

Liberdade de expressão e discurso de ódio na Internet

Mariana G. Valente

In Liberdade de Expressão e as Novas Mídias, José Eduardo Faria (org.), 2020

Em 2015, a Polícia Federal e o Ministério Público Estadual de São Paulo abriram uma investigação contra a página na internet chamada “Tio Astolfo”. Sem qualquer indicação pública de autoria, a página continha diversos conteúdos misóginos e racistas; ficou amplamente conhecida por publicar uma série de guias especificando instruções para cometer violência sexual contra alunas em diversas universidades do país, em “festas e baladas”, e especificamente contra mulheres homossexuais.¹ Em maio de 2018, foi deflagrada a Operação Bravata, da Polícia Federal, que expediu uma série de mandados de busca e apreensão e prendeu preventivamente um homem considerado um dos principais responsáveis pelo Tio Astolfo e por outros sites de conteúdo misógino, racista e neonazista.² O alvo da prisão já havia sido preso por motivos semelhantes em 2013, e durante anos ameaçou de morte a blogueira feminista e professora da Universidade Federal do Ceará Lola Aronovich.³ Pelas dificuldades em conseguir que a Polícia Federal investigasse seu caso, Lola tornou-se uma vocal ativista contra perseguições e manifestações de ódio na Internet. A Lei n. 13.642/18, aprovada pelo Congresso e sancionada pelo presidente Michel Temer em 2018, que estabelece que “crimes praticados por meio da rede mundial de computadores que difundam conteúdo misógino” devem ser investigados pela Polícia Federal e não pelas Polícias Cíveis, é entendida como fruto desse ativismo e leva seu nome.

Com a Lei Lola, a misoginia, definida como a propagação de ódio ou aversão às mulheres, passou a ser objeto do ordenamento jurídico brasileiro, ainda que em uma lei que apenas estabelece uma atribuição de investigação. O *discurso de ódio*, embora faça parte do vocabulário utilizado na esfera pública e na academia, não está previsto como tal no direito brasileiro; além da misoginia, a Lei Antirracista (Lei n. 7716/1989) prevê o crime de praticar, induzir ou incitar a discriminação ou preconceito de raça, cor, etnia, religião ou procedência nacional,⁴ e outros casos de ódio contra

¹ “Com dúvidas sobre autoria, polícia investiga site de apologia ao estupro ‘Tio Astolfo’”. *Revista Exame*, 31 de julho de 2015.

² MACEDO, Fausto, e AFFONSO, Julia. “‘Bravata’, da PF, ataca racismo virtual”, *in Estadão*, 10 de maio de 2018.

³ VIANNA, José e KANIAK, Thais. “PF prende uma pessoa em operação contra racismo, ameaça, incitação e terrorismo praticados na internet”. *RPC Curitiba e GI PR*, 10 de maio de 2018.
<https://g1.globo.com/pr/parana/noticia/pf-faz-operacao-contr-crimes-de-racismo-ameaca-e-incitacao-e-terrorismo-praticados-na-internet.ghtml>.

⁴ Art. 20. Praticar, induzir ou incitar a discriminação ou preconceito de raça, cor, etnia, religião ou procedência nacional. Pena: reclusão de um a três anos e multa. (Redação dada pela Lei nº 9.459, de 15/05/97).

outras populações poderiam ser enquadrados genericamente como crimes contra a honra, incitação ao crime ou apologia de crime, pelo Código Penal. Tramita na Câmara dos Deputados (PL 5003/01) e aguarda julgamento no STF desde 2012 (ADO 26) o tema da criminalização da homofobia.

Em outras jurisdições, existem proibições legais ao discurso de ódio como tal, mas, como explora Jeremy Waldron, com uma ampla gama de diferentes acepções, contornos e justificações (2012, p. 8): algumas proíbem ameaças, outras a difamação em si, outras o potencial do discurso de levar a consequências discriminatórias para além dele, ou seja, na existência de vínculo causal entre o discurso e um dano exterior a ele. No núcleo convergente estaria o uso de palavras deliberadamente abusivas, insultantes, ameaçadoras ou inferiorizantes, a membros de minorias vulneráveis, de forma a instigar o ódio contra elas (2012, p. 16). Além disso, discursos direcionados a uma pessoa em particular, mas direcionado a um grupo subalternizado em específico, são também parte do que se entende por discurso de ódio.⁵ Os sentidos dados ao termo, no debate na esfera pública brasileira, são relativamente variáveis.

A discussão em torno do discurso de ódio, em nível internacional, não é recente – pelo contrário, ainda que em outros termos, ela se dá ao menos desde as negociações na Conferência das Nações Unidas pela Liberdade de Informação em 1948 (BOYLE, 1992).⁶ Mas as transformações nas formas de se comunicar das últimas décadas renovaram o debate, agudizando preocupações e trazendo novos atores à baila. A possibilidade de se comunicar amplamente sem a barreira ou o crivo da mídia tradicional vem possibilitando a rápida difusão de conteúdos discriminatórios, violentos, ameaçadores, e colocando pressão nos intermediários das comunicações online, ou seja, as plataformas digitais.⁷

⁵ Adilson José Moreira (2017, p. 32) lembra, citando Kasper Lippert-Rasmussen, que a discriminação em si independe das características reais do sujeito. “Por exemplo, sabemos que a homofobia é um comportamento hostil contra homens e mulheres homossexuais. Porém, pessoas heterossexuais são vítimas frequentes dessa prática discriminatória. Isso acontece porque um dos seus propósitos é controlar os contornos sociais da heterossexualidade, pessoas heterossexuais estão sendo constantemente vigiadas e obrigadas a se adequar às expectativas sociais referentes a formas de comportamento desejáveis”.

⁶ A conferência falhou, mas o artigo 19 da Declaração Universal dos Direitos Humanos de 1948 foi escrito lá; as consequências se fizeram sentir também no Pacto Internacional dos Direitos Civis e Políticos de 1954 (arts. 19 e 20) e na Convenção contra o Genocídio de 1948 (BOYLE, 1992, p. 6).

⁷ Conforme aponta Juliano Cappi, em 2013, o Simon Wiesenthal Center, uma organização não governamental de proteção dos direitos humanos da comunidade judaica de Los Angeles, publicou um estudo indicando que teria ocorrido um aumento em 30% na disseminação de conteúdos discriminatórios no Twitter naquele ano, em relação aos anteriores, por meio do aumento de uso de hashtags para compartilhamento de conteúdo de ódio (20 mil hashtags em 2013, contra 5 mil em 2012), e perfis ligados a movimentos neonazistas com mensagens de ódio na própria descrição do perfil (2017, p. 51).

A relação entre internet e grupos sociais subalternizados, também chamados minoritários (em compreensão sociológica, e não numérica, ou seja, no sentido de não ser majoritário ou dominante), envolve paradoxos: as redes sociais são grande aliadas da organização desterritorializada e da possibilidade de desenvolvimento de discursos não dominantes, mas também permitem violências desintermediadas e potencializadas pela disseminação instantânea. E, é claro, na internet também vão se manifestar os preconceitos e estereótipos que existem na sociedade, e talvez de forma mais virulenta, associada ao caráter descorporificado da comunicação online.⁸

No livro “O que é discriminação?”, Adilson José Moreira argumenta que existem claras relações entre preconceitos, estereótipos e a discriminação:

Certos comportamentos podem ser motivo de violência em função da associação com o comportamento de minorias. A incessante circulação de estigmas que afirmam a inferioridade essencial de minorias corrobora a percepção de que todos os membros de grupos minoritários são inferiores a todos os membros do grupo majoritário, mesmo quando os primeiros têm um status cultural e material superior aos segundos (2017, p. 195).⁹

A discriminação, por sua vez, tem um papel central na estratificação social: “práticas discriminatórias acontecem dentro de uma cultura social que permite a construção e circulação de estigmas negativos” (2017, p. 197). Com isso, não se pressupõe um vínculo causal “forte” ou direto entre um discurso e a estratificação, mas a ideia, defendida por teóricos de estudos raciais críticos (*critical racial studies*)¹⁰ e por autoras feministas, de que o discurso produz efeitos no mundo

⁸ Apesar do lugar-comum de que a disseminação do ódio na internet estaria ligada ao anonimato, parece-nos que a experiência da descorporificação, e de uma conseqüente diminuição no sentimento de empatia, tem mais a dizer. Temos observado uma disseminação grande de conteúdos que poderiam ser classificados como odiosos sendo feita por pessoas que não se escondem atrás de perfis falsos ou páginas anônimas – pelo contrário, fazem suas afirmações utilizando o próprio nome.

⁹ “Preconceitos são avaliações sobre os membros de um segmento social baseadas em generalizações que podem ser verdadeiras em relação a alguns deles, mas que certamente não podem ser estendidas a todos os seus membros por causa da variedade existente entre os seres humanos” (MOREIRA, 2017, p. 40). Estereótipos, por sua vez, “designam os modelos mentais que dirigem a percepção das pessoas, expressando a internalização de valores e códigos culturais construídos por aqueles que possuem poder” (p. 41). Adilson José Moreira expressa também como esses estereótipos são internalizados no processo de socialização, e que as pessoas agem de acordo com eles mesmo sem ter consciência disso.

¹⁰ A *outsider jurisprudence*, traduzida como “teoria do direito marginal”, é um método jurídico defendido pelos estudos críticos raciais, de acordo com o qual as pretensões jurídicas de universalidade e universalidade não correspondem à experiência particular das pessoas que não pertencem a grupos dominantes, e que a descrição da experiência “from the bottom” (MATSUDA, 1989, p. 2324) é o que deve ser buscado. Para uma teoria do direito marginal feminista, ver Littleton, Equality and Feminist Legal Theory, 48 University of Pittsburgh Law Review, 1987.

(DELGADO, 1982; MATSUDA, 1989), de nível cultural e material (MOREIRA, 2017), desequilibrando a possibilidade de pessoas perseguirem sua dignidade no sentido realização de suas concepções de boa vida, de pertencimento à sociedade em boas condições e com status igualitário, ainda que a causalidade seja difusa (WALDRON, 2012).¹¹ O foco nos efeitos do discurso é uma das abordagens do problema, e que parece adequada por levar em conta as experiências concretas das pessoas que são vitimizadas por ele; há também quem defenda que o foco deva estar na intencionalidade, ou ainda no conteúdo, e há autores que ressaltam a necessidade de análise de contexto (FARIS *et alii*, 2016). Apesar das muitas disputas sobre os contornos do conceito, é relativamente assentado que o discurso de ódio é uma conduta, e não uma mera opinião – e é isso que faz com que faça sentido regulá-lo, ainda que a extensão dessa regulação seja objeto de importantes dissensos.

O efeito dessa conduta é intimidatório, e a intimidação em questão depende de muitos fatores contextuais. Quem diria que postar uma imagem de uma lâmpada fluorescente no perfil de uma pessoa nas redes sociais poderia ser uma mensagem intimidatória? Se a pessoa for LGBT e o post acontecer no Brasil, certamente é. Em 2010, um jovem foi espancado até ficar desacordado na Avenida Paulista em São Paulo, por outros 5 jovens, que portavam duas lâmpadas fluorescentes, em agressão com clara motivação homofóbica.¹² A lâmpada virou um símbolo de resistência da comunidade LGBT na cidade – formou-se um grupo chamado “A Revolta da Lâmpada”, que começou a organizar atos de rua anuais; enquanto isso, usuários de redes sociais começaram a denunciar o recebimento de fotos de lâmpadas como resposta a postagens suas.¹³ Muito provavelmente, uma pessoa estrangeira que não conhecesse esse contexto não seria capaz de entender a mensagem de ódio e a ameaça contidas nessa imagem. É evidente o efeito de intimidação e desincentivo à ocupação do espaço público por parte de membros da comunidade LGBT, ou seja, como a conduta de expressão de ódio tem efeitos no exercício de outros direitos; de forma menos direta e mais difusa, a presença desses discursos e sua normalização no espaço

¹¹ Um debate relevante sobre as razões e a legitimidade da proibição aos discursos de ódio é aquele que se deu entre Jeremy Waldron e Ronald Dworkin, e que gira em torno de um conceito de dignidade defendido por Waldron. Uma das análises sobre o debate no Brasil pode ser encontrada em PITERMAN, 2017.

¹² O bando teria dito: “V*** tem que morrer”. GOMES, Paulo. Acusados de ataque com lâmpada na Paulista são multados em R\$ 129 mi. Folha de S. Paulo, 17 de outubro de 2018. <https://www1.folha.uol.com.br/cotidiano/2018/10/agressores-sao-condenados-a-pagar-r-25-mil-por-homofobia-em-sp.shtml>.

¹³ De acordo com relatos de funcionários de empresas de aplicações de internet que trabalham com políticas de segurança, à autora, em 2017.

público podem trazer à população LGBT graves desvantagens de caráter cultural e material. O discurso de ódio não existe de forma dissociada dos fenômenos relacionados – racismo, sexismo, homofobia, transfobia – e faz assim parte de um sistema de dominação social, por influenciar muitos aspectos da vida dos indivíduos (MOREIRA, 2017, p. 33). Ligados que estão a estratificações sociais, assim, os discursos de ódio podem ser entendidos como tal quando são direcionados a grupos que se encontram em uma momentânea ou perene de desigualdade de status (e contribuem para sua manutenção). O quanto o efeito é sentido é ligado à situação de subalternização.

É bastante difundida a ideia de que o discurso de ódio coloca um conflito entre o direito à liberdade de expressão e outro direito – a igualdade, ou a dignidade. A discussão frequentemente é resolvida por um “balanceamento dos direitos” que não só é problemática em si, mas que também não corresponde ao problema. Em uma pesquisa realizada entre 1998 e 2010, Marta Machado, Marcia Lima e Natália Neris (2016) identificaram que, decidindo sobre os casos de injúria racial e racismo, que em grande parte referem-se a insultos raciais, os tribunais brasileiros muitas vezes desconsideram o caráter racial do conflito, e em outros muitos casos decidem que não havia intencionalidade específica de discriminar e ofender raça e etnia. Em suma, as pesquisadoras não puderam identificar critérios para o reconhecimento do conflito racial, e encontraram mesmo uma resistência em se considerar o insulto racista uma forma de racismo. “O sistema de justiça raramente emite declarações de que o insulto racial constitui um ilícito e não deve ser tolerado socialmente”, afirmam as autoras.¹⁴ É comum, também em decisões recentes envolvendo redes sociais, que o discurso racista sequer seja reconhecido como tal pelos magistrados brasileiros, o que faz com que a discussão não chegue a ser posta nos termos de um conflito entre direitos.

Assim, tratar do discurso de ódio de forma abstrata, sem observar como a categoria é mobilizada pelos movimentos sociais, por seus opositores, ou pelo judiciário coloca o risco de a discussão tornar-se etérea, apartada das nossas relações concretas de desigualdade e sofrimento.¹⁵ Isso leva à importante indagação sobre a densidade e a conveniência da própria categoria “discurso

¹⁴ As autoras reconhecem também as limitações do próprio direito penal para lidar com o fenômeno racista: a sua força vem de atos anteriores, de cumulação e repetição, mas, no processo, o conflito é individualizado (MACHADO; LIMA; NERIS, 2016, pp. 25-6). A dificuldade de responsabilizar um indivíduo e um ato singular por uma prática que é ritualizada é afirmada também por Judith Butler (1997), para quem o processo é uma redução de um embate político amplo à forma jurídica e a uma autoridade investida em decidir de forma definitiva.

¹⁵ Ver também MATSUDA, 1989; sobre o Brasil em particular, GABINA, 2015; PEREIRA, 2017.

de ódio” no contexto brasileiro. Embora ainda seja necessária mais pesquisa sobre o assunto, há bons indícios de que no Brasil os discursos que causam danos a grupos sociais subalternizados não tenham sido tratados por muito tempo sob a chancela do “discurso de ódio”, de forma que a maior parte das discussões importa versões estrangeiras do debate, sem muitas mediações. O Supremo Tribunal Federal, decidindo sobre o caso Ellwanger entre 2002 e 2004,¹⁶ caso considerado paradigma para a posição do STF sobre discurso de ódio (GROSS, 2017, p. 70), utilizou-se tanto de referências norte-americanas quanto alemãs (princípio da proporcionalidade e da ponderação de princípios),¹⁷ sem contanto que se desenvolvesse propriamente sobre a questão em contexto brasileiro de uma forma mais ampla.¹⁸

Hoje, para acrescentar uma camada de complexidade, as plataformas de internet “legislam” sobre o comportamento de seus usuários em seus termos de uso, que determinam o que eles podem ou não expressar ali. Como os termos de uso são globais, as definições que elas desenvolvem começam a ser também circuladas e vão se tornando lugares comum, mas o problema tem uma

¹⁶ HC 82424-RS. Siegfried Ellwanger foi condenado por racismo pelo Tribunal de Justiça do Rio Grande do Sul, por um livro que continha conteúdo discriminatório contra judeus, e que negava o holocausto; a condenação naquele momento caberia por tratar-se o racismo de um crime imprescritível. Ellwanger impetrou Habeas Corpus alegando que judeus não constituem uma raça e portanto o crime estaria prescrito, mas teve a ordem denegada.

¹⁷ É comum a crítica a como essa ponderação de princípios vem sendo feita de forma simplificada no Brasil. “Qualquer situação que possa ser identificada como um ‘caso difícil’ é rapidamente identificada como um conflito entre direitos fundamentais, capaz de ser solucionado racionalmente pelo método da ponderação de princípios, ou, no caso do mercado de ideias, por meio da definição do conteúdo da cláusula da liberdade de expressão e de seus ‘limites imanentes’, quais sejam, o perigo claro e iminente e a artificial diferenciação entre discurso/ação. Discurso de ódio, no Brasil, assume estes contornos: é um problema de direitos fundamentais; esses direitos - liberdade e igualdade - estão em conflito; ou, há certos tipos de discursos que não fazem parte da essência da liberdade de expressão; esse conflito se resolve por meio da ponderação de princípios ou de outra teoria constitucional que forneça o embasamento teórico capaz de garantir, aparentemente, o distanciamento do jurista das demais questões complexas que envolvem o caso, além de alçar a sua tarefa a uma espécie de atividade especial, inatingível pelos demais interessados em resolver o problema.” (PEREIRA, 2017, p. 103). Ver também críticas de GABINA, 2015, bem como Virgílio Afonso da Silva em *Comparing the Incommensurable: Constitutional Principles, Balancing, and Rational Decision*, em *Oxford Journal of Legal Studies* (2011) e Lênio Streck, *Ponderação de Normas no novo CPC? É o caos*, em *Conjur* (2015) (apud GROSS, 2017, p. 29).

¹⁸ Uma tentativa de fazer frente a esse vácuo jurídico é o Projeto de Lei n. 7582/14, proposto pela deputada Maria do Rosário (PT/RS), que “define crimes de ódio e intolerância e cria mecanismos para coibi-los”. O PL é amplo e estabelece crimes, medidas protetivas, e recomendações ao Poder Executivo e aos órgãos do sistema de justiça; define quais são as características protegidas (classe e origem social, condição de migrante, refugiado ou deslocado interno, orientação sexual, identidade e expressão de gênero, idade, religião, situação de rua e deficiência), e, dentre os crimes, estabelece o crime de discurso de ódio: “Art. 5º Praticar, induzir ou incitar a discriminação ou preconceito, por meio de discurso de ódio ou pela fabricação, comercialização, veiculação e distribuição de símbolos, emblemas, ornamentos, distintivos ou propaganda, por qualquer meio, inclusive pelos meios de comunicação e pela internet, em razão de classe e origem social, condição de migrante, refugiado ou deslocado interno, orientação sexual, identidade e expressão de gênero, idade, religião, situação de rua e deficiência”. Agradeço à Natália Neris pela lembrança.

natureza extremamente contextual: variam de local para local tanto a definição de quais são os grupos que podem ser atingidos, quanto a forma como se dão as desigualdades e, portanto, as agressões.¹⁹

A tentativa de criar um conceito para discurso de ódio passa por indiferenciar formas diferentes de discurso discriminatório contra grupos diferentes, o que nos coloca diante de questões sobre o reconhecimento da discriminação em questão,²⁰ mas também estratégicas. Afinal, o discurso de ódio englobaria características que já encontram uma proteção legal, ainda que haja graves problemas na aplicação desse arcabouço jurídico, e outras que não, como é o caso da orientação sexual.

Nos Estados Unidos, como se sabe, “o balanço é inequivocamente voltado à proteção da liberdade de expressão” (BOYLE, 1992, p. 5); são entendidos como protegidos discursos que na maioria dos outros países são sancionados. A União Europeia e países singulares da Europa vêm adotando uma abordagem muito menos complacente tanto com os discursos quanto com o papel das plataformas de internet em coibi-los. Em 2016, a Comissão Europeia assinou um código de conduta (não vinculante) com o Facebook, o Twitter, o YouTube (Google) e a Microsoft, estabelecendo um compromisso delas em explicitar aos usuários quais são os conteúdos não permitidos, em promover iniciativas de “contradiscorso” (ou seja, de combate ao discurso de ódio por meio de discursos em oposição a ele), e em remover o que foi chamado de “discurso de ódio ilegal” – o que assume, veja, que existe discurso de ódio não ilegal – em até 24 horas a partir de uma notificação de qualquer pessoa solicitando a remoção. O código de conduta tomou o cuidado de estabelecer os parâmetros do problema: discurso de ódio ilegal é definido²¹ como “toda conduta incitando publicamente à violência ou ao ódio dirigido a um grupo de pessoas ou a um membro desse grupo, definido por referência a raça, cor, religião, descendência ou origem nacional ou étnica”, e salienta-se também o respeito à liberdade de expressão, entendida inclusive como protegendo discursos que “ofendem, chocam e incomodam o Estado ou qualquer setor da população”. De todo modo, discursos de ódio são apontados como não somente danosos àqueles

¹⁹ Sobre o tema, ver também SELLARS, 2016.

²⁰ Em um texto clássico tratando de discurso racista, Mari Matsuda justificou-se por não tratar ali de discursos de ódio baseados em gênero: “while I believe these forms of hate speech require public restriction, these forms also require a separate analysis because of the complex and violent nature of gender subordination, and the different way in which sex operates as a locus of oppression” (1989, 2332).

²¹ A partir da “Framework Decision 2008/913/JHA of 28 November 2008 on combating certain forms and expressions of racism and xenophobia by means of criminal law and national laws transposing it”, que definiu que discurso de ódio é um crime também quando acontece no ambiente online.

a quem se orientam, mas também como tendo efeitos negativos sobre “aqueles que defendem publicamente a liberdade, a tolerância e a discriminação nas nossas sociedades abertas [*open societies*], e como tendo um efeito deletério [*chilling effect*] no discurso democrático nas plataformas online”.²² Vale apontar, também, que o código de conduta foi desenvolvido a partir do marco dos ataques terroristas em Bruxelas em 2016,²³ o que indica uma visão de associação entre discurso de ódio e terrorismo na Europa.

Depois de um ano, a Comissão Europeia informou por meio de um release de imprensa²⁴ que o acordo teria implicado uma mudança na proatividade das empresas em lidar com a questão (apesar de o resultado ter variado muito de empresa para empresa): em média, elas teriam respondido às notificações 59% das vezes, mais que o dobro em relação aos seis meses antes da existência do acordo, e aumentou de 40% para 51% a resposta em 24 horas às notificações – 100% no caso do Facebook.

As grandes empresas que oferecem aplicações na internet vêm sendo publicamente pressionadas, podem também ter preocupações genuínas com o tipo de comunicação que ocorre em suas plataformas (e com as reações que anunciantes podem ter ao verem seus produtos associados a determinados discursos), mas sobretudo percebem que compromissos negociados dessa forma podem ser mais favoráveis a elas que uma regulação unilateral mais pesada. E a impressão não está equivocada: em 2018, a Alemanha, que em vários setores tem apresentado uma estratégia de regulação forte às plataformas digitais, aprovou uma nova lei para combate ao discurso de ódio na internet.²⁵ Entre outras determinações, a lei estabelece multas milionárias caso as plataformas não removam em até 24h conteúdos “claramente ilícitos”, que tenham sido notificados: 5 milhões de euros, podendo chegar até 50 milhões, em casos extremos ou de repetição. Para os conteúdos cuja ilegalidade é menos clara, o prazo é de uma semana.

O foco nas plataformas de internet é uma estratégia que parte da percepção de que há algo de novo nas formas de se comunicar, e que não somente a responsabilização individual dos agressos

²² Comissão Europeia, “Code of conduct on countering illegal hate speech online”, em https://ec.europa.eu/info/files/code-conduct-countering-illegal-hate-speech-online_en.

²³ <http://www.consilium.europa.eu/en/press/press-releases/2016/03/24-statement-on-terrorist-attacks-in-brussels-on-22-march/>.

²⁴ European Commission – Press Release. Countering online hate speech – Commission initiative with social media platforms and civil society shows progress. http://europa.eu/rapid/press-release_IP-17-1471_en.htm.

²⁵ https://www.bmjv.de/SharedDocs/Gesetzgebungsverfahren/Dokumente/BGBl_NetzDG.pdf?__blob=publicationFile&v=2,

por parte do Estado torna-se uma atividade fragmentada e pouco eficaz, como também que as empresas que detêm grandes aplicações de internet estão empreendendo esforços insuficientes na coibição dessas condutas – na versão mais desconfiada, que elas estariam inclusive lucrando com o discurso de ódio, na medida em que ações contra usuários que se engajam nele podem ser impopulares, ou que conteúdos sensacionalistas podem atrair atenção. Existe também implícita uma disputa de Estados com plataformas, receosos do poder que elas adquirem, e muito mais se elas são imunes à regulação, e ainda de diferentes indústrias entre si.

No Brasil, o Marco Civil da Internet (Lei 12.965/14) estabeleceu uma limitação de responsabilidade às plataformas em relação a conteúdos ilícitos de terceiros: no seu artigo 19, determinou que os provedores de aplicações de internet são responsáveis por esse conteúdo apenas quando não agem para torná-lo indisponível depois de uma decisão judicial nesse sentido. Ou seja, antes de um juiz determinar a remoção de um determinado conteúdo de uma plataforma, ela não tem responsabilidade – pode removê-lo de acordo com seus termos de uso, mas, se não o fizer, não responde por ele. Essa regra de responsabilidade de intermediários foi amplamente debatida quando o Marco Civil da Internet esteve em consulta pública, e a racionalidade por detrás dela era que esse era o modelo mais conducente à liberdade de expressão, porque uma responsabilização dessas plataformas por conteúdo de outras pessoas levaria a políticas de censura prévia ou a incentivos para que elas removessem conteúdos excessivamente, receosas de responsabilização. Duas exceções foram estabelecidas na lei: o caso do direito autoral, que foi remetido para discussão futura, e o caso da disseminação não consentida de imagens íntimas – o art. 21 determinou que, nesse caso, a mera notificação da vítima leva à responsabilidade da plataforma, se o conteúdo não for disponibilizado rapidamente.²⁶ O Marco Civil da Internet não contém qualquer previsão específica sobre discurso de ódio. Uma argumentação em defesa do modelo vigente seria a de que cabe a um juiz a determinação do que é ódio, já que essa resposta nem sempre é simples, e que fazer das plataformas o árbitro dessa questão não é desejável, inclusive por lhes dar poder demasiado. A Alemanha seguiu na tendência oposta, colocando as plataformas na posição de ter de tomar decisões rápidas, em especial nos casos de ilicitude patente. Decidir se esse modelo conduziu ou não a um bloqueio indevido de conteúdos é algo que vai requerer observação e pesquisa.

²⁶ Fizemos uma discussão dessa política e de seus efeitos no capítulo 3 do livro *O Corpo é o Código*, sobre disseminação não consentida de imagens íntimas (VALENTE *et alii*, 2016).

As plataformas acabam servindo também como instâncias de decisão em si a partir de seus termos de uso, que são praticamente as leis que se aplicam a um determinado ambiente online, mas que são desenvolvidas de forma privada, por vezes com instâncias onde participam grupos da sociedade civil. As grandes plataformas como Facebook e Twitter contam com times de moderadores de conteúdo, que analisam as denúncias feitas por usuários e por vezes conteúdos reportados por processos automatizados; em 2017, a ProPublica, uma organização de jornalismo investigativo norte-americana, divulgou documentos internos do Facebook de treinamento de seus moderadores, o que permitiu entender em detalhes como as regras da plataforma, publicadas de forma bem mais genérica, estavam sendo aplicadas. Como observaram Oliva e Antonialli (2018), as regras continham pontos questionáveis: para além de definir as categorias (gênero, identidade de gênero, raça, afiliação religiosa, etnia, origem nacional, orientação sexual, séria deficiência ou doença), estabeleciam que, quando essas categorias eram combinadas com outras não protegidas, a recomendação era a não remoção do conteúdo. Seria o caso de “crianças negras”, ou de “motoristas mulheres” – já que idade e profissão ou ocupação não são categorias protegidas. Contraditoriamente, segundo esse critério, discursos dirigidos a “homem branco” poderiam ser considerados de ódio.

As plataformas têm falado publicamente sobre suas dificuldades em lidar com discurso de ódio e medidas específicas – o Facebook declarou, em 2017, que ia aumentar o número de moderadores de 4,5 mil para 7,5 mil,²⁷ e anunciou em 2018 políticas muito mais detalhadas sobre o que fica e o que sai da plataforma.²⁸ Em 2018, também, o Twitter eliminou uma série de contas de usuários norte-americanos de extrema direita, ligadas a discursos de ódio e outros tipos de abuso, inclusive a de Alex Jones e seu website Infowars.²⁹ Várias das pessoas por trás dessas contas suspensas, simpatizantes (e muitos outros usuários) passaram a se encontrar em uma rede social alternativa, chamada “gab”, que se intitula “uma rede social que professa a liberdade de expressão, a liberdade individual e o fluxo livre de informação online. Todos são bem vindos”.³⁰ No *gab*, não há qualquer controle de discurso.

²⁷ ALLAN, Richard. Hard Questions: Hate Speech. Facebook, 27 de junho de 2017.

<https://newsroom.fb.com/news/2017/06/hard-questions-hate-speech/>.

²⁸ <https://br.newsroom.fb.com/news/2018/04/estamos-tornando-nossos-termos-e-politica-de-dados-mais-claros-sem-solicitar-novos-dados-no-facebook/>.

²⁹ “Twitter bans Alex Jones and Infowars for abusive behavior”, *BBC News*, 6 de setembro de 2018.

<https://www.bbc.com/news/world-us-canada-45442417>.

³⁰ <https://gab.ai>.

Esse êxodo aponta para as complexidades envolvendo a regulação do discurso pelas plataformas, embora isso não signifique que elas não devam fazê-lo. A responsabilidade por combater discurso de ódio deve ser compartilhada, e não só por empresas e pelo Estado, mas também pela sociedade, principalmente por meio da estratégia do contradiscurso, sobre a qual mito se vem discutindo. O amplo compromisso é essencial para que se garanta um ambiente de expressão em que as pessoas mais vulnerabilizadas não sejam censuradas por condutas que atingem sua dignidade e autonomia. Os debates em torno dos três atores envolvem complexidades próprias. Não é simples, mas exige um compromisso que precisa passar por concretizar e compartilhar compreensões sobre o caráter das diferentes desigualdades e os efeitos dos discursos sobre elas.

Bibliografia

BOYLE, Kevin. Overview of a Dilemma: censorship versus racism, in COLIVER, Sandra (ed.), *Striking a Balance: hate speech, freedom of expression and non-discrimination*. Londres: ARTICLE 19 e Human Rights Centre, University of Essex, 1992.

Butler, Judith. *Excitable Speech: A Politics of the Performative*. Nova York; Londres: Routledge, 1997.

CAPPI, Juliano. *Internet, big data e discurso de ódio: reflexões sobre as dinâmicas de interação no Twitter e os novos ambientes de debate político*. Tese de doutorado defendida na PUC-SP, 2017.

CODING RIGHTS; INTERNETLAB. *Violências contra mulher na internet: diagnóstico, soluções e desafios*. Contribuição conjunta do Brasil para a relatora especial da ONU sobre violência contra a mulher. São Paulo, 2017.

DELGADO, Richard. Words That Wound: A Tort Action for Racial Insults, Epithets, and Name-Calling, 17 *Harvard Law Review*, 133, 1982.

DIAS OLIVA, Thiago; ANTONIALLI, Dennys. Estratégias de enfrentamento ao discurso de ódio na internet: o caso alemão. *Revista Direitos Culturais*, v. 13, n. 30, 2018. Em <http://srvapp2s.urisan.tche.br/seer/index.php/direitosculturais/article/view/2656/1293>.

GABINA, Lourenço Paiva. *Discurso de ódio e jurisdição constitucional: uma abordagem pragmática*. Dissertação de mestrado defendida no Instituto de Direito Público (IDP), 2015.

FARIS, Rob; GASSER, Urs; ASHAR, Amar; JOO, Daisy. Understanding harmful speech online. *Berkman Klein Center Research Publication No. 2016-21*, 2016. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2882824

GROSS, Clarissa Piterman. Poder dizer ou não?: discurso de ódio, liberdade de expressão e a democracia liberal igualitária. Tese de doutorado defendida na Universidade de São Paulo, 2017.

MACHADO, Marta; LIMA, Marcia. NERIS, Natália. Racismo e Insulto Racial na Sociedade Brasileira: Dinâmicas de reconhecimento e invisibilização a partir do direito. *Revista Novos Estudos CEBRAP*, II-28, 2016.

MATSUDA, Mari. Public Response to Racist Speech: Considering the Victim's Story *Michigan Law Review*, Vol. 87, No. 8, Legal Storytelling (Aug., 1989), pp. 2320-2381.

MEYER-PFLUG, Samantha Ribeiro. Liberdade de expressão e discurso de ódio. São Paulo: *Revista dos Tribunais*, 2009, p. 102.

OLIVA, Thiago Dias. ANTONIALLI, Dennys Marcelo. Estratégias de enfrentamento ao discurso de ódio na internet: o caso alemão. *Revista Direitos Culturais*, v. 13, n. 30 (2018).

PEREIRA, Lizando Mello. Descolonizar o pensamento jurídico sobre os discursos de ódio: desconstruindo a cultura da violência. Dissertação de mestrado defendida na UFRN, 2017.

SARMENTO, Daniel. Livres e Iguais: estudos de Direito Constitucional. Rio de Janeiro: Lumen Juris, 2006, p. 208.

SELLARS, Andrew. "Defining Hate Speech." Berkman Klein Center Research Publication, n. 2016-20. <https://cyber.harvard.edu/publications/2016/DefiningHateSpeech>.

SWIEBEL, Joke; van der VEUR, Dennis. Hate crimes against lesbian, gay, bisexual and transgender persons and the policy response of international governmental organizations. In: *Netherlands Quarterly of Human Rights*, Antuérpia, v. 27, n. 4, p. 6, dez. 2009.

VALENTE, Mariana Giorgetti; NERIS, Natália; RUIZ, Juliana Pacetta; BULGARELLI, Lucas. O Corpo é o Código: estratégias jurídicas de enfrentamento ao revenge porn no Brasil. InternetLab: São Paulo, 2016.

WALDRON, Jeremy. The harm in hate speech. Cambridge: Harvard University, 2012.