

ENTRE A REAÇÃO E A PREVENÇÃO: LIMITES E POTENCIALIDADES DO PL 2.628/2022 NA PROTEÇÃO DIGITAL DE CRIANÇAS E ADOLESCENTES

BETWEEN REACTION AND PREVENTION: LIMITS AND POTENTIALITIES OF BILL 2,628/2022 IN THE DIGITAL PROTECTION OF CHILDREN AND ADOLESCENTS

Rafael de Lima Kurschner¹

Lya Hadassa Sobral Viana²

Marina Lorenz Vernetti³

Resumo: A crescente digitalização da vida social ampliou de forma significativa as interações em plataformas digitais, mas também intensificou a exposição de crianças e adolescentes a riscos relacionados a conteúdos nocivos e violações de direitos fundamentais. Desse modo, a problematização central deste estudo consiste em analisar as limitações normativas e conceituais do Projeto de Lei nº 2.628/2022, que pretende estabelecer um marco regulatório voltado à proteção da infância e da adolescência no ambiente digital brasileiro. Metodologicamente, o trabalho utiliza abordagem qualitativa e crítico-comparativa, fundamentada em revisão bibliográfica, análise documental de legislações nacionais e internacionais. Parte-se da hipótese de que o PL 2.628/2022, em sua atual configuração, não assegura mecanismos suficientes para a proteção integral de crianças e adolescentes no ambiente digital. Em parte, conclui-se que apenas a adoção de um modelo híbrido e multissetorial poderá oferecer um marco regulatório eficaz e alinhado aos princípios internacionais de direitos humanos.

Palavras-chave: Empresas e Direitos Humanos; Direitos Humanos; Moderação de conteúdo; PL 2.628/2022.

¹ Mestrando pelo Programa de Pós-graduação em Direito da Pontifícia Universidade Católica do Paraná (PPGD/PUCPR); Bacharel em Direito pela Universidade Estadual do Oeste do Paraná (UNIOESTE).

² Mestranda pelo Programa de Pós-graduação em Direito da Pontifícia Universidade Católica do Paraná (PPGD/PUCPR); Bacharela em Direito pela Pontifícia Universidade Católica do Paraná (PUCPR).

³ Mestranda pelo Programa de Pós-graduação em Direito da Pontifícia Universidade Católica do Paraná (PPGD/PUCPR); Bacharela em Direito pela Pontifícia Universidade Católica do Paraná (PUCPR).

Abstract: The increasing digitalization of social life has significantly expanded interactions on digital platforms, but it has also intensified the exposure of children and adolescents to risks related to harmful content and violations of fundamental rights. Thus, the central issue of this study is to analyze the normative and conceptual limitations of Bill No. 2,628/2022, which seeks to establish a regulatory framework aimed at protecting childhood and adolescence in the Brazilian digital environment. Methodologically, the study employs a qualitative and critical-comparative approach, based on a literature review and documentary analysis of national and international legislation. The hypothesis is that Bill No. 2,628/2022, in its current form, does not provide sufficient mechanisms for the comprehensive protection of children and adolescents in the digital environment. It is partly concluded that only the adoption of a hybrid and multisectoral model could offer an effective regulatory framework aligned with international human rights principles.

Keywords: Business and Human Rights; Human Rights; Content Moderation; Bill No. 2,628/2022.

1 Introdução

A crescente digitalização da vida social ampliou de forma significativa as interações em plataformas virtuais, mas também intensificou a exposição de crianças e adolescentes a riscos relacionados a conteúdos nocivos e violações de direitos fundamentais.

O caso da plataforma Discord, inicialmente projetada para a comunidade gamer, ilustra com clareza tais desafios: a ausência de mecanismos eficazes de moderação e de verificação etária permitiu a circulação de materiais ilícitos e a ocorrência de crimes graves, envolvendo inclusive exploração sexual e violência. Diante desse cenário, torna-se urgente discutir o papel da moderação de conteúdo como instrumento de governança digital e a responsabilidade das empresas de tecnologia em garantir um ambiente online seguro e inclusivo.

A problematização central deste estudo consiste em analisar as limitações normativas e conceituais do Projeto de Lei nº 2.628/2022, que pretende estabelecer um marco regulatório voltado à proteção da infância e da adolescência no ambiente digital brasileiro. Embora represente um avanço, sua formulação ainda se mostra reativa e insuficiente diante da complexidade do ecossistema digital, ao não

incorporar devida diligência em direitos humanos, mecanismos preventivos de risco e participação efetiva das crianças nos processos de decisão.

Metodologicamente, o trabalho utiliza abordagem qualitativa e crítico-comparativa, fundamentada em revisão bibliográfica, análise documental de legislações nacionais e internacionais. A estrutura do artigo organiza-se em três eixos principais: (1) uma análise inicial sobre a moderação de conteúdo a partir do exemplo da plataforma Discord, destacando riscos e responsabilidades; (2) uma avaliação crítica do art. 22 do Projeto de Lei nº 2.628/2022, enfatizando suas fragilidades normativas e conceituais; e (3) a discussão sobre responsabilidade empresarial e devida diligência em direitos humanos, em diálogo com padrões internacionais e experiências estrangeiras.

Parte-se da hipótese de que o PL 2.628/2022, em sua atual configuração, não assegura mecanismos suficientes para a proteção integral de crianças e adolescentes no ambiente digital, sobretudo por adotar uma abordagem reativa de remoção de conteúdos em detrimento de estratégias preventivas, transparentes e participativas.

Ao final, conclui-se que apenas a adoção de um modelo híbrido e multissetorial, que envolve Estado, empresas, sociedade civil e os próprios sujeitos de direitos, poderá oferecer um marco regulatório eficaz e alinhado aos princípios internacionais de direitos humanos.

2 O Discord e a moderação de conteúdo

Para tratar da moderação de conteúdo, a título exemplificativo, inicia-se o presente eixo apresentando fatos relacionados à plataforma Discord, que nos últimos anos vem sendo alvo de denúncias por abrigar comunidades envolvidas em graves violações de direitos humanos e violência contra crianças e adolescentes.

O Discord foi desenvolvido em 2015, nos Estados Unidos da América, e inicialmente tinha como público-alvo a comunidade *gamer*, tornando-se um espaço propício para a interação entre usuários de diferentes idades, enquanto participavam de jogos online. Em 2021, a plataforma obteve um crescimento de usuários, sobretudo por conta do distanciamento social exigido no período de pandemia de COVID-19 (Bico; Da Silva, 2025, p. 7). Em tal ano, cerca de 70% dos usuários utilizavam a plataforma para atividades que não envolviam apenas jogos (Paz, 2023, p. 87).

Na prática, o uso da plataforma funciona da seguinte forma: o usuário pode escolher quais servidores (uma espécie de comunidade) quer participar, bem como quem vão ser os seus “amigos” no ambiente virtual, e são os “donos” ou “administradores” de cada servidor que estabelecem as regras específicas e instruções de comportamento para cada usuário participe (sem que haja um controle direto da plataforma sobre essas decisões). Além disso, qualquer pessoa, independentemente da idade, pode integrar um servidor e trocar mensagens de forma anônima e privada, ou de forma pública (Paz, 2023, p. 88).

Em seus Termos de Uso, o Discord delimita os 13 anos como idade mínima para a criação de uma conta. Contudo, nos Termos de Uso vigentes até 2022, durante o cadastro na plataforma, não era solicitada a idade do futuro usuário, apenas uma simples confirmação quanto a idade mínima indicada (Discord, 2024; Bico; Da Silva, 2025, p. 7).

Ora, a exigência mínima de idade associada à falta de controle por parte da plataforma no que diz respeito à moderação expõe crianças e adolescentes a vários tipos de violações. Como exemplo, o delegado Alesandro Barreto, coordenador-geral do Ciberlab, órgão do Ministério da Justiça de combate a crimes no ambiente digital, relata transmissão ao vivo no Discord, na qual um homem torturava um cachorro com as patas e focinho amarrados, enquanto os telespectadores solicitavam mais intensidade na crueldade contra o animal, que ao final foi morto a tiros em frente a câmera (Tavares, 2025).

Também por intermédio do Discord, criminosos lucram com a produção de conteúdo de violência contra crianças e adolescentes. Na plataforma, existem comunidades que são destinadas à prática de extorsão, tortura, automutilação e crimes sexuais, que ocorrem por intermédio de coação, chantagem e ameaças. Isso é possível pois, em grande parte dos relatos das vítimas, o agressor simulava um “web namoro” com a criança ou adolescente, que era convencido a enviar fotos íntimas, que posteriormente serviriam como moeda de troca para se submeter aos crimes, para não serem vazadas suas imagens íntimas (Hidalgo; Sant’ana, 2025).

Esse conteúdo produzido, em formato de fotos e vídeos, é catalogado e vendido por intermédio de links disponibilizados dentro do Discord. Parte desse conteúdo, livremente exposto dentro da plataforma, teve seu preço explícito em uma planilha que circulou entre comunidades, mostrando que os valores pagos dependem

do tipo de imagem e vídeo, bem como da parte do corpo da criança ou do adolescente exposto, indicando-se também pornografia infantil (Hidalgo; Sant'ana, 2025).

Tal cenário demonstra a urgência na discussão a respeito da moderação e dos conteúdos nas redes sociais, bem como a discussão sobre a responsabilidade das empresas. Conforme indica a TIC Kids Online Brasil (2022), 92% da população brasileira de 9 a 17 anos é usuária da internet, o que representa cerca de 24,4 milhões de crianças e adolescentes. Destes, 86% têm perfil em redes sociais, sendo esse percentual ainda maior entre adolescentes de 15 a 17 anos (96%).

A Observação Geral nº 25 do Comitê dos Direitos da Criança da ONU reforça que o ambiente digital deve apoiar, promover e proteger os direitos das crianças, sendo responsabilidade dos Estados assegurar legislação e políticas específicas que garantam sua segurança, inclusão e participação equitativa nesses espaços (Comitê dos Direitos da Criança, 2021, §3 e §7).

Embora seja possível adotar definições técnicas — como a prevista pela União Europeia no *Digital Services Act* (DSA), que será abordada nos próximos capítulos —, a moderação pode ser compreendida como o controle dos conteúdos disponibilizados aos usuários em determinada plataforma, associado à responsabilização pela hospedagem de conteúdos ilícitos. O DSA (Regulamento UE 2022/2065) define a prática como:

As atividades, automatizadas ou não, empreendidas por prestadores de serviços intermediários, destinadas, em especial, a detectar, identificar e combater os conteúdos ilegais ou informações incompatíveis com os seus termos e condições [...] incluindo as medidas tomadas que afetam a disponibilidade, visibilidade e acessibilidade desses conteúdos [...], bem como ações que impactam a capacidade dos destinatários de fornecerem essas informações (Regulamento dos Serviços Digitais, 2022, p. 43).

Na prática, a moderação constitui ferramenta essencial de governança digital. Conforme Grimmelmann (2015, p. 47), trata-se de “um mecanismo de governança que estrutura a participação em uma comunidade para facilitar a cooperação e evitar abusos”. Essa função é simultaneamente normativa e protetiva, pois garante a aplicação das regras internas das plataformas e protege os usuários contra interações nocivas. As motivações para a adoção da moderação são diversas, abrangendo

desde o alinhamento com os valores da comunidade até a contenção de conteúdos danosos sob as perspectivas social, emocional ou econômica (Almeida, 2022).

Gillespie (2020, p. 2) categoriza a moderação em três níveis operacionais: (1) equipes que elaboram políticas; (2) moderadores que aplicam essas diretrizes; (3) usuários que sinalizam possíveis violações. Essa estrutura funciona de forma híbrida, automatizada e manual e, embora *softwares* sejam eficazes para aplicar comandos padronizados em larga escala, falham diante de ambiguidades e contextos culturais específicos (Grimmelmann, 2015, p. 63).

Langvardt (2018, p. 1357) descreve três modelos regulatórios em debate nos EUA: (1) supervisão administrativa estatal; (2) responsabilização direta dos usuários; (3) imposição de transparência nas regras das plataformas. Todos apresentam desafios, do constitucionalismo à complexidade transnacional dessas corporações. Um exemplo recorrente é a despromoção algorítmica, prática que reduz a visibilidade de conteúdos sem removê-los (Gillespie, 2020, p. 2-3). Embora aplicada com a intenção de proteger crianças, ela suscita preocupações com transparência, prestação de contas e viés algorítmico.

A Estratégia Europeia BIK+ reconhece que a concepção de serviços digitais deve considerar ativamente os riscos específicos enfrentados pelas crianças. O DSA obriga as plataformas a adotarem medidas proporcionais para proteger os direitos das crianças, o que inclui desde a limitação de publicidade baseada em perfis até a adaptação da linguagem dos termos de uso. Interfaces que explorem a inexperiência dos menores ou promovam comportamentos aditivos são especialmente visadas, exigindo moderação de conteúdo que respeite os direitos da criança e garanta ambientes seguros, adequados à idade e transparentes (Comissão Europeia, 2022, p. 6-7)

O DSA avança ao exigir que grandes plataformas promovam transparência nos sistemas de recomendação e moderação. Contudo, segundo Common (2020), as plataformas ainda operam com regras vagas, incoerentes e de difícil compreensão, gerando “incertezas kafkianas” (Common, 2020, p. 9) e comprometendo o Estado de Direito.

Barrett (2020, p. 3) denuncia a precarização dos moderadores terceirizados, tratados como “funcionários de segunda classe”. Sem proteção psicológica adequada, esses trabalhadores enfrentam danos emocionais severos diante da exposição contínua a conteúdos extremos. Considera-se que essa terceirização

constitui uma “negação plausível” das empresas sobre o impacto real da moderação, ocultando uma função crítica sob estruturas precárias (Barrett , 2020, p. 4).

Adicionalmente, a moderação envolve interesses econômicos e políticos. A moderação visa preservar modelos de negócios baseados em engajamento e publicidade, além de que as decisões de moderação refletem a imagem que as empresas desejam projetar (Klonick, 2017, p. 1615-1616; Sander, 2020, p. 948). Adverte-se que esse poder regulatório exercido por empresas privadas demanda revisão crítica da separação entre esferas pública e privada. Quando tais decisões impactam diretamente os direitos fundamentais, surgem desafios éticos e jurídicos globais (Laidlaw, 2012, p. 3-4).

Internacionalmente, destacam-se três modelos regulatórios: (1) autorregulação (como nos EUA, amparada pela Seção 230 do *Communications Decency Act*); (2) moderação estatal (como no Brasil, via art. 19 do Marco Civil da Internet); e (3) autorregulação regulada, adotada pela União Europeia com o DSA. Cada modelo traz vantagens e limitações: agilidade, garantismo jurídico ou busca de equilíbrio entre eficiência e responsabilidade (Common, 2020; Smart; Coddou, 2022).

Diante do constatado na atuação do Discord, sem embargo de outros exemplos, é urgente a adoção de parâmetros públicos, democráticos e transparentes, e que estejam ancorados nos Princípios Orientadores sobre Empresas e Direitos Humanos (POs), visando assegurar simultaneamente a liberdade de expressão, a proteção de grupos vulneráveis e a responsabilidade social das plataformas (Smart; Coddou, 2022, p. 271).

A construção de um ambiente digital seguro e inclusivo para crianças e adolescentes exige a superação de modelos autorreferenciais de autorregulação empresarial. Como mostram os dados da *TIC Kids Online Brasil 2022*, as experiências infanto-juvenis na internet são intensas, frequentes e nem sempre mediadas por adultos, o que amplia os riscos de exposição a conteúdos nocivos, práticas discriminatórias e violações de direitos (CGI.br, 2023, p. 26).

Diante disso, torna-se imprescindível estabelecer mecanismos regulatórios híbridos e efetivamente fiscalizáveis, que combinem o dever estatal de proteção, a responsabilização das empresas e a escuta ativa dos próprios sujeitos de direitos: as crianças. Nesse sentido, a Observação Geral nº 25 do Comitê da ONU (2021) orienta os Estados a considerar, em todas as decisões que envolvam o ambiente digital, o interesse superior da criança, garantindo a inclusão digital equitativa, o respeito à

privacidade, a promoção de habilidades críticas e a participação significativa das crianças no desenho e implementação de políticas públicas (§12, §17).

Outro avanço relevante trazido pela Estratégia BIK+ é o reconhecimento do papel ativo das crianças na construção de uma Internet mais segura. A Comissão Europeia propõe envolver diretamente crianças e jovens em processos decisórios relacionados à governança digital, inclusive na formulação de códigos de conduta sobre moderação de conteúdo. Essa abordagem participativa busca garantir que as soluções adotadas não apenas protejam, mas também empoderem os sujeitos de direitos, valorizando suas experiências e percepções no espaço digital (Comissão Europeia, 2022, p. 18-19).

Portanto, verifica-se no exemplo da plataforma Discord que, embora criada com finalidade inicial voltada à comunidade *gamer*, tornou-se espaço de interação amplo, mas também vulnerável a graves violações, especialmente contra crianças e adolescentes. A ausência de mecanismos eficazes de moderação e verificação etária favoreceu a proliferação de práticas criminosas, como extorsão, exploração sexual e circulação de conteúdos ilícitos.

Nesse contexto, a moderação de conteúdo se revela não apenas um instrumento técnico, mas uma ferramenta de governança digital, essencial para equilibrar liberdade de expressão, proteção de grupos vulneráveis e responsabilidade das plataformas. Existe a necessidade de modelos regulatórios híbridos, que articulem Estado, empresas e sociedade civil, garantindo transparência, fiscalização e participação ativa das próprias crianças e adolescentes nas decisões sobre a segurança digital.

3 Análise crítica do artigo 22 do PL 2.628/2022

Considerando os problemas identificados a partir do modelo de negócio do Discord e as expectativas de proteção de crianças e adolescentes por meio da moderação de conteúdo, este eixo desenvolve uma análise crítica do Projeto de Lei nº 2.628/2022. É dada especial atenção às suas limitações normativas e conceituais, em particular no tocante à responsabilidade das empresas de tecnologia na moderação de conteúdo voltado a crianças e adolescentes no ambiente digital brasileiro.

O referido projeto propõe um marco normativo voltado à proteção infantojuvenil no meio digital. Entre seus dispositivos, verifica-se que o art. 6º estabelece que fornecedores de produtos e serviços de tecnologia da informação voltados ou acessíveis a crianças e adolescentes devem adotar medidas para prevenir e mitigar a exposição a conteúdos nocivos, especialmente os que envolvam: (1) exploração e abuso sexual de crianças e adolescentes; (2) violência, *bullying* virtual e assédio; (3) comportamentos de risco à saúde mental, como vício, ansiedade, depressão, transtornos alimentares, uso de substâncias e condutas suicidas; (4) promoção de jogos de azar, tabaco, álcool e drogas; (5) práticas publicitárias predatórias ou enganosas.

Por sua vez, o art. 21 pontua que os fornecedores de produtos ou serviços de tecnologia usados por crianças e adolescentes devem disponibilizar mecanismos de denúncia para violações de seus direitos. Quando notificados, têm o dever de comunicar às autoridades competentes para investigação. Já o art. 22 destaca como deve os fornecedores remover imediatamente conteúdos que violem direitos de crianças e adolescentes, assim que forem notificados, mesmo sem ordem judicial.

O §1º aponta que são consideradas violações os conteúdos previstos no art. 6º (exploração sexual, violência, vício, publicidade predatória etc.), e o §2º indica que a notificação deve ter informações claras que identifiquem o material, não sendo aceita denúncia anônima. Por fim, o §3º indica que o mecanismo de denúncia deve ser público, transparente e de fácil acesso. Desse modo, os artigos acima apresentados reforçam a responsabilidade ativa das plataformas: criar canais de denúncia, notificar autoridades e agir rapidamente para retirar conteúdos nocivos que afetam crianças e adolescentes.

Contudo, embora represente avanço ao reconhecer obrigações extrajudiciais de remoção, o art. 22 apresenta fragilidades normativas, sobretudo ao utilizar como parâmetro o Comentário Geral nº 25 da ONU sobre os Direitos da Criança, os CRBP, o *B-Tech UNICEF Briefing* e o DSA da União Europeia.

Observa-se que o PL falha ao não incorporar de forma sistemática a noção de devida diligência em direitos humanos. Tal ausência compromete a capacidade da norma de exigir das empresas um processo contínuo de identificação, mitigação e responsabilização pelos impactos negativos causados às crianças e adolescentes no ambiente digital. Diferente de um modelo centrado apenas na resposta a denúncias, a devida diligência pressupõe avaliações preventivas de risco, monitoramento regular

de impactos e engajamento significativo com os grupos afetados, dimensões ausentes do texto do PL.

Além disso, o PL não se alinha aos princípios contemporâneos de corresponsabilidade entre os diferentes atores da governança digital. A Estratégia Europeia BIK+ e os Princípios de Santa Clara (2020) enfatizam que a proteção da infância não deve recair exclusivamente sobre o usuário ou seu responsável legal, mas precisa ser partilhada entre Estado, empresas e sociedade civil, por intermédio de mecanismos estruturais e auditáveis de prevenção, resposta e prestação de contas.

A presente pesquisa está centrada no estudo da redação do art. 22 e a sua abordagem reativa, centrada na lógica de *notice and takedown*. Esse modelo é, em grande medida, insuficiente quando adotado de forma exclusiva, por não incorporar obrigações estruturais como avaliações de risco, medidas preventivas desde a concepção dos produtos (*safety by design*), sistemas automatizados de classificação etária e estruturas contínuas de governança interna. Além disso, essa lógica pode transferir o ônus da proteção às vítimas ou seus responsáveis, ao invés de impor obrigações preventivas claras às plataformas.

Os Princípios Orientadores sobre Empresas e Direitos Humanos da ONU, especialmente entre os 17 e 21, exigem que as empresas realizem avaliações contínuas de risco, monitorem o impacto de suas atividades e adotem medidas proativas de mitigação, sobretudo diante de grupos vulneráveis como crianças e adolescentes. A ausência dessa perspectiva impede que o PL exija das plataformas um compromisso sistemático e integrado com a prevenção de danos, limitando sua eficácia à resposta a violações já consumadas (ONU, 2011).

O projeto tampouco adota dispositivos estruturais que garantam que as plataformas assumam, desde o design de seus produtos, a responsabilidade pela proteção infanto-juvenil. Como exemplo, o DSA impõe obrigações concretas de design adequado à idade, transparência algorítmica e medidas técnicas como filtros, verificação de idade e canais de denúncia adaptados. O PL, ao não adotar o paradigma do *safety by design*, perpetua uma lógica de remediação que se mostra anacrônica diante dos riscos sistêmicos e do comportamento algorítmico das plataformas.

A Estratégia BIK+ evidencia que a proteção da infância em ambientes digitais exige medidas preventivas estruturais desde a concepção dos produtos e serviços,

especialmente no que diz respeito a riscos sistêmicos relacionados à usabilidade, design enganoso e exposição a conteúdos nocivos. O regulamento europeu exige que plataformas de grande porte adotem medidas apropriadas, como verificação de idade, design adaptado à faixa etária e interfaces acessíveis, de modo a impedir a exploração da inexperiência infantil e reduzir danos antes que ocorram (Comissão Europeia, 2022, p. 7-8)

A ausência de estrutura institucional também é refletida nas práticas de terceirização da moderação de conteúdo. Como denuncia Barrett (2020, p. 3), empresas como Facebook e Twitter (X) “marginalizam a função crítica de moderação, terceirizando-a para fornecedores terceirizados e afastando-a do núcleo decisório da empresa”, o que fragiliza a responsabilização e a qualidade das decisões. Essa precarização compromete a capacidade de resposta a riscos complexos como a exploração infantil e a disseminação de conteúdo nocivo a crianças. Nesse cenário, um aspecto não abordado no PL diz respeito à precarização da atividade de moderação de conteúdo.

Por sua vez, a ambiguidade do §1º, ao referir-se ao art. 6º como fonte dos conteúdos violadores, suscita dúvidas quanto à natureza taxativa ou exemplificativa do rol apresentado. Tal imprecisão pode implicar a exclusão de riscos digitais contemporâneos ainda não tipificados, como publicidade velada, coleta abusiva de dados sensíveis, manipulação algorítmica e conteúdos gamificados com práticas autolesivas.

Já o §2º, ao vedar denúncias anônimas, representa um entrave significativo à comunicação de violações, sobretudo em contextos de vulnerabilidade. Crianças e adolescentes, temendo exposição ou represálias, podem ser desencorajados a denunciar.

Nesse espaço, o §3º, ao exigir mecanismos públicos para denúncias, constitui avanço relevante, mas sua formulação genérica compromete sua acessibilidade efetiva. Para garantir que crianças e adolescentes possam utilizá-los adequadamente, é necessário desenvolver interfaces adaptadas à sua faixa etária, com linguagem simples, recursos audiovisuais, acessibilidade para pessoas com deficiência e possibilidade de mediação por adultos de confiança.

Além disso, o PL ignora a necessidade de envolvimento direto de crianças e adolescentes na formulação de políticas digitais que lhes digam respeito, conforme orienta o Comentário Geral nº 25 da ONU. Instrumentos como o *B-Tech UNICEF*

Briefing e os CRBP defendem a participação de crianças e adolescentes no desenho de produtos, sistemas de moderação e canais de denúncia. Sem esse engajamento, o risco é criar soluções normativas que, embora bem-intencionadas, se revelem ineficazes ou desconectadas das reais necessidades e vivências dos sujeitos que buscam proteger.

Outro ponto crítico é a ausência de exigências estruturais de governança e de prestação de contas. O projeto de lei não contempla a obrigatoriedade de relatórios de transparência, formação especializada de moderadores, avaliações de impacto nos direitos da criança (*Child Rights Impact Assessments* – CRIA), nem mecanismos de participação ativa de crianças na governança das plataformas.

O DSA incorpora tais diretrizes, reconhecendo a infância como grupo vulnerável e impondo às plataformas obrigações vinculantes quanto à avaliação de riscos, moderação algorítmica proporcional e publicação de relatórios periódicos. Parte-se do pressuposto de que os danos digitais não decorrem apenas de conteúdos explícitos, mas também da arquitetura das plataformas, dos algoritmos de recomendação e da ausência de mecanismos eficazes de mitigação (Comissão Europeia, 2022).

Ademais, a falta de definição legal clara sobre o conceito de “moderação de conteúdo” enfraquece a norma. A adoção de uma definição inspirada no art. 3º do DSA, que compreende moderação como o conjunto de ações, automatizadas ou não, voltadas à detecção, combate e remoção de conteúdos ilegais ou incompatíveis com os termos de uso, garantiria maior clareza sobre as obrigações das plataformas. Tal definição é crucial para evitar que práticas como despromoção algorítmica ou banimento silencioso (*shadow banning*) escapem ao escopo regulatório, mesmo quando afetam desproporcionalmente grupos vulneráveis (Comissão Europeia, 2022).

A inexistência de uma autoridade regulatória designada compromete a eficácia da norma. É fundamental prever um órgão com competência fiscalizatória, como a Autoridade Nacional de Proteção de Dados (ANPD), ou mesmo criar uma entidade especializada em governança digital da infância, dotada de capacidade para monitoramento, auditoria e sanção.

Além disso, a proteção digital da infância deve ser sustentada por uma estrutura regulatória clara, contínua e auditável. O DSA, nesse sentido, estabelece obrigações explícitas para que as plataformas publiquem relatórios periódicos,

avaliem riscos sistêmicos e implementem práticas de governança orientadas à proteção de direitos fundamentais. Tais exigências superam abordagens reativas ao promover uma cultura de responsabilidade institucional permanente, com foco na transparência, na prevenção e na corresponsabilidade entre empresas e autoridades públicas (Comissão Europeia, 2022, p. 6-9)

Torna-se imperativo situar a análise do PL no marco mais amplo da responsabilidade das empresas de tecnologia. Ao se estabelecerem em determinado Estado, as empresas transnacionais devem respeitar tanto a legislação interna (incluindo a Constituição) quanto os tratados internacionais, sobretudo os de direitos humanos dos quais o país seja parte. No caso brasileiro, isso inclui os direitos fundamentais constitucionais e os compromissos internacionais assumidos, cuja violação legitima a responsabilização jurídica das empresas, com direito à reparação (Olsen; Pamplona, 2019, p. 133-134).

Contudo, como alerta Richard-Carvajal (2024, p. 1), o avanço da regulação empresarial requer mais do que legislação: é necessário que os governos desenvolvam padrões substantivos de desempenho e sistemas de monitoramento efetivos. Isso implica criar indicadores específicos por setor, capacitar órgãos nacionais, incentivar a transparência e aplicar sanções proporcionais à conduta empresarial, medidas ausentes no PL em análise.

Além disso, a ausência de obrigações jurídicas internacionais vinculantes contribui para a perpetuação de um cenário de impunidade estrutural. Nesse sentido, ganha relevo a noção de responsabilidade extraterritorial dos Estados de origem das empresas, os quais devem adotar medidas que impeçam abusos em qualquer território sob sua influência (Kaleck; Saage-Maaß, 2008, p. 15-20).

As grandes plataformas digitais concentram poder normativo e comunicacional comparável ao de atores estatais, operando como agentes moduladores da esfera pública digital (Freitas; Sousa, 2022, p. 883). Sua responsabilidade deve, portanto, ser compreendida a partir de uma perspectiva ética, normativa e estrutural, em consonância com instrumentos como o *B-Tech UNICEF Briefing* (2024) e os CRBP (2012), que defendem que governos imponham padrões robustos, métricas claras e instituições reguladoras capacitadas.

Portanto, embora o art. 22 do PL represente um avanço normativo relevante, sua formulação atual revela-se insuficiente frente à complexidade do ecossistema digital. Para que a norma atenda aos princípios da proteção integral, da prevenção e

da participação, é essencial que o legislador adote uma abordagem holística, baseada em evidências, e articulada aos mais elevados padrões internacionais de responsabilidade das empresas e proteção dos direitos da criança.

4 Responsabilidade das empresas e devida diligência em direitos humanos

Como verificado no exemplo do Discord, a atuação transnacional das corporações tecnológicas pode impactar diretamente grupos vulneráveis, como as crianças e adolescentes. Torna-se, portanto, imprescindível analisar os deveres dessas empresas à luz tanto do ordenamento jurídico interno quanto dos compromissos internacionais assumidos pelos Estados (Silveira; Almeida, 2015, p. 359).

Ao se estabelecerem em um Estado, empresas devem respeitar a legislação nacional e os tratados internacionais de direitos humanos ratificados pelo país. Havendo violações, é possível buscar a responsabilização da empresa perante o sistema jurídico local, assegurando-se o direito à reparação. Tal possibilidade, no entanto, suscita questionamentos sobre os limites e a efetividade dessa responsabilização (Olsen; Pamplona, 2019, p. 133–134). Nesse cenário, a devida diligência em direitos humanos representa um dos pilares centrais da responsabilidade empresarial, conforme os Princípios Orientadores sobre Empresas e Direitos Humanos, especialmente os princípios 17 a 21.

A devida diligência em direitos humanos diz respeito a um processo que exige das empresas a identificação, prevenção, mitigação e prestação de contas em relação aos impactos que suas atividades possam causar, contribuir ou estar diretamente relacionadas, ainda que de forma indireta. Não se trata de prática voluntária, mas de uma exigência normativa internacional que afirma o dever universal de todas as empresas, independentemente de porte ou setor, de respeitar os direitos humanos (ONU, 2011, p. 17).

Em tais termos, o princípio 17 estabelece que a devida diligência deve ser integrada às operações e cadeias de valor da empresa. O princípio 18 orienta a avaliação de riscos reais e potenciais, especialmente para grupos vulneráveis como crianças, povos indígenas e minorias étnicas. O princípio 19 trata da integração dos resultados dessas avaliações nos processos internos e da adoção de medidas adequadas à gravidade dos impactos (ONU, 2011, p. 19-21).

Quando os danos são causados ou contribuídos diretamente pela empresa, ela deve cessar a violação e colaborar com sua remediação; se o dano estiver apenas relacionado a suas atividades comerciais, deve empregar sua influência para mitigá-lo (ONU, 2011, p. 19-21). Os princípios 20 e 21, por sua vez, ressaltam a importância do monitoramento contínuo e da prestação de contas públicas, equilibrando transparência e proteção de dados sensíveis (ONU, 2011, p. 22-24).

A diretriz da OCDE sobre devida diligência reforça esse entendimento, enfatizando que se trata de um processo contínuo, adaptável e proporcional ao risco. Essa abordagem abrange temas como direitos humanos, condições de trabalho, meio ambiente e relações de consumo, e exige integração dos direitos humanos às políticas corporativas, às cadeias produtivas e às decisões estratégicas (OECD, 2018, p. 16-17).

Um dos aspectos mais inovadores dessas diretrizes é a responsabilização por toda a cadeia de valor, o que implica que as obrigações empresariais devem se estender a fornecedores, licenciados, parceiros comerciais e investidores. Espera-se que as empresas exerçam influência sobre esses terceiros e adotem mecanismos de monitoramento, correção e, quando necessário, desengajamento responsável (OECD, 2018, p. 29-31).

Adicionalmente, destaca-se a exigência de engajamento significativo com as partes interessadas, especialmente com grupos potencialmente afetados, como crianças e adolescentes. Esse processo deve incluir consultas transparentes, comunicação acessível e prestação de contas contínua, respeitando o princípio da proporcionalidade e a sensibilidade cultural. Esse diálogo é crucial para mitigar riscos e consolidar a legitimidade social das empresas diante da sociedade e das autoridades reguladoras (OECD, 2018, p. 18-19).

No contexto brasileiro, esse cenário é agravado pelo limitado grau de letramento digital das crianças e pela opacidade dos modelos de negócio das plataformas digitais. Estudos indicam que crianças brasileiras enfrentam dificuldades em reconhecer conteúdos publicitários disfarçados, em compreender o uso comercial de seus dados e em lidar com estratégias algorítmicas de modulação de comportamento (CGI.br, 2023).

Apesar da crescente atuação transnacional das empresas, o direito internacional ainda carece de mecanismos eficazes e vinculantes para responsabilizá-las por violações de direitos humanos. Há um claro descompasso

entre o impacto real das corporações, especialmente no Sul Global, e a efetividade dos instrumentos jurídicos disponíveis para regulá-las. A ausência de obrigações jurídicas internacionais vinculantes contribui para a manutenção de um cenário de impunidade estrutural (Kaleck; Saage-Maaß, 2008, p. 15).

Diante disso, reforça-se a importância da responsabilidade extraterritorial dos Estados, sobretudo os de origem das corporações, que devem adotar medidas legislativas e administrativas para prevenir abusos tanto em seu território quanto no exterior (Kaleck; Saage-Maaß, 2008, p. 20). As redes sociais exercem enorme influência tanto no ciberespaço quanto na vida cotidiana, constituindo-se em canais centrais de debate público. Nesse cenário, plataformas acumulam poder normativo e comunicacional semelhante ao de atores estatais, tornando-se verdadeiros agentes reguladores da esfera digital (Freitas; Sousa, 2022, p. 883).

Com a comercialização de conteúdos inapropriados e/ou criminosos, somada à opacidade nos sistemas de recomendação e à ausência de mecanismos eficazes de controle parental, compromete-se o desenvolvimento saudável de crianças e adolescentes. Em resposta, o DSA estabeleceu restrições específicas à publicidade dirigida a menores e impôs obrigações de transparência algorítmica às grandes plataformas digitais, evidenciando a exigência de uma atuação empresarial proativa e responsável (Comissão Europeia, 2022, p. 8-9).

Isso inclui a exigência de termos de uso acessíveis, mecanismos de denúncia adaptados, verificação de idade, filtros de conteúdo e alertas preventivos. A Comissão Europeia destaca que essa responsabilidade deve ser compartilhada entre empresas, Estado e sociedade civil, por meio de sistemas de co-regulação que garantam supervisão pública e auditoria independente (Comissão Europeia, 2022, p. 11-13).

A proteção da infância deve ainda considerar riscos transversais identificados pelo Comentário Geral nº 25 do Comitê dos Direitos da Criança da ONU, como discriminação algorítmica, vigilância abusiva e manipulação de dados pessoais. O documento enfatiza que os Estados devem assegurar que os serviços digitais oferecidos pelas empresas sejam adaptados à idade e ao estágio de desenvolvimento das crianças, promovendo salvaguardas desde a concepção dos produtos (ONU, 2021).

O *B-Tech UNICEF Briefing* representa uma contribuição paradigmática ao reconhecer que, embora as tecnologias ampliem o acesso à informação, educação e participação cívica, também impõem riscos relevantes — como exposição a

conteúdos nocivos, violação de privacidade, exploração comercial e discriminação algorítmica (UNICEF, 2024). Para mitigá-los, o documento propõe a implementação de avaliações de impacto nos direitos da criança (CRIA), adoção de medidas preventivas desde o início do desenvolvimento tecnológico e criação de canais eficazes de participação infantojuvenil nas decisões que afetam suas vivências digitais.

Complementarmente, os CRBP estruturam-se em dois pilares: (1) o respeito obrigatório aos direitos da criança, com ênfase na prevenção e mitigação de impactos negativos, e (2) o apoio voluntário por meio de ações positivas, como filantropia, *advocacy* e parcerias estratégicas (UNICEF; UN Global Compact; Save the Children, 2012). Seus dez princípios abrangem temas como erradicação do trabalho infantil, segurança de produtos e serviços, marketing responsável, impactos ambientais e atuação empresarial em contextos de crise.

Essa abordagem amplia o entendimento de responsabilidade empresarial, ao reconhecer as crianças como sujeitos de direitos e integrar sua proteção às estratégias corporativas de forma contínua, estruturada e multissetorial. A convergência entre o *B-Tech UNICEF Briefing* e os CRBP sinaliza o surgimento de um novo paradigma de governança corporativa, orientado pela prevenção, transparência e participação. Ambos defendem uma atuação multissetorial, que envolve Estado, setor privado, sociedade civil e as próprias crianças, para construir um ecossistema digital que maximize os benefícios das tecnologias e reduza seus riscos.

Essa perspectiva se articula a normas emergentes como o DSA, aprovado pela União Europeia, que estabelece obrigações vinculantes às grandes plataformas digitais, em consonância com os princípios internacionais de proteção à infância. O DSA introduz avanços significativos ao reconhecer as crianças e adolescentes como grupos particularmente vulneráveis no ambiente digital. Seus artigos 45, 49 e 66 impõem às plataformas deveres como divulgar critérios de moderação (automatizados e humanos), publicar relatórios anuais e criar bases de dados acessíveis à sociedade civil e às autoridades reguladoras (União Europeia, 2022).

Ao institucionalizar práticas antes restritas à autorregulação, como avaliação de impacto e transparência algorítmica, o DSA estabelece um regime híbrido de regulação, combinando dispositivos legais e princípios éticos voltados ao bem-estar infantil. Empresas alinhadas a esse modelo mitigam riscos jurídicos e reputacionais,

fortalecem sua legitimidade social, conquistam a confiança dos stakeholders e contribuem para os Objetivos de Desenvolvimento Sustentável (ODS), especialmente os de número 4 (educação de qualidade), 10 (redução das desigualdades) e 16 (instituições eficazes e inclusivas).

Nesse cenário, a moderação de conteúdo se destaca como expressão concreta da responsabilidade empresarial. Mais do que uma função técnica, ela deve ser concebida como parte de uma estratégia corporativa orientada pela ética dos direitos humanos e pelo princípio do melhor interesse da criança (Barrett, 2020, p. 2).

Assim, observa-se que a responsabilidade das empresas na promoção e proteção dos direitos humanos, especialmente de crianças e adolescentes, exige uma atuação proativa, transparente e orientada por normas internacionais. A devida diligência não deveria ser opcional, mas um dever normativo que impõe às corporações o compromisso de identificar, prevenir e mitigar riscos em toda a sua cadeia de valor.

Em um cenário de crescente impacto digital, instrumentos como os POs, as diretrizes da OCDE e o DSA da União Europeia reforçam a urgência de uma governança multissetorial que garanta proteção efetiva à infância, fortalecendo a legitimidade social e ética das empresas.

Considerações finais

A pesquisa desenvolvida permitiu confirmar a hipótese de que o Projeto de Lei nº 2.628/2022, em sua formulação originária, não garante mecanismos suficientes para a proteção integral de crianças e adolescentes no ambiente digital.

Embora constitua um avanço ao reconhecer obrigações extrajudiciais para as plataformas, o texto permanece ancorado em uma lógica predominantemente reativa, sem incorporar instrumentos preventivos, participativos e de devida diligência em direitos humanos.

No primeiro eixo, observou-se que a experiência da plataforma Discord exemplifica de modo contundente os riscos enfrentados por crianças e adolescentes diante da ausência de moderação eficaz. A circulação de conteúdos ilícitos e práticas criminosas revela que a moderação de conteúdo transcende a função técnica, constituindo-se em verdadeira ferramenta de governança digital, essencial para

equilibrar liberdade de expressão, proteção de grupos vulneráveis e responsabilidade social das plataformas.

O segundo eixo destacou que, apesar de propor dispositivos relevantes, como mecanismos de denúncia e obrigação de remoção de conteúdos nocivos, o PL 2.628/2022 apresenta graves fragilidades. A ausência de conceitos claros, a vedação a denúncias anônimas, a inexistência de uma autoridade regulatória especializada e a falta de mecanismos estruturais de prevenção e transparência comprometem sua efetividade.

Assim, ao adotar uma lógica de *notice and takedown*, o projeto transfere parte do ônus de proteção às vítimas, deixando de responsabilizar as empresas por medidas preventivas como *safety by design*, classificação etária e governança contínua.

Por sua vez, o terceiro eixo reforçou que a responsabilidade empresarial deve se fundamentar na devida diligência em direitos humanos, entendida como processo permanente de identificação, mitigação e reparação de impactos negativos. Normas internacionais, como os POs, as diretrizes da OCDE, o *B-Tech UNICEF Briefing* e os CRBP, apontam para a necessidade de uma governança multissetorial que integre Estado, empresas, sociedade civil e, sobretudo, a participação ativa das próprias crianças.

Apenas por meio dessa corresponsabilidade é possível enfrentar riscos transversais, como exploração sexual, manipulação algorítmica e discriminação digital. Diante do exposto, conclui-se que a proteção da infância no espaço digital exige um modelo regulatório híbrido, preventivo e participativo, capaz de articular legislações nacionais às melhores práticas internacionais.

A problematização inicial revelou-se procedente: o PL, em sua forma atual, é insuficiente. A hipótese foi confirmada: somente a integração entre Estado, setor privado, sociedade civil e crianças pode construir um ecossistema digital inclusivo, seguro e alinhado ao interesse superior da criança.

Referências

ALMEIDA, Clara Leitão de. **Regulação da transparência em plataformas digitais e legitimidade na moderação de conteúdo**. 2022. Dissertação (Mestrado em Direito

da Regulação) – Fundação Getulio Vargas, Escola de Direito do Rio de Janeiro, Rio de Janeiro, 2022.

BARRETT, Paul. **Who moderates the social media giants?**. New York: NYU Stern Center for Business and Human Rights, 2020. Disponível em: https://issuu.com/nyusterncenterforbusinessandhumanri/docs/nyu_content_moderation_report_final_version. Acesso em: 10 jan. 2024.

BICO, Júlia da Rosa; DA SILVA, Rosane Leal. A violação de direitos fundamentais de crianças e adolescentes brasileiros no Discord: objetificação de corpos e abuso sexual. **Disciplinarum Scientia | Sociais Aplicadas**, Santa Maria (RS, Brasil), v. 21, n. 2, p. 1–16, 2025.

BRASIL. Senado Federal. **Projeto de Lei nº 2.628, de 2022**. Dispõe sobre a proteção de crianças e adolescentes em ambientes digitais. Brasília, DF: Senado Federal, 2022.

CGI.BR – COMITÊ GESTOR DA INTERNET NO BRASIL. **TIC Kids Online Brasil 2022**: pesquisa sobre o uso da Internet por crianças e adolescentes no Brasil. São Paulo: CGI.br, 2023. Disponível em: <https://cetic.br>. Acesso em: 10 jan. 2024.

COMISSÃO EUROPEIA. **Uma Década Digital para as Crianças e os Jovens**: a nova Estratégia Europeia para uma Internet Melhor para as Crianças (BIK+). Bruxelas, 11 maio 2022. Disponível em: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2022:212:FIN>. Acesso em: 20 jan. 2024.

COMMON, David. Kafka's moderation: content governance and the rule of law. **Journal of Free Speech Law**, v. 1, n. 1, p. 1–23, 2020.

DISCORD. Termos de Serviço do Discord. Disponível em: <https://discord.com/terms>. Acesso em: 21 maio 2024

FREITAS, Cinthia Obladen de Almendra; SOUSA, Devilson da Rocha. O contexto dos direitos humanos no ciberespaço e o papel das empresas de tecnologia a partir de

uma análise da ação das redes sociais. **Revista Jurídica Luso-Brasileira**, v. 8, n. 4, p. 871–894, 2022. Disponível em: https://www.cidp.pt/revistas/rjlb/2022/4/2022_04_0871_0894.pdf. Acesso em: 14 jun. 2025.

GILLESPIE, Tarleton. Content moderation, AI, and the question of scale. **Big Data & Society**, v. 7, n. 2, p. 1–5, 2020.

GRIMMELMANN, James. The virtues of moderation. **Yale Journal of Law & Technology**, v. 17, 2015.

HIDALGO, Rodrigo; SANT'ANA, Leandro. Tortura e crimes sexuais no Discord contra adolescentes: mercado criminoso por trás de transmissões ao vivo. *Jornal da Band*, São Paulo, 8 ago. 2025.

KALECK, Wolfgang; SAAGE-MAAß, Miriam. Corporate accountability for human rights violations amounting to international crimes: the status quo and its challenges. **Journal of International Criminal Justice**, v. 8, p. 699–724, 2010.

KALECK, Wolfgang; SAAGE-MAAß, Miriam. Transnational corporations on trial: on the threat to human rights posed by European companies in Latin America. **Publication Series on Democracy**, v. 4. Berlin: Heinrich-Böll-Stiftung, 2008.

KLONICK, Kate. The new governors: the people, rules, and processes governing online speech. **Harvard Law Review**, v. 131, 2017.

LIDLAW, Emily. **Internet gatekeepers, human rights and corporate social responsibilities**. 2012. Tese (Doutorado em Direito) – London School of Economics and Political Science, Londres, 2012.

LANGVARDT, Kyle. Regulating online content moderation. **Georgetown Law Journal**, v. 106, n. 5, p. 1353–1388, 2018.

OCDE. **Children in the digital environment**: revised typology of risks. Paris: OECD Publishing, 2021. Disponível em: <https://doi.org/10.1787/9b8f222e-en>. Acesso em: 7 jan. 2024.

OECD – ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. **OECD Due Diligence Guidance for Responsible Business Conduct**. Paris: OECD Publishing, 2018. Disponível em: <https://www.oecd.org/investment/due-diligence-guidance-for-responsible-business-conduct.htm>. Acesso em: 15 jan. 2024.

OLSEN, Ana Carolina Lopes; PAMPLONA, Danielle Anne. Violações a direitos humanos por empresas transnacionais na América Latina: perspectivas de responsabilização. **Revista Direitos Humanos e Democracia**, v. 7, n. 13, p. 129–151, 2019. DOI: 10.21527/2317-5389.2019.13.129-151.

ONU – ORGANIZAÇÃO DAS NAÇÕES UNIDAS. Objetivos de Desenvolvimento Sustentável. [S.l.]: ONU, 2015.

ONU – COMITÊ DOS DIREITOS DA CRIANÇA. **Comentário Geral nº 25 sobre os direitos das crianças em relação ao ambiente digital**. Genebra: ONU, 2021. Disponível em: <https://digitallibrary.un.org/record/3931979>. Acesso em: 10 jan. 2024.

ONU – ORGANIZAÇÃO DAS NAÇÕES UNIDAS. **Princípios Orientadores sobre Empresas e Direitos Humanos**: implementação do marco de “proteger, respeitar e reparar” das Nações Unidas. Genebra: ONU, 2011.

PAZ, Aline Amaral. **Explicação na internet**: violências digitais contra mulheres. 2023. Tese (Doutorado em Comunicação) - Universidade Federal de Santa Maria, Rio Grande do Sul, 2023.

REINO UNIDO. **Online Safety Act 2023**. Londres, 2023.

RICHARD-CARVAJAL, Fernando. La rendición de cuentas empresarial en contextos de impunidad: desafíos regulatorios para América Latina. **Revista Latinoamericana de Derechos Humanos**, v. 25, n. 1, p. 1–22, 2024.

ROBERTS, Sarah T. **Behind the screen: the hidden digital labor of commercial content moderation**. 2014. Tese (Doutorado) – University of Illinois, Chicago, IL, 2014.

ROBERTS, Sarah T. Digital refuse: Canadian garbage, commercial content moderation and the global circulation of social media's waste. **Wi: Journal of Mobile Media**, 2016.

SANDER, Barrie. Freedom of expression in the age of online platforms: the promise and pitfalls of a human rights-based approach to content moderation. **Fordham International Law Journal**, v. 43, 2020.

SANTANA, Anna Luisa Walter de. **El poder privado en Internet: retos y estrategias para la protección de la libertad de expresión en la era digital**. Bogotá: Tirant lo Blanch, 2023.

SMART, Paul; McMANUS, Jake. Moderação de conteúdo e responsabilidade das plataformas digitais: desafios normativos e caminhos regulatórios. **Revista de Direito Digital e Cidadania**, v. 4, n. 2, p. 267–289, 2022.

SMART, Paul; CODDOU, Francisco. Moderação de conteúdo e responsabilidade das plataformas digitais: entre autorregulação e regulação pública. **Revista de Direito Digital e Cidadania**, v. 4, n. 1, p. 255–276, 2022.

TAVARES, VITOR. Por que o Discord tem sido terreno fértil na internet para crimes de extremismo. BBC News Brasil, São Paulo, 9 mai. 2025.

THE SANTA CLARA PRINCIPLES. **The Santa Clara Principles on Transparency and Accountability in Content Moderation**. Santa Clara, 2020.

UNIÃO EUROPEIA. Regulamento (UE) 2022/2065 do Parlamento Europeu e do Conselho, de 19 de outubro de 2022, relativo a um mercado único para serviços digitais e que altera a Diretiva 2000/31/CE (**Digital Services Act**). Jornal Oficial da União Europeia, L 277, p. 1–102, 27 out. 2022.

UNICEF; ALTO COMISSARIADO DAS NAÇÕES UNIDAS PARA OS DIREITOS HUMANOS. **Taking a child rights-based approach to implementing the UNGPs in the digital environment:** a contribution to the B-Tech Project. [S.l.]: UNICEF; OHCHR, 2024.

UNICEF; UN GLOBAL COMPACT; SAVE THE CHILDREN. **Children's Rights and Business Principles.** [S.l.]: UNICEF; UN Global Compact; Save the Children, 2012.

WEST, Sarah Myers. Censored, suspended, shadowbanned: user interpretations of content moderation on social media platforms. **New Media & Society**, v. 20, n. 11, p. 4366–4383, 2018.