



ADVOCACIA-GERAL DA UNIÃO
SECRETARIA-GERAL DE CONTENCIOSO

Excelentíssimo Senhor Ministro DIAS TOFFOLI, Relator do Recurso Extraordinário nº 1.037.396 (Repercussão Geral - Tema 987)

A UNIÃO, representada por seu Advogado-Geral^[1], nos autos do recurso extraordinário em epígrafe, interposto por FACEBOOK SERVIÇOS ONLINE DO BRASIL LTDA (parte do grupo controlador META PLATFORMS, INC.^[2]), vem, na qualidade de *amicus curiae*, **disponibilizar o conteúdo integral da audiência pública realizada no âmbito desta Advocacia-Geral da União em 22 de janeiro de 2025** (Audiência Pública para o Debate Técnico das Políticas de Moderação de Conteúdo das Plataformas Digitais), além de **requerer prioridade no julgamento de mérito do presente recurso paradigma**, pelos fatos e fundamentos jurídicos a seguir expostos:

1. DO CONTEXTO ATUAL DOS AUTOS

1. Trata-se de recurso extraordinário, com repercussão geral reconhecida, no qual se discute a **interpretação de dispositivos do Marco Civil da Internet** - Lei Federal nº 12.965, de 2014^[3] -, notadamente quanto à extensão do **regime de responsabilidade dos provedores de aplicação por conteúdo ilícito gerado por terceiros**.

2. A questão em exame foi assim delimitada por essa Suprema Corte: "*discussão sobre a constitucionalidade do art. 19 da Lei nº 12.965/2014 que determina a necessidade de prévia e específica ordem judicial de exclusão de conteúdo para a responsabilização civil de provedor de internet, websites e gestores de aplicativos de redes sociais por danos decorrentes de atos ilícitos praticados por terceiros*" (Tema 987).

3. A controvérsia foi objeto de audiência pública, realizada nos dias 28 e 29.03.2023, em conjunto com o RE n. 1.057.258/MG, de relatoria do Min. Luiz Fux, considerando o debate quanto ao *dever de empresa hospedeira de sítio na internet fiscalizar o conteúdo publicado e de retirá-lo do ar quando considerado ofensivo, sem intervenção do Judiciário* (Tema 533).

4. O julgamento de mérito teve início em 27.11.2024, ocasião na qual o Min. Relator, Dias Toffoli, proferiu voto pelo desprovimento do recurso extraordinário e a inconstitucionalidade do artigo 19 do Marco Civil da Internet, fixando-se como regra a responsabilização nos termos do art. 21 do mesmo diploma. Em acréscimo, concluiu que os provedores de aplicação **respondem civilmente de forma objetiva e independentemente de notificação, pelos danos decorrentes de conteúdos gerados por terceiros, nas hipóteses que especifica -- como racismo e violência contra grupos vulneráveis -- estabelecendo-se, ainda, uma série de deveres anexos de segurança, prevenção, cuidado e transparência**. Ao final, propôs a fixação da seguinte tese, no que aqui importa:

Decálogo contra a violência digital e a desinformação

(...) 3. O provedor de aplicações de internet **responde civilmente de forma objetiva e independentemente de notificação, pelos danos decorrentes de conteúdos gerados por terceiros**, nas seguintes hipóteses:

3.1. **quando recomendem, impulsionem (de forma remunerada ou não) ou moderem tais conteúdos**, havendo responsabilidade solidária com o respectivo

anunciante ou patrocinador, quando se tratar de anúncios publicitários ou de material patrocinado; (...)

3.4. quando configurarem práticas previstas no seguinte rol taxativo: (a) crimes contra o Estado Democrático de Direito (CP, art. 296, parágrafo único; art. 359-L, art. 359-M, art. 359-N, art. 359-P, art. 359-R); (b) atos de terrorismo ou preparatórios de terrorismo, tipificados pela Lei nº 13.260, de 2016; (c) crime de induzimento, instigação ou auxílio a suicídio ou à automutilação (CP, art. 122); **(d) crime de racismo (Lei nº 7.716, de 1989, arts. 20, 20-A, 20-B e 20-C); (e) qualquer espécie de violência contra a criança, o adolescente e as pessoas vulneráveis, incluídos os crimes previstos nos arts. 217-A a 218- C do Código Penal, com redação dada pelas Leis nº 12.015, de 2009, e nº 13.718, de 2018, e na Lei nº 8.069, de 1990, e observada a Lei nº 13.257, de 2016, e a Res. CONANDA nº 245, de 2024;** (f) **qualquer espécie de violência contra a mulher, incluindo os crimes da Lei nº 14.192, de 2021;** (g) infração sanitária, por deixar de executar, dificultar ou opor-se à execução de medidas sanitárias em situação de Emergência em Saúde Pública de Importância Nacional, nos termos do art. 10 da Lei nº 6.437, de 1977; (h) tráfico de pessoas (CP, art. 149- A); (i) incitação ou ameaça da prática de atos de violência física ou sexual (CP, art. 29 c/c arts. 121, 129, 213, 215, 215-A, 216-A, 250 e 251 c/c art. 147); (j) **divulgação de fatos notoriamente inverídicos ou gravemente descontextualizados que levem à incitação à violência física, à ameaça contra a vida ou a atos de violência contra grupos ou membros de grupos socialmente vulneráveis;** (k) divulgação de fatos notoriamente inverídicos ou descontextualizados com potencial para causar danos ao equilíbrio do pleito ou à integridade do processo eleitoral (Res. nº 23.610/2019, arts. 9-C e 9-D); (...)

6. Os provedores de aplicações de internet **devem atuar de forma responsável, transparente e cautelosa, a fim de assegurar um ambiente digital seguro, previsível e confiável,** baseado nos **princípios gerais da boa-fé, da função social e da prevenção e mitigação dos danos;**

7. Os provedores de aplicações de internet possuem os seguintes **deveres anexos:** (a) **atualizar e manter atualizados os ‘termos e condições de uso’** (ou documento equivalente) e os regulamentos adicionais, dando-lhes ampla publicidade; (...) (c) **elaborar os respectivos códigos de conduta;** (d) **estabelecer regras claras e procedimentos padronizados para a moderação de conteúdos,** aos quais se dará ampla publicidade; (e) **proceder à constante atualização dos critérios e métodos empregados para a moderação de conteúdos,** dando ampla publicidade aos usuários; (f) **combater a difusão de desinformação e de notícias fraudulentas nos ambientes virtuais, adotando as providências necessárias** para a neutralização de redes artificiais de distribuição de conteúdo destinados a promovê-las, assim como para a identificação do perfil/conta do qual se originou a desinformação ou notícia fraudulenta para fins de encaminhamento dos dados às autoridades competentes para as providências cabíveis; (g) **monitorar os riscos sistêmicos de seus ambientes digitais,** produzindo relatórios semestrais de transparência, dos quais constem os riscos identificados e as medidas preventivas e/ou de neutralização adotadas, aos quais se dará ampla publicidade; (...)

8. Também é **dever anexo** dos provedores de aplicações de internet o de **atuar diligentemente para prevenir e mitigar práticas ilícitas no âmbito de seus serviços,** adotando as providências necessárias para combater a disseminação de conteúdos ilegais que configurem as condutas indicadas taxativamente no subtópico 3.4 desta tese e para a identificação do perfil/conta do qual se originou

a desinformação ou notícia fraudulenta, encaminhando esses dados às autoridades competentes para as providências cabíveis" (grifou-se)

5. Ato contínuo, na sessão de 11.12.2024, o Min. Luiz Fux proferiu seu voto, **acompanhando** o Relator, para negar provimento ao recurso, mas propondo a manutenção do dispositivo impugnado, mediante tese no sentido de que *"1. A disposição do art. 19 do Marco Civil da Internet **não exclui** a possibilidade de responsabilização civil de provedores de aplicações de internet por conteúdos gerados por terceiros nos casos em que, tendo ciência inequívoca do cometimento de atos ilícitos, seja porquanto evidente, seja porque devidamente informados por qualquer meio idôneo, não procederem à remoção imediata do conteúdo"*.

6. No que aqui mais importa, o item 2 de sua proposta enumera os **ilícitos de caráter evidente e, em tais situações, reconhece o dever de monitoramento ativo por parte das plataformas e de retirada imediata, sem necessidade de notificação**^[4]:

"(...) 2. Considera-se evidentemente ilícito (item 1) o conteúdo gerado por terceiro que veicule **discurso de ódio, racismo, pedofilia, incitação à violência, apologia à abolição violenta do Estado Democrático de Direito e apologia ao Golpe de Estado**. Nestas hipóteses específicas, **há para as empresas provedoras um dever de monitoramento ativo**, com vistas à preservação eficiente do Estado Democrático de Direito. 3. Nos casos de postagens ofensivas à honra, à imagem e à privacidade de particulares, a ciência inequívoca da ilicitude por parte das empresas provedoras, necessária à responsabilização civil, dependerá de sua prévia e fundamentada notificação pelos interessados, que poderá ser realizada por qualquer meio idôneo, cabendo às plataformas digitais o dever de disponibilizar meios eletrônicos eficientes, funcionais e sigilosos para o recebimento de denúncias e reclamações de seus usuários que se sintam lesados. 4. É presumido, de modo absoluto, o efetivo conhecimento da ilicitude do conteúdo produzido por terceiros por parte da empresa provedora de aplicações de internet, nos casos de postagens onerosamente impulsionadas" (grifou-se)

7. Na sessão de 18.12.2024, o Min. Roberto Barroso, por sua vez, antecipou seu voto^[5], anuindo com a premissa de que uma compreensão literal do art. 19 do Marco Civil da Internet, *"**não dá proteção suficiente a direitos fundamentais, como a dignidade da pessoa humana, e a valores importantes para a democracia**"*. No entanto, apresentou solução hermenêutica com nuances diversas, no sentido de que as plataformas *"**devem ser responsabilizadas por conteúdos de terceiros caso deixem de tomar as providências necessárias para remover postagens com teor criminoso**"*. Em outras palavras, entende que

basta a mera notificação extrajudicial nas hipóteses de práticas criminosas, salvo nos casos de crimes contra honra - para os quais deve ser mantida a necessidade de reserva de jurisdição, como forma de proteção da liberdade de expressão.

8. A seu ver, em complemento, no lugar de um *"monitoramento ativo, com responsabilidade, independentemente de notificação, por cada conteúdo individual, as empresas devem ter o chamado dever de cuidado, ou seja, devem trabalhar para reduzir os riscos sistêmicos criados ou potencializados por suas plataformas"*. As medidas, a seu ver, **devem minimizar esses riscos e seus impactos negativos sobre direitos individuais e coletivos**, segurança e estabilidade democrática:

Assim, as plataformas devem atuar proativamente para que seu ambiente esteja livre de conteúdos gravemente nocivos, como pornografia infantil e crimes graves contra crianças e adolescentes, indução, instigação ou auxílio a suicídio ou a automutilação, tráfico de pessoas, atos de terrorismo, abolição violenta do Estado Democrático de Direito e golpe de Estado." (grifou-se)

9. Na sequência, **o julgamento foi suspenso em virtude de pedido de vista** formulado pelo Min. André Mendonça. A devolução, para fins de prosseguimento da votação, submete-se ao prazo de 90 (noventa) dias, contados da publicação da ata, conforme art. 134 do RISTF -- na redação estabelecida pela Emenda Regimental n. 58/2022.

2. DA SUPERVENIÊNCIA DE FATO RELEVANTE. GRAVES PREOCUPAÇÕES GLOBAIS. DELIBERADO RETROCESSO NO COMBATE AO DISCURSO DE ÓDIO E À DESINFORMAÇÃO NO AMBIENTE DIGITAL. CONEXÃO DIRETA COM AS QUESTÕES ORA SOB EXAME DESSA SUPREMA CORTE. AUDIÊNCIA PÚBLICA

10. Conforme amplamente divulgado na mídia, no último dia 7.01.2025, o Diretor Executivo da empresa *Meta*, Mark Zuckerberg (responsável pelos serviços Facebook, Whatsapp, Instagram e Threads, entre outros) realizou pronunciamento público^[6], por meio do qual anunciou **significativas alterações na política de funcionamento de tais aplicações**.

11. Fundamentalmente, na esteira da reconfiguração do quadro político do país sede de sua matriz, essas foram as principais mudanças em curso: *i)* a que envolve o *término do programa de parceria com agências independentes de checagem de fatos e combate à desinformação (thirdparty fact-checking program)*, sistema criado em 2016 e que vinha sendo desenvolvido em mais de 80 países e 60 idiomas^[7], o qual será substituído por um modelo de "notas de comunidade", produzidas a partir dos próprios usuários - iniciando por sua matriz; *ii)* **a fragilização da política de moderação de conteúdo, mediante o afrouxamento das regras dos denominados "Termos de Uso" de seus serviços**, isto é, especialmente com a retirada de restrições ao que pode ser postado em casos de **discurso de ódio**, como, por exemplo, no que diz respeito a **grupos vulneráveis em questões de gênero, orientação sexual, imigração, raça e religião**, e; *iii)* retomada do mecanismo algorítmico de **recomendação de postagens com "conteúdo cívico" ou político**, ou seja, discussões relacionadas à política, governo, eleições e questões sociais -- **mesmo oriundas de contas que os usuários não seguem** -- que frequentemente geram debates nas plataformas digitais e **havia sido limitadas em 2021**.

12. Esta segunda modificação (<https://transparency.meta.com/pt-br/policies/communitystandards/hate-speech/> - cf. íntegra dos "registros das alterações" de 7.01.2025 em anexo - Doc. 03) - que substitui a política até então vigente sobre **"conduta de ódio"** e já foi incluída na seção em português do sítio oficial, com **aplicação imediata**^[8] - configura-se, neste momento, o objeto de preocupação mais imediata. Vale recordar que os serviços da referida empresa são largamente utilizados no país, especialmente Instagram e Facebook, respectivamente **com cerca de 134 milhões e 113 milhões de usuários**^[9].

13. Como não poderia deixar de ser, diante da **gravidade das alterações para a integridade da informação e os esforços globais democráticos pela manutenção de um ambiente online seguro e saudável**, bem com da ampla escala de difusão destes serviços, **tal guinada ocasionou perplexidade e preocupação** nas mais diversas autoridades, figuras públicas e atores relevantes da sociedade civil em diferentes níveis, desde o Alto Comissariado da Organização das Nações Unidas para os Direitos Humanos, Volker Türk^[10], à jornalista vencedora do Prêmio Nobel da Paz, a filipina Maria Ressa^[11], a representantes da Comissão

Europeia para Soberania Tecnológica, Defesa e Pesquisa^[12]. No plano interno, vale notar que **mais de 200 (duzentas) entidades subscreveram o manifesto de repúdio "Contra o Retrocesso na Moderação de Conteúdo da Meta e os Ataques à Regulação Democrática do Espaço Digital"**, organizado pela Coalizão Direitos na Rede^[13].

14. Não por outra razão, renomadas publicações classificam o movimento de reviravolta da empresa como **parte de uma aproximação na tentativa de se opor a avanços regulatórios em curso em jurisdições como a europeia e, justamente, a brasileira**^[14]. Nessa linha, sem a assunção dos devidos deveres e responsabilidades inerentes ao respectivo (e extremamente lucrativo) modelo de negócios, corre-se o risco de se aprofundar ainda mais a concentração e as assimetrias de poder que ameaçam direitos fundamentais na era digital, como há muito mapeado^[15].

15. Além disso, cumpre lembrar o ponto que o Min. Flávio Dino já teve a oportunidade de destacar, ou seja, que sob *"a imperativa moldura da SOBERANIA, não é possível a uma empresa atuar no território de um país e pretender impor a sua visão sobre quais regras devem ser válidas ou aplicadas (...)"* (PET-Ref n. 12.404/DF, DJ 04.09.2024).

16. A título de ilustração, especificamente em relação à nova política de conteúdo, tal empresa passou a permitir a **associações de pessoas LGBTIA+ a transtornos mentais**, autorizando *"alegações de doença mental ou anormalidade quando baseadas em gênero ou orientação sexual, considerando discursos políticos e religiosos sobre transgenerismo e homossexualidade"*^[16].

17. Foi retirado integralmente, ainda, trecho anterior das diretrizes da plataforma que citava que o discurso de ódio online tinha efeitos no meio físico. Bem assim, será permitida a defesa de limitações profissionais baseadas em **gênero**. A *Meta* também **eliminou trecho que proibia a desumanização de mulheres** com bases em comparações com objetos inanimados domésticos e estados não-humanos, de **pessoas negras** como "equipamentos agrícolas" e de pessoas **transgênero** como "coisas"^[17].

18. Além disso, como divulgado^[18], passaram a serem admitidas ofensas mediante “*linguagem insultuosa no contexto de discussão de tópicos políticos ou religiosos, como ao discutir direitos transgêneros, imigração ou homossexualidade*”^[19]. Foram suprimidas, igualmente, restrições que proibiam a “*admissão de intolerância com base em características protegidas [de discriminação], incluindo, mas não se limitando a homofóbica, islamofóbica, racista*”^[20].

19. O ato revela, assim, um **catálogo de colisões com os fundamentos e dispositivos já manifestados nos votos proferidos** durante o presente julgamento sobre o Marco Civil da Internet, **a exemplo dos que contemplam os mencionados deveres anexos**. Nessa linha, como será retomado, **referidas alterações violam garantias constitucionais, legais e jurisprudenciais vigentes no país** – bem como contrariam diretrizes internacionais de proteção da integridade de informação e diversos tratados de proteção de direitos humanos.

20. De plano, não é demais relembrar a advertência feita pelo Min. Ricardo Cuevas, do Superior Tribunal de Justiça, ao examinar tema correlato: “*Os termos de uso dos provedores de aplicação, que autorizam a moderação de conteúdo, devem estar subordinados à Constituição, às leis e a toda regulamentação aplicável direta ou indiretamente ao ecossistema da internet, sob pena de responsabilização da plataforma*” (REsp n. 2139749/SP (3ª Turma, DJ de 30.08.2024).

21. Nesse contexto, a Procuradoria-Nacional de Defesa da Democracia, órgão integrante desta Advocacia-Geral da União, por meio da NOTIFICAÇÃO EXTRAJUDICIAL n. 00002/2025/PGU/AGU (Doc. 01), requereu que a Meta esclarecesse “*quais as medidas proativas que vêm sendo e que serão adotadas a respeito da arquitetura digital da plataforma, especialmente sobre o desenho dos algoritmos, no sentido de inarredável promoção e proteção dos direitos fundamentais, com respeito à legislação infraconstitucional e à CF/88*”.

22. Ademais, a União solicitou o esclarecimento das seguintes questões:

Importa destacar, sobretudo, quais providências que vêm sendo e que serão adotadas a respeito do dever de cuidado com relação à coibição de violência de gênero, proteção contra crianças e adolescentes, prevenção contra racismo, homofobia e transfobia, prevenção contra suicídio, óbices a discurso de ódio, dentre outros temas de direitos fundamentais.

Ademais, requer-se seja esclarecido se existe algum novo canal específico por meio do qual se podem apresentar denúncias sobre eventuais violações a direitos fundamentais diante das novas diretrizes que afrouxaram a verificação de desinformação nas plataformas digitais.

Bem assim, requer-se que seja esclarecido se haverá divulgação de relatório de transparência sobre a checagem de desinformação realizada por notas da comunidade, sobre quem tem sido colaborador mais comum (se os colaboradores pertencem a esferas diversas e eventualmente divergentes), qual porcentagem de participação cidadã, sobre as contradições eventualmente existentes entre notas da comunidades divergentes, dentre outras informações que demonstrem os resultados das notas da comunidade para verificação da sociedade em geral sobre sua eventual eficácia.

23. Em resposta datada de 13.01.2024 (Doc. 02), a *Meta* afirma que, *"no momento, está encerrando seu Programa de Verificação de Fatos independente apenas nos Estados Unidos, onde testaremos e aprimoraremos as Notas da Comunidade antes de dar início a qualquer expansão para outros países"*.

24. Por outro lado, informa que as **"mudanças recentemente anunciadas aos Padrões da Comunidade da Meta limitam-se às atualizações da Política de Conduta de Ódio, e têm como objetivo garantir maior espaço para a liberdade de expressão. Tais atualizações procuram simplificar o conteúdo da política de modo a permitir um debate mais amplo e conversas sobre temas que são parte de discussões em voga na sociedade"**.

25. Com base nas informações fornecidas pela *Meta*, observa-se que **as alterações na Política de Conduta de Ódio da plataforma digital são aplicáveis, desde já, aos serviços prestados pela empresa no Brasil.**

26. Diante da repercussão desse conjunto de efeitos para a sociedade brasileira, em observância ao princípio da participação social e democrática, a Advocacia-Geral da União convocou **audiência pública** visando a colher subsídios e contribuições da sociedade civil, comunidade acadêmica, agências de checagem e de instituições públicas e privadas, conforme edital em anexo (Doc. 04). Notadamente, com foco nos seguintes temas:

- a) Política de conduta de ódio das plataformas digitais;
- b) Medidas para mitigação da circulação de conteúdos criminosos e ilícitos nas plataformas digitais;
- c) Impacto social e na sustentabilidade do jornalismo profissional com eventual fim no Brasil de programas de checagem de fatos e substituição por outras

metodologias, a exemplo de "notas de comunidade";

d) aprimoramento do canal de denúncias sobre enfrentamento da desinformação e violações de direitos fundamentais;

e) necessidade de relatórios de transparência sobre a promoção e proteção dos direitos fundamentais e o enfrentamento à desinformação; e

f) impacto das alterações na moderação de conteúdo sobre grupos historicamente marginalizados, como mulheres, pessoas LGBTQIA+, imigrantes e pessoas com deficiência, à luz do regime jurídico-constitucional brasileiro de igualdade, de não discriminação, da liberdade de expressão e dignidade da pessoa humana, em consonância com os princípios e normas do Direito Internacional dos Direitos Humanos, notadamente os compromissos assumidos pelo Brasil no âmbito do Sistema Interamericano e do Sistema das Nações Unidas, assegurando a proteção integral e o enfrentamento a práticas discriminatórias em ambientes digitais.

27. Em 22.01.2025, **trinta e quatro representantes especialistas**^[21] de diversos setores da sociedade civil, meio acadêmico e agências de checagem participaram dos debates técnicos, cuja transcrição integral segue anexa (Doc. 05). A gravação do evento encontra-se disponível (<https://www.youtube.com/live/uLYX5Nl2kjY>), assim como o conjunto de subsídios e sugestões posteriormente encaminhados na consulta via Plataforma "Participa + Brasil"^[22], majoritariamente contrários às modificações impostas^[23].

28. Nessa oportunidade, foram expostas **diversas preocupações e apresentadas valiosas contribuições em relação aos impactos imediatos das referidas alterações sobre os direitos fundamentais em tela**, bem como quanto a outros pontos do mérito e ao quadro de insuficiência regulatória em geral. Além de funcionarem como subsídios adicionais para a deliberação no âmbito do presente processo, tais documentos serão igualmente compartilhados com as demais autoridades competentes, para os devidos fins. Por pertinência, muitos destes elementos serão retomados ao longo desta petição.

29. Nesse contexto, **revela-se premente a conclusão do julgamento, por essa Suprema Corte, do mérito do presente recurso extraordinário**, a fim de que - ao definir **balizas seguras** para a responsabilização dos provedores de aplicações por danos decorrentes de atos ilícitos praticados por terceiros - se promova um **ambiente digital seguro e caracterizado pelo respeito aos direitos fundamentais e aos valores democráticos**.

3. DA DISSONÂNCIA ENTRE AS ALTERAÇÕES PROMOVIDAS PELA RECORRENTE EM SEUS "TERMOS DE USO" E A SOLUÇÃO ATÉ ENTÃO DELINEADA POR ESSE SUPREMO TRIBUNAL FEDERAL PARA A PRESENTE CONTROVÉRSIA CONSTITUCIONAL. DAS RAZÕES PARA O IMEDIATO JULGAMENTO DO FEITO

30. Verifica-se que as alterações promovidas pela recorrente em seus "Termos de Uso", notadamente quanto à maior permissividade para o discurso de ódio contra grupos vulneráveis (a exemplo das hipóteses de racismo, homofobia, transfobia, intolerância religiosa e violência contra mulheres), como visto, violam o conjunto de fundamentos e soluções até aqui delineadas por essa Suprema Corte para o deslinde da controvérsia constitucional em relação a práticas criminosas.

31. Isto vale tanto pela lógica do regime de responsabilidade civil que independe de notificação, nos casos de danos graves por conteúdos ilícitos de terceiros, como pela inobservância dos denominados deveres anexos de segurança, prevenção, cuidado e transparência. Tudo isto na forma reconhecida pelos votos proferidos, os quais são corolários de normas há muito vigentes no ordenamento jurídico nacional, seja na esfera constitucional, civilista ou consumerista, como limites conformadores juridicamente legítimos às liberdades de iniciativa e de expressão.

32. Pois bem, no rol de condutas graves elencadas nos votos proferidos - que, para além de um grau de responsabilização diferenciado, **demandam ainda das plataformas uma atuação proativa** (seja de prevenção e monitoramento, seja de cuidado ante riscos sistêmicos) -, não restam dúvidas que figuram as referidas práticas criminosas contra grupos vulneráveis que passaram a ser aceitas no âmbito dos serviços prestados pela recorrente.

33. Nessa perspectiva, por essencial, reitere-se o que consta expressamente das conclusões propostas pelo Min. Dias Toffoli, relator:

(...) 3. O provedor de aplicações de internet **responde civilmente de forma objetiva e independentemente de notificação, pelos danos decorrentes de conteúdos gerados por terceiros**, nas seguintes hipóteses: (...) 3.4. **quando configurarem práticas previstas no seguinte rol taxativo:** (a) crimes contra o Estado Democrático de Direito (CP, art. 296, parágrafo único; art. 359-L, art. 359-

M, art. 359-N, art. 359-P, art. 359-R); (b) atos de terrorismo ou preparatórios de terrorismo, tipificados pela Lei nº 13.260, de 2016; (c) crime de induzimento, instigação ou auxílio a suicídio ou à automutilação (CP, art. 122); **(d) crime de racismo (Lei nº 7.716, de 1989, arts. 20, 20-A, 20-B e 20-C); (e) qualquer espécie de violência contra a criança, o adolescente e as pessoas vulneráveis (...) (f) qualquer espécie de violência contra a mulher, incluindo os crimes da Lei nº 14.192, de 2021; j) divulgação de fatos notoriamente inverídicos ou gravemente descontextualizados que levem à incitação à violência física, à ameaça contra a vida ou a atos de violência contra grupos ou membros de grupos socialmente vulneráveis; (...)**

6. Os provedores de aplicações de internet **devem atuar de forma responsável, transparente e cautelosa, a fim de assegurar um ambiente digital seguro, previsível e confiável, baseado nos princípios gerais da boa-fé, da função social e da prevenção e mitigação dos danos; (...)**

7. Os provedores de aplicações de internet possuem os seguintes deveres anexos: (a) **atualizar e manter atualizados os “termos e condições de uso” (ou documento equivalente) e os regulamentos adicionais, dando-lhes ampla publicidade; (...)** (c) **elaborar os respectivos códigos de conduta; (d) estabelecer regras claras e procedimentos padronizados para a moderação de conteúdos, aos quais se dará ampla publicidade; (e) proceder à constante atualização dos critérios e métodos empregados para a moderação de conteúdos, dando ampla publicidade aos usuários;**

8. Também é **dever anexo** dos provedores de aplicações de internet o de **atuar diligentemente para prevenir e mitigar práticas ilícitas no âmbito de seus serviços**, adotando as providências necessárias para combater a disseminação de conteúdos ilegais que configurem as condutas indicadas.

34. Tal compreensão foi também referendada, no que aqui importa, pela manifestação exarada pelo Min. Lux. Nessa linha, como visto, assim consta do sítio oficial desse Supremo Tribunal Federal, a saber: ***"Se o conteúdo gerado por terceiros veicular discurso de ódio, racismo, pedofilia, incitação à violência e apologia à abolição violenta do Estado Democrático de Direito e ao golpe de Estado, Fux defende que as plataformas façam monitoramento ativo e retirem o conteúdo do ar imediatamente, sem necessidade de notificação"***.

35. Inclusive, mesmo sob a ótica da posição do Min. Roberto Barroso - com nuance mais branda quanto à responsabilidade por conteúdos individualmente considerados e à necessidade de notificação, mas reconhecendo o denominado **dever de cuidado, sobretudo diante de riscos sistêmicos e práticas gravemente criminosas** -, concluiu-se no sentido de que as *"empresas devem trabalhar para reduzir os riscos sistêmicos criados ou potencializados por suas plataformas. As medidas, a seu ver, **devem minimizar esses riscos e seus impactos negativos sobre direitos individuais e coletivos, segurança e estabilidade democrática**"*. Em

suma, a seu ver, *"as plataformas devem atuar proativamente para que seu ambiente esteja livre de conteúdos gravemente nocivos"*.

36. Pois bem, sob qualquer ângulo, as alterações sob escrutínio - operadas pela empresa *Meta* quanto à permissividade na moderação de conteúdo - devem ser tidas como afrontosas aos direitos e obrigações extraídos do ordenamento pátrio.

37. A propósito do debate, como se sabe, há muito essa Suprema Corte acolheu a doutrina da **eficácia dos direitos fundamentais nas relações privadas**. É dizer, *"as violações a direitos fundamentais não ocorrem somente no âmbito das relações entre o cidadão e o Estado, mas igualmente nas relações travadas entre pessoas físicas e jurídicas de direito privado. Assim, os direitos fundamentais assegurados pela Constituição vinculam diretamente não apenas os poderes públicos, estando direcionados também à proteção dos particulares em face dos poderes privados"*. (RE n. 201819, Rel. Min. Ellen Gracie, Red. p/ Acórdão: Gilmar Mendes, 2ª Turma, DJ 27.10.2006)

38. Tal qual salientado pelo Min. Edson Fachin, em outro debate, a presente tutela é mais uma *"controvérsia que deve ser compreendida e solucionada à luz dos direitos fundamentais, de sua eficácia horizontal e dos direitos da personalidade"* (ADPF 462, Pleno, DJ 22.08.2024). É dizer, na operação de seus serviços, os grandes provedores de aplicação **não podem simplesmente reduzir o nível de proteção do núcleo essencial de direitos já estabelecidos** (como, no caso, à não discriminação e à dignidade humana), a seu bel prazer, **tornando-a insuficiente**, o que poderia ser também compreendido pela mesma sistemática da vedação do retrocesso em matéria de direitos fundamentais de caráter prestacional (efeito *cliquet*).

39. Nas lições do Min. Celso de Mello, tal princípio *"impede (...) que sejam desconstituídas as conquistas já alcançadas pelo cidadão ou pela formação social em que ele vive. A cláusula que veda o retrocesso (...) traduz, no processo de efetivação desses direitos fundamentais individuais ou coletivos, obstáculo a que os níveis de concretização de tais prerrogativas, uma vez atingidos, venham a ser ulteriormente reduzidos"*. Não se trata aqui,

pois, de mera faculdade atribuída aos provedores de aplicação, mas sim de verdadeiro dever-poder.

40. A propósito, como destaca Balaguer Callejón, *"devido ao modo como estão configurados os processos comunicativos da era digital, estas empresas passaram a ocupar uma grande parte do espaço público, prestando serviços em um regime de monopólio ou oligopólio que obriga a reformular as categorias entre o público e o privado em relação ao exercício da liberdade de expressão"*^[24]. Diante da *"ampla capacidade de decisão sobre a liberdade de expressão (...) a tensão sobre as liberdades dos processos comunicativas não pode ser valorada desde a perspectiva exclusiva do direito privado"*.

41. Tal aspecto também foi ressaltado na contribuição de Carlos Affonso Souza, representando o Instituto de Tecnologia e Sociedade - ITS (CP 882299), entre outras reflexões. Nessa medida, afirma que, *"para além de uma questão teórica, reconhecer a característica de transversalidade entre o privado e o público das regras sobre moderação de conteúdo abre espaço para o debate sobre os contornos da autonomia privada, e revela os temas sobre os quais governos e demais autoridades precisam se debruçar para a proteção de direitos fundamentais e a repressão de práticas ilícitas"*. Não restam dúvidas, portanto, que atos regulamentares como os denominados Termos de Serviço, ora em xeque, necessitam ser alvo de especial escrutínio e conformidade, à luz do bloco de constitucionalidade aplicável.

42. Noutro giro, especificamente quanto à admissão de conteúdos e insultos homofóbicos e transfóbicos, em determinados contextos, além de permitir *"associar doenças mentais a gênero ou orientação sexual"*, as mudanças promovidas pela empresa recorrente **colidem, ainda, de forma direta com as conclusões do julgamento da ADO n. 26** (Rel. Min. Celso de Mello, Pleno, DJ de 06.10.2020), inclusive quanto à **densificação do discurso de ódio**.

43. Em particular, na medida em que o julgado equipara tais atos à prática de racismo, mesmo diante do exercício da liberdade religiosa, notadamente quando *"incitem a discriminação, a hostilidade ou a violência (física ou moral) contra pessoas em razão de sua orientação sexual ou de sua identidade de gênero"*. In verbis:

COMPATIBILIDADE CONSTITUCIONAL ENTRE A REPRESSÃO PENAL À HOMOTRANSFOBIA E A INTANGIBILIDADE DO PLENO EXERCÍCIO DA LIBERDADE RELIGIOSA – A repressão penal à prática da homotransfobia não alcança nem restringe ou limita o exercício da liberdade religiosa, qualquer que seja a denominação confessional professada, a cujos fiéis e ministros (sacerdotes, pastores, rabinos, mulás ou clérigos muçulmanos e líderes ou celebrantes das religiões afro-brasileiras, entre outros) é assegurado o direito de pregar e de divulgar, livremente, pela palavra, pela imagem ou por qualquer outro meio, o seu pensamento e de externar suas convicções de acordo com o que se contiver em seus livros e códigos sagrados, bem assim o de ensinar segundo sua orientação doutrinária e/ou teológica, podendo buscar e conquistar prosélitos e praticar os atos de culto e respectiva liturgia, independentemente do espaço, público ou privado, de sua atuação individual ou coletiva, **desde que tais manifestações não configurem discurso de ódio, assim entendidas aquelas exteriorizações que incitem a discriminação, a hostilidade ou a violência contra pessoas em razão de sua orientação sexual ou de sua identidade de gênero** (...)

O discurso de ódio, **assim entendidas aquelas exteriorizações e manifestações que incitem a discriminação, que estimulem a hostilidade ou que provoquem a violência (física ou moral) contra pessoas em razão de sua orientação sexual ou de sua identidade de gênero, não encontra amparo na liberdade constitucional de expressão** nem na Convenção Americana de Direitos Humanos (Artigo 13, § 5º), que expressamente o repele.

III - O conceito de racismo, compreendido em sua dimensão social, projeta-se para além de aspectos estritamente biológicos ou fenotípicos, pois resulta, enquanto manifestação de poder, de uma construção de índole histórico-cultural motivada pelo objetivo de justificar a desigualdade e destinada ao controle ideológico, à dominação política, à subjugação social e à negação da alteridade, da dignidade e da humanidade daqueles que, por integrarem grupo vulnerável (LGBTI+) e por não pertencerem ao estamento que detém posição de hegemonia em uma dada estrutura social, são considerados estranhos e diferentes, degradados à condição de marginais do ordenamento jurídico, expostos, em consequência de odiosa inferiorização e de perversa estigmatização, a uma injusta e lesiva situação de exclusão do sistema geral de proteção do direito (ADO 26, Rel. Min. Celso de Mello, Pleno, DJ de 06.10.2020, grifou-se)

44. Em igual rumo, a Organização das Nações Unidas, de acordo com a definição de sua Estratégia e Plano de Ação contra o Discurso de Ódio, entende que tal prática consiste em *"qualquer tipo de comunicação oral, escrita ou comportamento, que ataca ou usa linguagem pejorativa ou discriminatória com referência a uma pessoa ou grupo com base em quem eles são, ou seja, com base em sua religião, etnia, nacionalidade, raça, cor, descendência, gênero ou outro fator identitário"*^[25].

45. Conforme destacado pelo *"Manifesto de Movimentos Sociais à Meta #AcordaMeta"*, liderado pelo Conecta LGBTI+ (CP 875726), "seria inadmissível permitir que discursos de ódio fossem veiculados por essas mídias, uma vez que isto implicaria que crimes

passassem a ser permitidos contra a população LGBTI+". Em complemento, reforçam que tais *"medidas evidenciam a postura anticientífica da Meta. É um retrocesso admitir a associação de identidades de gênero e orientações sexuais dissidentes com questões de saúde mental. A luta por despatologização de pessoas LGBTI+ ao longo do último século culminou, na década de 1990 e no ano de 2019, respectivamente, na retirada da homossexualidade e da transexualidade da Classificação Internacional de Doenças pela Organização Mundial da Saúde"*.

46. Aliás, ao expressamente admitir **discursos que reforçam a estigmatização de pessoas com deficiências mentais**, associando indevidamente tal condição a debates sobre questões de gênero, as alterações dos "Termos de Uso" **vulneram, ainda, os ditames da Convenção Internacional sobre os Direitos das Pessoas com Deficiência** (Convenção de Nova York). Como se sabe, esse tratado foi incorporado ao ordenamento jurídico pátrio com *estatura de norma constitucional* (Decreto n. 6.949/2009), nos termos do § 3º do art. 5º da Lei Maior. Entre seus princípios gerais, constam *"a não-discriminação; a plena e efetiva participação e inclusão na sociedade e o respeito pela diferença e pela aceitação das pessoas com deficiência como parte da diversidade humana e da humanidade"*.

47. No mesmo rumo, cumpre recordar mandamentos da **Convenção Interamericana contra o Racismo, a Discriminação Racial e Formas Correlatas de Intolerância** (Decreto n. 10.932/2022), ora sob ameaça. Tal compromisso obriga os Estados-Parte a *"prevenir, eliminar, proibir e punir, de acordo com suas normas constitucionais e com as disposições desta Convenção, todos os atos e manifestações de racismo, discriminação racial e formas correlatas de intolerância"*. De igual modo, conceitua normativamente a **intolerância** como *"um ato ou conjunto de atos ou manifestações que denotam desrespeito, rejeição ou desprezo à dignidade, características, convicções ou opiniões de pessoas por serem diferentes ou contrárias. Pode manifestar-se como a marginalização e a exclusão de grupos em condições de vulnerabilidade da participação em qualquer esfera da vida pública ou privada ou como violência contra esses grupos"*.

48. Neste particular, o representante do Centro de Estudos de Segurança e Cidadania - CESeC, Pablo Nunes (Doc. 05), observou que *"as transformações digitais vem afetando de*

*forma desproporcional grupos historicamente marginalizados com a população negra A arquitetura algorítmica dessas plataformas amplifica conteúdos sensacionalistas, o que **reforça também estereótipos raciais** e contribui para a criminalização da população negra. A moderação de conteúdos não é apenas uma questão técnica ou regulatória, mas um imperativo ético e político. **Garantir que as plataformas adotem práticas de moderação responsáveis (...)** é essencial para que a gente possa enfrentar as desigualdades e construirmos um ambiente digital verdadeiramente seguro."*

49. De igual modo, quanto aos impactos para a **violência de gênero online**, a pesquisadora Beatriz Kira (CP 879258) trouxe relevantes aportes sobre o *papel catalisador das plataformas de Internet*. Nessa linha, enfatizou que as mudanças sob exame -- a par da própria recalibragem de moderação automatizada --, são *profundamente preocupantes*:

Inseridas em um contexto social de discriminação estrutural e violência de gênero, **a internet e as plataformas digitais não apenas refletem, mas intensificam essas desigualdades e violências**. Os algoritmos das plataformas priorizam o engajamento, contribuindo para a disseminação de conteúdos sexistas, misóginos, e violentos que, de outra forma, não alcançariam a mesma escala. **As consequências são graves e palpáveis: danos e sofrimento psicológicos, incitação à violência física, restrição aos direitos e a participação de mulheres na vida política, social e cultural do país.**

Os relatórios do observatório MonitorA expõem de maneira alarmante a prevalência de conteúdos de violência de gênero nas redes sociais, com destaque para a violência política de gênero. Por exemplo, os estudos demonstram que, embora as mulheres tenham representado apenas 15% das candidatas no segundo turno das eleições de 2024, elas receberam 68,2% dos comentários ofensivos feitos em transmissões de debates no YouTube

Nesse contexto, mudanças recentes nas políticas de moderação de conteúdo das plataformas, com o objetivo de relaxar as restrições sobre conteúdo inflamatório e nocivo, são profundamente preocupantes. Veículos de imprensa indicam, por exemplo, que as mudanças anunciadas em janeiro de 2025 pela Meta permitiriam comparar mulheres a “objetos domésticos” e “propriedade”. (grifou-se)

50. Em complemento, à luz das obrigações previstas no ordenamento jurídico do Reino Unido, onde leciona (*Online Safety Act 2023*) -- em processo de implementação --, referida especialista afirma que *"as mudanças anunciadas pela Meta dificilmente poderiam ser implementadas unilateralmente sem atrair escrutínio do regulador"*. Isto porque *"se um conteúdo é ilegal no Reino Unido, as plataformas não podem simplesmente flexibilizar ou*

remover as disposições correspondentes de seus termos de uso: a ilegalidade do conteúdo exige sua proibição nos termos de uso".

51. No contexto eleitoral, por sua vez, a Resolução TSE n. 23.724/2024 sobre a propaganda na Internet, de forma pioneira e com amplo suporte legal (art. 23 , X, do Código Eleitoral e arts. 57-J e 105 da Lei Federal nº 9.504/1997), igualmente tratou desse tema. É dizer, ao estabelecer em seu art. 9º que os provedores de aplicação serão solidariamente responsáveis, civil e administrativamente, quando não promoverem a indisponibilização imediata de conteúdos e contas *"nos seguintes casos de risco: (...) IV – de comportamento ou discurso de ódio, inclusive promoção de racismo, homofobia, ideologias nazistas, fascistas ou odiosas contra uma pessoa ou grupo por preconceito de origem, raça, sexo, cor, idade, religião e quaisquer outras formas de discriminação"*.

52. Além disso, conforme bem recordado pelo Min. Flávio Dino, é pertinente reforçar que **a soberania é fundamento essencial do Estado Democrático Brasileiro e princípio norteador da ordem econômica** (artigos 1º, inciso I, e 170, inciso I, da Lei Maior). Desse modo, à luz das transformações tecnológicas e dos papéis desempenhados por atores privados contemporâneos, cuja irradiação transnacional é manifesta, a sujeição ao escrutínio estatal de seus atos revela-se ainda mais premente para a adequada proteção de direitos fundamentais:

(.) É indubitável que é parte das obrigações essenciais a um Estado Soberano garantir que a legislação seja cumprida também no domínio dos atores não estatais. Destarte, os Estados incorrem em responsabilidade não apenas por abusos infligidos por eles próprios, mas também por aqueles que não conseguem prevenir ou sancionar causados por terceiros.

Ou seja, enquanto tradicionalmente havia um modelo dualista em que a efetivação de direitos passava quase que exclusivamente por uma relação entre o Estado e indivíduos, **no mundo de hoje - mediado por tecnologias de informação e comunicação - a função de concretizar direitos transita decisivamente pelo controle sobre esses novos intermediários privados.** Desta maneira, estes são destinatários inafastáveis da atenção da dimensão jurisdicional do Estado Soberano.

Não há democracia sem soberania, e a ausência de soberania significa o fim da própria democracia, destruindo a cidadania e os direitos humanos, entre os quais a garantia da liberdade (PET-Ref n. 12.404/DF, DJ 04.09.2024, grifou-se).

53. Na mesma oportunidade, a Min. Cármen Lúcia ponderou se *"haveria soberania de um povo quando, no espaço nacional, não houvesse como garantir o direito brasileiro, incluído*

*aquele afirmado na Constituição do Brasil?. Em arremate, nos moldes de pronunciamento recente do Min. Alexandre de Moraes, o Brasil "é uma terra que tem lei. **As redes sociais não são terras sem lei**"^[26].*

54. Parece evidente, mas não custa repisar, como salientou a pesquisadora Laura Schertel durante sua exposição em audiência pública (Doc. 05): *"as políticas das empresas de tecnologia só são válidas se estiverem de acordo com essas normas dos países em que elas operam". Isto é, "a internet não pode ser um espaço de vale-tudo, de violência, de desinformação, de crimes. **Tudo isso não é a garantia da liberdade de expressão, mas a sua mais clara e evidente violação**".*

55. Quanto às preocupações acima possuem magnitude global, as manifestações anteriores da União, bem como todos os debates travados, sustentações e votos atestam a relevância e a urgência de buscas por respostas regulatórias em diferentes contextos. A título de complemento, cumpre registrar que, no último dia 17.12.2024, na linha de outros compromissos, o Brasil aderiu à **Recomendação da Organização para a Cooperação e Desenvolvimento Econômico - OCDE sobre Integridade da Informação**^[27]. Tal ato *"visa estabelecer um quadro amplo e aplicável de políticas públicas para que os países aderentes, membros ou não membros, possam enfrentar as ameaças impostas pela manipulação da informação e implementar medidas que promovam a integridade da informação, em conformidade com os direitos humanos universais de liberdade de opinião e expressão"*.

56. Interessante notar, finalmente, que o conjunto de violações derivadas das alterações dos "Termos de Uso", na moderação de conteúdo sobre desinformação e discurso de ódio, **contradiz todas as promessas de proatividade e o esforço colaborativo até então narrado pela empresa recorrente no âmbito dessa Suprema Corte.**

57. Na audiência pública de 28.03.2023, por exemplo, seu representante destacou os investimentos que **vinham sendo feitos na direção contrária**, isto é, a fim de se aprimorar o combate a conteúdos nocivos, incitação à violência e discurso de ódio no país - **restando claro que existem ferramentas tecnológicas para tanto, inclusive mediante o emprego de**

inteligência artificial - com relato de expressivos números quanto a conteúdos nocivos.

Confira-se:

"(...) a Meta investe bilhões de dólares e trabalha no desenvolvimento de times de tecnologia, inclusive, inteligência artificial, **para fazer valer os termos de uso e as políticas**. A empresa tem um time com **milhares de colaboradores dedicados à segurança e à integridade** dos apps, incluindo pessoas no Brasil.

Falando sobre o primeiro turno das eleições, a Meta removeu 135 mil anúncios de natureza eleitoral e mais de **três milhões de conteúdos no Facebook e no Instagram por violação das políticas que vedam conteúdo violento, incitação à violência e discursos de ódio**. Esses três milhões são para o Brasil, no período entre agosto de 2022 e janeiro de 2023.

Evidentemente, reconhecemos que mais pode ser feito pelas plataformas. Diante da escala dos nossos serviços, proibir determinados conteúdos não vai significar incidência zero desses conteúdos, mas **a Meta está comprometida a seguir aprimorando esses sistemas**. Hoje, obtemos resultados muito superiores aos que obtínhamos na época dos fatos narrados nos recursos em análise.

Concluindo minha exposição, Excelências, **reitero a proatividade da Meta no combate a conteúdos nocivos e à desinformação**" (grifou-se)

58. Mesmo diante de problemas como a opacidade algorítmica e a pouca transparência tanto no procedimento de moderação como na curadoria de conteúdos -- como enfatizado nas exposições de Marie Santini do NetLab/UFRJ, Andre Boselli do Artigo 19 e Iná Jost, do Internet Lab (Doc. 05, CP 882280 e CP 881120) -- o modelo existente deveria ser aprimorado, na linha afirmada pela *Meta*. É dizer, jamais retrocedido para, na prática, passar a permitir a violações de direitos fundamentais.

59. Ademais, não se pode olvidar que, conforme reconhecido pela *Meta*, as alterações da política de moderação de conteúdo em tela, notadamente quanto à "Política de Conduta de Ódio", **já estão em vigor, contrariando frontalmente os deveres de prevenção, de segurança e de cuidado inerentes** aos riscos e às atividades prestadas pelos grandes provedores de aplicação no país - na forma delineada pelo voto do Min. Relator e, em diversas medidas, acompanhada pelos votos que lhes sucederam.

60. Ao assim fazê-lo, deliberadamente permitindo, na prática, a produção e disseminação de discursos relacionados a **graves ilícitos perante o ordenamento pátrio** - como os de cunho racista, homofóbico, transfóbico, xenófobo e misógino -, há potencial

impacto negativo para a segurança e a integridade do ambiente digital no país. Portanto, os riscos de dano se renovam seguidamente.

61. Em reforço ao que já mencionado, a contribuição apresentada pelo *Sleeping Giants* Brasil (CP 881179) demonstra que a gravidade dos impactos das alterações promovidas pela *Meta* para a população LGBTQIA+ é manifesta, na medida que *"expõem ainda mais esses grupos a ataques discriminatórios e ao aumento da violência"*:

(...) as novas políticas adotadas pela Meta geram **impactos gravíssimos à população LGBTQIA+, especialmente ao permitirem discursos que associam orientação sexual e identidade de gênero a transtornos mentais. Ao flexibilizar os critérios de combate a condutas de ódio, essas medidas expõem ainda mais esses grupos a ataques discriminatórios, perpetuando estereótipos desumanizantes** e comprometendo a dignidade humana, princípio fundamental protegido pela Constituição Federal.

O impacto dessas políticas é evidenciado por dados alarmantes. A experiência recente de mudanças nas políticas da plataforma X (antigo Twitter) resultou em um aumento de 58% no discurso de ódio homofóbico e de 62% no discurso transfóbico, conforme levantamento (...)"

A moderação insuficiente, decorrente dessas mudanças, não apenas facilita a disseminação de discursos de ódio como também pode levar ao aumento direto da violência contra pessoas LGBTQIA+. Essa postura contraria abertamente o ordenamento jurídico brasileiro (grifou-se).

62. Na mesma linha, o representante da Coalização Conecta LGBTI+, Vitor de Wolff (Doc. 05), salientou em audiência pública que *"nós ainda somos o país que mais mata e mais persegue a comunidade LGBTI, especialmente travestis e transexuais no mundo. Nós ainda somos o país que mais assassina e comete violações dos direitos. E as redes não estão fora disso. As redes, na verdade, são incitação a isso"* (grifou-se).

63. O pesquisador Tarcizio Silva, representando a ABONG (Doc. 5), igualmente manifestou preocupação quanto aos impactos nas mudanças de permissividade do modelo de moderação de conteúdo:

Nesse sentido, as recentes declarações das grandes Big Tech de mídias sociais e inteligência artificial **são, de fato, assustadoras e já têm novos impactos explícitos e imediatos**, como alguns colegas mencionaram relacionados à **moderação de conteúdo contra discurso de ódio ligado a racismo, misoginia, transfobia e homofobia, que já era falha**. E agora os grandes nomes dessas empresas **estão explicitamente autorizando e incentivando discurso nocivo** em mais uma nova reviravolta, onde o Brasil parece coadjuvante de problemas políticos estadunidenses. (...) Então, considerando os compromissos

constitucionais do Estado brasileiro em realizar pesquisas sobre a natureza, as causas e as manifestações do racismo, da discriminação racial e formas correlatas de intolerância, nos parece, na verdade, nos fica bastante evidente que hoje a postura das big tech, dos gestores de grandes mídias sociais, **vai contra as defesas mínimas constitucionais** (...) dos grupos vulnerabilizados no país (grifou-se).

64. A manifestação do InternetLab (CP 881120), centro independente de pesquisa interdisciplinar, reforça como a **violência de gênero** no debate público das plataformas digitais, agora agravada pelo relaxamento dos termos de uso em tela, **termina por impactar a representação desses grupos na política:**

(...) Vão abaixo alguns dados que ilustram a prevalência dessa violência:- No segundo turno das eleições municipais do ano passado, as mulheres eram só 15% das candidatas a prefeituras, mas foram elas que receberam mais de 68% das ofensas durante os debates de televisão.

Em todas as edições do MonitorA, ficou cristalino o índice muito alto de ataques transfóbicos às candidatas trans. Em 2020, a candidata Erika Hilton foi a que recebeu o maior número de ofensas de todos os perfis que a gente monitorou. Durante 2024, 85% dos ataques que identificamos contra a candidata Duda Salabert eram abertamente transfóbicos

Essa é a qualidade do debate público na internet brasileira. E **a etapa qualitativa da nossa pesquisa vem também mostrando como essa violência afasta as pessoas afetadas da política, o que afeta a própria representação dessas pessoas em órgãos eletivos. Em outras palavras, os efeitos são concretos.**

O anúncio da Meta na semana passada tem um impacto direto nessa situação, e ajuda a pensar o momento presente. Em primeiro lugar, a empresa anunciou e já praticou uma mudança nas regras da comunidade, que **permite várias coisas que antes eram proibidas.** Por exemplo, quando a Meta anuncia que vai começar a permitir várias formas ofensivas de se referir a pessoas LGBTQIA+, isso já vai permitir esses ataques que a gente mostrou que são tão prevalentes, e que antes ainda eram relativamente moderados. (...) **Esse recuo é muito alarmante da perspectiva de garantia de direitos de mulheres, pessoas LGBTQIA+, pessoas negras nos últimos anos** (grifou-se).

65. No âmbito da atividade jornalística, de acordo com o destacado pela jornalista Bia Barbosa e dados do relatório do Repórteres sem Fronteiras - RSF "*O jornalismo frente às redes de ódio no Brasil*" (Doc. 05 e CP 880599), tal violência online contra as mulheres se evidencia de maneira flagrante:

"Os diferentes casos já mencionados mostram como as mulheres jornalistas transformaram-se em alvos principais da violência online contra a imprensa. A tendência, que vem sendo apontada em estudos no Brasil e no mundo, foi reforçada nesta pesquisa. (...) quando olhamos para os 10 jornalistas mais

atacados – aqueles que receberam um maior número de mensagens ofensivas –, 7 são mulheres. (...)

Pesquisa lançada em 2022 pela Gênero e Número e a RSF mostrou que 19% das agressões contra jornalistas mulheres e LGBTQIA+ são misóginas ou tem conotação sexual, e visam, além de intimidar, causar dano à reputação dessas profissionais. Muitas vezes a violência está baseada no uso e manipulação da imagem das jornalistas (...) **Quando se trata de jornalistas negras, o racismo, estruturante da sociedade brasileira, também se manifesta fortemente.**

66. A contribuição da agência de checagem "Aos Fatos" (CP 879320), por sua vez, ratifica que *"as novas diretrizes da empresa passaram a permitir posts que reforçam estereótipos e e preconceitos raciais étnicos e religiosos contra pessoas com deficiência - condutas que são crimes de acordo com a lei brasileira"*:

(...)A documentação interna da Meta confirma que a omissão reflete uma mudança concreta na política da empresa, **orientando os moderadores a não remover declarações de ódio a minorias.**

Experimento executado pelo Aos Fatos confirma a nova diretriz. No dia 15, a reportagem apresentou, a partir de cinco perfis diferentes, denúncias ao Facebook contra um post que dizia “odeio negros” — publicado por uma página neonazista para, supostamente, testar as novas regras. Na tarde desta segunda (20), a publicação seguia no ar. **As regras da empresa também passaram a autorizar expressões de superioridade e comparações entre grupos protegidos, facilitando o discurso de supremacistas brancos** (grifou-se).

67. Por fim, informe da Relatoria Especial para Liberdade de Expressão da **Comissão Interamericana de Direitos Humanos** traz importantes reflexões sobre os **efeitos deletérios do discurso de ódio no ambiente digital**^[28], tal qual ocorre na hipótese. Ao lado de sua gravidade intrínseca, destaca-se que, não raro, tais atos migram para condutas de violência em face de grupos vulneráveis também fora das redes:

(...) 177. Complementariamente, la **propagación y viralización de discursos calificados como discurso de odio en internet, y su influencia en espacios físicos**, han impulsado um aumento en las propuestas normativas para restringir estos contenidos en línea. Dichos discursos amenazan derechos fundamentales como la libertad de expresión, la dignidad humana, la igualdad, la seguridad personal y la vida y **pueden desencadenar en violencia sistémica y de alto impacto en individuos y comunidades.**

178. Este tipo de discursos tienen la **capacidad de fomentar ambientes digitales discriminatorios, intolerantes y violentos**, coartando la participación en el espacio público digital y afectando la capacidad de las víctimas de buscar, recibir y difundir información en línea. Esto contribuye a desequilibrios en la pluralidad informativa y diversidad comunicativa en internet.

Al respecto, la Relatoría ha indicado que **la desigualdad frente a determinados grupos resulta en la exclusión de estos del debate público y, posteriormente, en mayor vulnerabilidad frente a la intolerancia, los prejuicios y la marginación.** Para la Relatoría de la ONU sobre la promoción y protección del derecho a la libertad de opinión y de expresión, **las consecuencias en la difusión del discurso de odio en línea le compete también a las empresas** que “[t]ienen un enorme impacto en los derechos humanos, en particular, pero no solo, en lugares en que son la forma predominante de expresión pública y privada, donde una limitación de la libertad de expresión puede equivaler a un silenciamiento público o a no **abordar una incitación que puede facilitar la violencia y la discriminación fuera de línea** (grifou-se).

68. Pelas mesmas razões, no último dia 20.01.2025, a **Comissão Europeia** anunciou a revisão de seu Código de Conduta sobre Discursos Ilegais de Incitação ao Ódio (*Code of Conduct on Countering Illegal Hate Speech Online+*), no âmbito do Regulamento de Serviços Digitais. Trata-se de expansão de um instrumento complementar que teve sua primeira versão em 2016 e, uma vez mais, **contou com a adesão de diversas empresas -- entre as quais a própria recorrente**^[29].

69. O primeiro artigo do referido compromisso dispõe exatamente que *"os Signatários devem fazer constar em seus termos e condições a informação aos usuários que eles **proíbem o discurso de ódio ilegal em seus serviços (...)** assim definidos pelas legislações aplicáveis"*^[30]. Em suma, o que se pleiteia na presente medida é, no mínimo, a mitigação da propagação do discurso de ódio e a prevenção de práticas criminosas, como claro risco sistêmico, com a **garantia da fiel observância das normas protetivas vigentes no país.**

70. Como dito, milhões de brasileiros fazem uso constante dos serviços da recorrente no Brasil e estão potencialmente expostos aos relatados riscos, afigurando-se, portanto, **premente a conclusão do julgamento deste recurso extraordinário**, a fim de que sejam estabelecidas balizas mais seguras para o funcionamento do ambiente digital em nosso país, **fazendo cessar** atos como o perpetrado pela recorrente, assegurando-se, desse modo, a soberania nacional e a plena vigência de nossa ordem constitucional.

4. DO PEDIDO

71. Ante todo o exposto, a União vem requerer *a)* a juntada do material produzido em audiência pública, conforme anexos, a fim de contribuir para o deslinde da controvérsia em exame e *b)* prioridade no julgamento do presente recurso paradigma.

Brasília, *na data da assinatura eletrônica.*

JORGE RODRIGO ARAÚJO MESSIAS

Advogado-Geral da União

ISADORA MARIA BELEM ROCHA CARTAXO DE ARRUDA

Secretária-Geral de Contencioso

RAPHAEL RAMOS MONTEIRO DE SOUZA

Advogado da União

GUSTAVO HENRIQUE CATISANE DINIZ

Advogado da União

RAFAEL FERNANDES DUBRA

Advogado da União

Anexos:

Doc. 01 - Notificação Extrajudicial;

Doc. 02 - Resposta empresa *Meta*;

Doc. 03 - Alteração Política de Conteúdo *Meta* - Conduta de ódio;

Doc. 04 - Edital Convocação Audiência Pública;

Doc. 05 - Transcrição Integral da Audiência Pública;

Doc. 06 - Contribuições Consulta - Anexos Plataforma Participa + Brasil.

Notas

1. [^] *Artigo 4º, inciso III, da Lei Complementar nº 73/1993*
2. [^] *Confira-se, a propósito, o Memorando de Entendimento TSE n. 29/2024 (SEI 2021.00.000012038-3)*
3. [^] *Art. 19. Com o intuito de assegurar a liberdade de expressão e impedir a censura, o provedor de aplicações de internet somente poderá ser responsabilizado civilmente por danos decorrentes de conteúdo gerado por terceiros se, após ordem judicial específica, não tomar as providências para, no âmbito e nos limites técnicos do seu serviço e dentro do prazo assinalado, tornar indisponível o conteúdo apontado como infringente, ressalvadas as disposições legais em contrário. (...) Art. 20. Sempre que tiver informações de contato do usuário diretamente responsável pelo conteúdo a que se refere o art. 19, caberá ao provedor de aplicações de internet comunicar-lhe os motivos e informações relativos à indisponibilização de conteúdo, com informações que permitam o contraditório e a ampla defesa em juízo, salvo expressa previsão legal ou expressa determinação judicial fundamentada em contrário. Parágrafo único. Quando solicitado pelo usuário que disponibilizou o conteúdo tornado indisponível, o provedor de aplicações de internet que exerce essa atividade de forma organizada, profissionalmente e com fins econômicos substituirá o conteúdo tornado indisponível pela*

motivação ou pela ordem judicial que deu fundamento à indisponibilização. Art. 21. O provedor de aplicações de internet que disponibilize conteúdo gerado por terceiros será responsabilizado subsidiariamente pela violação da intimidade decorrente da divulgação, sem autorização de seus participantes, de imagens, de vídeos ou de outros materiais contendo cenas de nudez ou de atos sexuais de caráter privado quando, após o recebimento de notificação pelo participante ou seu representante legal, deixar de promover, de forma diligente, no âmbito e nos limites técnicos do seu serviço, a indisponibilização desse conteúdo

4. [^] <https://noticias.stf.jus.br/postsnoticias/para-ministro-fux-e-inconstitucional-responsabilizar-plataformas-somente-em-casos-de-descumprimento-de-ordem-judicial/>
5. [^] <https://noticias.stf.jus.br/postsnoticias/marco-civil-da-internet-barroso-defende-que-plataformas-reduzam-riscos-de-postagens-criminosas/>
6. [^] <https://www.threads.net/@zuck/post/DEhgZJYJu5D> e <https://www.facebook.com/zuck/videos/1525382954801931/>
7. [^] <https://www.facebook.com/formedia/mjp/programs/third-party-fact-checking/partner-map>. No Brasil, a parceira foi estabelecida com a Agência Lupa, Estadão Verifica, Aos Fatos e AFP.
8. [^] <https://oglobo.globo.com/economia/tecnologia/noticia/2025/01/12/quais-mudancas-na-meta-javalem-para-o-brasil-entenda-em-5-pontos.ghtml#>
9. [^] <https://jornaldigital.recife.br/2024/04/10/brasil-dispara-como-um-dos-maiores-consumidores-de-midias-sociais-segundo-pesquisa/>
10. [^] <https://exame.com/mundo/regulamentar-conteudo-nocivo-na-internet-nao-e-censura-diz-onu/>
11. [^] <https://noticias.uol.com.br/ultimas-noticias/rfi/2025/01/08/nobel-da-paz-alerta-para-momento-extremamente-perigoso-apos-decisao-da-meta-de-acabar-com-chechagem-de-fatos.htm>
12. [^] <https://www.reuters.com/technology/we-do-not-censor-social-media-eu-says-response-meta-2025-01-08/>

13. [^] [^] <https://direitosnarede.org.br/2025/01/08/contra-o-retrocesso-na-moderacao-de-conteudo-dameta-e-os-ataques-a-regulacao-democratica-do-espaco-digital/>
14. [^] [^] <https://www.ft.com/content/917c9535-1cdb-4f6a-9a15-1a0c83663bfd>
15. [^] ZUBOFF, S. "Big other: surveillance capitalism and the prospects of an information civilization". *Journal of Information Technology*, v. 30, 2015, p. 75-89.
16. [^] "We do allow allegations of mental illness or abnormality when based on gender or sexual orientation, given political and religious discourse about transgenderism and homosexuality"
17. [^] "Certain objects (women as household objects or property or objects in general; Black people as farm equipment; transgender or non-binary people as "it")"
18. [^] <https://g1.globo.com/tecnologia/noticia/2025/01/09/meta-publica-em-portugues-nova-politica-que-associa-publico-lgbtqia-a-doencas-mentais.ghtml#mudancas>
19. [^] "Other times, they call for exclusion or use insulting language in the context of discussing political or religious topics, such as when discussing transgender rights, immigration, or homosexuality"
20. [^] No original: "Self-admission to intolerance on the basis of a protected characteristics, including but not limited to: homophobic, islamophobic, racist."
21. [^] [^] <https://www.gov.br/agu/pt-br/comunicacao/noticias/pesquisadores-e-sociedade-civil-queremregras-para-plataformas-digitais>
22. [^] [^] <https://www.gov.br/participamaisbrasil/consulta-publica-sobre-os-efeitos-das-novas-politicassimplementadas-no-brasil-por-plataformas-digitais-e-o-seu-impacto-sobre-o-fenomeno-dadesinformacao-e-a-protecao-dos-direitos-fundamentais>
23. [^] <https://www.gov.br/agu/pt-br/comunicacao/noticias/agu-recebe-78-contribuicoes-em-consultapublica-sobre-plataformas-digitais>
24. [^] CALLEJÓN, F. Balaguer. *La Constitución del Algoritmo*. 2. ed. Zaragoza : Manuel Giménez Abad. 2023, p. 121-122, t.l
25. [^] cf. https://www.un.org/en/genocideprevention/documents/advising-andmobilizing/Action_plan_on_hate_speech_EN.pdf

[https://brasil.un.org/sites/default/files/2023-](https://brasil.un.org/sites/default/files/2023-10/ONU_Integridade_Informacao_Plataformas_Digitais_Informe-Secretario-Geral_2023.pdf)

[10/ONU_Integridade_Informacao_Plataformas_Digitais_Informe-Secretario-Geral_2023.pdf](https://brasil.un.org/sites/default/files/2023-10/ONU_Integridade_Informacao_Plataformas_Digitais_Informe-Secretario-Geral_2023.pdf) (p. 5)

26. [^] <https://agenciabrasil.ebc.com.br/justica/noticia/2025-01/moraes-diz-que-plataformascontinuarao-no-brasil-se-respeitarem-lei>

27. [^] <https://www.gov.br/secom/pt-br/assuntos/noticias/2024/12/brasil-adere-a-recomendacao-daocde-sobre-integridade-da-informacao>

28. [^] https://www.oas.org/es/cidh/expresion/informes/Inclusion_digital_esp.pdf

29. [^] https://ec.europa.eu/commission/presscorner/detail/pt/ip_25_300

30. [^] <https://ec.europa.eu/newsroom/dae/redirection/document/111777>

Documento assinado eletronicamente por *.AGU.GOV.BR, de acordo com os normativos legais aplicáveis. A conferência da autenticidade do documento está disponível com o código 1841736690 no endereço eletrônico <http://sapiens.agu.gov.br>. Informações adicionais: Signatário (a): *.AGU.GOV.BR. Data e Hora: 06-02-2025 19:12. Número de Série: 24688056426646610828629120681. Emissor: Autoridade Certificadora do SERPRO Final SSL.

Documento assinado eletronicamente por *.AGU.GOV.BR, de acordo com os normativos legais aplicáveis. A conferência da autenticidade do documento está disponível com o código 1841736690 no endereço eletrônico <http://sapiens.agu.gov.br>. Informações adicionais: Signatário (a): *.AGU.GOV.BR. Data e Hora: 06-02-2025 10:02. Número de Série: 24688056426646610828629120681. Emissor: Autoridade Certificadora do SERPRO Final SSL.
