



## A AUTORREGULAÇÃO REGULADA E A MODERAÇÃO DE CONTEÚDO NO FACEBOOK.

### *REGULATED SELF-REGULATION AND CONTENT MODERATION ON FACEBOOK.*

Ana Paula Vieira Carvalho\*  
Ramon de Vasconcelos Negócio\*\*

**Resumo:** Nos últimos anos, a popularidade das redes sociais cresceu significativamente, de tal modo a se tornarem o principal palco da manifestação da liberdade de expressão na internet. Nesse contexto, tem gerado preocupação a moderação de conteúdo realizada por essas plataformas privadas com decisões tomadas, muitas vezes, visando aos interesses comerciais das redes sociais. Este estudo tem como objetivo analisar a autorregulação regulada como uma possibilidade estratégica para que os países lidem com a complexibilidade que envolve a dinâmica das redes sociais, destacando a experiência da Alemanha e da União Europeia, bem como o Projeto de Lei nº 2630/2020 brasileiro, que foi aprovado no Senado Federal, em 30 de junho de 2020. Utiliza-se o método dedutivo, com uma abordagem qualitativa, descritiva e teórica, realizada a partir de pesquisa bibliográfica em artigos jurídicos e legislações brasileira e estrangeira, para se concluir que a autorregulação regulada é uma alternativa oportuna para controlar o setor, pois busca compatibilizar a inovação tecnológica com os direitos fundamentais dos usuários.

**Palavras-chave:** Redes sociais; Liberdade de expressão; Moderação de conteúdo; Responsabilidade das plataformas, Autorregulação regulada.

**Abstract:** *In recent years, the popularity of social networks has grown significantly, so that they become the the main stage for the manifestation of freedom of speech on the internet. In this context, the moderation of content carried out by these private platforms has raised concern, with decisions taken, many times, aiming at the commercial interests of social networks. This study aims to analyze regulated self-regulation as a strategic possibility for countries to deal with the complexity that involves the dynamics of social networks, highlighting the experience of Germany and the European Union, as well as the Brazilian Bill of Law nº 2630/2020, which was approved by the Federal Senate on June 30, 2020. The deductive method is used, with a qualitative, descriptive, and theoretical approach, carried out through bibliographical research in legal articles and Brazilian and foreign legislation, to conclude that regulated self-*

\* Mestranda em Direito pela UNI7, com bolsa da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Especialista em Direito Público pela ESMEC. carvalhovieira.ap@gmail.com

\*\* Mestre em Direito do Estado pela Pontifícia Universidade Católica de São Paulo (2011) e doutorado (2019) na Goethe-Universität Frankfurt am Main. Professor da graduação e do programa de pós-graduação *stricto sensu* da UNI7. negocioramom@gmail.com



*regulation is a timely alternative to control the sector, as it seeks to reconcile technological innovation with the fundamental rights of users.*

**Keywords:** *Social networks; Freedom of speech; Content moderation; Platform responsibility; Regulated self-regulation.*

## **Introdução**

A cultura do compartilhamento de informações nas redes sociais tem crescido exponencialmente, revelando a força da liberdade de expressão na era digital. No entanto, juntamente com os benefícios dessa liberdade, surgem desafios significativos. O discurso de ódio, a difamação e a desinformação se tornaram problemas generalizados, impactando profundamente a sociedade. Embora essas questões existam independentemente das redes sociais, a popularidade e a velocidade com que as informações são propagadas nessas plataformas ampliam, de forma alarmante, o alcance e o impacto negativo do conteúdo compartilhado (Langvardt, 2021).

A necessidade de equilibrar a regulação das redes sociais e a liberdade de expressão tem gerado intenso debate diante das medidas adotadas por plataformas digitais. Redes sociais, como o Facebook, respaldadas por seus termos de uso, promovem a remoção e o bloqueio de publicações e, por vezes, até mesmo de usuários. Essas ações, que podem ser consideradas limitações à liberdade de expressão, têm sido alvo de discussões acaloradas, principalmente pela falta de transparência no processo de tomada de decisões.

Diante desse cenário, o presente trabalho visa investigar como a autorregulação regulada pode contribuir para controlar a atuação das redes sociais na moderação de conteúdo, buscando conciliar a inovação tecnológica, a liberdade de expressão e os direitos de personalidade dos usuários.

Objetiva-se, assim, analisar a autorregulação regulada como uma possibilidade estratégica para que os países lidem com a complexidade que envolve a dinâmica dessas plataformas. Para isso, utiliza-se o método dedutivo, por meio de uma abordagem de natureza qualitativa, descritiva e teórica, realizada mediante pesquisa bibliográfica em artigos jurídicos e em legislações brasileira e estrangeira.

O estudo se desenvolverá em três tópicos. Inicialmente, serão apresentadas as regras que regem a moderação de conteúdo do Facebook, ressaltando os aspectos da autorregulação. Após, serão explorados a necessidade de regulação estatal no contexto da



moderação de conteúdo e o conceito de autorregulação regulada. Por fim, serão analisadas normas que preveem a autorregulação regulada como forma de regular as redes sociais, a exemplo da *Netzwerkdurchsetzungsgesetz (NetzDG)* alemã e do *Digital Services Act (DSA)* da União Europeia, além do Projeto de Lei nº 2630/2020 brasileiro, que foi aprovado no Senado Federal, em 30 de junho de 2020.

## **1 A autorregulação das plataformas digitais e a moderação de conteúdo realizada pelo Facebook**

O surgimento da internet trouxe questionamentos sobre a regulação do ciberespaço. Entre o final dos anos 90 e início dos anos 2000, argumentava-se que a internet seria um espaço de igualdade universal, que não deveria se submeter a qualquer coerção hierárquica. Além disso, haveria questões morais, como a falta de legitimidade das estruturas governamentais convencionais no espaço virtual, e uma impossibilidade material de os Estados regularem o mundo digital, tendo em vista as características específicas da internet (Silva, 2022, p. 77). De acordo com Silva (2022, p. 77), esses debates “foram fundamentais para a construção argumentativa da autorregulação pelos setores privados na atualidade”.

Entende-se por autorregulação a capacidade dos atores econômicos de estabelecerem e manterem regulamentos e normas de conduta sem a necessidade da interferência estatal. No contexto da internet, a autorregulação também se verifica na própria arquitetura das plataformas digitais, que limitam e condicionam o comportamento dos seus usuários, por meio de escolhas feitas pelos desenvolvedores. Dessa forma, “[...] as redes sociais, por exemplo, criam estruturas normativas autônomas, dado que controlam a arquitetura interna de seu programa, de modo que os usuários são vinculados sem necessariamente haver participação do Estado” (Negócio, 2022, p. 2).

O Facebook possui seus Termos de Serviço e Padrões da Comunidade. Eles definem as regras sobre publicação de conteúdo e comportamento dos usuários. De acordo com a Meta, empresa responsável pelo Facebook, os Padrões da Comunidade consideram a expressão dos usuários, mas reconhecem que a internet cria oportunidades de abuso, razão pela qual proíbem a veiculação de alguns conteúdos com base em valores como autenticidade, privacidade, segurança e dignidade (Meta, 2022). Entre os conteúdos não permitidos, destacam-se aqueles que contenham violência e incitação, organizações e indivíduos perigosos, coordenação de danos e incitação ao crime, fraude e engano, suicídio e automutilação, exploração sexual,

abuso ou nudez infantil, *bullying* e assédio, discurso de ódio, violação de propriedade intelectual (Meta, 2022). Incorrendo em desvio dessas regras, o usuário pode ter suas publicações removidas ou sua conta suspensa. No caso de conteúdos considerados problemáticos, como publicações desinformativas ou sensacionalistas, a plataforma pode reduzir a frequência da postagem no *feed* de notícias (Moura; Rocha, 2022, p. 8).

Ressalta-se que, em regra, essas normas vinculam apenas as partes que as aderiram (Negócio, 2022, p. 5). Assim, ao aceitarem os termos de uso da Meta, os usuários se submetem à regulação privada e lhe dão legitimidade. Conforme aponta Marques Neto (2011, p. 89), “A regulação autônoma se viabiliza pela adesão dos agentes e pela preservação de sua funcionalidade (eficácia). A fonte de sua autoridade é a permanente adesão dos atores”.

Salienta-se que a autonomia da vontade dos usuários perante os termos de uso das plataformas suscita alguns questionamentos em razão da forma como se dá o consentimento. Em geral, os usuários não têm a oportunidade de influenciar o conteúdo desses termos, que podem refletir apenas os interesses das próprias plataformas. Estas, inclusive, possuem a capacidade de modificar unilateralmente o contrato firmado, que serve como base legal para justificar a intervenção na liberdade de expressão dos usuários (Koltay, 2022, p. 544). Além disso, em alguns casos, há apenas informações vagas e pouco claras sobre o escopo do que se está anuindo, sem especificar, por exemplo, o tratamento de dados pessoais (Hoffmann-Riem, 2019, p. 551). Embora seja possível argumentar que ninguém é obrigado a permanecer vinculado a uma rede social, é amplamente reconhecido que, nos tempos atuais, a ausência de conexão pode resultar, até mesmo, em exclusão social, de sorte que, muitas vezes, o usuário não possui a opção de não consentir com a política da plataforma.

No que se refere à moderação de conteúdo realizada pelas redes sociais, sabe-se que o Facebook, considerado a maior rede social do mundo, possui sistema de governança extremamente complexo, incorporando inteligência artificial alimentada com sua própria base de dados e participação em bancos de dados globais (Moura; Rocha, 2022, p. 4). Conquanto a utilização de técnicas automatizadas com o uso de inteligência artificial gere maior eficiência no que se refere à retirada de publicações danosas, Gorwa, Binns e Katzenbach (2020, p. 3) apontam que a moderação algorítmica tem o potencial de exacerbar, ao invés de amenizar, alguns problemas importantes relacionados à política de conteúdo das plataformas. Um desses problemas é a possível diminuição da transparência, tornando as tomadas de decisões das plataformas ainda mais difíceis de serem compreendidas e auditadas. Outrossim, a moderação



algorítmica pode obscurecer a política complexa que fundamenta as práticas de moderação de conteúdo.

Sobre esse último aspecto, os autores afirmam que o Facebook enaltece o fato de conseguir remover proativamente 99,6% da propaganda terrorista sem, no entanto, esclarecer questões como quais grupos podem ser considerados terroristas e que tipos de dados alimentam os classificadores da plataforma (Gorwa; Binns; Katzenbach, 2020, p. 12). Nesse cenário, causam preocupação a deficiência de informação entre os usuários sobre a operacionalização das restrições de conteúdo e a ausência de critérios claros para sua aplicação.

Destaca-se que inicialmente o Facebook realizava a moderação de conteúdo de forma isenta de responsabilidade, como um “bom samaritano”, amparado pela Seção 230 do *Communication Decency Act* de 1996 dos Estados Unidos. Entretanto, a expansão da plataforma para diferentes culturas e a pressão de países europeus, como a Alemanha, para adequação às diversas legislações nacionais de proteção à liberdade de expressão, exigiram que os padrões de moderação do Facebook fossem, cada vez mais, restritivos, também demandando uma análise regionalizada e casuística para sua aplicação (Moura; Rocha, 2022, p. 5).

Em resposta às pressões externas, especialmente após o episódio da Cambridge Analytica<sup>1</sup>, o Facebook anunciou, em 2018, a criação do Oversight Board, visando trazer maior transparência para a tomada de decisões sobre moderação de conteúdo (Moura; Rocha, 2022, p. 8). Criado em 2020, o Comitê de Supervisão – como é conhecido em português – consiste em um órgão de supervisão independente da plataforma e do Estado que fornece recomendações e orientações sobre políticas e práticas relacionadas à liberdade de expressão, revisando as decisões do Facebook sobre moderação de conteúdo, de modo a proporcionar uma instância de apelação para usuários (Koltay, 2022, p. 557). Ressalta-se ainda a possibilidade de o órgão interpretar “[...] os Padrões da Comunidade e as políticas de conteúdo, inclusive a fim de criar um sistema de precedentes a ser usado como base normativa para outras decisões e formulações de políticas do Comitê e do Facebook” (Moura; Rocha, 2022, p. 9).

Em um dos casos mais notórios, o Oversight Board validou a decisão do Facebook de suspender a conta do ex-Presidente dos Estados Unidos, Donald Trump, após a invasão ao

---

<sup>1</sup> Ocorrido em 2018, consistiu em um incidente em que a empresa Cambridge Analytica coletou dados de aproximadamente 87 milhões de usuários do Facebook via aplicativo de teste de personalidade, sem consentimento expresso. As informações foram usadas para criar perfis psicográficos e direcionar propaganda política. Trata-se de um marco na discussão sobre privacidade e proteção de dados (Kleinman, 2018).



Capitólio, em janeiro de 2021. O Comitê, no entanto, solicitou que o Facebook reavaliasse a decisão de suspender a conta de Trump por um período indefinido (Comitê, 2021).

Conquanto represente um avanço na busca por maior *accountability*, o Oversight Board é criticado, entre outras razões, porque o rol de casos e de legitimados para suscitar a análise, no Comitê, é bastante restrito, “[...] não permitindo, por exemplo, que o afetado por alguma postagem tenha oportunidade de acessar o board, limitando-se à denúncia singular (se usuário da rede social)” (Moura; Rocha, 2022, p. 9). Além disso, embora suas decisões sobre casos específicos sejam vinculativas, as recomendações de políticas são apenas consultivas (Facebook, 2019). Dessa forma, o Comitê de Supervisão pode emitir recomendações sobre mudanças na política de moderação de conteúdo do Facebook, mas a decisão final sobre a implementação cabe à própria plataforma, o que suscita questionamentos acerca do poder real do Oversight Board na moderação do conteúdo. Há também preocupações quanto à possibilidade de influência indireta da Meta na estrutura e nas decisões do Comitê. Isso porque seu financiamento decorre de um fundo fiduciário estabelecido pela empresa Meta (Securing..., 2022).

É importante enfatizar que o Comitê de Supervisão, apesar de merecer destaque pelo pioneirismo e pelo esforço na implementação de transparência e de *accountability*, não constitui uma solução definitiva para as questões próprias da moderação de conteúdo. As limitações em sua atuação e as críticas recebidas apontam para a necessidade de aprimoramento dos procedimentos adotados e para a instauração de novos mecanismos que possam atender, de forma eficiente, às necessidades crescentes no ambiente digital. Afinal, buscar um equilíbrio entre assegurar a liberdade de expressão dos usuários e combater práticas prejudiciais, como discurso de ódio, desinformação e abusos nas redes sociais, é um desafio de relevância global e demanda a contribuição de uma gama diversificada de agentes, incluindo legisladores, especialistas em tecnologia e em direitos digitais e o próprio público usuário dessas plataformas.

## **2 A necessidade de regulação estatal frente à regulação privada no contexto da moderação de conteúdo**

Uma das características da internet é “o desejo de muitos atores de agir de modo tão autônomo quanto possível. O que lhes interessa é a auto-organização e – se for preciso – a possibilidade mesmo da autorregulação” (Hoffmann-Riem, 2019, p. 531). No que se refere às plataformas digitais, elas têm se aproveitado da ausência de leis e das zonas de penumbra da





regulação dura, para atuarem regidas por uma regulação criada por elas mesmas (Frazão, 2018, p. 656). Um dos principais motivos para isso são seus próprios interesses comerciais. Mesmo quando defende a liberdade de expressão de seus usuários, o Facebook o faz com o objetivo subjacente de transparecer a ideia de espaço aberto e atrair novos inscritos. Ao mesmo tempo, remove publicações, a fim de permitir que os usuários se sintam seguros e não encontrem conteúdo ofensivo e perturbador que os afaste da plataforma (Koltay, 2022, p. 547).

Na conjuntura atual, a autorregulação predomina, nas plataformas digitais dirigidas por grandes empresas, sem a devida transparência, responsabilização ou filtros democráticos, beneficiando principalmente os interesses econômicos desses atores dominantes (Frazão, 2018, p. 657). É importante lembrar que o Facebook é pressionado por outros agentes econômicos, como os anunciantes da plataforma, para remover ou dar destaque a alguma publicação, de modo que gera preocupação ao Estado e à sociedade que interesses econômicos possam prevalecer em detrimento de direitos fundamentais, como a liberdade de expressão.

Marques Neto (2011, p. 89) destaca que a autorregulação é predominantemente uma forma de regulação econômica, com um impacto social mínimo ou meramente residual, apresentando limitações na prevenção de externalidades ou na implementação de políticas públicas. Um problema adicional da regulação privada é que ela é, algumas vezes, mais rigorosa e, outras vezes, mais branda do que a regulamentação estatal, o que pode tornar a atividade de moderação do conteúdo imprevisível (Koltay, 2022, p. 544).

Inicialmente, buscava-se oferecer proteções gerais de responsabilidade para as plataformas digitais, com o objetivo de fomentar a inovação. Em virtude do aumento significativo do uso das redes sociais e da dominação do mercado por poucos provedores poderosos, tornou-se necessário reavaliar essa postura (Eifert, 2021). A tendência de substituir a regulação jurídica estatal pela regulação tecnológica, ao invés de promover um ambiente mais equitativo, pode, na realidade, levar à dominação por parte de grandes corporações tecnológicas (Frazão, 2018, p. 657).

No contexto da moderação de conteúdo realizada pelas redes sociais, percebeu-se que estabelecer regras de “bom samaritano” e isentar os provedores de internet por conteúdo publicado por terceiros – como proposto pela Seção 230 do *Communication Decency Act* de 1996 dos Estados Unidos e pela Diretiva 2000/31/CE da União Europeia – foram importantes para o desenvolvimento das empresas de tecnologia, sendo justificáveis naquele momento, quando as redes sociais eram vistas como meros intermediários, abstendo-se de intervir no conteúdo

publicado pelos usuários. Hodiernamente, um tratamento menos rigoroso dado a grandes redes sociais, como o Facebook, não mais se justificaria, na medida em que a Meta, além de ter se tornado uma das maiores empresas de tecnologia do mundo, atua ativamente na seleção e no direcionamento do conteúdo publicado. O uso de algoritmos sofisticados para impulsionar ou deletar publicações evidencia um controle ativo sobre a visibilidade e o tipo de conteúdo permitido, o que desafia a ideia de que o Facebook seria apenas um condutor passivo de informações (Langvardt, 2021).

Observa-se que as redes sociais emergem, em certa medida, como formadoras de opinião e reguladoras do fluxo de informações. Isso porque passaram a funcionar como os novos guardiões do discurso público na internet, já que seus termos de serviço e algoritmos moldam o que pode ser dito e ouvido no espaço virtual (Eifert, 2021). Ademais, a possibilidade de remover conteúdos e de bloquear usuários concede a essas plataformas um significativo poder de influência sobre o discurso público (Eifert, 2021). De acordo com Fredes (2022, p. 148), “A forma como estruturam seus serviços influencia no livre mercado de ideias, que deixa de ser livre quando determinadas manifestações são prestigiadas e outras silenciadas, sem qualquer transparência, impactando diretamente na sociedade.” O autor prossegue apontando que

[...] a atuação das mídias sociais impacta diretamente não apenas a dimensão subjetiva da liberdade de expressão do pensamento, por exemplo, quando eventualmente alguma mídia social arbitrariamente censura determinada fala ou expulsa algum usuário, mas, mais do que isso, a atuação das mídias sociais fere a própria dimensão objetiva da liberdade de expressão em sua vinculação com a democracia, uma vez que o controle do conteúdo nas redes é realizado com finalidade lucrativa, o que resulta em uma tutela da manifestação do pensamento almejando alcançar os objetivos daqueles que patrocinam esses espaços (Fredes, 2022, p. 151).

Dessa forma, surgem questões relevantes sobre até que ponto as plataformas têm liberdade para definir seus próprios padrões e limites e até onde elas devem considerar o direito à liberdade de expressão de seus usuários (Eifert, 2021).

Por outro lado, não há como delegar apenas ao Estado a resolução dos problemas no âmbito da internet. Conforme observam Abboud e Campos (2019), a transformação do direito, nas sociedades centradas em redes e em plataformas digitais, cria desafios para a atuação jurídica tradicional, que, por vezes, “[...] decorre da falta de conhecimento do Estado para poder intervir de forma eficiente em ambientes digitais muito dinâmicos e muitas vezes constituídos por efeitos de redes e interação com inteligência artificial”. Dessa forma,



A eficiência das respostas dadas pelo direito e pela política padece de uma falha de conhecimento, da assimetria informacional trazida pela contradição entre a generalidade com que os problemas são tratados pelo ordenamento jurídico e a especialidade cognitiva própria dos subsistemas econômicos (Marques Neto, 2011, p. 82).

Diante dessas questões, buscam-se novos modelos para lidar com “problemas que não mais comportam soluções antigas”, tornando-se necessária “a criação de novas formas de geração do conhecimento dentro do direito regulatório estatal que incorpore o conhecimento advindo da sociedade” (Abboud; Campos, 2019). Discute-se, portanto, a ideia de autorregulação regulada, de modo que as autoridades estatais desempenhem um papel regulador na autorregulação, estabelecendo obrigações aos atores privados com o intuito de melhor conciliar a liberdade de expressão com os demais direitos dos usuários, considerando as complexidades do mundo digital.

Segundo Hoffmann-Riem (2019, p. 532), fala-se em

[...] autorregulamentação social regulada pelo Estado, ou, em termos breves, de autorregulação/autorregulamentação regulada/regulamentada quando órgãos estatais confiam, para a solução de problemas, nas ordens criadas com (relativa) autonomia pelos membros da sociedade, mas atuam regulatoriamente de modo que, quando isso é feito, (também) se observem ou persigam deliberadamente fins relacionados ao bem comum. A influência do Estado pode acontecer de forma extremamente diversificada, p. ex., na forma de normas ou estímulos comportamentais, por meio do estabelecimento de estruturas – como as de natureza corporativa – ou pela viabilização e apoio de sistemas funcionais da sociedade, como o mercado.

Nesse cenário, defende-se que os Estados devem exigir da normatividade interna das redes sociais pelo menos três elementos: transparência nos procedimentos, inclusive com a divulgação de relatórios sobre a moderação de conteúdo; princípios legais, “permitindo uma capacidade interna de adaptação na criação de regras de comportamentos” e princípio da boa-fé (Negócio, 2022, p. 19).

Como se verá, no próximo tópico, a Alemanha foi pioneira ao estabelecer obrigações de transparência nos procedimentos de moderação de conteúdo realizados por grandes redes sociais.

### **3 A autorregulação regulada e a experiência europeia e brasileira na regulação das redes sociais**

Compreendendo que a autorregulação precisava ser suplementada por uma estrutura de governança vertical, haja vista a ineficácia do regime regulatório existente para incentivar as redes sociais a protegerem os cidadãos do discurso de ódio, a Alemanha implementou, em 2017, a *Netzwerkdurchsetzungsgesetz* ou *NetzDG* (Pinho Filho, 2021, p. 112-114). O escopo da *NetzDG* não foi introduzir um novo regime de responsabilidade, na medida em que manteve a isenção de responsabilidade das plataformas digitais vigente na União Europeia, desde a Diretiva 2000/31/CE, quanto ao conteúdo publicado por terceiros nos casos em que as redes sociais agem imediatamente para remover ou bloquear o acesso assim que tomam conhecimento do conteúdo ilegal. Ao invés disso, a lei se destinou a regular como as plataformas gerenciam as reclamações dos usuários, a remoção e o bloqueio de publicações (Zurth, 2021, p. 1111-1112).

Nesse sentido, a *NetzDG* estabelece um conjunto de obrigações para as redes sociais com mais de dois milhões de usuários registrados na Alemanha, incluindo a obrigação de as plataformas realizarem procedimentos de remoção de publicações de forma transparente, que devem ocorrer dentro do prazo de 24 horas ou, no máximo, 7 dias, a depender da gravidade e da evidência de ilicitude (Pinho Filho, 2021, p. 114-115). Deve-se ressaltar que, de acordo com a lei, as redes sociais só são obrigadas a bloquear ou a remover conteúdo ilícito, ou seja, publicações que envolvam matérias elencadas no Código Penal alemão (Zurth, 2021, p. 1110). No caso de conteúdo ilegal que se refira a circunstâncias factuais, a rede social pode dar ao usuário a oportunidade de responder à reclamação e, assim, exceder o prazo de 7 dias. Não havendo manifestação da parte, a plataforma pode presumir que a reclamação é válida e remover o conteúdo. No caso de apresentação de defesa, a plataforma pode ponderar sobre a credibilidade das alegações (Zurth, 2021, p. 1113). O prazo de 7 dias também não se aplica se a rede social remeter a decisão para uma instituição autorreguladora, como o Comitê de Supervisão do Facebook. Nesse caso, entende-se que a rede social cumpriu sua obrigação ao repassar as queixas à referida instituição (Zurth, 2021, p. 1117).

Salienta-se, ainda, que

[...] como forma de controlar, em alguma medida, a qualidade da regulação pretendida, a *NetzDG* prevê a obrigatoriedade de fornecimento de relatórios periódicos que devem ser publicados semestralmente demonstrando não apenas os resultados do período, mas também o quão preparado está o provedor para cumprir os fins da legislação. Aqui, a lei demonstra que está preocupada não apenas com a remoção do conteúdo ilícito, mas com a transparência e responsividade dos sistemas de monitoramento (Pinho Filho, 2021, p. 115-116).



No contexto da União Europeia, também é importante destacar a aprovação do *Digital Services Act* (DSA) pelo Parlamento Europeu, que entrou em vigor em novembro de 2022. Essa legislação é voltada essencialmente para garantir o funcionamento apropriado dos serviços digitais, visando ao equilíbrio entre os poderes e responsabilidades dos provedores e os direitos fundamentais das partes envolvidas (Genç-Gelgeç, 2022, p. 31). Trata-se de regulamento que deve ser aplicado em toda a União Europeia e que atualiza as normas da Diretiva 2000/31/CE.

Entre suas inovações, destaca-se a diferenciação entre os diferentes tipos de serviços intermediários na internet, classificando-os em provedores de hospedagem, plataformas *online* – estas se diferenciam daqueles, pois, além de armazenar, disponibilizam as informações ao público –, plataformas *online* muito grandes e serviços intermediários, de modo a estabelecer obrigações diferenciadas de acordo com o tamanho da plataforma e os riscos do serviço (Buiten, 2021, p. 368). De acordo com o artigo 33º do DSA, plataformas *online* muito grandes são aquelas que “[...] têm um número médio mensal de destinatários ativos do serviço na União igual ou superior a 45 milhões” (Parlamento Europeu, 2022), o que inclui o Facebook<sup>2</sup>.

É relevante pontuar que o DSA preservou o regime que isenta as plataformas de responsabilidade pelo conteúdo publicado por terceiros, contudo detalhou o procedimento de notificação e retirada de conteúdo e estabeleceu obrigações de transparência e diligência aos provedores. A fim de ilustrar uma parte dessas obrigações, cabe ressaltar alguns dispositivos.

No que se refere à moderação de conteúdo, o *Digital Services Act* prevê, no Artigo 14º, a obrigação de as plataformas digitais incluírem, em seus termos e condições, informações sobre “[...] quaisquer políticas, procedimentos, medidas e instrumentos utilizados para efeitos de moderação de conteúdos, incluindo a tomada de decisões algorítmicas e a análise humana, bem como as regras processuais do respectivo sistema interno de gestão de reclamações” (Parlamento Europeu, 2022). Enfatiza ainda que referidas informações devem ser apresentadas ao público de forma acessível com uma linguagem clara. Além disso, estabelece, no Artigo 15º, o dever de as plataformas publicarem relatórios anuais sobre a moderação de conteúdo realizada no período, contendo, entre outras informações, apontamentos sobre moderação de conteúdo por iniciativa própria dos prestadores de serviço; medidas tomadas que impactam

<sup>2</sup> Segundo a Meta, “No período de seis meses encerrado em 31 de dezembro de 2022, havia aproximadamente 255 milhões de usuários ativos mensais no Facebook na União Europeia e aproximadamente 250 milhões de usuários ativos mensais no Instagram na União Europeia” (Meta, 2023, tradução nossa).



disponibilidade, visibilidade e acessibilidade das informações e restrições aos serviços; informações acerca de reclamações recebidas e ações tomadas em resposta; tempo de tomada dessas ações e ocorrências de reversões da decisão. A utilização de meios automatizados para moderação de conteúdo também deve ser registrada, fornecendo-se descrições qualitativas, metas específicas, indicadores de precisão, taxa de erros e garantias aplicadas (Parlamento Europeu, 2022).

O Artigo 17º dispõe que as plataformas, ao tomarem medidas como remover, suspender ou bloquear um conteúdo após notificação, devem fornecer aos destinatários dos serviços afetados uma exposição de motivos, que deve apresentar, entre outros detalhes, as justificativas da medida. Ademais, devem ser mencionados o fundamento jurídico invocado, no caso de conteúdo ilegal, e a referência contratual, se houver incompatibilidade com os termos do prestador de serviço (Parlamento Europeu, 2022). Cabe destacar também o Artigo 20º, que estabelece a criação de um sistema interno de gestão de reclamações. Por meio dele, as plataformas devem acolher reclamações sobre decisões que impliquem remoção de informações e suspensão de serviço, de conta ou de monetização, tratando-as de maneira imparcial, e reverter decisões infundadas sem atrasos desnecessários, sempre informando os autores sobre todas as resoluções e garantindo que as decisões sejam tomadas sob supervisão de colaboradores qualificados, e não apenas por meios automatizados (Parlamento Europeu, 2022).

Salienta-se, por fim, o Artigo 37º, que prevê que as plataformas *online* muito grandes devem ser submetidas a uma auditoria independente às suas próprias custas, pelo menos uma vez ao ano, para avaliar, entre outros, o cumprimento das obrigações de devida diligência (Parlamento Europeu, 2022). Para garantir o cumprimento das obrigações pelos provedores e a implementação do regulamento, o DSA estabelece a criação dos coordenadores dos serviços digitais, que são as principais autoridades nacionais designadas pelos Estados-membros para reforçar a aplicação do regulamento, além da criação do Comitê Europeu dos Serviços Digitais, um grupo consultivo independente de coordenadores dos serviços digitais (Parlamento Europeu, 2022).

No Brasil, o Marco Civil da Internet, Lei nº 12.965/2014, adotou uma abordagem semelhante à europeia, ao eximir os provedores de responsabilidade sobre as publicações de seus usuários, desde que não tenham conhecimento do conteúdo considerado ilícito ou tomem as medidas cabíveis após notificados sobre o fato. A lei brasileira, contudo, foi mais exigente ao prever que essa notificação deve ocorrer mediante ordem judicial específica, ressalvados os



casos do art. 21<sup>3</sup> (Brasil, 2014). Analisando a redação do art. 19 do Marco Civil, percebe-se a preocupação do legislador em proteger a liberdade de expressão. A adoção de uma estrutura de responsabilidade mais rígida poderia precipitar o aumento da remoção de conteúdo, possivelmente conduzindo a um cenário de censura privada, um resultado indesejado que a legislação buscou evitar para proteger o livre discurso. O mesmo nível de atenção à liberdade de expressão, todavia, não se estendeu para regulamentar, com mais rigor, a transparência necessária para a remoção de conteúdo realizada arbitrariamente pelas redes sociais. Atualmente, as plataformas têm autonomia para remover conteúdo sem uma ordem judicial ou notificação prévia, desde que considerem que houve violação dos seus termos de uso.

É importante mencionar o Projeto de Lei nº 2630/2020, aprovado no Senado, em 30 de junho de 2020, que estabelece “[...] normas, diretrizes e mecanismos de transparência para provedores de redes sociais e de serviços de mensageria privada a fim de garantir segurança e ampla liberdade de expressão, comunicação e manifestação do pensamento” (Brasil, 2020). De modo semelhante à *NetzDG* alemã, as disposições do PL 2630/2020 não se aplicam a redes sociais com menos de dois milhões de usuários registrados no Brasil (Brasil, 2020).

No texto original do Projeto, é importante sublinhar alguns pontos relevantes, como o art. 12, que impõe às plataformas o dever de garantir aos usuários acesso à informação e liberdade de expressão “nos processos de elaboração e aplicação de seus termos de uso, disponibilizando mecanismos de recurso e devido processo” (Brasil, 2020). Além disso, dispõe que os usuários devem ser notificados, caso suas publicações ou contas sejam alvo de medidas ou denúncias, recebendo informações sobre os fundamentos, processos de análise, aplicação de medidas, prazos e procedimentos para contestação<sup>4</sup>. O artigo também assevera o direito de defesa do usuário, quando houver remoção de conteúdo (Brasil, 2020).

<sup>3</sup> Consoante dispõe o art. 21 do Marco Civil da Internet, “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” (Brasil, 2014).

<sup>4</sup> Essa regra é excetuada pelo Projeto nas seguintes hipóteses: risco de dano imediato de difícil reparação; para a segurança da informação ou do usuário; violação a direitos de crianças e adolescentes; crimes tipificados na Lei nº 7.716, de 5 de janeiro de 1989; grave comprometimento da usabilidade, integridade ou estabilidade da aplicação (Brasil, 2020).

Ressalta-se ainda o art. 30, que aborda expressamente a autorregulação regulada, incentivando a criação de uma instituição com capacidade para autorregular a transparência e a responsabilidade no uso da internet, com atribuições mínimas definidas em lei (Brasil, 2020).

Percebe-se, dessa forma, que o Brasil, com o PL 2630/2020, busca alinhar-se com as experiências internacionais na regulação de plataformas digitais, “[...] no caminho da *due diligence obligations*, quer dizer, em deveres procedimentais e requisitos organizacionais”, não focando apenas nos aspectos da responsabilidade civil (Moura; Rocha, 2022, p. 22).

## Conclusão

A liberdade de expressão é um direito fundamental para muitos países e vem sendo testada à medida que as redes sociais se tornam, cada vez mais, importantes com o grande fluxo de informações e da comunicação ocorrendo nessas plataformas privadas.

Dessa forma, tem gerado preocupação a forma como grandes redes sociais, como o Facebook, realizam a moderação de conteúdo, em razão de sua discricionariedade e da falta de transparência no processo de tomada de decisões, especialmente quando utilizam sistemas de inteligência artificial.

Por outro lado, o Estado carece de mecanismos para regular totalmente o setor, de modo que permitir que as empresas privadas continuem se autorregulando consiste em uma alternativa oportuna. Nesse contexto, a Alemanha, com a aprovação da *NetzDG* em 2017, foi pioneira em estabelecer obrigações a serem cumpridas pela regulação privada no que se refere ao procedimento de remoção de conteúdo em grandes plataformas digitais. Destaca-se também a aprovação do *Digital Services Act (DSA)*, em julho de 2022, pelo Parlamento Europeu, como forma de autorregulação regulada, com obrigações de devida diligência e de transparência a provedores de internet, estabelecendo deveres diferenciados de acordo com o tamanho das plataformas e os riscos da atividade.

No Brasil, o Projeto de Lei 2630/2020 tenta estabelecer parâmetros para a atuação das plataformas digitais. Surgem, entretanto, questionamentos, entre os quais: a autoridade designada para controlar a autorregulação das plataformas está preparada para tal responsabilidade? É importante reconhecer que questões relativas a redes sociais enfrentam desafios próprios. A velocidade da disseminação de informações, a enorme quantidade de conteúdo gerado diariamente e a diversidade global de usuários são obstáculos significativos



para a efetiva moderação de conteúdo. A autoridade responsável precisa estar preparada para enfrentar o desafio de conciliar as liberdades individuais com a proteção do bem comum. Esse é um processo dinâmico e complexo que demandará discussões mais aprofundadas e mais pesquisas.

Em conclusão, fica evidente que a autorregulação regulada se apresenta como um instrumento promissor para reger as questões relacionadas à moderação de conteúdo nas grandes plataformas digitais, uma vez que busca harmonizar as especificidades da inovação tecnológica com os direitos fundamentais dos usuários. É importante perceber, contudo, que essa abordagem não está livre de obstáculos e de limitações, uma vez que não oferece, por exemplo, soluções aos problemas intrínsecos ao poder excessivo de mercado detido por algumas plataformas, seja no domínio sobre os dados, seja no controle sobre a atenção dos usuários. O desafio regulatório demanda um outro passo: uma integração regulatória entre moderação de conteúdo, proteção de dados e concorrência. Isso não exclui os primeiros passos dados com a autorregulação regulada.

## Referências

ABBOUD, Georges; CAMPOS, Ricardo. A autorregulação regulada como modelo do Direito proceduralizado: regulação das redes sociais e proceduralização. *In*: ABBOUD, Georges; NERY JUNIOR, Nelson; CAMPOS, Ricardo. (org.). **Fake News e Regulação**. 2. ed. São Paulo: Revista dos Tribunais, 2019. Disponível em:

<https://www.jusbrasil.com.br/doutrina/secao/capitulo-4-a-autorregulacao-regulada-como-modelo-do-direito-proceduralizado-regulacao-de-redes-sociais-e-proceduralizacao-fake-news-e-regulacao/1197132530#a-243083542> Acesso em: 6 nov. 2022.

BRASIL. Lei nº 12.965, de 23 de abril de 2014. **Marco Civil da Internet**. Estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2014/lei/112965.htm](https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/112965.htm) Acesso em: 2 dez. 2022.

BRASIL. Senado Federal. **Projeto de Lei nº 2630, de maio de 2020**. Institui a Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet. Disponível em: [https://www.camara.leg.br/proposicoesWeb/prop\\_mostrarintegra?codteor=1909983](https://www.camara.leg.br/proposicoesWeb/prop_mostrarintegra?codteor=1909983) Acesso em: 4 ago. 2023.

BUITEN, Miriam C., The Digital Services Act: From Intermediary Liability to Platform Regulation, **Journal of Intellectual Property, Information Technology and E-Commerce Law**, v. 12, 2021. HeinOnline.



COMITÊ de Supervisão mantém suspensão do ex-presidente Trump e conclui que Facebook falhou em impor a penalidade adequada. **Comitê de Supervisão**, 2021. Disponível em: <https://www.oversightboard.com/news/226612455899839-oversight-board-upholds-former-president-trump-s-suspension-finds-facebook-failed-to-impose-proper-penalty/> Acesso em: 3 jan. 2023.

EIFERT, Martin. Accountability of Social Media: A European Perspective. **Legal Grounds Institute**, 3 nov. 2021. Disponível em: <https://legalgroundsinstitute.com/blog/accountability-of-social-media-a-european-perspective/> Acesso em: 8 ago. 2023.

FACEBOOK. Oversight Board. **Oversight Board Charter**, set. 2019. Disponível em: [https://about.fb.com/wp-content/uploads/2019/09/oversight\\_board\\_charter.pdf](https://about.fb.com/wp-content/uploads/2019/09/oversight_board_charter.pdf) Acesso em: 2 ago. 2023.

FRAZÃO, Ana. Plataformas digitais e os desafios para a regulação jurídica. In: Parentoni, Leonardo (Coord.). **Direito, tecnologia e inovação**. Belo Horizonte: Editora D'Plácido, 2018. Disponível em: [https://www.academia.edu/37773223/Plataformas\\_digitais\\_e\\_os\\_desafios\\_para\\_a\\_regula%C3%A7%C3%A3o\\_jur%C3%ADdica](https://www.academia.edu/37773223/Plataformas_digitais_e_os_desafios_para_a_regula%C3%A7%C3%A3o_jur%C3%ADdica) Acesso em: 6 nov. 2022.

FREDES, Andrei Ferreira. **Liberdade de expressão, direito à informação e redes sociais: regulação constitucionalmente adequada sobre a moderação de conteúdo na construção de um espaço virtual democrático e plural**. Tese (Doutorado em Direito) - Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2022. Disponível em: <https://tede2.pucrs.br/tede2/handle/tede/10201>. Acesso em: 21 nov. 2022.

GENÇ-GELGEÇ, Berrak. Regulating digital platforms: will the DSA correct its predecessor's deficiencies? **Croatian Yearbook of European Law and Policy**, University of Zagreb, v. 18, nov. 2022. Disponível em: <https://www.cyelp.com/index.php/cyelp/article/view/485>. Acesso em: 21 nov. 2022.

GORWA, Robert; BINNS, Reuben; KATZENBACH, Christian. Algorithmic content moderation: Technical and political challenges in the automation of platform governance. **Big Data & Society**, v. 7, n. 1, p. 1-15, 2020. Disponível em: <https://journals.sagepub.com/doi/full/10.1177/2053951719897945> Acesso em: 3 dez. 2022.

HOFFMANN-RIEM, Wolfgang. Autorregulação, Autorregulamentação e autorregulamentação regulamentada no contexto digital. **Revista da AJURIS**, Porto Alegre, v. 46, n. 146, p. 530-553, 2019. Disponível em: <http://ajuris.kingghost.net/OJS2/index.php/REVAJURIS/article/view/1048> Acesso em: 3 dez. 2022.

KLEINMAN, Zoe. Cambridge Analytica: the story so far. **BBC News**, 21 mar. 2018. Disponível em: <https://www.bbc.com/news/technology-43465968> Acesso em: nov. 2022.

KOLTAY, Andras. The Protection of Freedom of Expression from Social Media Platforms. **Mercer Law Review**, v. 73, n. 2, p. 523-589, Spring 2022. HeinOnline.





LANGVARDT, Kyle. Can the First Amendment Scale? **J. Free Speech L.**, v. 1, p. 273-302, 2021. Disponível em: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3911521](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3911521) Acesso em: nov. 2022.

MARQUES NETO, Floriano de Azevedo. Regulação estatal e autorregulação na economia contemporânea. **Revista de Direito Público da Economia**. Editora Fórum. Belo Horizonte. v. 9, n. 33. jan./mar. 2011.

META. Os Padrões da Comunidade do Facebook descrevem o que é ou não permitido na plataforma. **Padrões de Comunidade do Facebook**. Disponível em: <https://transparency.fb.com/pt-br/policies/community-standards/?source=https%3A%2F%2Fpt-br.facebook.com%2Fcommunitystandards> Acesso em: 10 dez. 2022.

META. Digital Services Act - Information on Average Monthly Active Recipients in the European Union. **Transparency Center**, 14 fev. 2023. Disponível em: <https://transparency.fb.com/sr/dsa-report-feb2023/> Acesso em: 3 ago. 2023.

MOURA, Ariel Augusto Lira de; ROCHA, Leonel Severo. Direitos Fundamentais e redes sociais: Da moderação de conteúdo no Facebook ao Direito na cultura das redes. **Revista Direito Mackenzie**, v. 16, n. 2, 2022. Disponível em: <https://editorarevistas.mackenzie.br/index.php/rmd/article/view/15303> Acesso em: 3 dez. 2022.

NEGÓCIO, Ramon de Vasconcelos. Autorregulação digital: da normatividade excludente para o Diálogo normativo com o estado. **Revista Culturas Jurídicas**, Ahead of Print, v. 10, 2022. Disponível em: <https://periodicos.uff.br/culturasjuridicas/article/view/48188> Acesso em 2 dez. 2022.

PARLAMENTO EUROPEU, **Regulamento (UE) 2022/2065 do Parlamento Europeu e do Conselho de 19 de outubro de 2022 relativo a um mercado único para os serviços digitais e que altera a Diretiva 2000/31/CE (Regulamento dos Serviços Digitais)**. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:32022R2065> Acesso em: 12 dez. 2022.

PINHO FILHO, José Célio Belém de. **Desinformação e Regulação das Redes Sociais Digitais**. Dissertação (Mestrado) - Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa (IDP), São Paulo, 2021. Disponível em: <https://repositorio.idp.edu.br/handle/123456789/3391> Acesso em: 15 nov. 2022.

SECURING ongoing funding for the Oversight Board. **Comitê de supervisão**, 2022. Disponível em: <https://www.oversightboard.com/news/1111826643064185-securing-ongoing-funding-for-the-oversight-board/> Acesso em: 3 ago. 2023.

SILVA, Raphael Ferreira Santana. **Igualdade e governança de redes sociais: Interseção entre tecnologia, moderação de conteúdo e direito à igualdade**. Dissertação (Mestrado) -



Escola de Direito do Rio de Janeiro da Fundação Getúlio Vargas, Rio de Janeiro, 2022.

Disponível em:

<https://bibliotecadigital.fgv.br/dspace/bitstream/handle/10438/32485/Dissertac%cc%a7a%cc%83o%20Raphael%20Ferreira%20Santana%20Silva1.pdf?sequence=5&isAllowed=y> Acesso em: 15 nov. 2022.

ZURTH, Patrick. The German NetzDG as Role Model or Cautionary Tale? Implications for the Debate on Social Media Liability. **Fordham Intellectual Property, Media & Entertainment Law Journal**, v. 31, n. 4, p. 1084-1153, Summer 2021. HeinOnline.

