-tempo/>>. Acesso em: 22 fev. 2022).

as tecnologias digitais⁸⁵. Seu não alinhamento ao rigor científico não lhe retira o caráter de saber.

Algoritmos, processos e técnicas de moderação são alterados a todo momento, sendo possível que as descrições neste trabalho estejam defasadas já à época de sua apresentação, devido ao tempo decorrido desde a coleta de informações e a escrita. De todo modo, ainda que as técnicas de moderação de conteúdo tenham sido ajustadas pelo Instagram nos últimos meses, em decorrência das críticas que chegam à mídia, os fatos aqui descritos são ilustrativos de problemas maiores, que não são sanados pela mera adoção de comitês na estrutura societária do Instagram ou ajustes pontuais na forma de funcionamento de algoritmos: a desproporcional fiscalização dos corpos femininos sob um imperativo de combate à nudez e sexualidade (uma vez que são os seus corpos os sexualizados); o componente humano e os valores hegemônicos que estão por trás da inteligência artificial empregada nas redes sociais; a construção de um ideal homogêneo de usuário para a rede social; e o disciplinamento dos comportamentos dos usuários por técnicas de moderação de conteúdo. Esses aspectos serão abordados de forma mais aprofundada nos Capítulos 2 e 3 a seguir.

⁸⁵ PAIVA, Raquel, “Epistemologia compreensiva: garantia para pluralidade de vozes”. In: PRATA, Nair; PESSOA, Sônia Caldas; e ANDRADE, Ivanise Hilbig (Org.). *Um mundo e muitas vozes: da utopia à distopia?*. São Paulo: Intercom, 2021, p. 107.

2

Os valores subjacentes à moderação algorítmica

No Capítulo 1, foi possível perceber padrões discriminatórios em razão de raça e gênero revelados pelos efeitos de técnicas automatizadas de moderação de conteúdo. Esses padrões não são acidentais, nem meras *falhas* dos sistemas algorítmicos. Antes, indicam alguns dos valores que regem a construção desses sistemas no âmbito do Instagram – sejam objetivos comerciais perseguidos pela empresa, sejam classificações sociais transmitidas aos sistemas pelos humanos responsáveis pela programação dos sistemas.

Nesta seção, usarei as noções de “capitalismo de vigilância”⁸⁶ e “colonialidade automatizada”⁸⁷ de forma complementar com o objetivo de iluminar três aspectos da moderação de conteúdo do Instagram que têm influência sobre a distribuição de visibilidades: a lógica do lucro, a hierarquização de humanidades com base em raça e gênero e a adesão inflexível a uma epistemologia tecnocientífica centrada nos dados (que chamarei de datacêntrica), que descredibiliza as críticas à moderação algorítmica.

O conceito de capitalismo de vigilância explica o modelo negocial adotado pela Meta Platforms, Inc. (antigo Facebook) e sua controlada, o Instagram: a acumulação e mineração de quantidades massivas de dados comportamentais para o direcionamento publicitário e geração de lucros. É essa lógica do lucro que rege as escolhas do que permanece amplamente visível no Instagram. Já o conceito de colonialidade automatizada evidencia a lógica de hierarquização de humanidades com base em raça e gênero que é reproduzida e intensificada pela moderação algorítmica da rede social. Além disso, esse conceito ajuda a entender a adoção, por parte do Instagram, de uma epistemologia centrada na capacidade dos dados descreverem objetivamente a realidade e possibilitarem a otimização constante da rede social.

⁸⁶ ZUBOFF, Shoshana. *A Era do Capitalismo de Vigilância: A luta por um futuro humano na nova fronteira do poder*. Rio de Janeiro: Intrínseca, 2020.

⁸⁷ SILVA, Tarcizio, “Colonialidade automatizada: algoritmos e aprendizado de máquina nas plataformas de mídias sociais”. In: PRATA, Nair; PESSOA, Sônia Caldas; ANDRADE, Ivanize Hilbig de (Org.). *Um mundo e muitas vozes: da utopia à distopia?*. São Paulo: Intercom, 2021, pp. 139-155.

Em linha com a nomenclatura de Paola Ricaurte⁸⁸, chamo essa epistemologia de datacêntrica, a qual é tida pelo Instagram como a única capaz de entender de forma verdadeira o funcionamento da moderação; sob essa perspectiva, as experiências e as críticas dos usuários aos efeitos discriminatórios da moderação de conteúdo são desprezadas, tidas como incapazes de ter uma apreensão real e verdadeira da plataforma.

2.1.

Capitalismo de vigilância e a lógica do lucro

Nesta Seção, pretendo me valer do conceito de *capitalismo de vigilância* cunhado por Shoshana Zuboff para pensar de que modo o modelo de negócios do Instagram favorece o tratamento discriminatório dos seus usuários. Expandindo o conceito da autora, pretendo pensar como o compartilhamento da intimidade e a autoexposição espetacular – centros de gravidade dessa rede social – têm impacto sobre o modelo de negócios da plataforma. Além disso, argumento que o conceito de capitalismo de vigilância, no âmbito do Instagram, deve ser complementado para abranger não só um poder de vigilância que garante a capitalização sobre os dados alheios, como também um poder punitivo. Esse poder punitivo, consubstanciado nas técnicas de moderação de conteúdo, também almeja assegurar a continuidade da produção de dados pelos usuários – e, conseqüentemente, geração de lucros.

Shoshana Zuboff, estudando a nova lógica de acumulação que se tornou hegemônica no capitalismo digital, localiza na história do Google o seu nascimento. Trata-se do que ela chama de “capitalismo de vigilância”, uma lógica de superacumulação de dados de usuários com a finalidade de, por meio da inteligência artificial, gerar previsões de comportamentos e vendê-las a outros agentes do mercado⁸⁹. Esse modelo, gestado no Google, posteriormente veio a ser adotado por redes sociais como o Facebook (atualmente, Meta Platforms Inc.) – e, conseqüentemente, pelo Instagram, sua controlada⁹⁰.

⁸⁸ RICAURTE, Paola. “Data Epistemologies, The Coloniality of Power and Resistance”. *Television & New Media*, v. 20, n. 4, 2019, pp. 350-365.

⁸⁹ ZUBOFF, Shoshana. *A Era do Capitalismo de Vigilância*, pp. 81-86.

⁹⁰ FRIER, Sarah. *No Filter: The Inside Story of Instagram*. New York, London, Toronto, Sydney: Simon & Schuster, 2020, p. 52-65.

O elemento-chave para o surgimento desta nova lógica de acumulação foi a alteração na forma de uso dos dados coletados de usuários do Google. Segundo a autora, no final de 2000, com o estouro da bolha da internet e a crescente pressão de investidores de risco para que o Google gerasse lucros e retornos aos investimentos nele aportados, a empresa alterou seu modelo de negócios e a forma como usava os dados dos usuários para se salvar da crise. Os dados comportamentais passaram a ser usados não só para melhorar a experiência dos internautas, mas também para dirigir publicidade a usuários específicos, garantindo que tivessem acesso a anúncios que fossem “relevantes” para eles⁹¹. Os sinais colaterais das interações dos usuários com o sistema, antes descartados, passaram a ser acumulados e reaproveitados para melhorarem o direcionamento de publicidade, e, conseqüentemente, a lucratividade de anúncio para o Google e seus anunciantes.

Os dados comportamentais acumulados que excedem o necessário para o aprimoramento da experiência do usuário são o que Zuboff chama de “superávit comportamental”, onde a companhia encontrou sua maior fonte de geração constante e crescente de lucros. Com o tempo, a equipe do Google percebeu que o superávit comportamental, processado por meio de inteligência de máquina, levaria a predições cada vez mais precisas sobre os comportamentos dos usuários, que seriam úteis a anunciantes, na medida em que permitiam um direcionamento de publicidade mais ajustado aos interesses dos consumidores. Desse modo, o Google conseguiria atender às demandas de seus verdadeiros clientes, os anunciantes, e gerar cada vez mais lucros⁹².

Em síntese, na nova lógica de acumulação inaugurada pelo Google, o superávit comportamental é a matéria-prima mais valiosa de todas para a geração de lucro. Ele é processado, por meio de inteligência de máquina, em produtos de predição, cujo objetivo é antever nossos pensamentos, sentimentos e ações, e que são vendidos aos clientes das companhias de tecnologia – como anunciantes⁹³. No capitalismo de vigilância, a matéria-prima são os dados, o produto são as predições de comportamento, o meio de produção é a inteligência de máquina e os clientes são os anunciantes⁹⁴. Tudo isso gera o capital de vigilância.

⁹¹ ZUBOFF, Shoshana. *A Era do Capitalismo de Vigilância*, p. 93.

⁹² *Ibid.*, p. 100.

⁹³ *Ibid.*, p. 117.

⁹⁴ *Ibid.*, p. 114-118. A autora ressalta que, por mais que os anunciantes tenham sido os “jogadores principais” no início do capitalismo de vigilância, qualquer ator “com algum interesse em adquirir

Essa lógica de acumulação rapidamente se espalhou pelas empresas de tecnologia, e a primeira a adotá-la foi o Facebook, sociedade que atualmente, sob o nome de Meta Platforms Inc., controla o Instagram⁹⁵. O mesmo modelo de negócios baseado em publicidade foi implantado por Mark Zuckerberg, mas também potencializado por um componente particular do Facebook/Instagram que serviria como fonte inesgotável de superávit comportamental⁹⁶: o compartilhamento da intimidade e a autoexposição espetacular.

Além de os usuários do Facebook e do Instagram incluírem seus dados reais no site (como nome, gênero, localização e interesses), também compartilham avidamente seu dia a dia, percursos, atividades cotidianas, desejos e sentimentos. As informações são entregues de forma apaixonada e desimpedida às redes sociais em um culto ao compartilhamento da intimidade e do cotidiano⁹⁷. Esses mesmos compartilhamentos, por sua vez, atraem interações e engajamento de outros usuários, que podem curtir, comentar, salvar, encaminhar a um amigo ou ignorar uma postagem – tudo isso gera mais dados e sinais comportamentais que serão reaproveitados nas predições da plataforma.

Além do compartilhamento da intimidade, o Instagram é centrado no que chamo de “autoexposição espetacular” do indivíduo: o compartilhamento de imagens – sejam fotos ou vídeos – do próprio usuário que suscitem desejos nos demais. A autoexposição espetacular diz respeito à exposição de uma intimidade muitas vezes performada, fabricada para ser estética, apelativa e engajar outros usuários, aumentando as métricas de um perfil. No caso das mulheres, existem tipos de autoexposição espetacular que são preferidos pelo Instagram por seu potencial de atrair engajamento (e, portanto, lucro). Usualmente são materiais que se atêm a certas *imagens de feminilidade*.

Chamo de “imagens de feminilidade” aquelas construídas no seio do modo de vida capitalista que encerra mulheres em figuras como prostituta, santa, mãe de família, boa moça, traiçoeira e ser hipersexual. Geralmente portando características situadas em um extremo de virtuosismo ou de sua falta dele, e se apresentando como

informação probabilística sobre o nosso comportamento e/ou influenciar comportamento futuro pode pagar para jogar nos mercados onde os destinos comportamentais de indivíduos, grupos, corpos e coisas são narrados e vendidos”.

⁹⁵ Ibid., p. 112-113.

⁹⁶ Ibid., p. 113.

⁹⁷ HARCOURT, Bernard. *Exposed: Desire and Disobedience in the Digital Age*. Cambridge, Massachusetts/London, England: Harvard University Press, 2015, p. 13.

categorias aplicáveis às mulheres em geral, essas imagens nada mais são do que signos permeados por vieses de raça e classe, já consolidados pelo uso reiterado. São arquétipos da feminilidade atribuídos às mulheres como se fossem moldes prontos, objetivamente existentes no mundo e descritivos de categorias de mulheres.

As imagens de feminilidade priorizadas pelo Instagram, e, portanto, invisíveis aos sistemas automatizados de moderação, são aquelas que respeitam tanto as regras escritas quanto as regras ocultas da plataforma. Como visto no Capítulo 1, a sexualidade feminina, por exemplo, é admissível desde que conformada sobre determinados padrões e desde que performada dentro de um limite de “segurança” que impõe uma série de normas de como deve o corpo feminino se portar.

A empresa de Zuckerberg tem acesso a quantidades massivas de dados sendo gerados a cada compartilhamento, autoexposição, interação e ação na rede social, permitindo-lhe, por meio de inteligência de máquina e estatística, fazer correlações para prever e influenciar ações dos usuários. Essa dinâmica é extremamente benéfica aos anunciantes, permitindo que direcionem melhor sua publicidade e influenciem em comportamentos consumeristas dos usuários. O culto ao compartilhamento da intimidade e autoexposição espetacular sobre o qual o Instagram se baseia potencializou o modelo de negócios de acumulação de superávit comportamental para a geração de lucro.

Outra estratégia de marketing que se tornou a marca do Instagram e favoreceu a geração de superávit comportamental e de lucro foi aliar o compartilhamento da intimidade na rede social com a promoção de produtos e serviços. O time do Facebook percebeu que anúncios de produtos eram mais bem sucedidos quando transmitidos aos usuários sob a forma de conteúdo postado no próprio site, isto é, como “informações orgânicas que as pessoas estão produzindo no site”⁹⁸. Em outras palavras, a publicidade era mais bem recebida pelos consumidores e ganhava maior projeção quando, em vez de simplesmente exposta a eles, integrava uma postagem no site convidando os usuários para uma “conversa”, para interagirem e engajarem com o conteúdo do perfil da empresa. Isso poderia ser feito, por exemplo, em uma

⁹⁸ KIRKPATRICK, David. *The Facebook Effect: The Inside Story of the Company That Is Connecting the World*. New York, London, Toronto, Sydney: Simon & Schuster, 2010, p. 260.

publicação pedindo que os usuários comentem, marquem seus amigos, ou promovendo sorteios de produtos⁹⁹. O conteúdo postado na rede social pelos usuários, em si, passou a ser publicidade.

Esse modelo de distribuição de publicidade sob a forma de conteúdo publicado na rede social foi impulsionado pelo Instagram devido ao seu caráter imagético, de consumo rápido de estímulos visuais. A plataforma inaugurou uma dinâmica comunicacional baseada em conteúdo visual, uma nova estética, um novo olhar sobre a realidade, a ponto de indivíduos estarem sempre em busca de experiências, imagens e composições visuais que sejam “instagramáveis”, compartilháveis, que suscitem reações nos outros usuários da rede social e atraiam curtidas. A disputa, no Instagram, é pelo conteúdo cuja composição visual prenda o usuário, incite seu interesse, desperte desejos de ter algo, seja um estilo de vida, um corpo, uma estética ou uma roupa.

O apelo do visual, aspecto sempre explorado na publicidade, é encontrado em sua potência máxima no Instagram e passou a ser instrumentalizado para que os próprios usuários se tornassem anunciantes, publicando conteúdo monetizável. A figura do “Influenciador Digital”, consolidada pelo Instagram, é emblemática desse novo modelo de publicidade: uma forma de celebridade da Internet que normalmente começa como um usuário comum da rede social compartilhando imagens de seu cotidiano que por algum motivo suscitam interesse e desejos em outros usuários, até acumular base de seguidores grande suficiente para monetizá-la, produzindo publicidade para marcas¹⁰⁰.

O conteúdo postado no Instagram, portanto, é medido por seu valor comercial – a capacidade de gerar valor para a plataforma e seus clientes. Quando maior esse valor, maior a visibilidade concedida pelas técnicas automatizadas de curadoria e moderação. As regras não escritas de visibilidade dão primazia ao conteúdo que tenha maior probabilidade de gerar superávit comportamental – isto é, atrair interações, curtidas, comentários e compartilhamentos, a matéria prima do negócio do Instagram – ou destinado a promover determinado produto/serviço dos anunciantes.

⁹⁹ Idem.

¹⁰⁰ LEAVER, Tama; HIGHFIELD, Tim; ABIDIN, Crystal. *Instagram: Visual Social Media Cultures*. Cambridge, UK; Medford, MA, USA: Polity, 2020, p. 112.

É por isso que a lógica do superávit comportamental e do lucro presentes na moderação automatizada do Instagram tem influência sobre o seu caráter discriminatório. Os valores hegemônicos de estética, coerência de gênero e raça são reforçados pelos sistemas automatizados, que aprendem que são eles a agradar a maior parte dos usuários e atrair engajamento. A atribuição de valores comerciais a corpos e comportamentos importa em distribuição desigual de visibilidade, representação e discurso, porque os corpos de maior valor dentro do capitalismo sempre foram os brancos, cis e heterossexuais.

Quanto à repercussão dessa lógica do lucro especificamente sobre a visibilidade do corpo feminino no Instagram, se por um lado as técnicas de curadoria conferem grande visibilidade a conteúdos que hipersexualizam o corpo feminino para fins comerciais, as técnicas automatizadas de moderação impõem barreiras à circulação de imagens de corpos femininos não alinhados aos padrões de beleza, raça e gênero hegemônicos ou não voltados à comercialização de algo.

Como resultado, papéis sociais excludentes historicamente atribuídos às mulheres – o de objetos de consumo hipersexualizados, destinados ao olhar do outro – circulam com muito mais abrangência. Já as postagens dissidentes, que tentam expor outras possibilidades de existência, quando não censurados, têm pouca exposição. De todo modo, o corpo da “mulher” acaba sendo alvo de policiamento desproporcionalmente intenso: é preciso estar sempre atento para que o corpo feminino não ofenda, constranja e afaste usuários, mas é positivo que ele possa ser usado para seduzir, mostrar o belo e incitar desejos consumeristas ou estéticos em uma autoexposição espetacular, se estiver em conformidade com os padrões hegemônicos que atraem a geração de lucro.

Assim, por exemplo, a nudez feminina é tolerada quando exposta de forma comercial, apesar de proibida pelas Diretrizes da Comunidade e combatida por meio do *shadowban*. Publicações de celebridades como Kim Kardashian e Kylie Jenner seminuas são promovidas, enquanto as de *pole dancers* e pessoas transexuais são derrubadas por violar as Diretrizes da Comunidade que proíbem a nudez feminina – quando na realidade sequer estão mostrando seus corpos nus. Como explicam Are e Paasonen¹⁰¹, o conteúdo que não tem valor monetário óbvio para a rede social

¹⁰¹ ARE, Carolina; PAASONEN, Susanna. “Sex in the shadows of celebrity”, p. 7.

ou que é difícil de ser monetizado por ser considerado limítrofe às políticas de moderação da empresa, acaba caindo na malha da moderação do Instagram como violador das Diretrizes da Comunidade, ou próximo a isto, e é excluído ou tem sua visibilidade limitada (*shadowban*). Todo o valor cultural, social artístico ou político da publicação é desprezado diante de sua ausência de valor monetário óbvio.

Para além do poder de vigilância voltado à coleta de dados e geração de lucros evidenciado por Zuboff, o modelo de negócios do capitalismo de vigilância, no contexto do Instagram, também é acompanhado de um poder punitivo, o poder de punir os usuários da rede social que publiquem material que atrapalhe a geração de lucro, que ofenda a comunidade, que vá em sentido contrário às técnicas de manutenção da audiência. Esse poder é exercido pelas técnicas de moderação de conteúdo.

Segundo Irmgard Emmerlhainz, “a captura capitalista de dados para fins lucrativos não depende de policiamento de conteúdo; a geração de capital depende apenas da constante troca e circulação de informações”¹⁰². Reformulando a frase do autor, o policiamento só é útil na medida em que serve à mesma lógica de lucro, seja *ativamente* – suprimindo conteúdo que, em tese, poderia desagradar os valores hegemônicos, repelir anunciantes ou não gerar valor –, seja *passivamente* – ao deixar circular conteúdo que, apesar de odioso, atrai engajamento e geração de dados. A moderação e o policiamento de discursos de ódio a minorias passaram a ser uma preocupação maior somente quando se percebe o impacto que poderia ter na manutenção de anunciantes nas redes sociais. A título de exemplo, quando empresas supostamente ameaçaram retirar seus anúncios do Google depois de tomarem conhecimento que estavam sendo exibidos ao lado de publicações neonazistas, a empresa decidiu movimentar sua máquina de moderação e implementar mais controles de discurso de ódio¹⁰³.

Como visto no Capítulo 1, enquanto a moderação do Instagram se empenha em detectar e suprimir qualquer tipo de nudez – mas, por vieses dos sistemas algo-

¹⁰² EMMELHAINZ, Irmgard, “Authoritarianism and the Cybernetic Episteme, or the Progressive Disappearance of Everything on Earth”. *E-flux Journal*, n. 122, novembro 2021. Disponível em: <<https://www.e-flux.com/journal/122/430488/authoritarianism-and-the-cybernetic-episteme-or-the-progressive-disappearance-of-everything-on-earth/>>. Acesso em: 2 fev. 2022.

¹⁰³ CHANG, Emily. *Brotopia: Breaking Up the Boys’ Club of Silicon Valley*. New York: Penguin, 2018, p. 197.

rítmicos, resulta na supressão desproporcional de corpos, sequer desnudos, de pessoas racializadas, transexuais e portadoras de deficiências físicas – por outro lado, não demonstra a mesma preocupação em desenvolver estratégias para lidar com o racismo e as agressões misóginas interpessoais ou embutidas nos sistemas algorítmicos de moderação. A geração de capital, portanto, depende desses silenciamentos a conteúdo classificado como ofensivo a uma concepção de “comunidade diversa” preconizada pelas Diretrizes da Comunidade. Não basta apenas fomentar um espaço de produção e circulação de informações que serão utilizadas para a geração de lucro. Existe também o lado punitivo do capitalismo de vigilância, o lado em que é necessário exercer um poder de moderação para garantir que a vigilância, a coleta e o tratamento dos dados para a venda de predições terá êxito.

De todo modo, vale lembrar que, qualquer que seja o escândalo surgido do tratamento discriminatório do Instagram, a rede social emite manifestações públicas defendendo seus sistemas automatizados e dizendo que revisará os sistemas algorítmicos. Em pouco tempo, são feitos ajustes no *design* da plataforma. O capitalismo e a lógica do lucro rapidamente se adaptam a novas demandas de visibilidade, utilizando-as a seu favor. Mais vale promover a “inclusão” de corpos diferentes para acessar novos mercados, atrair outros usuários (e, conseqüentemente, mais dados) do que permanecer inflexível e repeli-los. O que é visto como “conquista” de direitos e visibilidade por parte de grupos sociais marginalizados, acaba vindo com capturas do sistema econômico¹⁰⁴. Nunca é simplesmente uma conquista de direitos.

Fato é que a lógica do lucro, o objetivo de angariar superávit comportamental e agradar anunciantes, é o que determina o direcionamento da moderação de conteúdo no Instagram. Com base nela, são atribuídos valores diversos a corpos diversos, conforme seu potencial de atrair anunciantes, engajamento e, conseqüentemente, geração de dados. Ocorre que essa hierarquia de valores corporais tem uma aparência física bem delimitada, reproduzindo uma classificação da população herdada da colonialidade. É o que será visto a seguir.

¹⁰⁴ Agradeço à Viviana Ribeiro pelas discussões a respeito da captura de pautas minoritárias pelo capitalismo digital e por ter introduzido essa ideia e sugerido a inclusão neste trabalho.

2.2.

Colonialidade automatizada e a hierarquização de humanidades

Nesta Seção, pretendo usar o conceito de *colonialidade automatizada* de Tarcízio Silva para pensar a ordenação social replicada pelos sistemas automatizados de moderação de conteúdo no Instagram. Argumento que essas tecnologias acabam por engessar seu componente humano responsável por discriminações sociais, na medida em que replicam em grande escala os valores embutidos na sua construção e funcionamento. Além disso, a partir da mesma categoria de Tarcízio Silva, exploro a epistemologia datacêntrica tida pelo Instagram como a única capaz de levar a uma apreensão legítima do funcionamento dos seus sistemas. A adesão inflexível a essa epistemologia importa na desconsideração da experiência de discriminação vivida pelos usuários com os sistemas algorítmicos e na negação do poder da plataforma.

Apesar de o capitalismo de vigilância apresentar-se como “uma força social democrática e libertadora”¹⁰⁵, as práticas e discursos de poder que o sustentam revelam a colonialidade que lhe é intrínseca. Como explica Quijano, a colonialidade, uma herança da estrutura de dominação colonial, não cessou de existir com a independência formal das colônias europeias¹⁰⁶. Pelo contrário, ela é constitutiva do padrão mundial de poder capitalista, tendo acompanhado as suas mutações¹⁰⁷, e vindo a se manifestar hoje no capitalismo de vigilância, nas tecnologias digitais usadas para a acumulação de dados e geração de lucro.

Tendo isso em vista, Tarcízio Silva chama atenção para a existência de uma nova forma de colonialidade, automatizada, em que as tecnologias digitais intensificam e reproduzem mecanismos da colonialidade, promovendo distribuições raci-

¹⁰⁵ ZUBOFF, Shoshana. *A Era do Capitalismo de Vigilância*, p. 85.

¹⁰⁶ Quijano explica o conceito de colonialismo, do qual a colonialidade se distingue: “Colonialidade é um conceito diferente de, ainda que vinculado a, Colonialismo. Este último refere-se estritamente a uma estrutura de dominação/exploração onde o controlo da autoridade política, dos recursos de produção e do trabalho de uma população determinada domina outra de diferente identidade e cujas sedes centrais estão, além disso, localizadas noutra jurisdição territorial. Mas nem sempre, nem necessariamente, implica relações racistas de poder. O colonialismo é, obviamente, mais antigo, enquanto a Colonialidade tem vindo a provar, nos últimos 500 anos, ser mais profunda e duradoira que o colonialismo.” (QUIJANO, Anibal, “Colonialidade do Poder e Classificação Social”. In: SANTOS, Boaventura de Sousa; MENESES, Maria Paula (Org.). *Epistemologias do Sul*. Coimbra: Almedina, 2009, p. 73)

¹⁰⁷ Idem.

alizadas e generificadas de “acesso, consumo, produção e representação no ambiente digital”¹⁰⁸⁻¹⁰⁹, que importam na concentração de poder e na dominação epistêmica¹¹⁰.

A título de exemplo, em 2021, com o vazamento dos “*Facebook Papers*” pela ex-funcionária do Facebook Frances Haugen, ficou claro que a empresa, agora Meta Platforms, Inc., deliberadamente ignorava o racismo reforçado pelos seus algoritmos e praticado pelos usuários entre si¹¹¹. Em outras palavras, escolhia ser “racialmente cega”, ou seja, não considerar questões de raça no desenvolvimento de seus produtos e na tomada de decisões negociais. Como visto no Capítulo 1, isso leva a produtos que perpetuam uma série de discriminações sociais.

Segundo Frances Haugen, a empresa “percebeu que se mudasse o algoritmo para ser mais seguro, as pessoas passariam menos tempo no site, clicariam em menos anúncios, [e o Facebook] ganharia menos dinheiro”¹¹². O episódio foi apenas um dos escândalos envolvendo a atual Meta Platforms, Inc. em que foi possível perceber que não se *supõe* que o funcionamento dos modelos algorítmicos e dinâmicas internas do Instagram sejam racistas. Pelo contrário, os vieses racistas dos algoritmos são conhecidos pela empresa, bem como as agressões entre usuários, e mesmo assim ignorados pela empresa, em prol da geração de capital de vigilância.

A indiferença da empresa a danos sociais provocados pelos sistemas automatizados do Instagram, sejam eles sofridos por mulheres brancas, negras, transexuais ou *pole dancers*, implica na adesão a uma determinada ordenação social, em que a

¹⁰⁸ SILVA, Tarcizio, “Colonialidade automatizada: algoritmos e aprendizado de máquina nas plataformas de mídias sociais”, p. 141.

¹⁰⁹ A discriminação algorítmica vem sendo crescentemente abordada na produção literária: O’NEIL, Cathy. *Algoritmos de Destruição em Massa: Como o Big Data Aumenta a Desigualdade e Ameaça a Democracia* (tradução de Rafael Abraham). Santo André: Editora Rua do Sabão, 2020; NOBLE, Safiya. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press, 2018. Para um estudo dos impactos da discriminação algorítmica nas classes econômicas mais baixas dos Estados Unidos: EUBANKS, Virginia. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York: St. Martin’s Press, 2017.

¹¹⁰ SILVA, Tarcizio, Op. Cit., p. 140.

¹¹¹ DIAS, Tatiana; RIBEIRO, Paulo Victor; LOPES, Débora, “O Grande Experimento. Facebook Papers: as provas que faltavam para mostrar como a rede manipulou você”. *The Intercept Brasil*, 01.12.2021. Disponível em: <<https://theintercept.com/2021/12/01/facebook-papers-provas-rede-manipulou-voce/#two>>. Acesso em: 9 mar. 2022.

¹¹² Tradução livre de “*Facebook has realized that if they change the algorithm to be safer, people will spend less time on the site, they’ll click on less ads, they’ll make less money*” (HAUGEN, Frances, “Facebook Whistleblower Frances Haugen: The 60 Minutes Interview”. Entrevista concedida a Scott Pelley. *60 Minutes*, 2021. (13:36). Disponível em: <<https://www.youtube.com/watch?v=Lx5VmAdZSI>>. Acesso em: 2 abr. 2022)

violência e discriminação experimentadas por certos segmentos de usuários são vistos com negligência, sem maior importância.

A experiência de usuários com os sistemas algorítmicos é considerada a partir de um ponto de vista “neutro”, que considera um usuário não situado socialmente: simplesmente usuário, universal, sem gênero, raça e classe. Ocorre que a falta de incorporação de perspectivas específicas na análise da experiência dos usuários importa na manutenção de uma hierarquia de humanidades já conhecida fora do mundo virtual, herdada da colonialidade.

Duas instâncias de manifestação da colonialidade no âmbito digital são a discriminação algorítmica com base na raça e gênero dos indivíduos e a supervalorização das epistemologias centradas em dados. Ambas são encontradas no Instagram, e sua ocorrência pode ser percebida nos efeitos da moderação algorítmica e nos pronunciamentos públicos da empresa quando confrontada com críticas à sua forma de conduzir o controle de conteúdo.

Na colonialidade, que forjou o pensamento ocidental hegemônico, o corpo é a base dos papéis, identidades e hierarquização sociais. Como explica Oyèrónké Oyewùmí, dentro da lógica colonial, as categorias sociais são baseadas na biologia¹¹³. Tanto as categorias de “mulher” quanto as de “negra(o)” são elaboradas com base na constituição corporal: órgãos sexuais e quantidade de melanina determinam a posição social de um indivíduo. Essas são noções ocidentais instaladas em outros povos – como os africanos e os das Américas – que não as conheciam nem estruturavam suas sociedades dessa maneira até o assalto dos colonizadores, e prevalecem até hoje na ordenação do mundo social.

Em linha com essa matriz de poder colonial, as técnicas de moderação automatizada de conteúdo do Instagram normalmente identificam como os corpos mais arriscados, violadores das Diretrizes da Comunidade e ofensivos, os racializados ou definidos por seu sexo. Como visto no Capítulo 1, esses corpos são desproporcionalmente invisíveis em sua humanidade aos sistemas algorítmicos de moderação e hipervisíveis em seu caráter transgressor.

¹¹³ OYEWÙMÍ, Oyèrónké. *A invenção das mulheres: construindo um sentido africano para os discursos ocidentais de gênero* (tradução de Wanderson Flor do Nascimento), 1ª edição. Rio de Janeiro: Bazar do Tempo, 2021, p. 15-16.

A hierarquização de humanidades com base em critérios coloniais fica evidente na cegueira dos sistemas algorítmicos para a humanidade de corpos de diferentes cores¹¹⁴, para a nudez feminina não objetificada e para sexualidades dissidentes, ao classificá-los como violações às Diretrizes da Comunidade quando não o são. A moderação da rede social é regida pela classificação racial/étnica dos seres humanos e pela concepção binária de homens e mulheres forjada dentro desse sistema de poder.

Como visto do exemplo dos *Facebook Papers*, a cegueira algorítmica à humanidade de indivíduos com determinadas características físicas não decorre do algoritmo em si, mas sim de seu *componente humano* – sejam os vieses incorporados aos sistemas pelo programador, seja o comportamento discriminatório dos usuários e moderadores humanos. Algoritmos que utilizem *machine learning* por exemplo são treinados com dados já existentes, os quais podem incorporar vieses de raça e gênero. Seu emprego em escala, para lidar com uma grande quantidade de conteúdo publicado na rede social, importa na replicação dos mesmos julgamentos sociais em larga escala. Além disso, a inteligência de máquina empregada na moderação de conteúdo do Instagram tem maior dificuldade de identificar contextos e nuances da comunicação em comparação com humanos – como o sarcasmo e a crítica. Assim, conteúdo não violador das Diretrizes da Comunidade relativas a nudez, por exemplo, pode ser suprimido por não ter seu contexto identificado.

De todo modo, o abuso de ferramentas algorítmicas para punir usuários que postem conteúdo [em tese] não tolerado pela rede social importa em um engessamento e replicação em larga escala dos valores humanos que regem essa atividade – os quais, no atual sistema econômico, ainda incorporam classificações coloniais dos seres humanos. Como ressalta Gillespie¹¹⁵, comunidades estão continuamente revendo seus valores e se reconstituindo, submetendo os valores amplamente aceitos a constante reconsideração. Seu engessamento pelo uso de sistemas algorítmicos para a classificação e supressão de grandes quantidades de conteúdo resulta não

¹¹⁴ Joy Buolamwini e Timnit Gebru verificaram disparidades de performance classificatória de imagens de seres humanos por produtos de inteligência artificial quando os rostos retratados fossem brancos ou negros (BUOLAMWINI, Joy; GEBRU, Timnit, “Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification”. *Proceedings of Machine Learning Research*”, v. 81, 2018, pp. 1-15. Disponível em: <<http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf>>. Acesso em: 28 fev. 2022).

¹¹⁵ GILLESPIE, Tarleton, “Content moderation, AI, and the question of scale”, p. 4.

só na manutenção como também na amplificação dos valores e classificações sociais incorporados a tais sistemas pelo componente humano.

Além da hierarquização de humanidades herdada da colonialidade, o Instagram também adere firmemente a uma epistemologia centrada nos dados (datacêntrica)¹¹⁶, que rechaça qualquer outra forma de entender o funcionamento dos sistemas algorítmicos empregados na moderação de conteúdo. Com o avanço do capitalismo de vigilância, da acumulação e mineração de dados e das capacidades preditivas de sistemas algorítmicos, a produção de conhecimento legítimo passou a estar fortemente atrelada aos dados, como se tivessem eles o potencial de descrever uma forma mais pura e imediata da realidade, livre de vieses humanos, ideologias e interpretações subjetivas que contaminam a apreensão do real¹¹⁷. O modelo hegemônico de produção de conhecimento confia cegamente nos dados e em sua imparcialidade.

Assim, se os sistemas baseados em dados geram efeitos prejudiciais a minorias étnicas e raciais, trata-se apenas de uma falha computacional, nada que uma reunião do conselho de administração do Instagram e ajustes feitos pelo setor de desenvolvimento não possam resolver. Basta adequar os modelos algorítmicos e a realidade voltará a ser descrita fielmente, desta vez considerando mais complexidades sociais incorporadas ao sistema. Sob essa epistemologia, os modelos algorítmicos jamais são discriminatórios. Qualquer alegação nesse sentido é mera impressão dos críticos, que não sabem o que falam.

A adesão inflexível à epistemologia datacêntrica é mais uma das manifestações da colonialidade automatizada, uma vez que consiste na imposição de saberes gestados em empresas de tecnologia do Norte Global a países periféricos. A colonialidade automatizada se manifesta na dimensão dos saberes quando há a imposição de um regime de verdades centrado na epistemologia datacêntrica e nos dados como medida universal de legitimação de conhecimento. Existe uma dominação epistêmica, um suposto monopólio da forma de conhecer o mundo e as experiências humanas em sua forma mais pura e objetiva, em exclusão à ciência produzida por

¹¹⁶ RICAURTE, Paola. “Data Epistemologies, The Coloniality of Power and Resistance”. *Television & New Media*, v. 20, n. 4, 2019, pp. 30-365.

¹¹⁷ ROUVROY, Antoinette, “L’art de ne pas changer le monde”. *La Revue Nouvelle*, n. 8, 2016, p. 45.

outros meios e com base na experiência de comunidades que não detêm os conhecimentos técnicos da ciência da computação, mas apenas vivem os seus efeitos. Suas experiências com a tecnologia são descartadas de pronto, por não estarem alinhadas à epistemologia datacêntrica.

No Instagram, a defesa ferrenha da epistemologia datacêntrica como a única capaz de chegar à verdade é encontrada nas manifestações públicas da empresa contra críticas à moderação discriminatória de conteúdo. Com o crescente número de usuários condenando a exclusão injustificada de conteúdo e o *shadowban* de grupos marginalizados, a empresa decidiu compartilhar uma série de postagens no blog do Instagram supostamente com o objetivo de oferecer aos usuários esclarecimentos e transparência sobre o funcionamento da rede social.

As postagens, no entanto, geralmente consistem em divulgações protocolares ao mercado dotadas de estratégias comunicacionais que reforçam a epistemologia datacêntrica, descredibilizam as críticas recebidas e evitam qualquer assunção de responsabilidade. Longe de ser uma especificidade do Instagram, esse é um padrão comunicativo de empresas de tecnologia já identificado por Tarcizio Silva.

O autor identificou três expedientes comumente utilizados por elas em resposta a críticas recebidas pelos danos sociais causados por seus sistemas algorítmicos. Em primeiro lugar, tais empresas costumam evocar um processo contínuo de “otimização” das tecnologias digitais utilizadas, como se bastassem alguns ajustes nos modelos algorítmicos para que os problemas de tratamento discriminatório fossem resolvidos, e não um remanejamento dos seres humanos que produzem a tecnologia. Além disso, também é comum alegarem que os danos sociais causados pelos seus sistemas apenas reproduzem os preconceitos e desigualdades da sociedade; assim, a distribuição mais ampla de imagens de mulheres brancas no Instagram, por exemplo, poderia ser atribuída à preferência dos usuários por esse tipo de foto. De modo algum a estrutura do sistema algorítmico e as decisões negociais da companhia têm alguma influência sobre isso. Por fim, as empresas de tecnologia costumam esquivar-se de responsabilidade pelas discriminações sociais operadas pelos seus sistemas fazendo referência à complexidade dos produtos tecnológicos desenvolvidos, sem qualquer menção às etapas prévias de treinamento dos sistemas e estudos de impactos em diferentes grupos sociais¹¹⁸.

¹¹⁸ SILVA, Tarcizio. “Controvérsias sobre Danos Algorítmicos: discursos corporativos sobre discriminação codificada”. *Revista Observatório*, v. 6, n. 4, julho-setembro 2020, pp. 14-15.

O Instagram segue à risca esse padrão, protegendo-se de críticas aos efeitos de sua moderação algorítmica com um discurso de “otimização”, complexidade dos sistemas e má apreensão de seu funcionamento pelos usuários. A empresa, por exemplo, em diversas ocasiões negou a existência do *shadowbanning*¹¹⁹ ou afirmou que a expressão era usada por usuários insatisfeitos com o fato de o seu conteúdo não atingir a amplitude de visualizações que gostariam¹²⁰ – deslocando a “culpa” pela experiência da discriminação para o público. Em pronunciamento publicado em 26 de abril de 2021¹²¹, a empresa explicava que o uso da expressão “*shadowbanning*” decorria de “falta de entendimento” sobre a redução de visibilidade de engajamento; “impressão” de que as políticas do Instagram não consideram nuances culturais e experiências diversas; “confusão” sobre a execução das políticas internas “por acharem que algumas comunidades estavam sujeitas a critérios que outras não estavam” ou que os sistemas algorítmicos removiam conteúdo sem motivo; e “problemas técnicos, interpretados por algumas pessoas como formas de repressão”.

Inflexível em seu discurso tecnoutópico, o Instagram impõe aos usuários a epistemologia datacêntrica como a única capaz de legitimar qualquer afirmação sobre os efeitos da moderação algorítmica, desestabilizando as críticas recebidas¹²². A empresa se mantém numa posição de autoridade a validar as preocupações e alegações dos usuários sobre sua moderação algorítmica, como se o acesso ao código-fonte ou informações sigilosas de design dos sistemas dessem amplo e irrestrito

¹¹⁹ COOK, Jesselyn, “Instagram’s CEO Says Shadow Banning ‘Is Not A Thing’. That’s Not True”. *HuffPost*, 25.02.2020. Disponível em: <https://www.huffpost.com/entry/instagram-shadow-banning-is-real_n_5e555175c5b63b9c9ce434b0>. Acesso em: 2 jan. 2022.

¹²⁰ MOSSERI, Adam. *How does Instagram work? Or more specifically – why do I see what I see on Instagram?*. 28 jun. 2021. Twitter: @mosseri. Disponível em: <https://twitter.com/mosseri/status/1409533965606821888?ref_src=twsrc%5Etfw%7Ctwcamp%5Etweetembed%7Ctwterm%5E1409533965606821888%7Ctwgr%5E%7Ctwcon%5E1_&ref_url=https%3A%2F%2Fblog.hootsuite.com%2Fsocial-media-updates%2Finstagram%2Finstagram-says-shadowbanning-does-not-exist%2F>. Acesso em: 1 fev. 2022.

¹²¹ BROOKS, Rachel; HOUSTON, Stacey; WADSWORTH, Christina, “Atualizações da Equipe de Equidade: o que realizamos e o que ainda está por vir”. *Instagram*, 26.04.2021. Disponível em: <<https://about.instagram.com/pt-br/blog/announcements/equity-updates-what-weve-been-up-to-and-what-comes-next>>. Acesso em: 08 mar. 2022.

¹²² COTTER, Kelley, “‘Shadowbanning is not a thing’: black box gaslighting and the power to independently know and credibly critique algorithms”. *Information, Communication & Society*, 2021, p. 14.

entendimento sobre todas as consequências e possibilidades dos sistemas algorítmicos¹²³. Numa postura paternalista¹²⁴ e condescendente, ela coloca seus usuários em posição de completa ignorância e inelutável erro, incapazes de avaliar a moderação algorítmica em suas próprias realidades: o desconhecimento técnico é tido como deslegitimador das críticas, por mais que elas sequer tenham a ver com a técnica, mas sim com vivências.

¹²³ Idem.

¹²⁴ PETRE, Caitlin; DUFFY, Brooke Erin; HUND, Emily, “‘Gaming the System’: Platform Paternalism and the Politics of Algorithmic Visibility”. *Social Media + Society*, outubro 2019. Disponível em: <<https://journals-sagepub-com.ez370.periodicos.capes.gov.br/doi/full/10.1177/2056305119879995>>. Acesso em: 23 mar. 2022.

O governo algorítmico das condutas

Vistas as discriminações perpetuadas pelos sistemas algorítmicos de moderação de conteúdo do Instagram e os valores que lhes são subjacentes, neste Capítulo pretendo aprofundar a ideia de que a moderação algorítmica da rede social integra uma nova forma de poder que tem efeitos de governo sobre as condutas dos usuários. Além disso, pretendo examinar especificamente como esse governo cobra e reitera determinada subjetividade das usuárias identificadas como mulheres e, por fim, onde se encontram as possibilidades de resistência a esse governo das condutas que parece não deixar brechas para a inventividade e a formulação de críticas.

Governar, para Foucault¹²⁵, consiste em “conduzir condutas”. O governo algorítmico das condutas que pretendo abordar nas páginas a seguir, portanto, consiste na condução de condutas dos usuários do Instagram por meio de sistemas algorítmicos que controlam o conteúdo circulante na rede social. Neste Capítulo, as discussões terão como base principalmente o conceito nefoucaultiano de Rouvroy e Berns de *governamentalidade algorítmica*¹²⁶, por meio do qual será possível apreender, na Seção 3.1., como a moderação algorítmica de conteúdo contribui para o governo das condutas dos usuários no Instagram. Na Seção 3.2., a partir das ideias de Lauren Bridges sobre o “bom sujeito de dados”¹²⁷, é explorada especificamente a subjetividade ideal esperada das usuárias do Instagram identificadas como mulheres e reforçadas pelos sistemas automatizados de moderação. Por fim, na Seção 3.3, em continuidade à tese da *governamentalidade algorítmica*, estudam-se quais são as dificuldades impostas à subjetivação dos indivíduos no Instagram e onde se

¹²⁵ CASTRO, Edgardo. *Vocabulário de Foucault* – Um percurso pelos seus temas, conceitos e autores (traduzido por Ingrid Müller Xavier). Belo Horizonte: Autêntica, 2009, pp. 188-193; FOUCAULT, Michel. *Nascimento da Biopolítica: Curso dado no Collège de France (1978-1979)*. São Paulo: Martins Fontes, 2008, p. 258; FOUCAULT, Michel. *Segurança, território, população: curso dado no Collège de France (1977-1978)*. São Paulo: Martins Fontes, 2008, pp. 117-180, 253-340.

¹²⁶ ROUVROY, Antoinette; BERNS, Thomas, “Governamentalidade algorítmica e perspectivas de emancipação: o dispar como condição de individuação pela relação?” (tradução de Pedro Henrique Andrade), *Revista Eco Pós*, dossiê Tecnopolíticas e Vigilância, v. 18, n. 2, 2015, pp. 36-56.

¹²⁷ BRIDGES, Lauren E., “Digital Failure: Unbecoming the ‘good’ data subject through entropic, fugitive, and queer data”. *Big Data & Society*, janeiro-junho 2021, pp. 1-17.

encontram as possibilidades de resistência ao governo algorítmico das condutas – isto é, onde se encontram as possibilidades de os usuários conduzirem as próprias condutas, em sentido diverso da subjetividade ideal deles esperada.

3.1.

A governamentalidade algorítmica no Instagram

Nesta Seção, a partir do conceito de governamentalidade algorítmica cunhado por Antoinette Rouvroy e Thomas Berns, proponho pensar a moderação automatizada do Instagram como integrante de uma nova forma de exercício do poder dentro da rede social. Pretendo analisar como a sua moderação algorítmica contribui para o governo das condutas dos usuários, especialmente como contribui para que os usuários se conduzam da forma mais produtiva possível dentro do modelo de negócios da rede social.

Com base no conceito de Foucault de “governamentalidade” como governo das condutas, Antoinette Rouvroy e Thomas Berns desenvolvem a tese de que, atualmente, existe uma nova forma de governo do mundo social, que se baseia mais no tratamento algorítmico de quantidades massivas de dados do que na política, no direito ou em normas sociais: a *governamentalidade algorítmica*¹²⁸. No entanto, ao contrário dos conceitos foucaultianos de poder disciplinar e biopolítica, esse novo modo de governar não tem como objetivo regulamentar as condutas dos indivíduos ou as populações¹²⁹. Na verdade, trata-se de um poder que tenta atuar sobre a *previsibilidade das ações dos indivíduos*¹³⁰, e, por meio do controle do ambiente digital em que eles interagem, cercando-os de estímulos que estatisticamente têm maior

¹²⁸ ROUVROY, Antoinette; BERNS, Thomas, “Governamentalidade algorítmica e perspectivas de emancipação: o dispar como condição de individuação pela relação?”, p. 37; ROUVROY, Antoinette, “La vie n’est pas donnée”, *Études digitales*, v. 2, n. 2, 2017, p. 195. Disponível em: <<https://pure.unamur.be/ws/portalfiles/portal/54932539/8162.pdf>>. Acesso em: 20 jun. 2021.

¹²⁹ WALTER, Bruno Eduardo Procopiuk; HENNIGEN, Inês, “Problematizando a Governamentalidade Algorítmica a Partir do Sistema de Recomendação da Netflix”. *Psicologia e Sociedade* [online], v. 33, e227258, 2021, p. 2. Disponível em: <<https://doi.org/10.1590/1807-0310/2021v33227258>>. Acesso em: 2 fev. 2022.

¹³⁰ ROUVROY, Antoinette. “De la surveillance au profilage”. *Philosophie Magazine*, hors-série n. 36. Foucault. Le courage d’être soi, 2018, p. 62. Disponível em: <https://www.academia.edu/36177927/De_la_surveillance_au_profilage_Philosophie_magazine_Hors-s%C3%A9rie_N_36_Foucault_Le_courage_d%C3%Aatre_soi_2018>. Acesso em: 20 jul. 2021.

probabilidade de desencadear reações já previstas, gera efeitos de governo sobre as condutas.

Como visto no Capítulo 2, o Instagram, seguindo o modelo de negócios por excelência do capitalismo de vigilância, apropria-se de informações extraídas dos usuários e de seus comportamentos, bem como de rastros, sinais colaterais de suas ações e interações na plataforma e de dados recolhidos ou cedidos de outras fontes – em suma, coleciona um superávit de dados comportamentais – para, mediante seu tratamento por inteligência artificial, gerar previsões cada vez mais precisas que beneficiarão seus anunciantes e servirão para “*otimizar*” os serviços oferecidos, tudo com o objetivo de aumentar sua eficácia comercial e geração de lucros.

Em uma publicação de 2021 cujo objetivo supostamente era o de conferir transparência ao funcionamento do Instagram, Adam Mosseri, diretor executivo da empresa, deixa claro esse mecanismo ao detalhar minimamente o processo de curadoria e moderação automatizadas de conteúdo¹³¹. O executivo explica que, tanto na curadoria quanto na moderação automatizadas do Instagram, são empregados modelos algorítmicos diversos conforme a seção do aplicativo (Feed, Stories, Reels, Explorar) e seu respectivo objetivo.

Segundo ele, em todas as seções do aplicativo, o processo invariavelmente se apoia sobre a coleta e processamento automatizado de conjuntos massivos de informações sobre o conteúdo postado e o usuário que o compartilhou (o que ele chama de “sinais”), as quais permitirão que sejam criados perfis atrelados aos usuários com a finalidade de eleger o que será visível ou não a cada um deles. Segundo Mosseri, “existem milhares” de sinais, que “incluem tudo, desde a que horas uma postagem foi compartilhada, se você está usando um celular ou a Internet, até com que frequência você curte vídeos”. A partir desses agregados de dados, por meio de inteligência de máquina, a plataforma faz uma série de previsões, isto é, “palpites instruídos sobre a probabilidade de interagir com uma postagem de diferentes maneiras”¹³². Por fim, com base em tais previsões, a rede social seleciona o conteúdo e a ordem do que será ou não exposto a cada usuário, e o que será suprimido ou

¹³¹ MOSSERI, Adam, “Shedding More Light on How Instagram Works”. *Instagram*, 08.06.2021. Disponível em: <<https://about.instagram.com/blog/announcements/shedding-more-light-on-how-instagram-works>>. Acesso em: 2 fev. 2022. O executivo explica: “*Instagram doesn’t have one algorithm that oversees what people do and don’t see on the app. We use a variety of algorithms, classifiers, and processes, each with its own purpose. We want to make the most of your time, and we believe that using technology to personalize your experience is the best way to do that.*”

¹³² Idem.

ocultado, tendo como objetivo *distribuir aquilo que atrairá maior engajamento* – e, conseqüentemente, maior geração de dados comportamentais – e *evitar a circulação do conteúdo que choque, afaste ou ofenda os usuários*.

O modelo descrito por Mosseri corresponde às etapas constitutivas da governamentalidade algorítmica segundo Rouvroy e Berns: (i) coleta de quantidades massivas de dados e constituição de *datawarehouses*, (ii) mineração dos dados – isto é, a aplicação de tecnologias como análise estatística e modelagem, para descobrir padrões ocultos e correlações entre os dados¹³³ e produzir, com base nessas correlações, saberes estatísticos e probabilistas; e (iii) uso dos saberes produzidos para fazer predições sobre comportamentos individuais, os quais são atrelados a perfis construídos a partir da base de correlações criada com a mineração de dados¹³⁴.

Baseado inteiramente em dados e correlações estatísticas, uma das características do governo algorítmico das condutas é a evitação do *sujeito reflexivo* em todas as suas etapas. Como será visto mais adiante, essa característica do governo algorítmico complica os processos de subjetivação dos indivíduos: os processos pelos quais um indivíduo reflete sobre sua condição na relação de poder com os sistemas algorítmicos, tomando a si mesmo como objeto de reflexão crítica, e age sobre si, determinando sua própria conduta.

Em primeiro lugar, na coleta de dados, o sujeito já é evitado, pois seu consentimento para a cessão de informações comportamentais é enfraquecido; muitas vezes ele não sabe quais dados estão sendo coletados, ou não dá maior importância a essa coleta, uma vez que se tratam de dados insignificantes em si mesmos, como cliques em conteúdo, horário de uso de determinada página, e assim por diante. Em segundo lugar, o sujeito é evitado justamente pelo fato de, na coleta e mineração, serem usados dados “infraindividuais”, isto é, desatrelados dos sujeitos e de qualquer traço que os individualize, insignificantes em si mesmos, avulsos e descontext-

¹³³ Por mineração de dados, entendo a seguinte concepção constante do relatório do *United States General Account Office*: “a aplicação de tecnologia e técnicas de banco de dados – tais como análise estatística e modelagem – para descobrir padrões ocultos e relacionamentos sutis nos dados e para inferir regras que permitam prever resultados futuros” (UNITED STATES GENERAL ACCOUNTING OFFICE. *Data Mining: Federal Efforts Cover a Wide Range of Uses*. Washington: GAO, 2014 *apud* WALTER, Bruno Eduardo Procopiuk e HENNIGEN, Inês, “Problematizando a Governamentalidade Algorítmica a Partir do Sistema de Recomendação da Netflix”, p. 3).

¹³⁴ ROUVROY, Antoinette; BERNS, Thomas, “Governamentalidade algorítmica e perspectivas de emancipação: o díspar como condição de individuação pela relação?”, p. 39-41.

tualizados. Em terceiro lugar, com os saberes estatísticos e probabilísticos produzidos na mineração de dados, são criados perfis ou modelos de comportamento “supraindividuais”, ou seja, que não dizem nada sobre um indivíduo específico, podendo ser atribuídos a diversos indivíduos distintos, pois nada mais são que correlações de dados e conjuntos de probabilidades.

A tese dos autores, portanto, é calcada na ideia de que não são sujeitos os objetos da governamentalidade algorítmica. O governo algorítmico das condutas é até mesmo indiferente aos sujeitos. Seu objeto, na verdade, seriam os *dados infra-individuais*: dados enquanto relações anteriores aos sujeitos, não atribuíveis a indivíduos singulares¹³⁵, desvinculados de qualquer resquício de humanidade que permita “relacioná-los a isto que faz a singularidade de uma vida”¹³⁶.

Por meio da coleta, mineração e tratamento algorítmico de dados infraindividuais, essa nova forma de exercício do poder modeliza, antecipa e *afeta por antecipação* os comportamentos individuais possíveis¹³⁷. Daí o motivo de ser ela uma nova forma de governo das condutas. Ainda que não atue *diretamente* sobre as ações dos usuários, a governamentalidade algorítmica gera efeitos sobre elas ao controlar o ambiente que os circunda, tentando prever e cercar de antemão todas as possibilidades de agência, de modo a tornar a desobediência às suas previsões altamente improvável¹³⁸.

A moderação automatizada do Instagram, seja pela remoção de conteúdo ou pela restrição de visibilidade (*shadowbanning*), contribui com a governamentalidade algorítmica do Instagram, porque integra um conjunto mais amplo de artifícios algorítmicos da rede social cujo objetivo é manter o ambiente digital sob controle e

¹³⁵ Ibid., p. 49: “A incitação a abordar a governamentalidade algorítmica sob a perspectiva simondoniana provém do fato de que esse modo de governo parece não ter mais, por apoio e por alvo, os sujeitos, mas as relações enquanto sendo anteriores aos seus termos, isto é, não apenas as relações sociais, intersubjetivas na medida em que constroem os indivíduos, de modo que todo indivíduo seria considerado como a soma dessas relações, mas sobretudo as relações elas mesmas, independentemente de toda individuação simples e linear: as relações enquanto permanecem inatribuíveis aos indivíduos que elas vinculam”.

¹³⁶ ROUVROY, Antoinette, “Des données sans personne: le fétichisme de la donnée à caractère personnel à l’épreuve de l’idéologie des Big Data”. Les rapports du Conseil d’État (ancienne collection Étude et documents du Conseil d’État, 2014, pp. 409. Disponível em: <<https://researchportal.unamur.be/fr/publications/des-donn%C3%A9es-sans-personne-le-f%C3%A9tichisme-de-la-donn%C3%A9e-%C3%A0-caract%C3%A8re>>. Acesso em: 2 mar. 2022.

¹³⁷ ROUVROY, Antoinette; BERNS, Thomas, “Governamentalidade algorítmica e perspectivas de emancipação: o dispar como condição de individuação pela relação?”, pp. 41-42.

¹³⁸ Segundo Rouvroy e Berns, o objeto da governamentalidade algorítmica são as relações – mais precisamente, os dados, que constituem relações. Trata-se, portanto, de um governo das relações.

as possibilidades de ações e interações dos usuários dentro de comportamentos previstos. Como será visto, no entanto, existem brechas nesse controle ambiental. Usuários e usuárias conseguem apreender o funcionamento dos sistemas algorítmicos a partir de suas experiências e responderem a eles, determinando suas condutas fora das previsões.

Todas as seções do Instagram são arquitetadas para prender a atenção dos usuários o máximo de tempo possível¹³⁹ e influenciar na adoção de comportamentos lucrativos para a rede social: o engajamento com o conteúdo alheio, a autoexposição e a incitação a que outros usuários façam o mesmo, pois são esses comportamentos que geram dados, a matéria prima da empresa. Os sistemas algorítmicos de curadoria e moderação da rede social são direcionados a incitarem nos usuários a tomada de uma ação, a avidez por consumir mais material gráfico, expor mais intimidade, produzir conteúdo que obedeça às regras [expressas e ocultadas] da plataforma para atrair engajamento de outros usuários – ações que alimentam os sistemas internos de recompensa hormonal e, ao mesmo tempo, são produtivos para o aplicativo¹⁴⁰. Nesse processo, os usuários têm suas condutas moldadas.

Portanto, para atingir à finalidade de manter os usuários consumindo e produzindo conteúdo e atendendo às previsões comportamentais, o ambiente de exposição de estímulos visuais do Instagram é controlado por meio dos sistemas algorítmicos de curadoria e moderação: as técnicas de curadoria objetivam exibir a cada usuário publicações selecionadas para atenderem a desejos estatisticamente prováveis¹⁴¹, e as técnicas de moderação pretendem ocultar o que é julgado pelo Instagram como violador das políticas internas ou inapropriado, capaz de suscitar nos usuários reações e comportamentos indesejados.

Em resumo, a moderação algorítmica da rede social contribui para o *governo algorítmico das condutas* de seus usuários, porque ajuda a delimitar o campo de agência dentro da plataforma, exercendo um poder punitivo para desincentivar

¹³⁹ BENTES, Anna, “A Gestão Algorítmica da Atenção: enganchar, conhecer e persuadir”. In: POLIDO, Fabrício; ANJOS, Lucas; BRANDÃO, Luíza. *Políticas, Internet e Sociedade*. Belo Horizonte: IRIS, 2019, pp. 222-234.

¹⁴⁰ LINDSTRÖM, Björn; BELLANDER, Martin; SCHULTNER, David T; CHANG, Allen; TOBLER, Philippe; AMODIO, David M., “A computational reward learning account of social media engagement”. *Nature Communications*, v. 12, n. 1311, 2021. Disponível em: <<https://www.nature.com/articles/s41467-020-19607-x#citeas>>. Acesso em: 22 mai. 2022.

¹⁴¹ Segundo Rouvroy e Berns, o objetivo é menos adaptar as ofertas exibidas aos desejos dos usuários do que adaptá-los às ofertas. (ROUVROY, Antoinette; BERNS, Thomas, “Governamentalidade algorítmica e perspectivas de emancipação: o dispar como condição de individuação pela relação?”, p. 44).

publicações de conteúdo improdutivo ou que possa desencadear comportamentos indesejados em outros usuários comprometendo os objetivos comerciais da empresa.

Como abordado no Capítulo 1, o problema está em que esse governo algorítmico das condutas trata de forma diferenciada os usuários. O controle do ambiente imagético do Instagram é voltado a reter e fazer crescer a “comunidade diversa” que as Diretrizes da Comunidade dizem proteger, sua fonte primária de dados. Em razão dessa finalidade, a plataforma, por meio de sistemas algorítmicos de moderação, suprime ou restringe a circulação de materiais que são identificados como sensacionalistas, provocativos ou violadores das regras internas, capazes de incitar nos demais usuários comportamentos indesejados, como a radicalização de ideias, o choque e o abandono da rede social. Ocorre que as imagens e ideias identificadas dessa forma são justamente aquelas de pessoas não contidas ou diminuídas dentro de um espectro de humanidade, que, como visto, incorpora classificações coloniais.

A imagem de um corpo não compreensível dentro da binariedade de gênero – não classificável como mulher ou homem – ou de um corpo de pele e traços fenotípicos tidos como inferiores dentro da lógica colonial que desde a Modernidade ordena o mundo social, é justamente o que pode impactar a audiência. Em suma, pessoas cuja aparência não obedeça determinadas regras sociais de atribuição de qualidades. São essas imagens, ou ideias dissidentes relacionadas a elas, que podem suscitar a “radicalização” da “comunidade diversa” do Instagram, “ofender” e afetar o acesso de usuários, produção de dados e de lucro.

Não existe, a princípio, valor monetário óbvio atrelado à imagem desses corpos. Pelo contrário, eles ameaçam a geração de lucro, e, em razão disso são excluídos ou ocultados. De todo modo, essa não é uma conclusão absoluta. O Instagram, como qualquer outra companhia, precisa flexibilizar sua estratégia negocial tanto para manter seu mercado existente quanto para conquistar novos mercados, adaptando-se à mentalidade dos consumidores. Ou melhor, tanto para manter seus geradores de dados antigos quanto para atrair novos.

No caso, vale mais criar comitês de diversidade e inclusão de pessoas LGBTQIA+, anunciar mudanças algorítmicas que corrijam o racismo do sistema (jamais reconhecido como racismo, mas sim como “falha” dos sistemas) e contratar mem-

bro do conselho de administração de identidades sociais diversas. Estas são medidas que sinalizam ao mercado o compromisso raso da companhia com a inclusão de outros seres humanos em sua lógica, e, desse modo, lhe permitem manter e conquistar novos usuários. Ao mesmo tempo, a empresa divulga comunicados ao mercado que descredibilizam críticas dos usuários, reafirmam a epistemologia datacêntrica como a única legítima e evitam a própria responsabilidade por danos sociais causados ou piorados por seus sistemas algorítmicos.

Feitas essas considerações, é possível concluir que os sistemas algorítmicos de moderação do Instagram integram uma ampla arquitetura de controle ambiental da plataforma que tem *efeitos de governo* sobre as condutas dos usuários. O controle ambiental da rede social pelos sistemas algorítmicos visa a estimulá-los a se comportarem de determinada maneira, adotando condutas lucrativas para a rede social e *punindo*, por meio da moderação automatizada, aqueles que façam o contrário. Trata-se de um governo algorítmico das condutas que privilegia determinadas aparências físicas em detrimento de outras, sob a máxima de manter o engajamento e consequentemente a geração da matéria-prima da rede social: os dados.

3.2.

A subjetividade feminina produtiva no Instagram: Quem é a “boa sujeita de dados”?

Como visto, o governo algorítmico das condutas no Instagram reforça um tipo de usuário produtivo dentro do modelo de negócios da rede social, que se autoexponha, interaja com o conteúdo alheio e incite outros usuários a fazerem o mesmo, mas esse governo ganha contornos específicos quando tem por objeto usuárias identificadas como mulheres.

Estudando a resistência às relações de poder calcadas na vigilância e coleta de dados, Lauren Bridges identifica a provocação de falhas digitais no Instagram como uma tática para desconstituir-se um “bom sujeito de dados”¹⁴². Apesar de a autora não definir o termo, ela dá a entender que o “bom sujeito de dados” seria um tipo de subjetividade ideal a determinado modelo de negócios dentro do capitalismo de vigilância, isto é, um modo de se conduzir dentro de certo ambiente digital que

¹⁴² BRIDGES, Lauren E., “Digital Failure: Unbecoming the ‘good’ data subject through entropic, fugitive, and queer data”, pp. 1-17.

seja produtivo e lucrativo à respectiva empresa de tecnologia. Nesta Seção, pretendo retomar o conceito de Bridges e expandi-lo no contexto do Instagram, especificamente no que se refere à subjetividade produtiva esperada das usuárias identificadas como mulheres.

Considerando-se a dinâmica do capitalismo de vigilância, os comportamentos esperados dos “bons sujeitos de dados” em redes sociais como o Instagram em geral estão atrelados à sua capacidade de produzir e estimular a produção de superávit comportamental por parte de outros usuários. Essa capacidade está diretamente atrelada ao valor do usuário para a rede social: um perfil é mais lucrativo e valioso para o Instagram quanto mais ele se atenha às previsões algorítmicas, interaja com o conteúdo do aplicativo gerando dados e estimule outros usuários a engajar com o conteúdo e produzir dados. Inversamente, um perfil é menos lucrativo e valioso quanto menos atenda a tais padrões de produtividade. Além disso, como visto, o Instagram gira ao redor do compartilhamento da intimidade e da autoexposição espetacular. Assim, a produtividade de um usuário – sua capacidade de atrair atenção e fomentar a geração de dados – acaba estando atrelada em grande parte à sua imagem ou à estética do conteúdo que produz¹⁴³.

A partir dessas percepções, multiplicam-se os estudos na academia sobre o trabalho de usuários do Instagram que gira em torno da visibilidade do conteúdo postado¹⁴⁴. Crystal Abidin usa a expressão “trabalho de visibilidade” para se referir ao trabalho de usuários comuns no Instagram que engajam com anúncios da plataforma, inadvertidamente *trabalhando* para produzir ou amplificar a circulação de conteúdo comercial¹⁴⁵. Do mesmo modo é o trabalho dos influenciadores digitais no Instagram: aqueles que acumulam uma base de seguidores grande suficiente para monetizá-la¹⁴⁶. Trabalhando com visibilidade do seu próprio perfil e/ou do seu

¹⁴³ Ressalvado o limiar de conteúdo que choque ou radicalize os usuários da plataforma, que, como visto no Capítulo 1, foi identificado pelo Facebook como um problema a ser combatido dentro da rede social, dando impulso à criação de inteligência artificial destinada a reduzir a visibilidade desse tipo de conteúdo. A propósito, v. Seção 1.3. deste trabalho.

¹⁴⁴ ABIDIN, Crystal, “Visibility labour: Engaging with Influencers’ fashion brands and #OOTD advertorial campaigns on Instagram”. *Media International Australia*, v. 161, n. 1, 2016, pp. 86-100; COTTER, Kelley, “Playing the visibility game: How digital influencers and algorithms negotiate influence on Instagram”. *New Media & Society*, v. 21, n. 4, 2019, pp. 895-913; AR-RIAGADA, Arturo; IBÁÑEZ, Francisco, “‘You Need At Least One Picture Daily, if Not, You’re Dead’: Content Creators and Platform Evolution in the Social Media Ecology”. *Social Media + Society*, julho-setembro 2020, pp. 1-12; DUFFY, Brooke Erin, “Algorithmic precarity in cultural work”, *Communication and the Public*, v. 5, n. 3-4, 2020, pp. 103-107.

¹⁴⁵ ABIDIN, Crystal, Op. Cit., p. 86.

¹⁴⁶ LEAVER, Tama; HIGHFIELD, Tim; ABIDIN, Crystal. *Instagram*, p. 112.

corpo, eles se veem pressionados a desenvolver e manter em permanente inovação uma marca própria (*self-branding*) de seu estilo de vida, rotina ou personalidade, de onde poderá ser extraído valor¹⁴⁷, bem como a cumprirem com determinados valores estéticos na produção de seu conteúdo caso queiram empreender seu perfil com sucesso – isto é, visibilidade e engajamento. Em suma, o trabalho de visibilidade, segundo Abidin, é aquele que “os indivíduos fazem quando se auto-posicionam e curam suas autorrepresentações para que sejam notáveis e positivamente proeminentes entre potenciais empregadores, clientes, a imprensa ou seguidores e fãs, entre outras audiências”¹⁴⁸.

Portanto, os usuários do Instagram, advertidamente ou não, trabalham para a rede social na condição de produtores de dados e/ou de conteúdo: empreendendo sua imagem, quando aceita, apelativa e não ofensiva, ou apenas consumindo conteúdo, produzindo dados comportamentais e deixando rastros de suas ações – tudo isso integrante do superávit comportamental a ser aproveitado pelo Instagram. Há, contudo, diferenças na cobrança de comportamentos lucrativos dos usuários quando se tratam de usuárias identificadas como mulheres¹⁴⁹.

O governo das condutas das trabalhadoras mulheres do Instagram é em muito atrelado ao estímulo de uma sensibilidade “feminina” neoliberal¹⁵⁰, em que a mulher é vista como empresária de si, independente, livre e responsável por se gerir e

¹⁴⁷ BANET-WEISER, Sarah, “Branding the Post-Feminist Self: Girls’ Video Production and YouTube”, 2011, p. 13. Disponível em: <https://socyeventsucsc.files.wordpress.com/2011/02/banet-weiser_reading.pdf>. Acesso em: 24 mai. 2022.

¹⁴⁸ ABIDIN, Crystal, Op. Cit., p. 90.

¹⁴⁹ ABIDIN, Crystal, “Visibility labour: Engaging with Influencers’ fashion brands and #OOTD advertorial campaigns on Instagram”, p. 89; DUFFY, Brooke Erin, “The Gendered Politics of Digital Brand Labor”. *Antenna*, 18.03.2015. Disponível em: <<https://blog.com-marts.wisc.edu/2015/03/18/the-gendered-politics-of-digital-brand-labor/>>. Acesso em: 24 mai. 2022; DUFFY, Brooke Erin; HUND, Emily, “Gendered Visibility on Social Media: Navigating Instagram’s Authenticity Bind”. *International Journal of Communication*, v. 13, 2019, pp. 4983-5002.

¹⁵⁰ Sarah Banet-Weiser chama de pós-feminismo o conjunto de práticas e discursos que ligam ideias tradicionais de feminismo liberal com um engajamento contemporâneo com mídia, merchandising e consumo (BANET-WEISER, Sarah. *Authentic TM: the politics of ambivalence in a brand culture*. New York and London: New York University Press, p. 56). A propósito, Gill explica: “[...] *postfeminism is best understood as a distinctive sensibility, made up of a number of interrelated themes. These include the notion that femininity is a bodily property; the shift from objectification to subjectification; an emphasis upon self-surveillance, monitoring and self-discipline; a focus on individualism, choice and empowerment; the dominance of a makeover paradigm; and a resurgence of ideas about natural sexual difference. Each of these is explored in some detail, with examples from contemporary Anglo-American media. It is precisely the patterned articulation of these ideas that constitutes a postfeminist sensibility.*” (GILL, Rosalind, “Postfeminist media culture: elements of a sensibility”. *European Journal of Cultural Studies*, v. 10, n. 2, 2007, pp. 147-166)

educar de modo a obter sucesso. Componentes da subjetividade neoliberal sobre- põem-se às cobranças comportamentais e de estilização corporal impostas às mu- lheres. Se tradicionalmente o corpo tido como feminino já era apropriado pelo mer- cado, sendo objetificado e hipersexualizado para promover produtos e serviços¹⁵¹, na economia neoliberal, do sujeito como “empresa de si”¹⁵², associada ao uso de uma rede social imagética como plataforma de trabalho, essa exposição do corpo feminino é autoinfligida. A auto-objetificação das mulheres na exposição do pró- prio corpo, muitas vezes hipersexualizado, é vista como empoderamento e usada para obter “sucesso” em redes sociais. Outra marca dessa sensibilidade “feminina” neoliberal é a cobrança de contínuo aprimoramento de si, de auto-disciplina- mento¹⁵³, o que se traduz em cobranças relacionadas à aparência física quando se trata de mulheres e na ideia de responsabilidade pessoal por seus sucessos e fracas- sos, independentemente de condições materiais e sociais¹⁵⁴. A sujeita neoliberal disciplina seu corpo para obter sucesso, que só depende de si mesma.

No Instagram, a usuária ideal é a empreendedora individual que deve investir em seu *self-branding* para ter êxito econômico com sua autoexposição na rede so- cial. Este *self-branding* é muito ligado ao cumprimento de padrões estéticos, à ma- nutenção de algum grau de autenticidade e naturalidade no conteúdo produzido¹⁵⁵, à exposição de um estilo de vida estético e consumível e à auto-objetificação cor- poral que garantam visibilidade – e, portanto, maiores chances de lucratividade.

Portanto, uma particularidade nas exigências das trabalhadoras mulheres do Instagram é que, caso queiram empreender sua própria imagem dentro da rede so- cial, geralmente se veem restringidas por imagens de feminilidade às quais devem se enquadrar por meio da autodisciplina. Isto é, imagens relativamente fixas atribu- ídas à mulher socialmente percebida como tal no seio da economia capitalista neo- liberal: imagens da mulher como um objeto sexual, posando sensualmente nas fotos compartilhadas, seja qual for o contexto – para performar um evento cotidiano de

¹⁵¹ ABIDIN, Crystal; THOMPSON, Eric C., “Buymylife.com: Cyber-femininities and commercial intimacy in blogshops”. *Women’s Studies International Forum*, v. 35, 2012, pp. 474.

¹⁵² FOUCAULT, Michel. *Nascimento da Biopolítica*: Curso dado no Collège de France (1978-1979), p. 311.

¹⁵³ ABIDIN, Crystal; THOMPSON, Eric C., “Buymylife.com: Cyber-femininities and commercial intimacy in blogshops”, p. 474.

¹⁵⁴ SCHARF, Christina, “The Psychic Life of Neoliberalism: Mapping the Contours of Entrepre- neurial Subjectivity”. *Theory, Culture & Society*, v. 33, n. 6, 2016, pp. 107-122.

¹⁵⁵ DUFFY, Brooke Erin; HUND, Emily, “Gendered Visibility on Social Media: Navigating Insta- gram’s Authenticity Bind”, p. 4996.

sua vida para os seus seguidores ou promover um produto; imagens da mulher empoderada, em um tipo de feminismo neoliberalizado, em que importa mais engajar os usuários com imagens agradáveis de *body positivity*¹⁵⁶ e produtos com frases de efeito feministas¹⁵⁷; imagens de mulheres de biquini ou roupas de ginástica, promovendo um estilo de vida saudável, apelativo e consumível. Tudo isso é vendável para o Instagram, motivo pelo qual tais figuras de feminilidade são priorizadas na distribuição de conteúdo do Instagram e não na moderação. Invisíveis aos sistemas algorítmicos de moderação, as usuárias que incorporem as imagens de feminilidade têm maiores chances de deter plena visibilidade diante dos demais usuários.

Além disso, a posição de consumidoras é indissociável da “boa sujeita de dados” no âmbito do Instagram. O consumo é um ato de empoderamento dentro do pós-feminismo, o que é perceptível a partir da publicidade de produtos de beleza feminina, que passou a atrelar a aquisição de bens materiais à autoaceitação corporal, à pluralidade de formas de ser bela e ao autoempoderamento das mulheres. Como argumentam Abidin e Thompson, estando as mulheres nessa qualidade de consumidoras, a base das relações comerciais é o corpo feminino desprovido, portador de alguma *falta*, que precisa de algo para completá-lo ou consertá-lo. Assim, os sentimentos de baixa autoestima, complexo de inferioridade, ódio ao próprio corpo e idealização de padrões de beleza feminina são apropriados ou fomentados pelas estratégias de marketing no Instagram – sejam elas implementadas pelos perfis dos anunciantes ou pelas próprias usuárias quando postam conteúdo promocional – para estimular desejos materialistas e o consumo¹⁵⁸.

Em razão disso, a lucratividade no Instagram muitas vezes está atrelada ao ódio da mulher ao próprio corpo ou à performance de um tipo de feminilidade que atenda a signos já domesticados pelo capitalismo atrelados ao consumo de bens materiais. O conteúdo que promova chás emagrecedores, cirurgias plásticas, harmonizações faciais e procedimentos estéticos, por mais prejudicial que seja a meninas e mulheres, continua no ar e impulsionado pelos sistemas algoritmos. Do

¹⁵⁶ VENDEMIA, Megan A.; DEANDREA, David C.; BRATHWAITE, Kyla N., “Objectifying the body positive movement: The effects of sexualizing and digitally modifying body-positive images in Instagram”. *Body Image*, v. 38, 2021, pp. 137-147; CWYNAR-HORTA, Jessica, “The Commodification of the Body Positive Movement on Instagram”. *Stream: Culture, Politics, Technology*, v. 8, n. 2, 2016, pp. 36-56.

¹⁵⁷ BANET-WEISER, Sarah, “Popular Feminism and Popular Misogyny”. In: RIBEIRO, Nelson; HENRICHSEN, Jennifer R. *Media and Populism: 1st Lisbon Winter School for the Study of Communication*. 15-19 de Janeiro de 2019, p. 72.

¹⁵⁸ BANET-WEISER, Sarah. *Authentic TM*, p. 56.

mesmo modo, fotos e vídeos em que uma influenciadora claramente identificada como mulher exiba orgulhosamente seu modo de vida burguês, rodeado de bens materiais, patrocinadores e comodidades, é rapidamente circulado, por mais que haja nudez envolvida – o que a princípio violaria as Diretrizes da Comunidade. Esse é o tipo de conteúdo *rentável*, que atrai engajamento de outros usuários, e, portanto, *superávit comportamental*, além de atrair anunciantes para a rede social.

Existe, portanto, uma produção de trabalhadores do Instagram, que assume particularidades quando se trata de usuárias mulheres. No Instagram, a “boa sujeita de dados” ou bem tem um corpo legível pela moderação e de valor econômico óbvio, e por isso pode expor-se à vontade para atrair usuários, anunciantes e gerar superávit comportamental (portanto, lucro) à rede social, ou bem tem um corpo ilegível, às margens da inteligibilidade da moderação, e deve permanecer não vista, para não chocar, ofender, ou violar o regime ocultado de visibilidade, mas sempre consumindo o conteúdo e anúncios da plataforma e entregando dados, de modo a sustentar a geração de lucro.

Essa subjetividade produtiva associada à usuária feminina é resultado de processos de normalização/normatização dentro do Instagram, conduzido em grande parte por seus sistemas algorítmicos de moderação de conteúdo, que desempenham um papel punitivo às usuárias destoantes dessa subjetividade. Por meio da promoção do conteúdo que exponha corpos femininos enquadrados em certos padrões estéticos [além de raciais e de gênero] e do ocultamento daqueles que destoem das normas sociais presentes na rede social, o controle ambiental do Instagram forja a subjetividade ideal ao seu modelo de negócios. As usuárias às quais são oferecidas maiores oportunidades de autoexposição aprendem a se conduzir de forma produtiva para a rede social, a forma que atrai visibilidade e engajamento, que se apoia sobre o consumismo “feminino” e a atualização de padrões estéticos e comportamentos atrelados a figuras de feminilidade já conhecidas no âmbito do capitalismo. Já as usuárias que são punidas por sua autoexposição desconforme às regras ocultadas de visibilidade aprendem quais são os padrões e comportamentos valorizados pela rede social e se adequam para obter visibilidade ao seu conteúdo ou mantêm-se simplesmente consumindo conteúdo alheio na rede social. Todas, sempre, produzindo dados.

Se por um lado os estímulos do Instagram, articulados por meio de sistemas algorítmicos de recomendação e moderação de conteúdo, encaminham a conduta

das usuárias para que elas se conduzam como “boas sujeitas de dados”, da forma mais lucrativa possível, por outro, elas não são simplesmente subjugadas pelos sistemas algorítmicos. Pelo contrário, estão em constante relação com eles, aprendendo seu *modus operandi* e desenvolvendo estratégias para obter visibilidade para conteúdo crítico ou improdutivo dentro das tendências dos sistemas. É a partir dos erros e inconsistências do governo algorítmico das condutas que as usuárias encontram brechas para se constituírem fora da “boa sujeita de dados”, isto é, para se conduzirem em seus próprios termos.

3.3.

Complicação dos processos de subjetivação e possibilidades de resistência

Nesta seção, a partir de uma crítica às reflexões de Berns e Rouvroy sobre a complicação dos processos de subjetivação na governamentalidade algorítmica, pretendo pensar onde se encontram as oportunidades de os usuários do Instagram se perceberem como submetidos a uma relação de poder algorítmica, iniciarem uma reflexão crítica sobre sua condição e agirem sobre si mesmos, ainda que o governo algorítmico das condutas imponha dificuldades a essa atividade. Ou seja, onde se encontram as oportunidades de subjetivação dos usuários, de constituição de si mesmos como sujeitos, não meramente assujeitados ao governo algorítmico das condutas. Onde, afinal, é possível encontrar resistência.

Segundo Rouvroy e Berns, o governo algorítmico das condutas complica as possibilidades de subjetivação dos indivíduos. Isso porque tal governo resulta em um “declínio da reflexividade subjetivante” dos indivíduos e no “distanciamento das ocasiões de contestação das produções de ‘saber’ fundadas no *datamining* e na elaboração de perfis”. Para os autores, portanto, as possibilidades de subjetivação estariam ligadas à *reflexividade* e à *crítica*: à propensão do indivíduo de pensar a si mesmo como um sujeito moral, em sua condição dentro do governo algorítmico; de pensar de forma crítica como se é conduzido, como não se gostaria de ser conduzido; de adotar uma postura *crítica* e determinar a própria conduta em sentido contrário à governamentalidade dos sistemas algorítmicos.

Com essas expressões, portanto, os autores parecem referir-se ao conceito foucaultiano de subjetivação. Mais especificamente, ao sentido de subjetivação desenvolvido pelo filósofo a partir de 1978, nos cursos ministrados no Collège de France¹⁵⁹, em muito relacionado às suas reflexões sobre governamentalidade, condutas e postura crítica a práticas de governo, desenvolvidas na mesma época¹⁶⁰. Segundo esse sentido, a subjetivação seria uma prática de si, um modo de se constituir sujeito em contínua relação crítica com o governo de suas condutas, em vez de ser simplesmente sujeitado por essa relação de poder. A subjetivação adviria de exercícios do sujeito sobre si mesmo, da capacidade de ele tomar-se a si como objeto de conhecimento e agir sobre si mesmo¹⁶¹. Por meio desses exercícios, o sujeito transforma sua relação consigo e se conduz em seus próprios termos.

Levando em consideração esse conceito de subjetivação, para Rouvroy e Berns, os processos de subjetivação dentro da governamentalidade algorítmica seriam complicados porque o funcionamento desse governo das condutas, a princípio, não abre espaços para que o sujeito reflita e aja sobre si mesmo. É por isso que os autores falam em “declínio da reflexividade subjetivante”.

Como visto, a governamentalidade algorítmica evita os sujeitos reflexivos em todas as suas etapas, pois estes servem apenas como fontes de dados infraindividuais usados para a construção de perfis supraindividuais, que guiarão as previsões de comportamento. A governamentalidade algorítmica não reconhece o sujeito como agente moral, não o interpela nem convoca a refletir sobre o que ele é dentro desse sistema de poder ou o que poderia vir a ser¹⁶². Antes, trata os indivíduos como meros agregados de dados que podem levar a previsões comportamentais por correlações estatísticas. O campo de possibilidades de agência dos indivíduos já lhes

¹⁵⁹ Segundo Edgardo Castro, na obra de Foucault, é possível distinguir dois sentidos para os “modos de subjetivação” por ele estudados. No primeiro sentido, os modos de subjetivação consistiriam em “modos de objetivação do sujeito, isto é, modos em que o sujeito aparece como objeto de uma determinada relação de conhecimento e de poder”. No segundo sentido, os modos de subjetivação, seriam “formas de atividade sobre si mesmo”, abrangendo as técnicas de si, os procedimentos por meio dos quais o sujeito se constitui como sujeito, se transforma a si mesmo (CASTRO, Edgardo. *Vocabulário de Foucault*, pp. 408-409).

¹⁶⁰ NETO, João Leite Ferreira, “A Analítica da Subjetivação em Michel Foucault”, *Revista Polis e Psique*, v. 7, n. 3, 2017, p. 11-12.

¹⁶¹ CASTRO, Edgardo. Op. Cit., pp. 408-409.

¹⁶² ROUVROY, Antoinette. “La vie n’est pas donnée”, p. 209; ROUVROY, Antoinette e BERNIS, Thomas, “Governamentalidade algorítmica e perspectivas de emancipação: o dispar como condição de individuação pela relação?”, p. 42.

seria entregue de antemão, sem necessidade de hesitação, de ponderação, formulação de desejos e tomada de decisão¹⁶³. As oportunidades de subjetivação seriam complicadas porque são escassas as situações nas quais os indivíduos se deparam com algo que os faz refletir sobre sua condição de sujeitos digitais. Tais oportunidades de subjetivação geralmente surgem quando a lacuna entre realidade e descrição da realidade se torna aparente: onde não há significado, onde existe o encontro com o imprevisível e a necessidade de formulação de respostas. É neste momento que surge a reflexividade crítica que permite a subjetivação.

É possível perceber a redução das possibilidades de subjetivação no Instagram quando se olha para o funcionamento dos sistemas de recomendação e moderação de conteúdo. Os usuários são bombardeados com vídeos e fotos semelhantes àqueles que têm sido procurados por eles dentro e fora da plataforma ou relacionados a interesses prováveis. Os estímulos audiovisuais lhes são expostos para que respondam prontamente com alguma ação dentro da rede social: curtir, compartilhar, salvar, clicar. Cada ação deixa sinais comportamentais que serão aproveitados pela empresa em sua geração de receitas. Por isso, o objetivo é mostrar mais do mesmo, mais do que já é interessante, do que já se encontra no raio de pensamento do usuário, do que estatisticamente tem mais chances de lhe incitar alguma reação pré-calculada.

Se o objetivo da rede social é “otimizar” a experiência dos usuários mostrando-lhes apenas conteúdo em que eles têm maior probabilidade de ter interesse e engajar (para gerar dados e lucro), não há lugar para o encontro com imagens e textos *alheios* ao que já estão em nossa esfera de interesse e na de indivíduos com interesses semelhantes. Os usuários, portanto, dificilmente entram em contato com fotos e vídeos que jamais pensariam em ter interesse ou que sejam radicalmente diferentes de sua realidade. Além disso, a plasticidade das recomendações de conteúdo no Instagram, que se adequam aos interesses momentâneos do usuário com

¹⁶³ A plasticidade da governamentalidade algorítmica, em que os sistemas algorítmicos continuamente se adaptam ao tempo real, às condutas e interações flutuantes, também dificulta o momento de crítica. Não há espaço para a crise, para a reflexão sobre como se gostaria de ser governado [ou não governado]. (FOUCAULT, Michel. *The Politics of Truth*. New York: Semiotext(e), 1997, p. 28; HARCOURT, Bernard. “Epilogue 2/13”. *Critique & Praxis* 13/13, 27.10.2018. Disponível em: <<http://blogs.law.columbia.edu/praxis1313/bernard-e-harcourt-epilogue-2-13/>>. Acesso em: 17 nov. 2021; ROUVROY, Antoinette, “The end(s) of critique: data behaviourism versus due process”. In: HILDEBRANDT, Mireille; VRIES, Katja de. *Privacy, Due Process and the Computational Turn: The philosophy of law meets the philosophy of technology*. New York: Routledge, 2017, p. 168).

base em seu comportamento, os mantém dentro de uma bolha de desejos previsíveis. Desse modo, não há encontro com o desconhecido, o imprevisto, o inexplorado, aquilo que faz surgir um momento de expansão de horizontes, de ausência de respostas e significados, de reflexividade crítica e elaboração de desejos. A lacuna entre o encontro com o conteúdo e a formulação de uma resposta é abreviado: o conteúdo exposto já tem, estatisticamente, maior probabilidade de suscitar uma resposta calculada.

No Instagram, portanto, dificilmente há o encontro com o que poderia causar afetos imprevistos e, conseqüentemente, nos causar uma transformação, ou melhor, desencadear um processo de subjetivação, de reflexão crítica a propósito da própria condição dentro da rede social e de trabalho sobre si mesmo. Os sistemas algorítmicos de moderação de conteúdo, em conjunto com as técnicas de curadoria e recomendação organizam o ambiente imagético e informacional do Instagram para provocar afetos antecipadamente calculados.

Assim, as possibilidades de suspensão de significados sociais dominantes, autorreflexão e formulação de críticas são constantemente postergadas. O usuário normalmente não se confronta com situações nas quais seja chamado a refletir sobre sua condição de assujeitado pela rede social, de indivíduo cuja conduta é moldada pelos sistemas algorítmicos sobre os interesses do Instagram. Afinal, essa é apenas uma rede social voltada ao entretenimento fugaz, momentâneo, que simplesmente organiza conteúdo visual que será usado como passatempo. Nada no Instagram imediatamente diz sobre a relação de poder entabulada pela empresa com os usuários do aplicativo por meio dos sistemas algorítmicos. Pelo contrário, o que é dito pela empresa em meios oficiais visa sempre negar pensamentos como esse.

Apesar disso, indo além do que Rouvroy e Berns dizem, é possível conceber as práticas de resistência de determinados usuários à arquitetura algorítmica como processos de tomada de consciência sobre o seu posicionamento individual dentro dessa relação de poder. Como será ilustrado no Capítulo 4, alguns usuários, percebendo-se conduzidos por sistemas algorítmicos de moderação de conteúdo em redes sociais, interagem criticamente com eles, seja testando seu funcionamento, produzindo dados falsos ou desenvolvendo estratégias para burlar a lógica algorítmica e os significados atualizados por esses sistemas. Essas ações consistem em uma tomada de consciência sobre a condição de indivíduo *assujeitado* a certa relação de

poder digital – no caso, com o Instagram. São ações de resistência a partir do próprio ambiente que assujeita, expondo o governo das condutas que se nega como tal.

Ainda a respeito da supressão das oportunidades de subjetivação, segundo Rouvroy e Berns, existe um paralelo importante entre o modo pelo qual as condutas são moldadas em sistemas jurídicos, isto é, por normas legais, e o modo como o são em sistemas algorítmicos e quais são as implicações para as oportunidades de subjetivação. No primeiro caso, o sujeito é confrontado diretamente com um comando ou vedação e, ponderando os riscos de inobservância, opta ou não por seguir aquele ditame: existe uma oportunidade de reflexão e agência. Na governamentalidade algorítmica, por outro lado, essa margem de escolha não é mantida, porque sua atuação ocorre de dentro do sujeito para fora. Os estímulos geram afetos internos, que resultam na adoção ou não de determinado comportamento já previsto. Não há confronto com afetos imprevistos e necessidade de escolha entre adotar uma conduta ou outra, porque os afetos e comportamentos possíveis já foram calculados e o ambiente foi arquitetado para que tudo seguisse as previsões. Assim, o governo algorítmico minimizaria a incerteza dos afetos e da agência humanos, justamente o campo onde as possibilidades de subjetivação surgem.

Ao contrário do que os autores dizem, no entanto, é possível pensar que o governo algorítmico das condutas *umenta* a incerteza dos afetos e da agência humanos. Rouvroy e Berns extraem suas conclusões de um ponto de vista que parte do funcionamento dos sistemas algorítmicos, mas, se pensarmos a governamentalidade algorítmica no Instagram a partir do *comportamento dos usuários* cujos dados são o objeto desse governo, outras reflexões vêm à tona.

A princípio, as regras do aplicativo que vedam determinado tipo de conteúdo estão nas Diretrizes da Comunidade e os usuários podem optar por segui-las ou não, assumindo os riscos de desobediência, como a exclusão da imagem ou vídeo postados. Ainda assim, existem regras *não escritas* de visibilidade, perceptíveis a partir do funcionamento da moderação automatizada, como abordado no Capítulo 1. Portanto, por mais que um usuário deliberadamente opte por seguir as regras escritas a fim de manter seu conteúdo no ar, pode ser que seja punido pela moderação automatizada.

Por um lado, a moderação algorítmica minimiza a incerteza da agência humana quando pune a exposição de imagens que a princípio seriam permitidas pelas Diretrizes da Comunidade, porque ajuda a conduzir os usuários a se comportarem

de determinada maneira. Ao mesmo tempo, no entanto, a moderação algorítmica *aumenta* a incerteza da agência humana dentro da rede social. Dado que o funcionamento das técnicas algorítmicas de moderação de conteúdo é padronizado, automático, segue uma fórmula que se alimenta de dados, os usuários pouco a pouco aprendem seu funcionamento – por mais efêmero que seja, devido a constantes ajustes feitos pelas equipes de tecnologia da empresa – e se apropriam de um campo de incerteza não abrangido pelos sistemas, até que esse campo de incerteza seja incorporado por eles.

Os sistemas algorítmicos de moderação de conteúdo buscam estabilizar o ambiente digital, suprimindo e ocultando de forma padronizada determinados tipos de conteúdo. De forma paralela, os usuários continuamente desenvolvem táticas para burlar as regras que apreendem a partir de sua experiência com a tecnologia e tornam visível conteúdo que a princípio seria suprimido ou ocultado. Isso é feito até que o ciclo se reinicie, os algoritmos sejam ajustados para identificar tentativas de os burlarem e as regras de moderação e distribuição de conteúdo sejam alteradas. Esse movimento é contínuo: o governo algorítmico do Instagram se nutre da resistência dos usuários e vice-versa.

Existe, afinal, encontro com o improvável: o improvável para determinados usuários é a restrição de visibilidade de um conteúdo formalmente em conformidade com as Diretrizes da Comunidade, mas identificado pela moderação algorítmica como violação. Como visto no Capítulo 1, por mais que um perfil só publique imagens que estejam dentro dos limites das regras escritas de visibilidade do Instagram, pode ser punido pela moderação automatizada com base regras ocultadas de visibilidade: aquelas que replicam significados sociais hegemônicos herdados da colonialidade.

O improvável para eles, portanto, não é o encontro com conteúdo que destoe de interesses já identificados pela rede social, mas sim o encontro com a punição que não era esperada, uma punição “ilegal”, desconforme às regras escritas de moderação e conforme às regras ocultadas. É neste improvável que reside a oportunidade de defrontar-se com a própria condição de assujeitado dos sistemas algorítmicos, de agregado de dados conduzido pelo governo das condutas do Instagram. É nele que surge a possibilidade de reflexividade crítica: tomar-se a si mesmo, na condição de assujeitado digital, como objeto de análise e trabalho. A possibilidade de subjetivação dos usuários advém justamente da tentativa de impossibilitá-la.

Portanto, existem diferentes ordens de imprevisível dentro do Instagram. De um lado, existe o conteúdo não previsto pelos sistemas de moderação de conteúdo que não viola as Diretrizes da Comunidade, mas é identificado como violador ou como inapropriado, e, por isso, punido com a supressão ou ocultamento. Por outro, existe o imprevisível que parte dessa punição “ilegal”: aprende o funcionamento do regime ocultado de visibilidades, encontra suas brechas e o engana.

Um sistema que busca continuamente suprimir o imprevisível acaba suscitando o imprevisível em suas lacunas. É neste imprevisível que se encontra escape aos afetos calculados e ao adiamento da reflexividade crítica. São as insuficiências dos sistemas algorítmicos que suscitam a crítica, o espaçamento entre realidade e suposta descrição por meio dos números e dados. No caso da curadoria e moderação de conteúdo do Instagram, as oportunidades de crítica, de reflexão, de elaboração de desejos, muitas vezes surgem dos silenciamentos e das insuficiências dos sistemas algorítmicos, que são vividos em grande parte por aqueles que são hipervisíveis aos modelos algorítmicos como violações por não cumprirem com os seus pressupostos. Lá, abundam os exemplos de resistência à governamentalidade algorítmica do Instagram e processos de subjetivação, como será visto no Capítulo a seguir.

Autoexposição e resistência das usuárias do Instagram

Em *Autodefesa: Uma filosofia da violência*, Elsa Dorlin faz uma genealogia das técnicas alternativas de defesa de si, não abrangidas pelo conceito jurídico de “legítima defesa” nem pelos discursos e práticas sociais que o sustentam¹⁶⁴. Seu projeto parte da constatação de que existe uma distribuição desigual da defesa legítima de si, baseada nas relações sociais forjadas pela colonialidade. A autora faz uma diferenciação entre legítima defesa e *autodefesa*. A primeira seria um privilégio de determinados segmentos da população, vistos como sujeitos de direito cuja reação a uma violência é sempre legítima, mera reação de proteção do próprio corpo, sua propriedade. A autodefesa, por outro lado, é o que resta às “subjetividades desarmadas”, aqueles indivíduos que foram desarmados pelas práticas sociais e ordenamento jurídico-político, cuja defesa de si é vista sempre como “ato primário de violência”, nunca como resposta legítima a uma agressão injusta¹⁶⁵.

Assim, a proposta de Dorlin, por meio do estudo de estratégias de resistência de minorias sociais, é formular uma ética da autodefesa subterrânea¹⁶⁶, das técnicas de defesa de si praticadas por minorias sociais na experiência diária da violência colonial ou neocolonial. Ela adota o que chama de uma abordagem feminista, que analisa a resistência – ou, em suas palavras, a autodefesa – fora do que é tradicionalmente situado no político. Antes, sua pretensão é explorar uma concepção de autodefesa contínua, experimentada não apenas no evento de uma violência física e imediata ou em contextos tradicionalmente tidos como políticos, mas sim na vivência cotidiana, onde o político não é localizado.

Isso porque a questão da defesa, para muitos indivíduos, não se encerra nos espaços coletivos de enfrentamento, mas é levada para dentro de casa, para as relações interpessoais, para espaços e relações tidos como exclusivamente privados, mas que são constituídos pela política e pelo coletivo. Assim, sua proposta é pensar

¹⁶⁴ DORLIN, Elsa. *Autodefesa: Uma filosofia da violência*. São Paulo: Ubu Editora, 2020, p. 52.

¹⁶⁵ *Ibid.*, p. 27.

¹⁶⁶ COELHO, Bruna Martins, “A Filosofia a Golpes de Martelo de Elsa Dorlin: ‘se defender’ contra o terrorismo de Estado – resenha crítica”. *Revista TEL*, v. 10, n. 2, jul.-dez. 2019, p. 286.

“não na escala dos sujeitos políticos constituídos, e sim na politização das subjetividades: no cotidiano, na intimidade, da raiva presa em nós mesmo(a)(s), na solidão das experiências de violência vividas, diante da qual praticamos continuamente uma autodefesa que não se rotula”¹⁶⁷.

A autodefesa de que Dorlin fala não pressupõe a existência de um sujeito que se defende. Esta é uma condição da legítima defesa, mas não da autodefesa. Na autodefesa, o indivíduo que resiste a uma agressão não é um sujeito de direito, não tem a posse sobre o seu próprio corpo e, portanto, não pode protegê-lo legitimamente, com respaldo do Estado e de sua estrutura jurídica e política. Em razão disso, a tais indivíduos restam constituírem-se como sujeitos – se subjetivarem – na explosão de sua resposta à violência sofrida. É nessa resposta que agem sobre si mesmos como indivíduos assujeitados em uma relação de poder e instauram outra relação consigo mesmos, com o seu meio e com o outro: subjetivam-se.

Além disso, em sua tese a autodefesa não consiste em um meio para chegar a determinado fim, mas sim em uma prática de politização de “órgãos sem mediação, sem delegação, sem representação”¹⁶⁸. Às subjetividades desarmadas muitas vezes sequer é possível obter qualquer conquista política por meio da reação a uma violência. A autodefesa pode então ser vista como prática cotidiana que evita a aniquilação de si, como um ato de reafirmação da própria existência, de constituir-se a si mesmo como um sujeito político.

Proponho a expansão deste conceito de Dorlin para conceber atos individuais de resistência de mulheres no Instagram ao governo algorítmico de suas condutas. Em vez de pensar a autodefesa em espaços a princípio tidos como apolíticos mas físicos, pretendo pensá-la em um espaço politizado em alguma medida – o Instagram –, mas cujo caráter político se esconde sob argumentos de ser uma rede social voltada ao entretenimento, bem como sob discursos tecnoutópicos que se recusam a admitir os danos sociais causados pelos sistemas algorítmicos de moderação.

Trata-se de pensar em uma *autodefesa digital*. Por meio da autodefesa digital, usuárias do Instagram conseguem subjetivar-se em seus próprios termos, apesar das condições desfavoráveis do governo algorítmico das condutas, bem como expor o poder algorítmico como poder. Em vez de técnicas marciais de si, “técnicas digitais

¹⁶⁷ DORLIN, Elsa, *Autodefesa*, p. 30.

¹⁶⁸ *Ibid.*, p. 108.

de si”¹⁶⁹, empregadas contra a imposição de significados do governo algorítmico, a complicação dos processos de subjetivação, a colonialidade inerente ao funcionamento dos sistemas de moderação, a lógica do lucro que dita os padrões estéticos autorizadores de visibilidade, e assim por diante.

Por meio do conceito de *autodefesa* de Elsa Dorlin, neste Capítulo analiso a possibilidade de usuárias do Instagram que tenham sido alvo de tratamento discriminatório pela moderação algorítmica *se defenderem e responderem* a essa prática. Em linha com as reflexões da autora, argumento que essa forma de autodefesa digital não apenas é uma tentativa de se defender de práticas discriminatórias da moderação automatizada do Instagram, como também é um processo de subjetivação das blogueiras em um governo [algorítmico] das condutas que ignora sua existência como sujeitos digitais e complica os processos de subjetivação.

Apesar da governamentalidade algorítmica em funcionamento no Instagram complicar os processos de subjetivação e favorecer a atualização de uma imagem apolítica das mulheres, há estratégias, como as analisadas nos exemplos a seguir, que escapam pelas falhas e inconsistências do sistema algorítmico para constituir um sujeito político. Além disso, tais estratégias de resistência operam uma exposição do poder que se apoia sobre a autoexposição das usuárias: apesar de a governamentalidade algorítmica se camuflar sob uma pretensa objetividade dos dados, a descrição pura da realidade e discursos técnicos, há usuárias que o expõem como poder, jogando luz sobre a condução das condutas operada por ele e sobre as discriminações perpetuadas.

Neste Capítulo, são abordados dois casos reais em que usuárias do Instagram usaram seus perfis na rede social para resistir a algum aspecto discriminatório da moderação automatizada e se afirmarem como sujeitas políticas. Em primeiro lugar, tratarei do exemplo de Polly Oliveira, em que o humor e a ironia são usados em performances visuais para contestar o regime de visibilidades da plataforma. Em segundo lugar, analisarei o exemplo de Morgan Barbour, em que o digilantismo e a fotografia artística são empregados para criticar a inércia do Instagram diante das violências misóginas que sofria na rede social.

¹⁶⁹ Agradeço ao professor Antonio Pele pela sugestão dessa nomenclatura e pelas observações sobre a aplicabilidade desse conceito.

4.1.

Polly Oliveira: Performances humorísticas

Na análise do caso de Polly Oliveira, pretendo explorar o uso do humor como autodefesa ao silenciamento do *shadowban* do Instagram. Argumento que, ao expor o seu corpo em composições irônicas, a blogueira consegue avocar legitimidade discursiva para que suas críticas sejam ouvidas na rede social. Desse modo, ela não apenas adota um processo de subjetivação de si, tentando determinar sua conduta em seus próprios termos, como também expõe traços do governo algorítmico das condutas de usuárias no Instagram.

Polly Oliveira é uma influenciadora digital brasileira que usa seu perfil no Instagram (@pollyoliveirareal) para publicar imagens, vídeos e textos que discutem temas caros a muitas mulheres, como a autoaceitação corporal e a violência física e psicológica contra a mulher, além de problemas relacionados aos sistemas algorítmicos da rede social, como a distribuição ampla de conteúdo que promova ideais estreitos de beleza e produtos prejudiciais à autoestima e saúde femininas em detrimento dos conteúdos que externalizem um posicionamento crítico.

Em 31 dezembro de 2020, o perfil de Polly Oliveira foi suspenso pelo Instagram após ela ter republicado o vídeo de outra blogueira, criticando-a por propagar padrões de beleza e comportamento irreais às mulheres. No dia seguinte à postagem, sua conta foi desativada, e assim permaneceu por 20 dias, vindo a ser reestabelecida apenas em janeiro de 2021, depois de Polly Oliveira entrar em contato com funcionários da plataforma, que alegaram ter sido um erro a suspensão da conta¹⁷⁰.

Esse, contudo, foi apenas o ápice de uma série de medidas de silenciamento da moderação contra a influenciadora digital. Segundo ela, há dois anos suas publicações eram recorrentemente excluídas pelo Instagram, a maioria das quais retratavam a blogueira em biquínis ou roupas íntimas, em postagens voltadas a promover

¹⁷⁰ Por mais que não se saiba ao certo o que levou à suspensão da conta, algumas medidas anteriores do Instagram podem ser indicativas. Além de Polly Oliveira ter tido suas postagens recorrentemente apagadas nos últimos dois anos por violar as regras que vedam a nudez feminina (ainda que ela estivesse de biquíni ou lingerie nas imagens), o vídeo da outra blogueira que havia sido republicado continha nudez, além de o fato de republicar o conteúdo de outro usuário, por si só, poder ter sido considerado uma violação das Diretrizes da Comunidade, que dispõe: “Compartilhe somente as fotos que você tirou e os vídeos que gravou ou conteúdo que você tem direito de compartilhar. [...] Publique conteúdo autêntico e não publique nada copiado ou obtido da internet sem ter o direito”. (INSTAGRAM, “Diretrizes da Comunidade”. Disponível em: <<https://www.facebook.com/help/instagram/477434105621119>>. Acesso em: 9 abr. 2022)

a naturalização de traços físicos desvalorizados por padrões estéticos hegemônicos, como cicatrizes, manchas de pele e dobras cutâneas. Esse padrão de censura teve seu auge na suspensão da conta da blogueira, mas, mesmo após sua restauração, o conteúdo que ela publicava passou a ter sua visibilidade radicalmente reduzida – Polly Oliveira passou a ser alvo do *shadowban*.

Cada vez mais inconformada com a opacidade e o duplo padrão da moderação algorítmica do Instagram, ela percebeu que, enquanto suas postagens eram derrubadas, aquelas nas quais o corpo feminino era exposto de forma comercial eram livremente circuladas. A partir dessa percepção, a blogueira repensou a forma como passaria a abordar os seus temas e os meios que usaria para fazer suas mensagens chegarem até as seguidoras do seu perfil e outros usuários da rede social: passaria a usar a lógica dos algoritmos a seu favor, vestindo os atributos físicos que tornassem seu corpo aceitável para a moderação – ou melhor, invisível à moderação e visível aos demais usuários.

Assim, como estratégia de resistência ao silenciamento que vinha sofrendo, ela iniciou um projeto batizado de “O Experimento”. Nele, Polly Oliveira passou a postar imagens e vídeos de si mesma que autorizariam a sua plena visibilidade, ou seja, em que seu corpo estivesse alinhado à estética do corpo feminino usualmente promovido pela rede social ou em que ela anunciava produtos falsos de beleza feminina, contribuindo com a economia do Instagram. As publicações, no entanto, eram acompanhadas de legendas e composições irônicas criticando o silenciamento e os valores subjacentes à moderação algorítmica de conteúdo. Dessa maneira, ela conseguiria burlar o *shadowban* e fazer a sua crítica ser ouvida, garantindo que chegasse às suas seguidoras e a outros usuários.

Em alguns dos experimentos (Figuras 1 e 2), a blogueira postou fotos suas, editadas no *photoshop* para que parecesse mais magra, performando uma personagem recém submetida a múltiplas cirurgias estéticas, que compartilhava a experiência com suas seguidoras enquanto as incentivava a amarem seus corpos do jeito que são. Em uma das imagens (Figura 1), Polly Oliveira posa sorridente com um curativo no nariz, vestindo uma cinta de compressão e segurando orgulhosamente duas bolsas ligadas a drenos que levavam à sua barriga; em outra (Figura 2), ela posa sensualmente em um maiô, usando uma peruca de cabelos longos e lisos expressando satisfação com seus falsos procedimentos estéticos.

Com publicações como essas, às quais se seguiram muitas outras semelhantes, Polly Oliveira parodiava influenciadoras digitais brasileiras proeminentes no Instagram, repetindo, na legenda, frases comumente atribuídas a elas que já se tornaram clichês no ramo, como: “Estou aqui linda e magra porque eu me esforcei muito e fiz isso porque me amooooo [...] Eu acho que todas as mulheres precisam se aceitar e amar seu corpo do jeitinho que ele é [...] Meu cirurgião e a equipe dele foram maravilhosos comigo”. Seu intuito com as composições era criticar a predileção dada pelo Instagram a influenciadoras que, apesar de adotarem um discurso de autoaceitação corporal feminina, promovem intervenções estéticas e cosméticos desnecessários que entreguem corpos alinhados a padrões hegemônicos de beleza.

Com a performance exagerada de clichês associados a usuárias do Instagram que têm ampla visibilidade, ela evidencia e critica a lógica do lucro que rege a curadoria e moderação automatizada de conteúdo na rede social, além de suscitar reflexões sobre a hierarquia de humanidades, baseada em gênero e raça, que é reproduzida pelos sistemas algorítmicos do Instagram. Os artificios que ela elege para montar uma publicação que agradará a moderação – ou melhor, se tornará invisível a ela – e alcançará ampla visibilidade perante os demais usuários, afinal, têm os seus propósitos.

A padronização do seu corpo com o *photoshop*, para torná-lo mais magro e seus traços mais finos – mudanças obtidas por meio de uma falsa cirurgia – chama atenção para o tipo de corpo feminino que é valorizado pela moderação da rede social. Como visto nos relatos e nos relatórios de pesquisa abordados no Capítulo 1, o corpo branco cis e heterossexual, sobretudo conformado a padrões de beleza heteronormativos, passa facilmente pela malha da moderação algorítmica do Instagram, ainda que seminu ou em posições sensuais – o que, a princípio, é vedado pela literalidade das Diretrizes da Comunidade. Enquanto isso, imagens de mulheres trans totalmente vestidas, principalmente se racializadas ou gordas, são ocultadas ou excluídas por um “erro” dos sistemas algorítmicos, que as identificam como conteúdo sexual, violações a serem punidas.

Tanto o afinamento de seus traços quanto o alisamento do cabelo de Polly Oliveira também suscitam reflexões sobre a estética racista e eurocêntrica priorizada no controle de conteúdo da rede social – seja pelos sistemas automatizados ou pela resposta destes às denúncias feitas por outros usuários. Os filtros que tornam

peles mais claras¹⁷¹ e o *shadowban* a produtores de conteúdo negros¹⁷² são algumas das manifestações dessa prioridade dentro do Instagram. A mensagem subjacente à performance de Polly Oliveira é que a atribuição de valores e oportunidades discursivas é diferente conforme a cor e os traços fenotípicos do corpo exposto na plataforma. Existe uma maior valorização de corpos femininos brancos, que podem participar da economia expondo-se aos demais usuários e lidando com o público, mas os corpos femininos negros, de cabelos afro e traços fenotípicos não-brancos só podem participar da economia digital do Instagram nas encolhas, no *backoffice*, a partir das margens, fornecendo e consumindo dados.

Não pretendo sugerir que todos os corpos de mulheres negras recebam o mesmo tratamento pela moderação do site, mas apenas que há uma desvalorização da negritude feminina em comparação com a branquitude, a quem a visibilidade é presenteada com muito mais facilidade no Instagram. Existe, é claro, a sobreposição de opressões e silenciamentos, sendo corpos pretos e *queer*, gordos ou portadores de deficiências, por exemplo, mais escondidos do que aqueles que não os sejam.

Por fim, a retórica de Polly Oliveira, satirizando o discurso de incentivo a cirurgias estéticas, chama a atenção para o fato de que o corpo feminino que detém passe livre na plataforma é aquele exposto de forma comercial – afinal, o Instagram é uma empresa de publicidade. As produtoras de conteúdo priorizadas são as que promovem produtos, serviços e ideais de beleza, auxiliando os anunciantes na rede social, ou expondo corpos que mantenham a audiência, atraiam engajamento e produção de dados comportamentais: o norte do controle de conteúdo.

Nas suas sátiras, portanto, Polly Oliveira performa a boa sujeita de dados¹⁷³ que incorpora o máximo de produtividade para a rede social: seu corpo está dentro de todas as margens de inteligibilidade da moderação, podendo circular desimpedidamente para promover anúncios, empreender sua imagem, atrair usuários, engajamento e superávit comportamental – não importa se prega ideais de beleza nocivos às mulheres. Fosse ela um corpo feminino hipervisível para a moderação, gordo, racializado, transgênero, ou portador de deficiências físicas, poderia ser uma boa

¹⁷¹ RYAN-MOSLEY, Tate, “How digital beauty filters perpetuate colorism”. *Artificial Intelligence*, 15.08.2021. Disponível em: <<https://www.technologyreview.com/2021/08/15/1031804/digital-beauty-filters-photoshop-photo-editing-colorism-racism/>>. Acesso em: 11 Nov. 2021.

¹⁷² V. nota de rodapé nº 75, no Capítulo 1, Seção 1.3 deste trabalho.

¹⁷³ BRIDGES, Lauren, “Digital failure: Unbecoming the ‘good’ data subject through entropic, fugitive, and queer data”, pp. 1-17.

sujeita de dados de outra maneira: consumindo o conteúdo alheio, fornecendo seus dados e sua atenção, mas não expondo ideias ou corpos dissidentes.

As performances humorísticas são a estratégia que a blogueira escolhe para se autodefender do silenciamento que vinha sofrendo – por força do *shadowban* – e para expor o governo algorítmico das condutas no Instagram.

Os sistemas algorítmicos de moderação não conseguem detectar a ironia da autoexposição de Polly Oliveira. Apesar de ela se expor seminua – o que a moderação não vinha permitindo –, a edição do seu corpo com *photoshop* e a promoção de produtos de beleza que não existem não eram identificados. Antes, camuflaram a blogueira dos olhos do sistema automatizado de moderação e a tornaram uma usuária relevante para o sistema de recomendação de conteúdo, que passou a veicular mais amplamente suas publicações. Sendo assim, por meio de sua *autoexposição* performática, ela se *autodefende* do poder algorítmico do Instagram, desviando dos sistemas automatizados de moderação de conteúdo, e, mais ainda, enganando os sistemas automatizados de recomendação, que passaram a veicular mais amplamente suas fotos.

A autodefesa de Polly Oliveira gira ao redor da criação de autoridade discursiva para si mesma, para que suas críticas possam ser ouvidas. Em outras palavras, sua técnica de defesa consiste em desviar do *shadowban* e fazer-se visível para ser escutada. Antes de suas paródias, a blogueira era ocultada pelas técnicas algorítmicas de moderação, que a identificavam como uma usuária que sistematicamente publicava conteúdo sensível, tendo até mesmo tido sua conta suspensa. Ela era hipervisível aos algoritmos de moderação, não como Polly, mas como conjuntos de dados em desacordo com as regras do Instagram. Não havia sujeito que podia legitimamente falar – ou melhor, falar a uma audiência. Com as performances humorísticas, no entanto, ela avoca a legitimidade discursiva necessária para fazer-se ser ouvida¹⁷⁴, ter visibilidade: ser invisível aos algoritmos de moderação e visível aos usuários.

Além de se defender do silenciamento do *shadowban*, Polly Oliveira expõe o poder algorítmico do Instagram. Em sua autoexposição, ela “contraexpõe” o poder circulante na rede social, recorrentemente ocultado pelos discursos da empresa, que,

¹⁷⁴ BUTLER, Judith. *Excitable Speech: A politics of the performative*. New York/London: Routledge, 1997, p. 146-147.

como visto, nega ou diminui críticas aos vieses dos sistemas algorítmicos de moderação sob argumentos tecnoutópicos que confiam quase cegamente na capacidade da tecnologia. Ao performar a boa sujeita de dados a quem a visibilidade é concedida – uma personagem de cabelos lisos, pele clara e corpo artificialmente editado –, ela revela aquelas a quem a visibilidade é negada. É mostrando os limites do visível que ela indica o invisível, contestando os limites das identidades reconhecidas pelos sistemas algorítmicos de controle de conteúdo no Instagram.

Quando mostra as falhas e inconsistências dos sistemas algorítmicos, ela não os mostra como meramente *acidentais*. Na verdade, sua crítica revela tais falhas e inconsistências como vieses úteis ao modelo de negócios do Instagram: a distribuição de visibilidades dentro da rede social é instrumental para atingir determinado fim (o lucro), ainda que às expensas de determinados grupos sociais. Esse é o maior argumento incutido na performance da boa sujeita de dados. Seu corpo identificado como de mulher é útil, consumível, atrativo, capaz de incitar a geração de dados alheios, contribui para o fomento de indústrias paralelas, como a farmacêutica e a de cirurgia plástica. O corpo da boa sujeita de dados digna de visibilidade, performado ironicamente por alguém que não detém os mesmos atributos, é usado para indicar aqueles que são desprivilegiados ou até mesmo ocultados: os corpos femininos cuja exposição não tem valor monetário óbvio dentro do governo algorítmico, não são mostrados para promover produtos ou serviços ou, de outro jeito, que são mostrados para promover ideias irrelevantes, como a libertação feminina de padrões estéticos.

É desse modo que Polly Oliveira expõe o governo algorítmico das condutas no Instagram: apontando para a gestão de visibilidades e de corpos femininos voltada ao lucro. Dessa maneira ela revela o regime *ocultado* de visibilidades – evidenciado pelos efeitos da moderação algorítmica de conteúdo, mas ocultado pelo discurso do Instagram. Sua performance indica o tipo de subjetividade feminina premiado e reforçado pelos sistemas algorítmicos da rede social em detrimento de outras. Existe o reconhecimento de que tais sistemas conduzem o comportamento das usuárias mulheres por meio do controle ambiental do Instagram, seja suscitando desejos, premiando autoexposições alinhadas aos valores estéticos priorizados pela rede social ou ocultando aquelas que estejam em desacordo com eles. A gestão do corpo feminino no Instagram é voltada ao lucro.

A autoexposição vista neste primeiro exemplo constitui um processo de subjetivação dentro do Instagram, uma forma de Polly Oliveira reconhecer-se dentro do governo algorítmico das condutas na rede social e de agir sobre si mesma, conduzindo sua própria conduta em seus termos, em uma tentativa de divergir do que é reforçado pelo poder algorítmico. Por mais que ela não seja interpelada a refletir sobre si, o tratamento diferenciado que recebe da moderação automatizada faz com que ela o faça. A oportunidade de crítica e reflexão sobre a própria condição de assujeitada digital surge desta ocasião: perceber a própria visibilidade fora de seu controle, a incapacidade de ter suas palavras ouvidas – pois as postagens eram alvo de *shadowban*. Depois perceber que vinha sendo silenciada por essa técnica algorítmica de moderação, ela se percebe a si mesma na condição de assujeitada digital, conduzida pelos sistemas automatizados de moderação de conteúdo. Diante disso, articula uma estratégia para lidar com o governo algorítmico. Suas performances, portanto, resgatam a reflexividade que é dificultada pelo governo algorítmico.

Além do pensamento reflexivo, Polly Oliveira age sobre a sua conduta e se conduz nas performances humorísticas. Em vez de assujeitada, conformada ao molde da boa sujeita de dados, ela se indigna e desenvolve artifícios para criar suas práticas de ser e estar na rede social da própria forma. Quando usa o humor para burlar o *shadowban* e tornar visível um conteúdo formalmente em conformidade com a moderação, mas materialmente inconformado, ela provoca o que Lauren Bridges chama de “fracasso digital”¹⁷⁵: adota uma conduta destinada a produzir dados falsos, que resultam em identificações erradas por parte de marketeiros, codificadores e moderadores de conteúdo. Em suma, nas palavras de Bridges, ela se “desfaz uma boa sujeita de dados”¹⁷⁶, recusando-se a ser realmente a usuária produtiva para o capitalismo digital que o governo algorítmico da rede social pretende forjar. Em vez de autoexposição voltada primariamente ao lucro, a autoexposição voltada à crítica. Em vez de se deixar conformar com a subjetividade fomentada pelo governo algorítmico das condutas, ela opta por adotar práticas humorísticas a princípio improdutivas, não voltadas a realmente promover produtos e procedimentos estéticos nem atrair seguidoras com discursos de beleza, mas sim a criticar a moderação algorítmica do Instagram.

¹⁷⁵ BRIDGES, Lauren E., “Digital Failure: Unbecoming the ‘good’ data subject through entropic, fugitive, and queer data”, pp. 1-17.

¹⁷⁶ Idem.

4.2.

Morgan Barbour: Digilantismo feminista

Na análise do caso de Morgan Barbour, pretendo focar na exposição de agressões sofridas no Instagram como estratégia de autodefesa à moderação – ou melhor, à falta de moderação – da rede social. A usuária expõe, por meio do nu artístico, o silêncio do Instagram diante de suas denúncias a assédios sexuais de outros usuários. Sustento que a estratégia de Morgan Barbour consiste em usar o seu corpo para relembrar a humanidade e materialidade das usuárias por trás das representações digitais no Instagram e criticar a negligência da rede social, que as enxerga como mero agregados de dados, cuja proteção não é necessária até que as receitas da empresa sejam impactadas.

Morgan Barbour, artista de circo, diretora de coreografia e modelo¹⁷⁷, usa o seu perfil no Instagram e em outras redes sociais para promover e desenvolver o próprio trabalho, tendo conseguido uma sólida base de seguidores e apreciadores de sua arte após anos de presença digital. A admiração e gentileza dos seguidores, no entanto, é contrabalanceada com abusos e agressões que sofre nas mesmas redes sociais, os quais se intensificaram a partir do início do isolamento social obrigatório decorrente da pandemia de Sars-Cov2.

Desde março de 2020, quando se agravaram as contaminações pelo vírus e teve início a pandemia, Morgan passou a receber no Instagram e em outras redes sociais e aplicativos de encontro uma enxurrada de mensagens privadas contendo assédio, ameaças de estupro e morte violenta, fotos íntimas não solicitadas e xingamentos. A intensificação do assédio não foi acompanhada da capacidade do Instagram tomar medidas para proteger a usuária. As denúncias que Morgan fazia à plataforma eram respondidas com uma mensagem automática informando que o Instagram não conseguiria analisar a sua denúncia, pois, devido à escassez de funcionários disponíveis para fazer a moderação do conteúdo, só seria possível focar naquilo que tivesse “maior potencial de dano”.

¹⁷⁷ BARBOUR, Morgan. “About Morgan”. Disponível em: <<https://www.morganbarbour.com/about-1>>. Acesso em: 26 abr. 2021.

Diante dessa situação de crescentes agressões e desamparo, Morgan Barbour criou um perfil na rede social especificamente voltado a criticar a moderação do Instagram: a página “*Cummunity Standards*” (@cummunitystandards). Nele, a artista se vale da fotografia artística, do próprio corpo e da exposição das agressões e agressores para resistir à indiferença da rede social diante dos seus pedidos de ajuda.

Isso porque suas denúncias eram retornadas com uma resposta *automática* e genérica do Instagram dizendo que o comportamento do respectivo usuário não infringia as Diretrizes da Comunidade ou que a moderação estava ocupada com denúncias de maior potencial de dano, o que ela sentia como se fosse quase uma resposta subliminar da plataforma de que “homens são homens”, não há nada que se possa fazer, “esse é o preço que se paga por ser uma mulher na Internet”¹⁷⁸.

As publicações do perfil *Cummunity Standards* geralmente consistiam em fotografias de Barbour posando artisticamente sem roupa, as quais ela editava para incluir, sobrepostas ao seu corpo, os insultos que recebia dos outros usuários. Em uma das publicações (Figura 3), ela é retratada em preto e branco, deitada de bruços com uma das mãos repousando delicadamente sobre as costas. Contra a imagem, se destacam as inscrições em tinta na pele do seu torso “*fucking nasty slut whore bitch*” – uma série de insultos que, traduzidos literalmente para o português, em sua maioria significam “vadia”. A frase era uma dentre muitas outras mensagens recebidas de um mesmo usuário que repetidamente a assediava, cuja captura de tela foi publicada pela artista na mesma postagem sem tarjar o nome do usuário (Figura 4). Em outra postagem (Figura 5), Morgan Barbour posa sem roupa com as mãos na cabeça expressando ira ou desespero e sobre a imagem outra mensagem de assédio recebida é escrita em letras brancas: “Você pode me mostrar seus peitos? Eu acho que você tem belos peitos”¹⁷⁹. Na mesma postagem, o autor da mensagem tem seu nome de usuário exposto em uma captura de tela.

Na era do “cancelamento”¹⁸⁰ como estratégia de punição digital de indivíduos após uma conduta julgada como reprovável, a exposição de agressores como resposta ao assédio é especialmente relevante, diante dos efeitos que sua existência

¹⁷⁸ BARBOUR, Morgan, “Cummunity Standards”. *GenderIT.org*, 25.05.2020. Disponível em: <<https://genderit.org/feminist-talk/cummunity-standards>>. Acesso em: 13 mar. 2022.

¹⁷⁹ Tradução livre de “Can you show me your boobs? I think you have nice boobs.”

¹⁸⁰ NG, Eve, “No Grand Pronouncements Here...: Reflections on Cancel Culture and Digital Media Participation”. *Television and New Media*, v. 21, n. 5, 2020, pp. 621-627; HORA, HORA, Bianca Dias da; MARTINS, Milena Fagundes, “A Cultura do Cancelamento e Suas Facetas: Justiça Social, Intransigências e Disputas Narrativas”. *Revista Iniciacom*, v. 10, n. 3, 2021. Disponível

digital pode sofrer. Segundo Morgan Barbour, no entanto, é uma das únicas medidas punitivas que restam diante da inércia do Instagram e de outros poderes institucionais¹⁸¹. Emma Jane¹⁸² identifica estratégias ativistas como essa dentro do que chama de “digilantismo feminista”, um neologismo que faz referência ao vigilantismo feminista em meios digitais. O digilantismo feminista abrange práticas extrajudiciais e politicamente motivadas, colocadas em prática por mulheres em espaços virtuais, cuja intenção é punir ou responsabilizar terceiros diante de uma ausência de respostas institucionais a agressões sofridas *online*.

O digilantismo feminista da artista não tem como alvo somente os agressores expostos, mas também, e principalmente, o Instagram: ela se *autoexpõe*, nua em fotografias artísticas, para contestar a rede social e sua política de moderação de conteúdo. Carla Gomes e Bila Sorj desenvolvem a noção de “corpo-bandeira”, ao analisarem o uso do corpo por protestantes na Marcha das Vadias do Rio de Janeiro¹⁸³, para se referir ao uso do corpo como objeto de reivindicação e como suporte de determinada mensagem em um protesto político. O corpo, afinal, é central na análise das dinâmicas de poder, pois é forjado por discursos e práticas de poder: o corpo é construído socialmente porque significado socialmente¹⁸⁴. Daí porque sua ressignificação visual é tão impactante.

Morgan Barbour usa o seu corpo como corpo-bandeira para falar sobre nudez e violência – ou melhor, para falar sobre o corpo feminino constituído, no Instagram, pela nudez e pela violência legitimada. Como visto no Capítulo 1, em um espelho do que ocorre fora do virtual, a identidade social de mulher no Instagram também é constituída pela nudez e pela condição de alvo de um tipo de violência sancionada por uma sensibilidade coletiva que naturaliza certas agressões a pessoas que carreguem tal identidade. O regime de visibilidades da rede social – tanto

em: <<https://revistas.intercom.org.br/index.php/iniciacom/article/view/4110/pdf>>. Acesso em: 28 fev. 2022.

¹⁸¹ BARBOUR, M. “[Cum]munity Standards: Resisting Online Sexual Harassment and Abuse”. In: POWELL, Anastasia; FLYNN, Asher; SUGIURA, Lisa (Ed.). *The Palgrave Handbook of Gendered Violence and Technology*. Switzerland: Palgrave Macmillan, 2021, p. 28. Morgan diz: “se as plataformas não vão moderar adequadamente, e se a polícia argumenta que é demais para eles gerenciarem, o que podem fazer as vítimas de abuso online?”.

¹⁸² JANE, Emma, “Online misogyny and feminist digilantism”, *Continuum*, v. 30, n. 3, 2016, pp. 284-297. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/10304312.2016.1166560>>. Acesso em: 20 fev. 2022.

¹⁸³ GOMES, Carla; e SORJ, Bila, “Corpo, geração e identidade: a Marcha das vadias no Brasil”. *Revista Sociedade e Estado*, v. 29, n.2, maio-agosto 2014, p. 437.

¹⁸⁴ PRECIADO, Paul B. *Testo Junkie: sexo, drogas e biopolítica na era farmacopornográfica*. São Paulo: N-1, 2018, pp. 76-88.

aquele expresso nas Diretrizes da Comunidade quanto o ocultado na moderação automatizada de conteúdo – restringe o corpo identificado como feminino a um lugar de sexualização, que precisa ser minuciosamente regulado e fiscalizado para não ofender o público com sua sexualidade, e de violência legitimada, uma vez que agressões misóginas não são vistas como discurso de ódio com impactos reais sobre as vidas das vítimas, mas sim como meras injúrias de ordem privada e inofensivas. A moderação algorítmica do Instagram, como poder punitivo que fiscaliza as regras de inteligibilidade e visibilidade dos corpos, contribui para atualizar esses significados sociais em larga escala.

A autoexposição de Morgan Barbour é uma estratégia de autodefesa a essa imposição de significados. Ela critica a construção social do corpo da mulher tal como é e reivindica a significação da sua nudez e de seu corpo, sempre sexualizado contra a sua vontade e visto como destinatário de violências legitimadas, por usuários e pelo Instagram. Para tanto, ela usa a arte.

A potência da arte como resistência está em sua capacidade de vincular significantes a novos significados, de “revolucionar o campo simbólico e social”¹⁸⁵ e criar experiências estéticas não normalizadas¹⁸⁶. Pode-se falar até mesmo em “artivismo”¹⁸⁷, uma forma de usá-la para fazer intervenções sociais e políticas. Esse uso da arte foi favorecido pela intensificação do uso de redes sociais e novas tecnologias, que possibilitaram a criação visual e artística por qualquer um e seu compartilhamento a baixos custos (não obstante a diferença de acesso a tais ferramentas devido a desigualdades materiais, especialmente no Brasil)¹⁸⁸.

¹⁸⁵ TVARDOVSKAS, Luana Saturnino, “Cópia e paródia: práticas de si na arte contemporânea de mulheres e a crítica feminista à História da Arte”. In: SILVA, Edilene; OLIVEIRA, Susane Rodrigues de; e ZANELLO, Valeska (Org.). *Gênero, subjetivação e perspectivas feministas*. Brasília: TechnoPolitik, 2019, p. 311.

¹⁸⁶ LESSA, Patrícia, “Práticas de resistência e poéticas lesbianas nas obras de Karin Schwarz e Alma López”. *ANPUH SP – XX Encontro Estadual de História*, 2016, p. 3. Disponível em: <http://www.encontro2016.sp.anpuh.org/resources/anais/48/1462818428_ARQUIVO_PatriciaLes-saunesprtf2016.pdf>. Acesso em: 3 mar. 2022.

¹⁸⁷ “Artivismo é um neologismo conceptual ainda de instável consensualidade quer no campo das ciências sociais, quer no campo das artes. Apela a ligações, tão clássicas como prolixas e polémicas entre arte e política, e estimula os destinos potenciais da arte enquanto ato de resistência e subversão. Pode ser encontrado em intervenções sociais e políticas, produzidas por pessoas ou coletivos, através de estratégias poéticas e performativas [...]. A sua natureza estética e simbólica amplifica, sensibiliza, reflete e interroga temas e situações num dado contexto histórico e social, visando a mudança ou a resistência.” (RAPOSO, Paulo, “‘Artivismo’: Articulando dissidências, criando insurgências”. *Cadernos de Arte e Antropologia*, v. 4, n. 2, 2015, p. 5.)

¹⁸⁸ COLLING, Leandro, “A emergência dos artivismos das dissidências sexuais e de gêneros no Brasil da atualidade”. *Revista sala preta*, v. 18, n. 1, 2018, p. 162.

As redes sociais, principalmente aquelas focadas na imagem como o Instagram, oportunizaram às mulheres serem produtoras visuais de arte que quebre com significações hegemônicas de seus corpos, de peças que se valham do próprio corpo nu para colocarem a artista na posição de criadora do significado de seu corpo material. Apesar de as produções visuais que obedeçam aos significados hegemônicos usualmente serem as que recebem maior atenção e promoção, blogueiras como Morgan Barbour e Polly Oliveira seguem produzindo conteúdo cuja proposta é contestar tais significados e abrir espaço para novos. As redes sociais como o Instagram são mais um espaço de naturalização de estilizações corporais – especialmente por sistemas algorítmicos, que replicam o mesmo significado social atribuído a um corpo em larga escala. A resposta a essa repetição de significados atribuídos ao corpo pode ser feita por meio da autoexposição que pretenda mostrar a instabilidade do corpo e a efemeridade de significados.

As montagens fotográficas do perfil *Cummunity Standards* são relativamente simples, uma vez que consistem em fotografias de Morgan Barbour com a edição de insultos em texto por cima ou sua inscrição diretamente no corpo da artista, mas mesmo assim têm sua importância artística e semântica. Ao contrário de Polly Oliveira, ela não performa personagens nem estilizações corporais diversas da sua própria, com trajes, edições visuais e legendas encenadas, para enganar os algoritmos ou tecer uma crítica ao que é considerado “normal”. A mensagem de Morgan Barbour vem do contraste entre seu corpo-bandeira nu e os assédios, insultos e ameaças sobrepostos a ele. Segundo a artista, essa composição das imagens deixa evidente que é um ser humano o alvo dos xingamentos e ameaças gratuitos.

Com as inscrições de palavras agressivas sobre o seu corpo, ela o particulariza, em uma autobiografia que resgata sua condição de pessoa humana e quebra com a expectativa de vê-la como sexual ou erótica. Seu corpo nu é mostrado sem a intenção de ser sexualizado, mas para afirmar a violência que palavras digitadas em uma rede social infligem sobre um corpo bem real. Ainda que somente sua expressão digital apareça aos outros usuários, e ainda que seja ela tratada como agregado de dados pelo Instagram, existe uma materialidade por trás das relações travadas em plataformas virtuais, onde são sentidos os impactos das agressões praticadas à distância.

É nisto que reside sua autodefesa ao governo algorítmico das condutas: em reclamar o sentido de seu corpo humano, retirando-o das mãos do Instagram. Por

mais que ela não enderece expressamente a moderação automatizada da rede social, suas postagens funcionam como um protesto aos valores atribuídos à condição de mulher que são replicados em larga escala pelos algoritmos de curadoria e moderação. Seu corpo nu de “mulher” não é exposto de forma sexual, comercial, nem por um culto à estética valorizada pelo Instagram. O corpo é veículo de dor e raiva, não objeto de inscrição do desejo alheio ou objetivos comerciais. Seu principal intuito não é atrair engajamento de outros usuários ou promover produtos de anunciantes, mas punir os usuários que a assediam e o Instagram pela negligência, conscientizar, expressar sua dor, atribuir seus próprios significados ao seu corpo.

A autodefesa de Morgan Barbour também consiste na exposição do governo algorítmico das condutas em certa medida. Com o digilantismo feminista e a arte, ela chama atenção para a indiferença do Instagram às violências misóginas sofridas em seus espaços e aos efeitos reais sofridos pelas usuárias. Além de lembrar que a rede social as enxerga como meras expressões digitais, sem considerar sua constituição material, ela problematiza o lugar que o corpo feminino ocupa na economia do Instagram. Se Polly Oliveira mostra que a exposição comercial e sexualizada do corpo feminino é útil enquanto gerar receitas e que deve ser ocultado caso interfira negativamente na geração de lucro, o caso de Morgan Barbour mostra, semelhantemente, que a proteção especificamente do corpo feminino obedece à mesma lógica. A proteção do corpo feminino contra agressões misóginas de outros usuários só será útil quando as receitas da empresa forem afetadas; até que isso aconteça, não há necessidade de atribuir tanta importância a tais problemas, que são meros insultos de ordem privada.

Portanto, Morgan Barbour, ao se deparar com a falta de respostas por parte do Instagram e com a situação na qual o assédio online é visto como uma parte inexorável da experiência feminina em redes sociais, reconhece sua condição como assujeitada dentro de uma relação de poder digital que tem o monopólio sobre o conceito de violência – conceito este que não abrange o discurso de ódio misógeno praticado nos espaços “privados” do Instagram. Ela age como sujeito reflexivo, pensa na sua presença digital no Instagram submetida a um regramento que ignora agressões recebidas.

O caso de Morgan Barbour também evidencia seu privilégio de circulação na rede social como mulher branca – apesar de não necessariamente ser definida nesses

termos nos Estados Unidos, país de residência, devido à sua ascendência. Relembrando os relatos abordados no Capítulo 1, corpos de mulheres trans, racializadas ou portadoras de deficiências físicas são desproporcionalmente ocultados e excluídos pela moderação do Instagram, em comparação com aquelas que não sejam atravessadas por esses marcadores. As fotografias de Morgan na maior parte das vezes conseguem fugir ao controle de nudez da rede social, esquivando-se inclusive dos erros da moderação, que vê outros tipos de corpos como desvios da norma, mas não o dela. O escape pelas malhas da moderação por parte de Morgan Barbour é tanto um ato de resistência quanto um privilégio social – ou melhor, pode ser um ato de resistência por ser um privilégio social. Talvez os modos de articulação de uma autoexposição corporal crítica por mulheres trans, não binárias, negras, indígenas, portadoras de deficiências físicas, por exemplo, tenha que se valer de outros artifícios para avocar uma legitimidade para sua fala que Morgan Barbour já detém em alguma medida.

4.3

Autodefesa Digital: Autoexposição e contraexposição

Dorlin analisa a defesa de si como ato contínuo, cotidiano, praticado nas esferas e eventos da vida tradicionalmente colocados do lado de fora da política. Em sintonia com essa escolha, a prática política nas redes sociais não pode ser apreendida com a métrica da política tradicional¹⁸⁹.

Não se deve descartar o caráter político de determinado conteúdo publicado no Instagram meramente por ter se valido de um meio que propaga sistemas de opressão; por consistir, à primeira vista, em mera comunicação cotidiana; ou por girar em torno de experiências pessoais. Algumas postagens fazem enunciações políticas mais diretas, outras não. Por mais que algumas vezes o conteúdo publicado por uma blogueira não tenha a expressa intenção de fazer política, pode funcionar como uma instância de enunciação política na medida em que afirma e dissemina valores contrários a uma narrativa dominante e excludente, clamando pela visibilidade de outros corpos e por outras significações ao corpo.

¹⁸⁹ CALDEIRA, Sofia P; RIDDER, Sander De; BAUWEL, Sofie, “Between the Mundane and the Political: Women’s Self-Representations on Instagram”, p. 2.

Diante dos exemplos analisados, é possível estender as reflexões de Dorlin para pensar em uma *autodefesa digital*: técnicas de defesa de si desenvolvidas em ambientes digitais contra a experiência de violências ou discriminações ocorridas nesse âmbito – sejam elas perpetradas por outros indivíduos ou pelo próprio ambiente digital, como são os casos tratados neste Capítulo.

Neste trabalho, no entanto, faço uma distinção entre tipos de ativismo coletivo *online* e as manifestações individuais de resistência a determinada relação de poder. A autodefesa digital diz respeito à resistência no nível individual, porque envolve técnicas [digitais] *de si*, processos em que um indivíduo se percebe como sujeito (ou melhor, assujeitado) a determinado poder exercido por meios digitais e age sobre si mesmo para determinar a própria conduta. Em outras palavras, do mesmo modo que Dorlin, penso a autodefesa digital como um *processo de subjetivação*, de tornar-se sujeito, não meramente assujeitado, e que, por isso mesmo, é individual.

A categoria de autodefesa digital não é aplicável a todo e qualquer ato de enunciação política na rede social. Segundo Jota Mombaça, “há estratégias, técnicas e ferramentas que somente uma corporalidade e subjetividade capaz de habitar a fragilidade consegue desenvolver”¹⁹⁰. A autodefesa digital é o que sobra às subjetividades desarmadas do Instagram, usuários mirados por determinada prática discriminatória do poder algorítmico da rede social. Diante de práticas do poder algorítmico calcadas na normatização e homogeneização do corpo segundo critérios hegemônicos, alguns usuários desproporcionalmente apagados, silenciados, ou que sofrem imposições de significados – como visto no Capítulo 1 – pensam em formas de reclamar novamente a visibilidade necessária para suas críticas ou de ressignificar seus corpos materiais dentro do Instagram. Os demais usuários, por mais que também estejam submetidos ao governo algorítmico das condutas, escapam de classificações sociais discriminatórias devido à sua aparência física.

Na governamentalidade algorítmica nos termos em que existente no Instagram, há uma *sobreposição* de desarmamento das subjetividades com relação a certos usuários. Determinados indivíduos são duplamente despojados de suas subjetividades. Além do desarmamento operado pelo governo algorítmico das condutas, que prescinde dos sujeitos e os enxerga como conjuntos de dados asinificantes,

¹⁹⁰ MOMBAÇA, Jota. *Não vão nos matar agora*. Rio de Janeiro: Cobogó, 2021, p. 79-80.

também existe o desarmamento decorrente da imposição de significados sociais hegemônicos aos corpos – como os de gênero e raça. É a estas “subjetividades duplamente desarmadas” que sobra a autodefesa digital.

Nessa condição, as subjetividades duplamente desarmadas desenvolvem técnicas digitais de si, estratégias para se constituírem sujeitos dentro da governamentalidade algorítmica, e não meramente assujeitados, escaparem de efeitos discriminatórios ou reaverem o sentido de sua presença digital.

Tanto Polly Oliveira quanto Morgan Barbour usam seus perfis no Instagram para autodefender-se do governo algorítmico das condutas de usuárias identificadas como mulheres – principalmente a manifestação desse governo por meio dos sistemas de moderação. Diante da percepção de si mesmas como assujeitadas a um poder da rede social, em um movimento de reflexividade crítica, elas colocam em prática técnicas digitais de si e se constituem como sujeitas, alterando sua relação consigo mesmas e com o Instagram.

As estratégias de autodefesa das duas blogueiras têm semelhanças e diferenças, mas ambas podem ser decompostas em dois eixos em comum: (i) a autoexposição corporal crítica e (ii) a exposição do poder algorítmico do Instagram sobre as usuárias identificadas como mulheres.

Polly Oliveira performa ironicamente a boa sujeita de dados premiada com visibilidade para que possa burlar a moderação algorítmica e ter suas críticas ouvidas. Morgan Barbour expõe o seu corpo nu coberto de agressões em desacordo com os significados reiterados pela moderação do Instagram que constituem a boa sujeita de dados. A primeira se autoexpõe estrategicamente para avocar legitimidade discursiva e a segunda para ressignificar o próprio corpo e punir tanto assediadores quanto a rede social pelo descaso com a violência misógina. Nos dois casos a autoexposição corporal é o ato de resistência imediato, ainda que se valha do mesmo ambiente que exerce poder sobre elas.

É possível pensar em resistência não em termos de intervenção direta na alteração das estruturas da rede social, mas como condutas individuais adotadas por usuários da rede social que se valem da própria arquitetura e lógica que as oprimem. Como ressaltam Velkova e Kaun, em uma análise do conceito de resistência algorítmica, embora a resistência à opressão seja comumente pensada em termos de ação coletiva organizada, de consequências substanciais e revolucionárias que neguem as fundações da própria dominação, a realidade da resistência é que muitas

vezes se mantém dentro dos limites do sistema de poder, testando suas brechas e fraquezas sob uma aparência de conformidade¹⁹¹.

A defesa das duas blogueiras ao governo algorítmico das condutas, em certa medida, se mantém dentro dos seus limites, apesar de testá-los. Não há articulação coletiva para derrubar esse poder – pois isto seria até mesmo ingênuo, diante da magnitude da Meta Platforms, Inc.. Não há alteração da forma de governo algorítmico de suas condutas nem interlocução direta com a empresa – não há revolução no sentido em que é usualmente apreendida. O que elas fazem é autoexpor seus corpos de forma crítica, em descompasso com a subjetividade feminina fomentada pelo Instagram. São atos de “insubordinação solitários”¹⁹² contra aquilo que as fere na moderação da rede social.

Nos dois casos, Polly Oliveira e Morgan Barbour usam a autodefesa para se recusarem a ser consideradas meras expressões digitais, e, mais ainda, já significadas pelo Instagram por sua corporalidade. Ambas negam ser a boa sujeita de dados: o corpo feminino dócil gerido pelo governo algorítmico. Se Polly não pode expor seu corpo tal como é, então o expõe disfarçado para avocar legitimidade discursiva e criticar a gestão dos corpos femininos pelo Instagram (gestão esta que é ocultada pela ideia de autogestão do corpo feminino). Se Morgan pode expor o seu corpo, desde que suporte sozinha a sexualização por parte de outros usuários e o assédio naturalmente decorrente, ela o exporá coberto de insultos e jogando luz sobre a inércia do Instagram. A gestão dos corpos femininos pelo Instagram é evocada pelas duas.

Assim, a autodefesa de ambas também consiste na exposição do poder algorítmico do Instagram: uma *contraexposição*. Como visto, trata-se de um poder que se camufla sob discursos tecnoutópicos. A rede social, em uma manifestação da colonialidade automatizada, impõe uma epistemologia datacêntrica como a única capaz de ter uma apreensão *real* do funcionamento dos algoritmos de curadoria e moderação de conteúdo. Desse modo, tratam as críticas dos usuários como impressões, equívocos, sensações ruins não fundamentadas – qualquer outra categoria que não saberes produzidos a partir de sua experiência concreta com a rede

¹⁹¹ VELKOVA, Julia; KAUN, Anne, “Algorithmic resistance: media practices and the politics of repair”, p. 534.

¹⁹² COELHO, Bruna Martins, “A Filosofia a Golpes de Martelo de Elsa Dorlin: ‘se defender’ contra o terrorismo de Estado – resenha crítica”, p. 284.

social. Soma-se a isso o fato de que o governo algorítmico atua de dentro dos sujeitos para fora: organiza o ambiente informacional dos usuários de forma a suscitar-lhes afetos e comportamentos pré-calculados altamente prováveis. Desse modo, torna-se difícil localizar o poder algorítmico do Instagram, e, mais ainda, responder a ele, considerando o caráter difuso das comunicações virtuais.

Expor o governo algorítmico do Instagram consiste em um ato de autodefesa na medida em que nomeia o poder. A resposta a um poder ocultado e sutil não deixa de ser uma interpelação a esse poder: o ilumina, o chama a se manifestar, marca a sua existência. É preciso nomear a norma antes de combatê-la, e é isso que a exposição de aspectos do governo algorítmico por Polly Oliveira e Morgan Barbour faz. A crítica à gestão dos corpos e comportamentos femininos pelo Instagram; ao tratamento das usuárias como se fossem meros conjuntos de dados e representações virtuais desvinculados de uma materialidade que sofre os efeitos do assédio e do ódio online; ao reforço de uma subjetividade ideal (boa sujeita de dados), dócil, consumista, que autoexpõe seu corpo quando aceitável dentro dos interesses da rede social e sempre engaja com o conteúdo alheio gerando dados. A iluminação de todos esses aspectos pela autoexposição de Polly Oliveira e Morgan Barbour é uma defesa contra o governo algorítmico na medida em que o localiza e a ele responde, passando por cima das manifestações públicas do Instagram e de sua epistemologia datacêntrica. Trata-se de uma *contraexposição*: em uma rede social que lucra da autoexposição espetacular e que esconde sua governamentalidade atrás de discursos datacênicos, as usuárias usam sua autoexposição não para promover conteúdo comercial, mas para *contraexpor* o poder difuso e silencioso exercido por meio dos sistemas algorítmicos do Instagram.

Por fim, tanto o uso do humor quanto da arte e digilantismo funcionam como técnicas de subjetivação: práticas [digitais] de si, exercícios das blogueiras sobre si mesmas como sujeitas em uma relação de poder digital. A performance humorística e o nu artístico, ambos, são o uso de um corpo material para instaurar outra relação com o digital. O corpo – sua imagem compartilhada no Instagram – propõe outras formas de ser dentro do governo algorítmico das condutas, formas não inteiramente conformadas à boa sujeita de dados, formas de serem governadas não deste jeito,

não por esses meios, não em nome destes princípios¹⁹³ (a hierarquização colonial de humanidades e o lucro a todo custo, por exemplo).

Ainda que os exemplos analisados não digam respeito a usuárias do Instagram que reúnam outros marcadores sociais, como usuárias trans, portadoras de deficiências físicas ou racializadas, elas ocupam um lugar dentro da rede social que lhes permite experimentar o imprevisto de alguma forma e, a partir disso, perceber-se como governadas no Instagram. Com tal percepção, resgatam a reflexividade humana que é descartada e complicada pela governamentalidade algorítmica e determinam suas condutas em seus termos. Isso implica em conduzirem as suas condutas como queiram, governar a si mesmas, ainda que não inteiramente de forma alheia ao poder que criticam, ao menos não inteiramente conformadas à boa sujeita de dados. Procuram outros modos de ser dentro do governo por meio da criatividade das técnicas digitais de si.

Na relação de poder algorítmica, não é exatamente a violência da resposta das usuárias que constitui sujeitos, mas a inventividade no uso do próprio corpo e da arquitetura digital do Instagram. Ainda é um uso revoltado do corpo, mas não necessariamente violento. É a capacidade de inventividade, de encontrar meios de multiplicar os imprevistos que a moderação algorítmica não consegue identificar, de aproveitar as brechas do funcionamento automatizado e usar o próprio meio que lhe governa para criticá-lo.

É preciso reconhecer, contudo, que há limitações à concepção de autodefesa digital tal como estudada aqui. Em primeiro lugar, considerando o modelo de negócios do Instagram, que capitaliza sobre o compartilhamento de ideias, sejam quais forem, isso importaria em que quanto mais as usuárias resistem, mais alimentam o sistema contra o qual se insurgem. Afinal, a publicação de suas fotos, por mais que críticas ao Instagram, dependem de engajamento de outros usuários para estabelecerem uma interlocução, para que a crítica não seja proferida ao vento. A comunicação, ainda que em espaços virtuais, depende da interlocução com outros agentes para que exista, e interlocução, no Instagram, significa engajamento e geração de dados. As críticas ao Instagram, portanto, apesar de se revoltarem contra

¹⁹³ FOUCAULT, Michel. *The Politics of Truth*, p. 28.

o modelo de negócios da rede social, também o alimentam. Nas palavras de Rouvroy, “indignação online é o que alimenta o capitalismo algorítmico”¹⁹⁴.

A crença exacerbada no poder transformador do ativismo online contra os malefícios sociais causados pela governamentalidade algorítmica pode levar ao mesmo erro cometido por aqueles que sustentam essa forma de governo das condutas: adotar cegamente a epistemologia datacêntrica e a esperança incondicional no potencial transformador do mundo via tecnologias digitais. É preciso cuidar para não romantizar a autodefesa digital nem cair novamente na passividade digital que impede a análise crítica do emprego crescente de processos algorítmicos na moderação do Instagram.

Em segundo lugar, considerando a autoexposição das usuárias como estratégia de autodefesa digital no Instagram, é difícil, se não impossível, escapar do exibicionismo e do individualismo típicos da subjetividade predominante no sistema econômico atual. Poderia ser argumentado que as críticas de Polly Oliveira e Morgan Barbour são individualistas, publicações de exposição de si mesmas criadas para defender pautas permeadas de um feminismo embranquecido. Afinal, não há, em suas publicações, menção expressa a divisões de classe e raça nem reflexões sobre as condições materiais de existência das mulheres, que variam drasticamente conforme seu posicionamento social (sobretudo na realidade brasileira). No máximo, a composição das imagens de Polly Oliveira denuncia o racismo da plataforma ao incluir elementos fenotípicos da branquitude valorizados pelos sistemas de curadoria e moderação de conteúdo, mas não existe a problematização socialmente situada das opressões por elas vividas no Instagram.

Sem pretender ignorar as limitações de uma manifestação de autodefesa digital em que as discussões podem ser simplificadas, restritas ao binarismo de gênero e a pautas do feminismo branco, é possível afirmar que pode essa mesma manifestação servir como um meio para a formação de espaços de trocas de experiência e, acima de tudo, a abertura de uma oportunidade de criticar e complementar as insuficiências de um discurso que, falando em nome das “mulheres” ou do “feminismo”

¹⁹⁴ ROUVROY, Antoinette, “Algorithmic Governmentality and the Death of Politics: An Interview with Antoinette Rouvroy”. *Green European Journal*, 27.03.2020, p. 4. Disponível em: <<https://www.greeneuropeanjournal.eu/algorithmic-governmentality-and-the-death-of-politics/>>. Acesso em: 2 nov. 2021.

abstratamente, se além a pautas brancas, sem considerar a pluralidade de mulheres que reúnem em si diversos graus de opressão social.

Existe uma diferença entre o mero exibicionismo valorizado pelo Instagram e a autoexposição com propósitos ético-político maiores. Por outro lado, como explica Banet-Weiser, atualmente as redes sociais como o Instagram, o TikTok, o Facebook e muitas outras voltadas à autoexposição espetacular dos usuários, promovem a hipervisibilidade de qualquer experiência, de modo que a visibilidade se tornou um fim, não um meio para um fim. Assim, a política transmuta-se em visibilidade pura: política é vestir uma camisa expondo seu status de feminista, uma postagem com muitas visualizações, uma foto compartilhada milhões de vezes¹⁹⁵, sem uma continuidade de críticas e intervenções nas estruturas que sustentam as divisões sociais opressivas. Por isso, até mesmo as publicações de viés mais ativista precisam ser avaliadas sob lentes críticas.

Por fim, outra limitação às manifestações de resistência à governamentalidade algorítmica como analisadas é a dependência da mídia tradicional. Como Treré resalta, diante das práticas de poder do capitalismo de vigilância, os ativistas digitais têm poucas armas à sua disposição para combatê-las com a mesma potência. Por mais que usuários se valham de seu conhecimento sobre a plataforma – aprendido a partir da observação, já que as informações técnicas são ocultadas e mascaradas sob discursos tecnoutópicos – e da criatividade para denunciar práticas de poder, comumente suas estratégias dependem de veículos tradicionais da mídia para terem alguma repercussão, além de raramente chegarem a ter algum impacto significativo na opinião pública e, mais difícil ainda, a nível internacional¹⁹⁶.

¹⁹⁵ BANET-WEISER, Sarah, “Popular Feminism and Popular Misogyny”, p. 75.

¹⁹⁶ TRERÉ, Emiliano. “The Dark Side of Digital Politics: Understanding the Algorithmic Manufacturing of Consent and the Hindering of Online Dissidence”. *IDS Bulletin: Transforming Development Knowledge*, v. 47, n. 1, Janeiro 2016, p. 135.

5

Considerações Finais

O presente trabalho se resume nos seguintes principais argumentos.

Em primeiro lugar, partindo de um exame das relações entre o Instagram e os seus usuários a partir da perspectiva destes últimos, este trabalho se propôs a examinar o regime de visibilidades não apenas positivado nas Diretrizes da Comunidade da rede social como também experimentado pelos usuários em sua interação com o aplicativo. Diante dos efeitos inconsistentes da moderação algorítmica da rede social relatados pelos usuários – em especial, das duas técnicas examinadas, a exclusão automatizada de conteúdo e o *shadowban* – foi possível perceber um padrão discriminatório que revela as insuficiências e vieses dos sistemas algorítmicos de moderação. O funcionamento da moderação muitas vezes importa na imposição de significados hegemônicos aos corpos cujas imagens circulam na rede social, bem como no silenciamento seletivo sob a justificativa de violação das Diretrizes da Comunidade, quando, na realidade, ora não houve violação, ora o mesmo tipo de conteúdo é aceito quando publicado por corpos não dissidentes dos significados normativos que informam os sistemas algorítmicos.

Em segundo lugar, os efeitos discriminatórios da moderação algorítmica revelam alguns dos valores que subjazem à criação dessas tecnologias, além de estarem presentes em outros lugares, como no comportamento dos usuários que treina tais modelos e no julgamento dos moderadores humanos. Neste trabalho, foram estudados a lógica do lucro decorrente do modelo de negócios escolhido pelo Instagram, a hierarquização colonial de humanidades e a epistemologia datacêntrica reforçada pelo Instagram em seus pronunciamentos públicos, que desconsidera a experiência dos usuários. São esses valores que acabam por determinar o que pode ou não ser visível dentro da plataforma. O descumprimento às suas normas leva à punição do respectivo usuário pelo sistema automatizado de moderação de conteúdo.

É inelutável que o Instagram está inserido em um sistema econômico que se apoia sobre determinadas relações de opressão, como relações racistas, sexistas, capacitistas e transfóbicas, que se guia por determinados valores estéticos e comerciais. As normas que perpetuam essas opressões são propagadas em larga escala

pelas técnicas automatizadas de moderação de conteúdo do Instagram. Reconhecer esse fato não consiste em assumir uma visão derrotista que não vê escapatória para a discriminação reproduzida por companhias como o Instagram, mas sim de reconhecer a complexidade das relações materiais humanas que precedem e constituem os sistemas automatizados de moderação, não sendo um problema resolúvel simplesmente com um ajuste nos algoritmos e reuniões do conselho de administração de uma empresa de tecnologia.

Em terceiro lugar, os sistemas automatizados de moderação de conteúdo integram uma nova forma de exercício do poder dentro do Instagram baseada em dados: a governamentalidade algorítmica. Trata-se de um governo das condutas que se apoia sobre dados comportamentais para gerar previsões que tentam intervir nas possibilidades de ação do usuário, encaminhando seus comportamentos para serem o mais produtivos possível. Os efeitos de governo sobre as condutas, no entanto, são diferentes quando se trata de usuárias identificadas como mulheres. Argumentou-se que a subjetividade ideal cobrada das usuárias do Instagram é a da “boa sujeita de dados”, aquela que se expõe, caso sua aparência seja aceitável, e, caso não o seja, se mantém interagindo com a plataforma e gerando dados – sempre cumprindo com a produtividade esperada pelo Instagram. A princípio, a forma de operação do governo algorítmico das condutas deixa poucas brechas para a inventividade humana, uma vez que o campo de ações já é, de antemão, calculado e oferecido aos usuários por meio de estímulos lançados no ambiente digital do Instagram. No entanto, argumentamos que o encontro com o improvável muitas vezes está na experiência daqueles que sofrem discriminação pelo governo algorítmico das condutas, que se veem punidos “ilegalmente” sem sequer terem descumprido as Diretrizes da Comunidade, mas sim por terem corpos inconformados com as normas de visibilidade ocultadas. Diante dessas situações, de encontro com o imprevisto, podem eles criar novos significados, resistirem, subjetivarem-se.

Por fim, com base em dois estudos de caso realizados no Capítulo 4, argumentamos pela existência de um tipo de *autodefesa digital*, estratégias digitais de si utilizadas por aqueles usuários que têm suas subjetividades duplamente desarmadas. Por meio da autodefesa digital, aqueles que sequer são considerados sujeitos, seja pelo governo algorítmico ou pelas normas ocultadas que o informam, respondem ao poder e alteram sua relação com o governo das condutas e consigo mesmos: se subjetivam. Com a autodefesa digital, atos individuais de resposta a um governo

das condutas levado a cabo no digital, é possível refletir criticamente sobre outras formas de ser governado.

Referências Bibliográficas

ABIDIN, Crystal; THOMPSON, Eric C., “Buymylife.com: Cyber-femininities and commercial intimacy in blogsops”. *Women’s Studies International Forum*, v. 35, 2012, pp. 467-477.

ABIDIN, Crystal, “Visibility labour: Engaging with Influencers’ fashion brands and #OOTD advertorial campaigns on Instagram”. *Media International Australia*, v. 161, n. 1, 2016, pp. 86-100.

ANDERSSON, Erin, “Breastfeeding on social media: How one picture led to a Toronto mother’s ban from Instagram”. *The Globe and Mail*, 13.05.2014. Disponível em: <<https://www.theglobeandmail.com/life/parenting/a-breastfeeding-mom-on-instagram-call-in-the-censors/article18644323/>>. Acesso em: 2 fev. 2022.

ARE, Carolina, “How Instagram’s algorithm is censoring women and vulnerable users but helping online abusers”. *Feminist Media Studies*, v. 20, n. 5, 2020, pp. 741-744.

ARE, Carolina, “Instagram Quietly Admitted Algorithm Bias, But How Will It Fight It”. *Blogger on Pole*, 25.06.2020. Disponível em: <<https://bloggeron-pole.com/2020/06/instagram-quietly-admitted-algorithm-bias-but-how-will-it-fight-it/>>. Acesso em: 23 fev. 2022.

ARE, Carolina, “The Shadowban Cycle: an autoethnography of pole dancing, nudity and censorship on Instagram”. *Feminist Media Studies*, maio 2021, pp. 1-18. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/14680777.2021.1928259>>. Acesso em: 20 nov. 2021.

ARRIAGADA, Arturo; IBÁÑEZ, Francisco, “‘You Need At Least One Picture Daily, if Not, You’re Dead’: Content Creators and Platform Evolution in the Social Media Ecology”. *Social Media + Society*, julho-setembro 2020, pp. 1-12.

BANET-WEISER, Sarah. *Authentic TM: the politics of ambivalence in a brand culture*. New York and London: New York University Press.

BANET-WEISER, Sarah, “Popular Feminism and Popular Misogyny”. In: RIBEIRO, Nelson; HENRICHSEN, Jennifer R. *Media and Populism: 1st Lisbon Winter School for the Study of Communication*. 15-19 de Janeiro de 2019, pp. 68-80.

BANET-WEISER, Sarah, “Branding the Post-Feminist Self: Girls’ Video Production and YouTube”, 2011, p. 13. Disponível em: <https://socyeventsucsc.files.wordpress.com/2011/02/banet-weiser_reading.pdf>. Acesso em: 24 mai. 2022.

BARBOUR, M. “[Cum]munity Standards: Resisting Online Sexual Harassment and Abuse”. In: POWELL, Anastasia; FLYNN, Asher; SUGIURA, Lisa (Ed.). *The Palgrave Handbook of Gendered Violence and Technology*. Switzerland: Palgrave Macmillan, 2021, pp. 19-39.

BARBOUR, Morgan, “Cummunity Standards”. *GenderIT.org*, 25.05.2020. Disponível em: <<https://genderit.org/feminist-talk/cummunity-standards>>. Acesso em: 13 mar. 2022.

BARBOUR, Morgan, “About Morgan”. Disponível em: <<https://www.morganbarbour.com/about-1>>. Acesso em: 26 abr. 2021.

BECK, Willa, “Why are Certain Bodies being Censored on IG? : Uye Surana”. *Office*, 18.11.2021. Disponível em: <<http://officemagazine.net/why-are-certain-bodies-being-censored-ig-uye-surana>>. Acesso em: 22 abr. 2022.

BENTES, Anna, “A Gestão Algorítmica da Atenção: enganchar, conhecer e persuadir”. In: POLIDO, Fabrício; ANJOS, Lucas; BRANDÃO, Luíza (Org.). *Políticas, Internet e Sociedade*. Belo Horizonte: IRIS, 2019, pp. 222-234.

BRIDGES, Lauren E., “Digital Failure: Unbecoming the ‘good’ data subject through entropic, fugitive, and queer data”. *Big Data & Society*, janeiro-junho 2021, pp. 1-17.

BROOKS, Rachel; HOUSTON, Stacey; WADSWORTH, Christina, “Atualizações da Equipe de Equidade: o que realizamos e o que ainda está por vir”. *Instagram*, 26.04.2021. Disponível em: <<https://about.instagram.com/pt-br/blog/announcements/equity-updates-what-weve-been-up-to-and-what-comes-next>>. Acesso em: 08 mar. 2022.

BUOLAMWINI, Joy; GEBRU, Timnit, “Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification”. *Proceedings of Machine Learning Research*, v. 81, 2018, pp. 1-15. Disponível em: <<http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf>>. Acesso em: 28 fev. 2022.

BUTLER, Judith. *Excitable Speech: A politics of the performative*. New York/London: Routledge, 1997.

BUTLER, Judith. *Problemas de Gênero: Feminismo e subversão da identidade* (tradução de Renato Aguiar), 18ª edição. Rio de Janeiro: Civilização Brasileira, 2019.

CALDEIRA, Sofia P.; RIDDER, Sander De; BAUWEL, Sofie Van. “Between the Mundane and the Political: Women’s Self-Representations on Instagram”. *Social Media + Society*, julho-setembro 2020, pp. 1-14.

CARVALHO, Luana. *Último teste*. 11 out. 2020. Instagram: @lxccarvalho. Disponível em: <https://www.instagram.com/p/CGOB_2DHHMO/>. Acesso em: 2 fev. 2022.

CASTRO, Edgardo. *Vocabulário de Foucault* – Um percurso pelos seus temas, conceitos e autores (traduzido por Ingrid Müller Xavier). Belo Horizonte: Autêntica, 2009.

CHANG, Emily. *Brotopia: Breaking Up the Boys’ Club of Silicon Valley*. New York: Penguin, 2018.

CHRISTIE, Lacey-Jade, “Instagram censored one of these photos but not the other. We must ask why”. *The Guardian*, 19.10.2020. Disponível em: <<https://www.theguardian.com/technology/2020/oct/20/instagram-censored-one-of-these-photos-but-not-the-other-we-must-ask-why>>. Acesso em: 17 nov. 2021.

CODING RIGHTS; INTERNETLAB. Violências de gênero na internet: diagnóstico, soluções e desafios. Contribuição conjunta do Brasil para a relatora especial da ONU sobre violência contra a mulher. São Paulo, 2017. Disponível em: <https://www.internetlab.org.br/wp-content/uploads/2017/11/Relatorio_ViolenciaGenero_UNU.pdf>. Acesso em: 20 nov. 2021.

COELHO, Bruna Martins, “A Filosofia a Golpes de Martelo de Elsa Dorlin: ‘se defender’ contra o terrorismo de Estado – resenha crítica”. *Revista TEL*, v. 10, n. 2, julho-dezembro 2019, pp. 282-297.

COLLING, Leandro, “A emergência dos ativismos das dissidências sexuais e de gêneros no Brasil da atualidade”. *Revista sala preta*, v. 18, n. 1, 2018, pp. 152-167.

COOK, Jesselyn, “Instagram’s CEO Says Shadow Banning ‘Is Not A Thing’. That’s Not True”. *HuffPost*, 25.02.2020. Disponível em: <https://www.huffpost.com/entry/instagram-shadow-banning-is-real_n_5e555175c5b63b9c9ce434b0>. Acesso em: 2 jan. 2022.

COTE, Amanda, “‘I Can Defend Myself’: Women’s Strategies for Coping With Harassment While Gaming Online”. *Games and Culture*, v. 12, n. 2, 2017, pp. 136-155.

COTTER, Kelley, “‘Shadowbanning is not a thing’: black box gaslighting and the power to independently know and credibly critique algorithms”. *Information, Communication & Society*, 2021, pp. 1-18.

COTTER, Kelley, “Playing the visibility game: How digital influencers and algorithms negotiate influence on Instagram”. *New Media & Society*, v. 21, n. 4, 2019, pp. 895-913.

CWYNAR-HORTA, Jessica, “The Commodification of the Body Positive Movement on Instagram”. *Stream: Culture, Politics, Technology*, v. 8, n. 2, 2016, pp. 36-56.

D’IGNAZIO, Catherine; KLEIN, Lauren F.. *Data feminism*. Cambridge, Massachusetts; London, England: The MIT Press, 2020.

DIAS, Tatiana; RIBEIRO, Paulo Victor; LOPES, Débora, “O Grande Experimento. Facebook Papers: as provas que faltavam para mostrar como a rede manipulou você”. *The Intercept Brasil*, 01.12.2021. Disponível em: <<https://theintercept.com/2021/12/01/facebook-papers-provas-rede-manipulou-voce/#two>>.

Acesso em: 9 mar. 2022.

DICKSON, E.J., “Why Did Instagram Confuse These Ads Featuring LGBTQ People for Escort Ads?”. *Rolling Stone*, 11.07.2019. Disponível em: <<https://www.rollingstone.com/culture/culture-features/instagram-transgender-sex-workers-857667/>>. Acesso em: 20 nov. 2021.

DORLIN, Elsa. *Autodefesa: Uma filosofia da violência*. São Paulo: Ubu Editora, 2020.

DOVE, Rain. *I am NOT “Female”*. 27 jun. 2018. Instagram: @raindovemodel. Disponível em: <<https://www.instagram.com/p/BkibsngHp73/>>. Acesso em: 2 fev. 2022.

DUFFY, Brooke Erin, “Algorithmic precarity in cultural work”, *Communication and the Public*, v. 5, n. 3-4, 2020, pp. 103-107.

DUFFY, Brooke Erin, “The Gendered Politics of Digital Brand Labor”. *Antenna*, 18.03.2015. Disponível em: <<https://blog.commarks.wisc.edu/2015/03/18/the-gendered-politics-of-digital-brand-labor/>>. Acesso em: 24 mai. 2022.

DUFFY, Brooke Erin; HUND, Emily, “Gendered Visibility on Social Media: Navigating Instagram’s Authenticity Bind”. *International Journal of Communication*, v. 13, 2019, pp. 4983-5002.

DUGUAY, Stefanie; BURGESS, Jean; SUZOR, Nicolas, “Queer women’s experiences of patchwork platform governance on Tinder, Instagram, and Vine”. *Convergence: The International Journal of Research into New Media Technologies*, v. 26, n. 2, 2020, pp. 237-252.

ELKIN-KOREN, Niva, “Contesting algorithms: Restoring the public interest in content filtering by artificial intelligence”. *Big Data & Society*, julho-dezembro 2020, pp. 1-13.

EL-WARDANY, Salma, “Like Our Society, Instagram In Biased Against Women of Colour”. *Refinery29*, 10.12.2020. Disponível em: <<https://www.refinery29.com/en-gb/2020/12/10150275/shadow-ban-instagram-censorship-women-of-colour>>. Acesso em: 17 nov. 2021.

EMMELHAINZ, Irmgard, “Authoritarianism and the Cybernetic Episteme, or the Progressive Disappearance of Everything on Earth”. *E-flux Journal*, n. 122, novembro 2021. Disponível em: <<https://www.e-flux.com/journal/122/430488/authoritarianism-and-the-cybernetic-episteme-or-the-progressive-disappearance-of-everything-on-earth/>>. Acesso em: 2 fev. 2022.

EUBANKS, Virginia. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York: St. Martin's Press, 2017.

EUROPEAN DIGITAL RIGHTS, "How online ads discriminate: Unequal harms of online advertising in Europe", 2021. Disponível em: <https://edri.org/wp-content/uploads/2021/06/EDRi_Discrimination_Online.pdf>. Acesso em: 3 fev. 2022.

FEDERICI, Silvia. *Calibã e a Bruxa: mulheres, corpos e acumulação primitiva*. São Paulo: Elefante, 2017.

FERRIER, Morwenna, "How to beat Instagram and Facebook's female nipple ban? Use male nipples instead". *The Guardian*, 07.07.2015. Disponível em: <<https://www.theguardian.com/lifeandstyle/shortcuts/2015/jul/07/instagram-facebook-female-nipple-ban-use-male-nipples-instead>>. Acesso em: 3 fev. 2022.

FOUCAULT, Michel. *The Politics of Truth*. New York: Semiotext(e), 1997.

FOUCAULT, Michel. *Nascimento da Biopolítica: Curso dado no Collège de France (1978-1979)*. São Paulo: Martins Fontes, 2008.

FOUCAULT, Michel. *Segurança, território, população: curso dado no Collège de France (1977-1978)*. São Paulo: Martins Fontes, 2008.

FRIER, Sarah. *No Filter: The Inside Story of Instagram*. New York, London, Toronto, Sydney: Simon & Schuster, 2020.

GERRARD, Ysabel; THORNHAM, Helen, "Content Moderation: Social Media's Sexist Assemblages". *New Media & Society*, v. 22, n. 7, 2020, pp. 1266-1286.

GILL, Rosalind, "Postfeminist media culture: elements of a sensibility". *European Journal of Cultural Studies*, v. 10, n. 2, 2007, pp. 147-166.

GILLESPIE, Tarleton, "Content moderation, AI, and the question of scale". *Big Data & Society*, julho-dezembro 2020, pp. 1-5.

GILLESPIE, Tarleton, "Governance of and by platforms". In: BURGESS, Jean; POELL, Thomas; MARWICK, Alice (Eds.). *The SAGE Handbook of Social Media*. London: SAGE Publications Ltd., 2018, pp. 254-278.

GOLDBECK, Jennifer; ASH, Summer; CABRERA, Nicole, "Hashtags as online communities with social support: A study of anti-sexism-in-science hashtag movement". *First Monday*, v. 22, n. 9, 2017. Disponível em: <<https://firstmonday.org/ojs/index.php/fm/article/view/7572/6529>>. Acesso em: 2 fev. 2022.

GOMES, Carla; e SORJ, Bila, "Corpo, geração e identidade: a Marcha das vadias no Brasil". *Revista Sociedade e Estado*, v. 29, n.2, maio-agosto 2014, p. 433-447.

GORWA, Robert; BINNS, Reuben; KATZENBACH, Christian, "Algorithmic content moderation: Technical and political challenges in the automation of platform governance". *Big Data & Society*, janeiro-junho 2020, pp. 1-15.

GRADY, Lora, “Women are Calling Out Instagram for Censoring Photos of Fat Bodies”. *Flare*. 18.10.2018. Disponível em: <<https://www.flare.com/news/instagram-censorship-fat-plus-size/>>. Acesso em: 17 nov. 2021.

GRETCHEN, Faust, “Hair, Blood and the Nipple: Instagram Censorship and the Female Body”. In: KÖHN, Steffen; FOX, Samantha; TERRY, Mike; FRÖMMING, Urte Undine (Coord.). *Digital Environments*. Ethnographic Perspectives across Global Online and Offline Spaces. Bielefeld: Verlag, 2017, pp 159-170.

HAIMSON, Oliver, “Social Media Sites are Censoring Transgender, Black and Conservative Users – But for Very Different Reasons”. *Salty*. Disponível em: <<https://saltyworld.net/social-media-sites-are-censoring-transgender-black-and-conservative-users-but-for-very-different-reasons/>>. Acesso em: 7 mar. 2022.

HAN, Byung-Chul. *Psicopolítica: O neoliberalismo e as novas técnicas de poder* (tradução de Maurício Liesen). Belo Horizonte: Editora Áyiné, 2018.

HARCOURT, Bernard. “Epilogue 2/13”. *Critique & Praxis* 13/13, 27.10.2018. Disponível em: <<http://blogs.law.columbia.edu/praxis1313/bernard-e-harcourt-epilogue-2-13/>>. Acesso em: 17 nov. 2021.

HARCOURT, Bernard. *Exposed: Desire and Disobedience in the Digital Age*. Cambridge, Massachusetts/London, England: Harvard University Press, 2015.

HAUGEN, Frances, “Facebook Whistleblower Frances Haugen: The 60 Minutes Interview”. Entrevista concedida a Scott Pelley. *60 Minutes*, 2021. (13:36). Disponível em: <https://www.youtube.com/watch?v=_Lx5VmAdZSI>. Acesso em: 2 abr. 2022.

HEBRON, Micol, “Nipples! Articles about my digital pasty designed to fight censorship of female nipples”. Disponível em: <<http://micolhebron.artco-deinc.com/pages/nipples-articles-about-my-digital-pasty-designed-to-fight-censorship-of-female-nipples/>>. Acesso em: 2 jan. 2022.

HORA, Bianca Dias da; MARTINS, Milena Fagundes, “A Cultura do Cancelamento e Suas Facetas: Justiça Social, Intransigências e Disputas Narrativas”. *Revista Iniciacom*, v. 10, n. 3, 2021. Disponível em: <<https://revistas.inter-com.org.br/index.php/iniciacom/article/view/4110/pdf>>. Acesso em: 28 fev. 2022.

HORA, Nina da. “Aula Aberta: Racismo e Algoritmos com Nina da Hora”. Disponível em: <<https://www.ninadahora.dev/>>. Acesso em: 3 fev. 2022.

HORNE, Julian Van, “Shadowbanning is a Thing – and It’s Hurting Trans and Disabled Advocates”. *Salty*. Disponível em: <<https://saltyworld.net/shadowbanning-is-a-thing-and-its-hurting-trans-and-disabled-advocates/>>. Acesso em: 7 mar. 2022.

HUNT, Elle, “Male or female? Genderless Nipples account challenges Instagram’s sexist standards”. *The Guardian*, 19.01.2017. Disponível em: <<https://www.theguardian.com/world/2017/jan/19/genderless-nipples-generates-50000-followers-on-instagram>>. Acesso em: 2 fev. 2022.

INSTAGRAM, “An update on our work to tackle abuse on Instagram”. *Instagram*, 11.02.2021. Disponível em: <<https://about.instagram.com/blog/announcements/an-update-on-our-work-to-tackle-abuse-on-instagram>>. Acesso em: 3 fev. 2022.

INSTAGRAM, “Community Guidelines”. Disponível em: <https://help.instagram.com/477434105621119/?helpref=hc_fnav>. Acesso em: 2 fev. 2022.

INSTAGRAM, “Diretrizes da Comunidade”. Disponível em: <<https://www.facebook.com/help/instagram/477434105621119>>. Acesso em: 9 abr. 2022.

INSTAGRAM, “I don’t think Instagram should have taken down my post”. Disponível em: <https://www.facebook.com/help/instagram/280908123309761?helpref=uf_permalink>. Acesso em: 2 fev. 2022.

INSTAGRAM, “Feed”. Disponível em: <<https://business.instagram.com/instagram-feed>>. Acesso em: 10 fev. 2022.

IT FOR CHANGE, India; INTERNETLAB, Brazil, “General Assembly Report on Gender Justice: Submission to the Special Rapporteur on the right to freedom of opinion and expression”, junho 2021. Disponível em: <<https://itforchange.net/sites/default/files/1738/ITforChangeInternetLab-SubmissionToGARReportGenderJustice-June2021.pdf>>. Acesso em: 20 mar. 2022.

JACOBS, Julia, “Will Instagram Ever ‘Free the Nipple’?”. *The New York Times*, 22.11.2019. Disponível em: <<https://www.nytimes.com/2019/11/22/arts/design/instagram-free-the-nipple.html>>. Acesso em: 22 fev. 2022.

JANE, Emma A, “‘Dude...stop the spread’: Antagonism, agonism and #manspreading on social media”. *International Journal of Cultural Studies*, v. 20, n. 5, 2016, pp. 459-475.

JANE, Emma, “Online misogyny and feminist digilantism”, *Continuum*, v. 30, n. 3, 2016, p. 284-297. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/10304312.2016.1166560>>. Acesso em: 20 fev. 2022.

JANE, Emma. *Misogyny Online: A short (and brutish) history*. Los Angeles, London, New Delhi, Singapore, Washington DC, Melbourne: SAGE, 2017.

JONES, Phil. *Work Without the Worker: Labour in the Age of Platform Capitalism*. Londres: Verso, 2021.

JOSEPH, Chanté, “Instagram’s murky shadow bans just serve to censor marginalized communities”. *The Guardian*, 08.11.2019. Disponível em: <<https://www.theguardian.com/commentisfree/2019/nov/08/instagram-shadow-bans-marginalised>>

[communities-queer-plus-sized-bodies-sexually-suggestive](#)>. Acesso em: 17 nov. 2021.

KIRKPATRICK, David. *The Facebook Effect: The Inside Story of the Company That Is Connecting the World*. New York, London, Toronto, Sydney: Simon & Schuster, 2010.

KURASAWA, Fuyuki; RONDINELLI, Elisabeth; KILICASLAN, Gulay, “Evidentiary activism in the digital age: on the rise of feminist struggles against gender-based online violence”. *Information, Communication & Society*, v. 24, n. 14, 2021, pp. 2174-2194.

RODIS, Paulina d. C. Inara, “Let’s (re)tweet about racism and sexism: responses to cyber aggression toward Black and Asian women”. *Information, Communication & Society*, v. 24, n. 14, 2021, pp. 2153-2173.

LAPLANCHE, Jean; PONTALIS, Jean-Bertrand. *Vocabulário da psicanálise*. Santos: Martins Fontes, 1970 *apud* GONZALEZ, Lélia. *Por um Feminismo afro-latino-americano: ensaios, intervenções e diálogos* (Organização: Flávia Rios e Márcia Lima). Rio de Janeiro: Zahar, 2020, p. 115.

LEAVER, Tama; HIGHFIELD, Tim; ABIDIN, Crystal. *Instagram: Visual Social Media Cultures*. Cambridge, UK; Medford, MA, USA: Polity, 2020.

LESSA, Patrícia, “Práticas de resistência e poéticas lesbianas nas obras de Karin Schwarz e Alma López”. *ANPUH SP – XX Encontro Estadual de História*, 2016. Disponível em: <http://www.encontro2016.sp.anpuh.org/resources/anais/48/1462818428_ARQUIVO_PatriciaLessaunesprtf2016.pdf>. Acesso em: 3 mar. 2022.

LIM, Merlyna; ALRASHEED, Ghadah, “Beyond a technical bug: Biased algorithms and moderation are censoring activists on social media”. *The Conversation*, 16.05.2021. Disponível em: <<https://theconversation.com/beyond-a-technical-bug-biased-algorithms-and-moderation-are-censoring-activists-on-social-media-160669>>. Acesso em: 7 abr. 2022.

LINDSTRÖM, Björn; BELLANDER, Martin; SCHULTNER, David T; CHANG, Allen; TOBLER, Philippe; AMODIO, David M., “A computational reward learning account of social media engagement”. *Nature Communications*, v. 12, n. 1311, 2021. Disponível em: <<https://www.nature.com/articles/s41467-020-19607-x#citeas>>. Acesso em: 22 mai. 2022.

LLANSÓ, Emma J., “No amount of ‘AI’ in content moderation will solve filtering’s prior-restraint problem”, *Big Data & Society*, Janeiro-junho 2020, pp. 1-16.

MAJSIAK, Brianna, “Instagram and Facebook, Please Stop Sexualizing Our Mastectomy Photos”. *Everyday Health*, 22.10.2020. Disponível em: <<https://www.everydayhealth.com/breast-cancer/instagram-and-facebook-please-stop-sexualizing-our-mastectomy-photos/>>. Acesso em: 2 fev. 2022.

MOMBAÇA, Jota. *Não vão nos matar agora*. Rio de Janeiro: Cobogó, 2021.

MANOVICH, Lev. *Instagram and Contemporary Image*. Disponível em: <<http://manovich.net/index.php/projects/instagram-and-contemporary-image>>. Acesso em: 10 out. 2021.

MORALES, Eduardo, “How Instagram Upholds the Censoring of Black Bodies and the Suppression of POC-Run Accounts”. *An Injustice*, 03.08.2020. Disponível em: <<https://aninjusticemag.com/how-instagram-upholds-the-censoring-of-black-bodies-and-the-suppression-of-poc-run-accounts-cd1d33ff4dc5>>. Acesso em: 17 nov. 2021.

MOSSERI, Adam, “Introducing New Ways to Protect Our Community from Abuse”. *Instagram*. 10.08.2021. Disponível em: <<https://about.instagram.com/blog/announcements/introducing-new-ways-to-protect-our-community-from-abuse>>. Acesso em> 3 fev. 2022.

MOSSERI, Adam. *How does Instagram work? Or more specifically – why do I see what I see on Instagram?*. 28 jun. 2021. Twitter: @mosseri. Disponível em: <https://twitter.com/mosseri/status/1409533965606821888?ref_src=twsrc%5Etfw%7Ctwcamp%5Etweetembed%7Ctwterm%5E1409533965606821888%7Ctwgr%5E%7Ctwcon%5Es1_&ref_url=https%3A%2F%2Fblog.hootsuite.com%2Fsocial-media-updates%2Finstagram%2Finstagram-says-shadowbanning-does-not-exist%2F>. Acesso em: 1 fev. 2022.

MOSSERI, Adam, “Shedding More Light on How Instagram Works”. *Instagram*, 08.06.2021. Disponível em: <<https://about.instagram.com/blog/announcements/shedding-more-light-on-how-instagram-works>>. Acesso em: 30 jul. 2021.

MYERS WEST, Sarah, “Censored, suspended, shadowbanned: user interpretations of content moderation on social media platforms”. *New Media & Society*, v. 20, n. 11, novembro 2018, pp. 4366-4383.

MYERS-WEST, Sarah, “Censored, suspended, shadowbanned: User interpretations of content moderation on social media platforms”. *New Media & Society*, v. 20., n. 11, 2018, pp. 4366-4383.

NETO, João Leite Ferreira, “A Analítica da Subjetivação em Michel Foucault”, *Revista Polis e Psique*, v. 7, n. 3, 2017, pp. 7-25.

NG, Eve, “No Grand Pronouncements Here...: Reflections on Cancel Culture and Digital Media Participation”. *Television and New Media*, v. 21, n. 5, 2020, pp. 621-627.

NOBLE, Safiya. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press, 2018.

OLIVEIRA, Triscila. *Vamos ver...* Niterói. 12 out. 2020. Instagram: @soulanja. Disponível em: <<https://www.instagram.com/p/CGOk-xiBx1e/>>. Acesso em: 8 fev. 2022.

O'HARA, Mary Emily, "Queer and Feminist Brands say they are Being Blocked from Running Ads on Instagram and Facebook". *MTV News*, 19.07.2019. Disponível em: <<https://www.mtv.com/news/3131929/queer-and-feminist-brands-say-they-are-being-blocked-from-running-ads-on-instagram-and-facebook/>>. Acesso em: 2 fev. 2022.

O'HARA, Mary Emily, "Queer and Feminist Brands Say They Are Being Blocked from Running Ads on Instagram and Facebook". *MTV News*, 19.07.2019. Disponível em: <<http://www.mtv.com/news/3131929/queer-and-feminist-brands-say-they-are-being-blocked-from-running-ads-on-instagram-and-facebook/>>. Acesso em: 23.02.2022.

O'NEIL, Cathy. *Algoritmos de Destruição em Massa: Como o Big Data Aumenta a Desigualdade e Ameaça a Democracia* (tradução de Rafael Abraham). Santo André: Editora Rua do Sabão, 2020.

OYEWÙMÍ, Oyèrónké. *A invenção das mulheres: construindo um sentido africano para os discursos ocidentais de gênero* (tradução de Wanderson Flor do Nascimento), 1ª edição. Rio de Janeiro: Bazar do Tempo, 2021.

PAASONEN, Susanna; ARE, Carolina, "Sex in the shadows of celebrity". *Porn Studies*, v. 8, n. 4, setembro 2021, pp. 411-419.

PAASONEN, Susanna; JARRETT, Kylie; LIGHT, Ben. *NSFW: Sex, Humor and Risk in Social Media*. Cambridge, MA: The MIT Press, 2019.

PAIVA, Raquel, "Epistemologia compreensiva: garantia para pluralidade de vozes". In: PRATA, Nair; PESSOA, Sônia Caldas; e ANDRADE, Ivanise Hilbig (Org.). *Um mundo e muitas vozes: da utopia à distopia?*. São Paulo: Intercom, 2021, pp. 103-114.

PARSONS, Vic, "Instagram accused of 'sexualising, policing and censoring trans bodies'". Disponível em: <<https://www.pinknews.co.uk/2021/04/22/instagram-trans-bodies-censorship-we-deserve-to-be-here/>>. Acesso em: 20 abr. 2022.

PELE, Antonio, "Data Necropolitics". *Abolition and Democracy*, v. 13, n. 13. Disponível em: <<https://blogs.law.columbia.edu/abolition1313/antonio-pele-data-necropolitics-2/>>. Acesso em: 20 mai. 2022.

PERRIGO, Billy, "Inside Facebook's African Sweatshop". *Time*, 17.02.2022. Disponível em: <<https://time.com/6147458/facebook-africa-content-moderation-employee-treatment/>>. Acesso em: 21 mai. 2022.

PETRE, Caitlin; DUFFY, Brooke Erin; HUND, Emily, "'Gaming the System': Platform Paternalism and the Politics of Algorithmic Visibility". *Social Media + Society*, outubro-dezembro 2019, pp. 1-12.

PLAN INTERNATIONAL. “Free to be online? Girls’ and young women’s experiences of online harassment”, 2020. Disponível em: <<https://plan-international.org/publications/freetobeonline>>. Acesso em: 9 nov. 2020.

PRECIADO, Paul B. *Testo Junkie: sexo, drogas e biopolítica na era farmacopornográfica*. São Paulo: N-1, 2018.

QUIJANO, Aníbal, “Colonialidade do Poder e Classificação Social”. In: SANTOS, Boaventura de Sousa; MENESES, Maria Paula (Org.). *Epistemologias do Sul*. Coimbra: Almedina, 2009, pp. 73-118.

RADSCH, Courtney, “Shadowban/Shadow Banning”. In: BELLI, Luca; ZINGALES, Nicolo; CURZI, Yasmin (Eds.). *Glossary of Platform Law and Policy Terms*. Rio de Janeiro: FGV Direito Rio, 2021, pp. 295-296.

RAPOSO, Paulo, “‘Artivismo’: Articulando dissidências, criando insurgências”. *Cadernos de Arte e Antropologia*, v. 4, n. 2, 2015, pp. 3-12.

REID, Rebecca, “Plus-size model says she’s been shadow banned by Instagram because of her body”. *Metro*, 11.06.2018. Disponível em: <<https://metro.co.uk/2018/06/11/plus-size-model-says-shadow-banned-instagram-body-7621367/>>. Acesso em: 17 nov. 2021.

RIBEIRO, Viviana Mendes, “O controle da reprodução está no coração da organização sócio-política. A discussão sobre a descriminalização da prática da interrupção da gravidez voluntária no Brasil. *Revista da Defensoria Pública da União*, n. 15, janeiro-junho 2021, pp. 231-250.

RICAURTE, Paola. “Data Epistemologies, The Coloniality of Power and Resistance”. *Television & New Media*, v. 20, n. 4, 2019, pp. 350-365.

ROBERTS, Sarah T. *Behind the Screen: Content Moderation in the Shadows of Social Media*. New Haven: Yale University Press, 2019.

ROUVROY, Antoinette, “Des données sans personne: le fétichisme de la donnée à caractère personnel à l’épreuve de l’idéologie des Big Data”. Les rapports du Conseil d’État (ancienne collection Étude et documents du Conseil d’État, 2014, pp. 407-422. Disponível em: <<https://researchportal.unamur.be/fr/publications/des-donn%C3%A9es-sans-personne-le-f%C3%A9tichisme-de-la-donn%C3%A9e-%C3%A0-caract%C3%A8re->>. Acesso em: 2 mar. 2022.

ROUVROY, Antoinette. “De la surveillance au profilage”. *Philosophie Magazine*, hors-série n. 36. Foucault. Le courage d’être soi, 2018. Disponível em: <https://www.academia.edu/36177927/De_la_surveillance_au_profilage_Philosophie_magazine_Hors-s%C3%A9rie_N_36_Foucault_Le_courage_d%C3%AAtre_soi_2018>. Acesso em: 20 jul. 2021.

ROUVROY, Antoinette, “The end(s) of critique: data behaviourism versus dua process”. In: HILDEBRANDT, Mireille; VRIES, Katja de. *Privacy, Due Process*

and the Computational Turn: The philosophy of law meets the philosophy of technology. New York: Routledge, 2017, p. 143-168.

ROUVROY, Antoinette, “La vie n’est pas donnée”, *Études digitales*, v. 2, n. 2, 2017, pp. 195-217. Disponível em: <<https://pure.unamur.be/ws/portalfiles/portal/54932539/8162.pdf>>. Acesso em: 20 jun. 2021.

ROUVROY, Antoinette, “L’art de ne pas changer le monde”. *La Revue Nouvelle*, n. 8, 2016, pp. 44-50.

ROUVROY, Antoinette, “Algorithmic Governmentality and the Death of Politics: An Interview with Antoinette Rouvroy”. *Green European Journal*, 27.03.2020. Disponível em: <<https://www.greeneuropeanjournal.eu/algorithmic-governmentality-and-the-death-of-politics/>>. Acesso em: 2 nov. 2021.

ROUVROY, Antoinette; BERNS, Thomas, “Governamentalidade algorítmica e perspectivas de emancipação: o dispar como condição de individuação pela relação?” (tradução de Pedro Henrique Andrade), *Revista Eco Pós*, dossiê Tecnopolíticas e Vigilância, v. 18, n. 2, 2015, PP. 36-56.

RYAN-MOSLEY, Tate, “How digital beauty filters perpetuate colorism”. *Artificial Intelligence*, 15.08.2021. Disponível em: <<https://www.technologyreview.com/2021/08/15/1031804/digital-beauty-filters-photoshop-photo-editing-colorism-racism/>>. Acesso em: 11 nov. 2021.

SALTY, “Exclusive: Victoria’s Secret Influence on Instagram’s Censorship Policies”. *Salty*, 22.11.2019. Disponível em: <<https://saltyworld.net/exclusive-victorias-secret-influence-on-instagrams-censorship-policies/>>. Acesso em: 28 jan. 2022.

SALTY. *Unlike other big brands and media outlets, Salty is unable to advertise on @Instagram or @Facebook because the algorithms assume that Salty is “promoting escort services”*. 8 jul. 2019. Instagram: @salty.world. Disponível em: <<https://www.instagram.com/p/BzrYz9jF4Qt/>>. Acesso em: 23 fev. 2022.

SARMENTO, Rayza, “Ativismo Feminista Online: mapeando eixos de atuação”. *Revista Sul-Americana de Ciência Política*, v. 7, n. 1, 2021, pp. 19-37.

SCHARF, Christina, “The Psychic Life of Neoliberalism: Mapping the Contours of Entrepreneurial Subjectivity”. *Theory, Culture & Society*, v. 33, n. 6, 2016, pp. 107-122.

SIBILIA, Paula. *O show do eu: a intimidade como espetáculo*. Rio de Janeiro: Nova Fronteira, 2008.

SILVA, Tarcizio, “Colonialidade automatizada: algoritmos e aprendizado de máquina nas plataformas de mídias sociais”. In: PRATA, Nair; PESSOA, Sônia Caldas; ANDRADE, Ivanize Hilbig de (Org.). *Um mundo e muitas vozes: da utopia à distopia?*. São Paulo: Intercom, 2021, pp. 139-155.

SILVA, Tarcizio. “Controvérsias sobre Danos Algorítmicos: discursos corporativos sobre discriminação codificada”. *Revista Observatório*, v. 6, n. 4, julho-setembro 2020, pp. 1-17.

SILVA, Tarcizio, “Linha do Tempo do Racismo Algorítmico: casos, dados e reações”, 2019. Disponível em: <<https://tarciziosilva.com.br/blog/destaques/posts/racismo-algoritmico-linha-do-tempo/>>. Acesso em: 22 fev. 2022.

SMITH, Adam. “Instagram Boss Says It Will Change Algorithm To Stop Mistreatment Of Black Users, Alongside Other Updates”. *Independent*, 16.06.2020. Disponível em: <<https://www.independent.co.uk/life-style/gadgets-and-tech/news/instagram-black-lives-matter-racism-harassment-bias-algorithm-a9567946.html>>. Acesso em: 17 nov. 2021.

SMITH, Shakira; HAIMSON, Oliver L; FITZSIMMONS, Claire; BROWN, Nikki Echarte (Salty Algorithmic Bias Collective in collaboration with Michigan University). “Censorship of Marginalized Communities on Instagram”, 27.09.2021. Disponível em: <<https://saltyworld.net/product/exclusive-report-censorship-of-marginalized-communities-on-instagram-2021-pdf-download/>>. Acesso em: 2 dez. 2021.

SOUZA, Emanuella de; LEÃO, Odilanir; BOAKARI, Francis, “Nós por nós: protagonismo e resistência de mulheres afrodescendentes no Instagram”. *Educação, Cultura e Comunicação*, v. 12, n. 24, 2021, pp. 78-94. Disponível em: <<http://publicacoes.unifatea.edu.br/index.php/ECCOM/article/view/1358>>. Acesso em: 9 nov. 2021.

SPISÁK, Sanna; PIRJATANNIEMI, Elina; PAALANEN, Tommi; PAASONEN, Susanna; VIHLMAN, Maria, “Social Networking Sites’ Gag Order: Commercial Content Moderation’s Adverse Implications for Fundamental Sexual Rights and Wellbeing”. *Social Media + Society*, abril-junho 2021, pp. 1-9.

SURINGAR, G.A. Fighting Instagram Censorship of the Female Nipple: @exotic.cancer and @genderless_nipples Seeking the Boundaries of the Instagram Guidelines to Critique the Sexualisation of the Female Body. Dissertação de Mestrado. Universidade de Utrecht. 29f. TCS Postcolonial and Gender Studies, 2020.

SWIFT, Skylar, “I Modeled Menstrual Products as a Trans Man and Instagram Let Neonazis Harass Me”. *Salty*. Disponível em: <<https://www.saltyworld.net/i-modeled-menstrual-products-as-a-trans-man-and-instagram-let-neonazis-harass-me/>>. Acesso em: 23 dez. 2021.

THOMPSON, Laura, “I can be your tinder nightmare”: Harassment and misogyny in the online sexual marketplace”. *Feminism & Psychology*, v. 28, n. 1, 2018, pp. 69-89.

TRERÉ, Emiliano. “The Dark Side of Digital Politics: Understanding the Algorithmic Manufacturing of Consent and the Hindering of Online Dissidence”. *IDS Bulletin: Transforming Development Knowledge*, v. 47, n. 1, Janeiro 2016, pp. 127-138.

TVARDOVSKAS, Luana Saturnino, “Cópia e paródia: práticas de si na arte contemporânea de mulheres e a crítica feminista à História da Arte”. In: SILVA, Edilene; OLIVEIRA, Susane Rodrigues de; e ZANELLO, Valeska (Org.). *Gênero, subjetivação e perspectivas feministas*. Brasília: TechnoPolitik, 2019, pp. 309-328.

UN WOMEN. “Online and ICT facilitated violence Against women and girls during COVID-19”, 2020. Disponível em: <<https://www.unwomen.org/en/digital-library/publications/2020/04/brief-online-and-ict-facilitated-violence-against-women-and-girls-during-covid-19>>. Acesso em: 26 set. 2021.

UNITED STATES GENERAL ACCOUNTING OFFICE. *Data Mining: Federal Efforts Cover a Wide Range of Uses*. Washington: GAO, 2014.

UNITED STATES, SECURITIES AND EXCHANGE COMMISSION. Form 10-K for the fiscal year ended December 31, 2021, Meta Platforms, Inc. Disponível em: <<https://d18rn0p25nwr6d.cloudfront.net/CIK-0001326801/14039b47-2e2f-4054-9dc5-71bcc7cf01ce.pdf>>. Acesso em: 3 abr. 2022.

VALENTE, Mariana; NERIS, Natália, “Elas vão feminizar a Internet?”. *Revista SUR*, v. 15, n. 27, 2018, p. 105-119.

VELKOVA, Julia; KAUN, Anne, “Algorithmic resistance: media practices and the politics of repair”. *Information, Communication & Society*, v. 24, n. 4, 2021, pp. 523-540.

VENDEMIA, Megan A.; DEANDREA, David C.; BRATHWAITE, Kyla N., “Objectifying the body positive movement: The effects of sexualizing and digitally modifying body-positive images in Instagram”. *Body Image*, v. 38, 2021, pp. 137-147.

VOLKERS, R., “Perverse Media: How Instagram limits the potential of feminist art”. *EUC Student Academic Journal – ESAJ*, n. 2. Disponível em: <<https://esaj.nl/Roos-Volkers-2>>. Acesso em: 2 fev. 2022.

WALTER, Bruno Eduardo Procopiuk; HENNIGEN, Inês, “Problematizando a Governamentalidade Algorítmica a Partir do Sistema de Recomendação da Netflix”. *Psicologia e Sociedade* [online], v. 33, e227258, 2021, pp. 1-15. Disponível em: <<https://doi.org/10.1590/1807-0310/2021v33227258>>. Acesso em: 2 fev. 2022.

WATSON, Imogen, “Tackling LGBT+ blocklists: publishers and agencies rally against digital censorship”. *The Drum*, 05.03.2020. Disponível em: <<https://www.thedrum.com/news/2020/03/05/tackling-lgbt-blocklists-publishers-and-agencies-rally-against-digital-censorship>>. Acesso em: 20 abr. 2022.

WILLETT, Cynthia; WILLETT, Julie e SHERMAN, Yael D., “The Seriously Erotic Politics of Feminist Laughter”. *Social Research: An International Quarterly*, v. 79, n. 1, primavera 2012, pp. 217-246.

WORLD WIDE WEB FOUNDATION. “*Women’s Rights Online: closing the digital gender gap for a more equal world*”, 2020. Disponível em: <<https://web-foundation.org/research/womens-rights-online-2020/>>. Acesso em: 24 set. 2021.

ZOLIDES, Andrew, “Gender moderation and moderating gender: Sexual content policies in Twitch’s community guidelines”. *New Media & Society*, v. 23, n. 10, 2021, pp. 2999-3015.

ZUBOFF, Shoshana. *A Era do Capitalismo de Vigilância: A luta por um futuro humano na nova fronteira do poder*. Rio de Janeiro: Intrínseca, 2020.

ZUCKERBERG, Mark, “A Blueprint for Content Governance and Enforcement”. *Facebook*, 2018 (última edição em 05.05.2021). Disponível em: <<https://www.facebook.com/notes/751449002072082/>>. Acesso em: 2 fev. 2022.

Anexo



Figura 1 – Postagem de @pollyoliveirareal em 20.02.2021



Figura 2 – Postagem de @pollyoliveirareal em 04.03.2021



Figura 3 – Postagem de @cummunitystandards em 20.12.2020



Figura 4 – Postagem de @cummunitystandards em 20.12.2020



Figura 5 – Postagem de @communitystandards em 26.03.2020