Contribuições para uma regulação democrática das grandes plataformas que garanta a liberdade de expressão na internet

Uma perspectiva latino-americana para alcançar processos de moderação de conteúdo compatíveis com os padrões internacionais de direitos humanos

Elaboração

intervozes coletivo brasil de comunicação social

OBSERVACOM Observatorio Latinoamericano de Regulación Medios y Convergencia

desarrollo digital

idec Instituto Brasileiro de Defesa do Consumidor

Apoios / Adesões

TEDIC

PROLEDI

USUARIOS DIGITALES

CAinfo

O documento contém medidas elaboradas por várias organizações, ativistas e especialistas durante vários meses, esperando que seja uma boa contribuição para o debate sobre o tema. A proposta trata de dar respostas integrais e o mais concretas possível, mas segue 'em construção', e por essa razão a colocamos em consulta pública. Uma solução em um assunto tão polêmico e sensível como este necessita de processos mais amplos e abertos com a participação das múltiplas partes interessadas, para que possa se converter em um marco de referência democrático que acreditamos necessário ter neste momento.

ACTUALIZACIÓN

La versión final del producto del proceso de consulta pública iniciado con este documento puede encontrarse acá:

Estándares regulación grandes plataformas

En el documento final nuevas organizaciones han sumado su apoyo, y ya no está firmado por Valeria Betancourt

Colaboraram na redação da proposta (a título pessoal):

- **Javier Pallero**
  Coordenador de Políticas Públicas para América Latina, Access Now
- **Joan Barata**
  Membro da Plataforma de Defesa da Liberdade de Informação
- **Valeria Betancourt**
  Associação para o Progresso das Comunicações (APC)
- **Andrés Piazza**
  Consultor, ExLACNIC, ExLACTLD
- **Guillermo Mastrini**
  Professor pesquisador na Universidade Nacional de Quilmes (UnQ), Universidade de Buenos Aires (UBA), Conicet
- **Martín Becerra**
  Universidade Nacional de Quilmes (UnQ), Universidade de Buenos Aires (UBA), Conicet
- **Juan Ortiz Freuler**
  Pesquisador associado ao Berkman Klein Center (2019-2020)

RESUMO

Este documento inclui recomendações sobre princípios, padrões e medidas específicas para estabelecer formas de co-regulação e de regulação pública que limitem o poder das grandes plataformas de Internet (tais como redes sociais e mecanismos de busca) para proteger a liberdade de expressão de seus usuários e garantir uma Internet livre e aberta.

Isso se justifica dada a crescente intervenção desses intermediários no conteúdo on-line, através da adoção de termos de serviço e aplicação de políticas empresariais de moderação que se tornaram formas de regulação privada de espaços públicos que hoje são vitais para a deliberação democrática e o exercício dos direitos fundamentais.

A proposta procura ser compatível com as normas internacionais de direitos humanos e levar em conta as assimetrias existentes, alcançando grandes plataformas de Internet sem ser um freio ou obstáculo à inovação, competição ou desenvolvimento de *start-ups*, pequenas empresas ou comunidade, iniciativas educacionais ou sem fins lucrativos.

INTRODUÇÃO

A crescente intervenção de plataformas de Internet no conteúdo de seus usuários se tornou um tópico de preocupação em todo o mundo. De fato, o "*controle privado*" tem sido considerado um dos três principais desafios para a próxima década e uma "*ameaça à liberdade de expressão*", nada menos do que pelos Relatores da Liberdade de Expressão[1]. Para eles, "*uma característica transformadora do ambiente de comunicação digital é o poder das empresas privadas e, particularmente, das redes sociais, plataformas de busca e outros intermediários, em comunicações, com enorme poder concentrado em poucas empresas*".

Essa preocupação não é nova, já que em muitas ocasiões tanto organizações internacionais quanto organizações de direitos digitais questionaram essas práticas e formularam recomendações para que as empresas fizessem uma mudança nas políticas e práticas em conformidade com os padrões internacionais de direitos humanos[2].

Por sua vez, a Relatoria de Liberdade de Opinião e Expressão das Nações Unidas publicou diversos relatórios sobre o assunto[3] e a Relatoria para a Liberdade de Expressão da Comissão Interamericana de Direitos Humanos (CIDH) observou durante anos que "*intermediários devem evitar que suas atividades causem ou contribuam para consequências negativas sobre o direito à liberdade de expressão*" em suas ações voluntárias de moderação de conteúdo, que "*só devem ser consideradas legítimas quando tais restrições não impeçam ou obstaculizem arbitrariamente as possibilidades de expressão de uma pessoa na Internet*"[4].

Também é crescente o interesse de governos e parlamentos - em ambientes autoritários, mas também em democracias consolidadas - por regular sua atividade e a distribuição de conteúdo através de diferentes tipos de regulamentação, particularmente regulando o conteúdo disseminado nas redes sociais. Mas a maioria dessas iniciativas legais configura soluções que são ilegítimas ou desproporcionais, atribuindo responsabilidades e obrigações que tornam as plataformas juízes ou polícias privadas sobre os conteúdos de terceiros que podem circular na Internet.

Os signatários deste documento se opusemos a essas propostas e continuaremos a fazê-lo. Mas acreditamos que o modelo de autorregulação que tem prevalecido até agora foi exaurido no atual desenvolvimento da Internet, onde algumas poucas corporações centralizaram e concentraram a circulação, troca ou busca de informações e opiniões, distorcendo a ideia de uma Internet descentralizada, livre e aberta pela qual lutamos.

[1] Declaração Conjunta: Desafios para a liberdade de expressão na próxima década do Relator Especial das Nações Unidas (ONU) para a Liberdade de Opinião e Expressão, o Representante da Organização para a Segurança e Cooperação na Europa (OSCE) para o Liberdade dos meios de comunicação, o relator especial da Organização dos Estados Americanos (OEA) para a Liberdade de Expressão e o Relator Especial sobre Liberdade de Expressão e Acesso à Informação da Comissão Africana dos Direitos Humanos e dos Povos (CADHP), 2019
[2] Entre elas, os Princípios de Santa Clara
[3] Regulamento do conteúdo na Internet, Relatoria Especial sobre a Promoção e Proteção do Direito à Liberdade de Opinião e Expressão, 2018
[4] Liberdade de Expressão e Internet, Relatoria Especial para a Liberdade de Expressão CIDH, 2013, par. 111

Diante desse cenário polarizado de "autorregulação corporativa versus regulação autoritárias ou excessiva", várias organizações latino-americanas acreditamos que é necessário e possível construir um terceiro caminho. O de construir uma proposta de regulação democrática, adequada e inteligente, capaz de assegurar ambientes normativos adequados para proteger os direitos humanos das ações dos gigantes tecnológicos, respeitando os padrões internacionais de direitos humanos.

O poder de *gatekeepers* que essas empresas de Internet têm exige que as sociedades democráticas sejam capazes de estabelecer limitações a seus poderes para garantir a efetividade dos direitos e liberdades historicamente reconhecidos, bem como a predominância do interesse geral e público.

A proposta não pretende atingir todos os intermediários da Internet, mas sim determinados tipos de plataformas e aplicativos cujo principal serviço é permitir ou facilitar o acesso à informação disponível na Internet e/ou ser suporte para expressão, comunicação e troca de conteúdo entre seus usuários (Isso inclui redes sociais, mecanismos de pesquisa e plataformas de compartilhamento de vídeo, por exemplo, embora não para serviços de mensagens).

Um princípio de "regulação progressiva" é proposto com base no impacto que as medidas tomadas pelos intermediários têm no exercício dos direitos fundamentais na Internet, especialmente a liberdade de expressão. Ou seja, o regulamento deve ser mais rigoroso no caso de grandes plataformas que, devido ao seu porte e abrangência, tenham se tornado espaços públicos de deliberação e/ou principais vias de acesso para o acesso à informação, com um nível de concentração excessivo.

Tendo em vista as características especiais da internet, busca-se criar um ambiente normativo adequado ao funcionamento e às características da rede, que inclua mecanismos de autorregulação, co-regulação e regulação pública, com o entendimento de que os desafios apresentados pelo novo cenário digital (dentre outros, a velocidade e volume de informações) não permitem a aplicação de soluções únicas e iguais do que em outros sistemas de informação e comunicação.

O documento não propõe uma legislação que determine qual conteúdo pode ser divulgado na Internet e qual não, nem obriga as plataformas a moderar o conteúdo. Mas, se decidirem fazê-lo, uma série de condições será estabelecida para que os direitos fundamentais de seus usuários não sejam violados. Para isso, são incluídas propostas sobre quais são os limites para a moderação dos conteúdos que estas plataformas já realizam, para que seus termos de serviços, seus critérios e seus procedimentos sejam compatíveis com os padrões internacionais de direitos humanos, levando especialmente em consideração a proteção de minorias e grupos vulneráveis.

Uma regulamentação democrática e equilibrada também deve proteger as plataformas das pressões ilegítimas de governos e outros atores. Devido ao seu papel como intermediários, elas são fundamentais para facilitar o exercício desses direitos e, portanto, a proposta inclui recomendações para que os marcos regulatórios lhes permitam desempenhar esse papel de

maneira apropriada: a não responsabilidade legal pelo conteúdo de terceiros ou a proibição de obrigá-los a acompanhamento genérico ou supervisão de conteúdo são algumas delas.

A regulamentação privada na Internet é produzida e agravada por um contexto de forte concentração de poder em poucas corporações internacionais. A regulação pública sobre as atividades das plataformas deveria adotar medidas antimonopólicas para enfrentar o cenário de concentração e ausência de concorrência, mas elas não estão incluídas nesta oportunidade. A simples ideia de que os principais espaços públicos para a circulação de informações e opiniões são todos controlados pela mesma empresa deveria forçar os órgãos antitruste dos Estados Unidos e de outros países a agir.

Tampouco se incluem nesta proposta questões importantes como mecanismos para garantir o pluralismo e a diversidade na Internet ou o que diz respeito às questões tributárias. O documento se concentra em questões relacionadas à moderação de conteúdo, com princípios de aplicação geral. As especificidades de determinados serviços justificam também abordagens específicas. Por exemplo: para serviços de intermediação e publicação de bens culturais, poderiam ser adotadas obrigações positivas para a proteção e promoção da diversidade cultural, em conformidade com a Convenção sobre a Proteção e Promoção da Diversidade das Expressões Culturais da UNESCO.

Finalmente, as normas e desenhos institucionais adotados devem ser adequadamente desenvolvidos, levando em conta as necessidades da regulação do mercado em contínuo desenvolvimento, as características particulares do ambiente digital em cada país e as necessidades específicas da América Latina, no marco dos padrões internacionais de direitos humanos.

O documento está organizado nos seguintes capítulos:

1. Alcance e caráter da regulação
2. Termos e condições de serviço
3. Transparência
4. Aplicação de políticas e devido processo
5. Direito à defesa e reparação
6. Prestação de contas
7. Aprovação e aplicação da regulação

1. ALCANCE E CARÁTER DA REGULAÇÃO

1.1 Esta proposta de regulação atinge os provedores de serviços online quando atuam como intermediários ou plataformas de armazenamento, busca ou troca de informações, opiniões, expressões e outros conteúdos gerados por seus usuários e que realizam algum tipo de curadoria ou moderação desses conteúdos (neste documento: "plataformas de conteúdo"). Entre eles: mecanismos de busca, redes sociais e outras plataformas para troca de textos, imagens e vídeos.

1.2 Os limites ao poder das grandes plataformas de conteúdos devem ser estruturados com base num modelo de co-regulação, onde as estruturas de auto-regulação e regulação pública são complementadas para formular soluções legais, contratuais e técnicas que garantam a liberdade de expressão online, em equilíbrio com outros direitos fundamentais[5]. Os instrumentos regulatórios e regulatórios devem ser o resultado de um processo de governança multissetorial que leve em consideração os contextos locais e regionais.

1.3 As plataformas devem incorporar diretamente em suas condições de serviço e seus padrões de comunidade os princípios pertinentes em matéria de direitos humanos que garantam que as medidas relacionadas ao conteúdo serão guiadas pelos mesmos critérios que governam a proteção da expressão por qualquer meio[6]. Entre esses princípios se incluem: transparência, prestação de contas, devido processo legal, necessidade, proporcionalidade, não discriminação e direito à defesa e reparação. As plataformas devem também garantir o respeito total pelos direitos do consumidor.

1.4 As plataformas de conteúdo que possuem poder de mercado significativo em um determinado âmbito de atuação[7] ("grandes plataformas de conteúdo") devem ter regulamentação assimétrica em relação a outros provedores, tendo em vista a importância e o impacto que suas decisões empresariais podem ter sobre o intercâmbio de informações, opiniões e bens culturais, bem como o exercício da liberdade de expressão e do debate público.

1.5 Uma regulação inteligente é aquela que considera de maneira adequada e diferenciada plataformas de conteúdo que não atendem aos critérios acima e não têm fins lucrativos, têm fins científicos ou educacionais, bem como aquelas que reúnem grupos de usuários pequenos e fechados de caráter privado e homogêneo.

[5] O desenho institucional e a divisão de responsabilidades são desenvolvidos no capítulo 7 desta proposta
[6] Regulação do conteúdo online gerado por usuários, Relatoria Especial sobre a Promoção do Direito ao Liberdade de Opinião e Expressão das Nações Unidas, 2018
[7] A definição de poder de mercado significativo, feita por um órgão regulador independente, deve levar em conta as definições de mercados relevantes atualizadas para o cenário atual, considerando a função específica do serviço, a disputa por receita de publicidade e o tempo de atenção do usuário, o número absoluto de usuários e a substituibilidade de um serviço por outros.

2. TERMOS E CONDIÇÕES DE SERVIÇO

2.1 Os termos de serviço (TOS) de todas as plataformas de conteúdo, bem como outros documentos complementares (como guias ou diretrizes de aplicação de conteúdo) devem ser escritos de forma clara, precisa, inteligível e acessível para todos os usuários. Plataformas grandes também devem apresentá-las no idioma nacional do usuário.

2.2 Todas as plataformas de conteúdo devem estabelecer e implementar TOS que sejam transparentes, claros, acessíveis e de acordo com os padrões e princípios internacionais de direitos humanos, incluindo as condições sob as quais se pode gerar a interferência no direito à liberdade de expressão ou privacidade dos usuários[8]. Em particular, o usuário deve ser informado sobre as condições que podem levar à rescisão do contrato (remoção de conta, por exemplo), bem como a eliminação, desindexação ou redução significativa do escopo de suas expressões e conteúdos de modificações unilaterais de algoritmos de curadoria.

2.3 Nenhuma plataforma de conteúdo deve ser capaz de alterar unilateralmente os termos de serviço e condições de uso, ou aplicar novos termos, sem informar claramente o usuário da justificativa e sem dar, sob aviso razoável, a possibilidade de cancelar o contrato, sem consequências econômicas ou legais para o usuário[9].

2.4 As plataformas não devem estabelecer condições abusivas ou assimétricas sobre o uso e propriedade do conteúdo gerado e publicado por seus usuários, respeitando seus direitos autorais da mesma forma que os usuários devem cumprir as regulamentações referentes ao conteúdo gerado por terceiros. A este respeito, as restrições decorrentes da proteção de direitos autorais devem considerar as limitações e exceções reconhecidas em tratados internacionais e leis nacionais.

2.5 Os termos de serviços não devem conceder poder ilimitado e discricionário às plataformas para determinar a idoneidade do conteúdo gerado pelo usuário[10]. Em particular, os TOS que possam implicar limitações no exercício do direito à liberdade de expressão e de acesso às informações de seus usuários não devem ser formulados de maneira vaga ou ampla, de tal forma que permitam uma interpretação e aplicação arbitrárias por parte das plataformas.

2.6 Em relação à curadoria/priorização da visualização dos conteúdos gerados por seus usuários (em *feeds* de notícias, resultados de busca, serviços de acesso a notícias e similares), as grandes plataformas devem:

[8] Liberdade de expressão e Internet, Relatoria Especial para a Liberdade de Expressão CIDH, 2013, par. 112
[9] Acordo da União Europeia com Facebook, Google e Twitter em 2018 "*Better social media for European consumers*"
[10] Última frase tomada do Acordo da União Europeia com Facebook, Google e Twitter em 2018 "*Better social media for European consumers*"

A. Tornar transparentes os critérios utilizados pelos algoritmos para a sua ordenação ou direcionamento, se possível explicitando os efeitos para o usuário.
B. Não usar critérios arbitrários ou discriminatórios que possam afetar ilegitimamente a liberdade de expressão e o direito à informação de seus usuários.
C. Fornecer mecanismos de filtragem personalizados de forma clara, transparente, explícita, revogável/editável e sob o controle do usuário, para que eles sejam os que decidem o conteúdo que desejam priorizar e como desejam fazê-lo (por exemplo: ordem cronológica).
D. Respeitar o direito do usuário de conhecer e definir quais de seus dados pessoais são coletados e armazenados, e como são utilizados no endereçamento de conteúdo, respeitando o princípio da autodeterminação informacional.

2.7 No caso em que as grandes plataformas decidam, por si só, incorporar em seus TOS certas restrições e até mesmo proibições à publicação de conteúdos ou expressões geradas por seus usuários, elas só poderiam fazê-lo com as seguintes limitações, para que sejam compatíveis com padrões internacionais de direitos humanos:

A) Poderiam proibir, mesmo através da filtragem automática, os conteúdos clara e manifestamente ilegais que, ao mesmo tempo, sejam reconhecidos como limitações legítimas à liberdade de expressão nas declarações ou tratados internacionais de direitos humanos, como o abuso sexual ou exploração de menores ou propaganda para a guerra e qualquer apologia de ódio nacional, racial ou religioso que constitua incitação à violência ou qualquer outra ação ilegal semelhante contra qualquer pessoa ou grupo de pessoas, por qualquer motivo, incluindo os de raça, cor, religião, língua ou origem nacional[11].

B) Poderiam restringir, como medida cautelar não definitiva, conteúdos que, ainda que não reconhecidos como ilegais, causem danos sérios, iminentes e irreparáveis ou de difícil reparação a outras pessoas, tais como: divulgação não autorizada de conteúdo sexual ou atos de natureza ilícita; atos de violência ou crueldade explícita e excessiva ou aberrante. Nesses casos, a lista e as definições de conteúdo restrito devem ser incluídas nos Termos de Uso de forma taxativa, clara e precisa e considerar-se, na análise da medida a ser tomada, o contexto da expressão publicada e que não sejam utilizado no âmbito de expressões legítimas (conteúdo educacional, informativo ou outro).

C) Conteúdos como o cyberbullying ou o uso explícito e abusivo de drogas podem ser restritos a públicos específicos, como crianças e adolescentes.

[11] Convenção Americana de Direitos Humanos, art. 13, inc. 5

D) Para qualquer outra medida de priorização ou restrição a expressões e outros conteúdos gerados por seus usuários que a plataforma possa considerar – por razões comerciais ou outras – “ofensivas”, “inapropriadas”, “indecentes” e definições vagas ou amplas semelhantes que possam afetar ilegitimamente a liberdade de expressão, as grandes plataformas devem fornecer mecanismos e avisos para outros usuários - voluntariamente e com base em suas preferências morais, religiosas, culturais, políticas ou outras - que decidem a que conteúdo querem ter acesso e quais não. Esses conteúdos não devem ser proibidos, excluídos ou reduzidos em escopo por padrão, afetando desproporcionalmente o direito à liberdade de expressão de seus usuários.

3. TRANSPARÊNCIA

3.1 As plataformas devem publicar on-line suas políticas de restrição de conteúdo, em linguagem clara e em formatos acessíveis, mantê-las atualizadas à medida que evoluem e notificar os usuários sobre alterações, conforme apropriado[12].

3.2 Quando o conteúdo é restrito em um produto ou serviço do intermediário que permite exibir um aviso ao tentar acessar o referido conteúdo, o intermediário deve exibir um aviso claro explicando que conteúdo foi removido e por quê[13].

3.3 Nas ações de priorização de conteúdos online acessíveis ao usuário (feeds, resultados de busca e outros), a natureza comercial das comunicações, o conteúdo patrocinado, bem como a propaganda eleitoral ou política devem ser claramente identificados, identificando a parte contratante e sem gerar dúvidas sobre o seu significado[14], além de ser transparente sobre os metadados do conteúdo (preços, etc.).

3.4 As plataformas devem informar seus usuários, de forma clara, explícita e acessível[15], pelo menos sobre:

A. Que tipos de conteúdo e atividades são proibidos em seus serviços?
B. Quais são os critérios e mecanismos de curadoria e moderação de conteúdo? Quais são controlados diretamente pelo usuário e quais não são?
C. Em que casos, quando e como a automação de análise de conteúdo se aplica[16]?
D. Em que casos, quando e como a revisão de conteúdo humano se aplica? Em particular, referência aos critérios para a tomada de decisões de não afetar os direitos humanos, levando em consideração o contexto, a ampla variação de nuances idiomáticas e o significado e as peculiaridades linguísticas e culturais dos conteúdos sujeitos a uma possível restrição[17].
E. Quantos moderadores tem, descrevendo em detalhes seu perfil profissional (experiência, especialização ou conhecimento), sua localização espacial e sua distribuição de tarefas (em termos de temas, áreas geográficas etc.)?
F. Quais são os direitos dos usuários em relação ao conteúdo gerado e publicado por eles mesmos e as políticas aplicadas pela empresa a esse respeito?

[12] Princípios de Manila
[13] Idem
[14] Texto base: Acordo da União Europeia com Facebook, Google e Twitter em 2018 "*Better social media for European consumers*"
[15] Para "permitir aos usuários prever com razoável certeza qual conteúdo os coloca no lado perigoso da linha" (Relatoria de Liberdade de Expressão das Nações Unidas, 2018)
[16] Regulação do conteúdo online gerado por usuários, Relatoria Especial sobre a Promoção do Direito ao Liberdade de Opinião e Expressão das Nações Unidas, 2018
[17] Idem anterior.

4. APLICAÇÃO DAS POLÍTICAS E DEVIDO PROCESSO

4.1 Na concepção e aplicação das suas políticas comunitárias de tratamento de conteúdos, as plataformas devem procurar que qualquer restrição derivada da aplicação dos termos de serviço não restrinja de forma ilegal ou desproporcionada o direito à liberdade de expressão[18], que deve respeitar os requisitos de busca de uma finalidade imperativa, bem como a necessidade, adequação e proporcionalidade da medida para atingir a finalidade pretendida[19].

4.2 Os critérios para a tomada de decisões, para não afetar os direitos humanos, devem levar em conta o contexto, a ampla variação de nuanças idiomáticas e o significado e as peculiaridades linguísticas e culturais dos conteúdos sujeitos a uma possível restrição[20].

4.3 Além disso, na análise das medidas de restrição de conteúdo aplicáveis em cada caso, os princípios de proporcionalidade e progressividade devem ser respeitados, ponderando a gravidade dos danos, a recorrência das violações e o impacto que tal restrição poderia ter sobre a capacidade de Internet para garantir e promover a liberdade de expressão em relação aos benefícios que a restrição traria para a proteção de outros direitos[21].

4.4 Os usuários devem sempre ter o direito de que as decisões de restrição de conteúdo das próprias plataformas sejam cumpridas com respeito ao devido processo, em particular quando se trata de medidas que possam afetar seu direito à liberdade de expressão. Como princípio geral, e salvo casos excepcionais devidamente justificados, as pessoas afetadas por uma medida de restrição ou interferência por parte das plataformas e, se for caso disso, do público em geral, devem ser previamente notificadas das medidas restritivas que as afetam[22].

4.5 Atendendo aos princípios de necessidade e proporcionalidade acima mencionados, em caso de possíveis violações dos TOS, as plataformas devem adoptar medidas menos onerosas do que a remoção ou outras de efeitos semelhantes, optando por mecanismos de aviso ou notificação, sinalizando, vinculando com informações contraditórias ou outras.

4.6 As medidas unilaterais mais drásticas que forem adotadas sem aviso prévio ou devido a um processo anterior, como remoção de conta ou perfil, remoção de conteúdo ou

[18] Liberdade de expressão e Internet, Relatoria Especial para a Liberdade de Expressão CIDH, 2013, par. 112
[19] Liberdade de expressão e Internet, Relatoria Especial para a Liberdade de Expressão CIDH, 2013, par. 55
[20] Extraído do Relatório de 2018 David Kaye (no original refere-se a recomendações de transparência para as plataformas)
[21] Liberdade de expressão e Internet, Relatoria Especial para a Liberdade de Expressão CIDH, 2013, par. 54
[22] Liberdade de expressão e Internet, Relatoria Especial para a Liberdade de Expressão CIDH, 2013, par. 115

outras medidas que tenham um impacto semelhante de exclusão das possibilidades de participação na plataforma devem ser tomadas pelas grandes plataformas apenas sob as seguintes condições:

A. Quando se tratar de intervenções de gerenciamento técnico não arbitrárias ou discriminatórias (como *spam*, contas falsas, *bots* maliciosos ou similares).
B. No caso de duplicatas ou reiterações idênticas (não comentado ou editado para fins jornalísticos ou informativos ou outros fins legítimos) de outros conteúdos e expressões de ilegalidade óbvia que já estavam restritos após avaliação humana seguindo as normas acima mencionadas.
C. Quando as seguintes situações foram identificadas:
    a. Os fundamentos apresentados na seção 2.7 A;
    b. o cumprimento de ordens de autoridades competentes de retirada imediata e a consumação de crimes comuns já tipificados na legislação nacional;
    c. danos sérios, iminentes e irreparáveis ou de difícil reparação aos direitos de outras pessoas como nos casos listados no ponto 2.7 literais B e C.

Em todos esses casos, exceto no caso de ordens de autoridades competentes, a plataforma deve proceder à imediata notificação subsequente, com a possibilidade de recurso para uma possível revisão da medida nos termos do capítulo 5 deste documento.

4.7 A filtragem e o bloqueio de *uploads* só são legítimos e compatíveis com os padrões internacionais de direitos humanos quando se trata de pornografia infantil[23] ou nas duas primeiras situações descritas no ponto anterior. Caso contrário, deve ser considerado um ato de censura, nos termos estabelecidos pela Convenção Americana sobre Direitos Humanos.

4.8 Qualquer outra medida de restrição de conteúdos ou expressões que a plataforma pretenda adotar em caso de uma possível violação dos Termos do Serviço ou a denúncia de terceiros (por exemplo, sobre uma afetação de direitos autorais), o conteúdo questionado deve ser mantido publicado na plataforma até que uma decisão definitiva decorra de um devido processo onde, após a notificação a) promova-se a retirada voluntária do conteúdo questionado; b) garanta-se o exercício do direito de defesa do usuário permitindo uma contra-notificação com contrarrazões antes de tomar uma decisão.

4.9 Nenhuma plataforma de conteúdo deve ter responsabilidade legal pelo conteúdo gerado por terceiros, desde que não intervenha modificando ou editando esses

[23] Convenção Americana de Direitos Humanos, art. 13, inc. 4

conteúdos, nem se recusem a executar ordens judiciais ou de autoridades oficiais competentes e independentes que cumpram com garantias adequadas correspondentes do devido processo.

4.10 Plataformas de conteúdo só podem ser consideradas responsáveis em função de suas ações ou negligência na priorização ou promoção ativa de expressões que possam afetar os direitos de terceiros se afastarem-se dos princípios estabelecidos no ponto 2.6. Nesses casos, a responsabilidade legal que pode corresponder às plataformas para expressões de terceiros ou atividades de curadoria não deve ser do tipo objetivo.

5. DIREITO À DEFESA E REPARAÇÃO

5.1 Todas as plataformas de conteúdo devem explicar claramente aos usuários por que seu conteúdo foi restrito, limitado ou removido; ou sua conta ou perfil suspenso, bloqueado ou excluído:

A. As notificações devem incluir, pelo menos, a cláusula específica das regras da comunidade que se supõe que o usuário tenha violado.
B. A notificação deve ser detalhada o suficiente para permitir que o usuário identifique especificamente o conteúdo restringido e inclua informações sobre como o conteúdo ou a conta foi detectada, avaliada e excluída ou restrita.
C. As pessoas devem ter informações claras sobre como recorrer da decisão[24] [25].

5.2 As plataformas de conteúdo não devem excluir publicações ou outros conteúdos gerados por usuários sem notificá-los, sem fornecer justificativa clara e sem dar aos usuários a possibilidade de apelar[26], para que possam exercer seu direito de defesa e evitar abusos. A esse respeito, as plataformas devem oferecer aos usuários a oportunidade de apelar às decisões de moderação de conteúdo, nas seguintes condições:

A. Os mecanismos de recurso devem ser muito acessíveis e fáceis de usar.
B. Os recursos devem estar sujeitos a revisão por uma pessoa ou painel de pessoas que não estiveram envolvidas na decisão inicial.
C. Os usuários devem ter o direito de propor novas evidências ou materiais a serem considerados em sua opinião.
D. Apelações devem resultar em rápida determinação e resposta ao usuário.
E. Qualquer exceção ao princípio dos recursos universais deve ser claramente divulgada e compatível com os princípios internacionais de direitos humanos[27].

5.3 Os usuários afetados por qualquer medida de restrição de sua liberdade de expressão como resultado das decisões das plataformas, dependendo das regulamentações específicas da legislação nacional, devem ter o direito de acessar recursos legais para

[24] Princípios de Santa Clara
[25] Princípios de Manila: "A notificação sobre as decisões de restrição de conteúdo adotadas por uma plataforma deve, no mínimo, ter as seguintes informações:
As razões pelas quais o conteúdo em questão viola as políticas de restrição de conteúdo do intermediário.
O identificador da Internet e uma descrição da suposta violação das políticas de restrição de conteúdo.
Os detalhes de contato do emissor ou seu representante, a menos que isso seja proibido por lei.
Uma declaração de boa fé de que as informações fornecidas são precisas"
[26] Acordo da União Europeia com Facebook, Google e Twitter em 2018 "*Better social media for European consumers*"
[27] Princípios de Santa Clara

contestar tal decisão e mecanismos de reparação em relação à eventual violação de seus direitos[28]. Nesse sentido, as plataformas de conteúdo não podem impedir seus usuários de entrar com uma ação legal contra elas em seu país de residência, o que implicaria uma negação de seu direito de acessar a Justiça[29] de maneira subsidiária ou paralela às reivindicações através dos mecanismos. de apelo interno. Para este efeito, o contrato celebrado entre o utilizador com uma plataforma de conteúdos deve incluir expressamente que os litígios serão regidos pela lei e pela justiça do país em que o utilizador tem a sua residência habitual e não pelo local onde estão localizados os escritórios da plataforma[30].

[28] Liberdade de expressão e Internet, Relatoria Especial para a Liberdade de Expressão CIDH, 2013, par. 115
[29] Acordo da União Europeia com Facebook, Google e Twitter em 2018 "*Better social media for European consumers*"
[30] Acordo da União Europeia com Facebook, Google e Twitter em 2018 "*Better social media for European consumers*"

6. PRESTAÇÃO DE CONTAS

6.1 As plataformas de conteúdo devem publicar relatórios de transparência que forneçam informações específicas sobre todas as restrições de conteúdo adotadas pelo intermediário, incluindo ações tomadas antes de solicitações do governo, ordens judiciais, requisitos privados e sobre a implementação de suas políticas de restrição de conteúdo[31].

6.2 As plataformas de conteúdo devem emitir relatórios periódicos de transparência sobre a aplicação de suas regras comunitárias, que incluam pelo menos:

A. Dados completos descrevendo as categorias de conteúdo do usuário que são restritas (texto, foto ou vídeo; violência, nudez, violações de direitos autorais, etc.), bem como o número de publicações que foram restringidas ou removidas em cada categoria, detalhadas por país[32].

B. Dados sobre quantas ações de moderação de conteúdo foram iniciadas por informação (*flag*) de um usuário, por um programa de informação por usuários de confiança (*trusted flagger program*) ou pela aplicação proativa de padrões da comunidade (por exemplo, por meio de uso de um algoritmo de aprendizado de máquina)[33].

C. Dados sobre o número de decisões que foram efetivamente apeladas ou que foram consideradas que foram tomadas por engano[34].

D. Dados que refletem se a empresa realiza uma auditoria proativa de suas decisões de moderação não apeladas, bem como as taxas de erro encontradas pela empresa[35].

E. Dados agregados que ilustram tendências no campo do monitoramento do cumprimento de normas e exemplos de casos reais ou casos hipotéticos detalhados que esclarecem as nuances da interpretação e aplicação de normas concretas[36].

[31] Princípios de Manila
[32] Princípios de Santa Clara
[33] Princípios de Santa Clara
[34] Princípios de Santa Clara
[35] Princípios de Santa Clara
[36] Regulação do conteúdo online gerado por usuários, Relatoria Especial sobre a Promoção do Direito ao Liberdade de Opinião e Expressão das Nações Unidas, 2018

7. APROVAÇÃO E APLICAÇÃO DA REGULAÇÃO

7.1 Considerando que se trata de medidas que podem afetar os direitos fundamentais, os aspectos substantivos da regulamentação proposta neste documento devem ser adotados de maneira prévia e por lei formal, ou seja, uma lei aprovada pelo órgão legislativo (Congresso, Parlamento, Assembleia Nacional ou similar), após consulta pública e aberta. Quando necessário, as delegações de regulamentação para os órgãos de aplicação devem ser cuidadosamente estabelecidas por lei.

7.2 As plataformas de conteúdo não devem depender de licenças para sua operação em um determinado país, mas deve haver uma obrigação de identificar responsáveis legais e formas eficazes de comunicação e resposta para os usuários e as respectivas autoridades, como uma conta de e-mail, um formulário eletrônico ou meios equivalentes[37].

7.3 As plataformas de conteúdo não devem ser obrigadas a monitorar ou supervisionar o conteúdo gerado por terceiros, de forma genérica, a fim de detectar supostas violações da lei ou impedir futuras violações[38].

7.4 O funcionamento das plataformas de conteúdos deve enquadrar-se num ambiente de co-regulação adequado às características do ambiente digital:

A. Os princípios e padrões incluídos nesta proposta devem ser incluídos pelas plataformas de conteúdo em seus termos de serviço e outros documentos complementares (como *guidelines*);
B. As plataformas devem aplicar estes princípios e padrões sem prévia intervenção de organismos estatais;
C. As plataformas devem ter mecanismos internos e eficazes de apelação, bem como instâncias externas independentes às empresas para a revisão de casos e políticas adotadas, com o entendimento de que a regulação estatal deve atuar somente quando as instâncias de autorregulação não funcionam;
D. Deve haver um órgão regulador especializado que opere com garantias suficientes de independência, autonomia e imparcialidade e que tenha a capacidade de avaliar os direitos em jogo e oferecer as garantias necessárias ao usuário[39] em relação às políticas e práticas de implementação dos termos de serviços das plataformas, e indicar a adequação das mesmas quando apropriado.

[37] Este último tomado do Projeto de Lei de Intermediários da Argentina, art. 7
[38] Projeto de Lei de Intermediários da Argentina, art. 5
[39] Nos termos expressados em Liberdade de Expressão e Internet, Relatoria Especial para a Liberdade de Expressão CIDH, 2013, par. 56

Em caso de não cumprimento das obrigações de transparência, devido processo legal, direito à defesa e outros, a agência deve ter capacidade de *enforcement* suficiente, podendo aplicar sanções, se necessário. Entretanto, não deve avaliar nem ter uma decisão vinculante em casos individuais;

E. Casos individuais em que haja violação dos direitos dos usuários e que não sejam satisfatoriamente resolvidos dentro dos escopos internos e mecanismos de resolução de disputas devem ser resolvidos por órgãos judiciais, Defensorias Públicas ou órgãos públicos independentes similares e especializados – do país onde o usuário tem sua residência habitual – por meio de procedimento abreviado, de trâmite digital e de notificação eletrônica (*fast track*) com garantias de revisão por uma autoridade imparcial. Outras autoridades ou órgãos estatais não devem poder obrigar as plataformas a remover ou dar tratamento específico a conteúdos.

Elaboração

Apoios / Adesões