

Padrões para uma regulação democrática das grandes plataformas que garanta a liberdade de expressão online e uma Internet livre e aberta

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

Julio 2020

Organizações brasileiras
que apoiam a proposta:



Organizações Latino-Americanas
que apoiam a proposta:



Os seguintes colaboradores colaboraram na elaboração da proposta (a título pessoal):

- **Javier Pallero**
Director de Políticas, Access Now

- **Joan Barata**
Miembro de la Plataforma de Defensa de la Libertad de Información (España)

- **Andrés Piazza**
Consultor, ExLACNIC, ExLACTLD

- **Guillermo Mastrini**
Profesor Investigador en la Universidad Nacional de Quilmes (UnQ), Universidad de Buenos Aires (UBA), Conicet

- **Martín Becerra**
Universidad Nacional de Quilmes (UnQ), Universidad de Buenos Aires (UBA), Conicet

- **Juan Ortiz Freuler**
Investigador afiliado al Berkman Klein Center (2019-2020)

RESUMO

O presente documento contém recomendações sobre princípios, padrões e medidas específicas de co-regulação e regulação pública para proteger as liberdades de expressão, informação e opinião¹ dos usuários de plataformas e garantir uma Internet livre e aberta. A proposta inclui tanto limitações ao poder das grandes plataformas de Internet (tais como redes sociais e mecanismos de busca) como proteções para dotar os intermediários de instrumentos adequados para facilitar o exercício da liberdade de expressão.

Esta proposta é compatível com os padrões internacionais de direitos humanos e leva em conta as assimetrias existentes, alcançando as grandes plataformas de Internet sem ser um freio ou obstáculo à inovação, competição ou desenvolvimento de start-ups, pequenas empresas ou iniciativas comunitárias, educacionais ou sem fins lucrativos.

¹ Neste documento, deve-se entender o conceito de liberdade de expressão incluindo o direito a expressar-se e difundir tanto ideais como informações e opiniões, assim como o direito a buscar e receber informações, ideia e opiniões de toda índole.

INTRODUÇÃO

A crescente intervenção das grandes plataformas de Internet no conteúdo de seus usuários mediante a adoção de termos de serviço e aplicação de políticas de moderação se tornou um tema de preocupação em todo o mundo, uma vez que foram constituídas formas de regulação privada de espaços públicos que hoje são vitais para a deliberação democrática e o exercício de direitos fundamentais.

De fato, o “controle privado” em Internet é considerado um dos três principais desafios para a próxima década e uma “ameaça à liberdade de expressão” pelos Relatores de Liberdade de Expressão². 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”³.

Essa preocupação não é nova. Em muitas ocasiões tanto organismos internacionais quanto organizações de direitos digitais questionaram essas práticas e formularam recomendações para que as empresas fizessem uma mudança nas políticas e práticas para se ajustarem aos padrões internacionais de direitos humanos⁴.

Por sua vez, a Relatoria de Liberdade de Opinião e Expressão das Nações Unidas publicou diversos relatórios sobre o assunto⁵ e a Relatoria Especial 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”⁶.

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 por meio de diferentes tipos de regulamentação, particularmente regulando o conteúdo disseminado nas redes

² 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.

³ Idem.

⁴ Entre elas, os Princípios de Santa Clara.

⁵ Regulação de conteúdo na Internet, Relatoria Especial sobre a Promoção e Proteção do Direito à Liberdade de Opinião e Expressão, 2018.

⁶ Liberdade de Expressão e Internet, Relatoria Especial para a Liberdade de Expressão CIDH, 2013, par. 111

sociais. Mas a maioria dessas iniciativas legais configura soluções que são desproporcionais e ampliam o risco de violações ao direito à livre expressão, uma vez que atribuem 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 nos opusemos a essas propostas e continuaremos a fazê-lo. Mas acreditamos que o modelo de autorregulação que tem prevalecido até agora apresenta riscos similares ao efetivo exercício de direitos humanos básicos.

Algumas poucas corporações centralizaram e concentraram o poder de administrar a circulação, troca ou busca de informações e opiniões e exercem esse poder de forma arbitrária e sem nenhum mecanismo de *accountability* às instituições que garantem direitos. Esse cenário representa um risco ao efetivo exercício de direitos, um risco que vai crescendo à medida que tal punhado de empresas reforça sua posição dominante no mercado e desenvolve tecnologias opacas para a governança da informação. A trajetória natural desse processo eleva os níveis de alarme e requer ações urgentes para atenuar os riscos que isso implica para o efetivo exercício de direitos e por distorcer a ideia de uma Internet descentralizada, livre e aberta pela qual lutamos.

Diante desse cenário polarizado de “autorregulação corporativa versus regulação autoritária”, 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 regulatórios 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 têm sobre o fluxo de informação na Internet exige que as sociedades democráticas adquiram capacidades de garantir à população o efetivo exercício 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 busca e plataformas de compartilhamento de vídeo, por exemplo, embora não inclua serviços de mensagens⁷.

⁷ É sabido que, em alguns casos, os serviços de mensagem funcionam como serviços de comunicação massiva, excedendo sua função original de comunicação interpessoal. Mesmo assim, por suas características e funções primordiais e pela ausência de moderação de conteúdo por parte das empresas, esses serviços não estão incluídos na presente proposta.

O documento propõe um princípio de “regulação progressiva” com base no impacto que as medidas tomadas pelos intermediários tenham no exercício dos direitos fundamentais na Internet, especialmente a liberdade de expressão. Ou seja, a regulação deve ser mais rigorosa no caso de grandes plataformas que, por seu tamanho, alcance ou posição de mercado tenham se tornado espaços públicos cuja massividade é quase monopolística e, assim, dominam as opções de deliberação e/ou as principais vias de acesso à informação nos ambientes digitais.

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-regulação⁸ 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 como em outros sistemas ou suportes de informação e comunicação.

Este 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 estabelece uma série de condições para que não se violem direitos fundamentais de seus usuários no processo de moderação privada que essas empresas já realizam de forma inconsistente e opaca.

Para isso, são incluídas propostas sobre quais são os limites para a moderação dos conteúdos que essas 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 deveria 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 regulação privada na Internet é produzida e agravada por um contexto de forte concentração de poder em poucas corporações multinacionais. A regulação pública sobre as atividades das plataformas deveria adotar medidas antimonopólio para enfrentar o cenário de concentração e ausência de

⁸ Definida como o estabelecimento, por lei formal aprovada por um Congresso constituído democraticamente, de diretrizes e resultados que deverão ser alcançados pelas empresas, com aplicação direta por parte das mesmas e um processo de supervisão feito por um organismo idôneo, com garantias de autonomia, independente de governos e empresas e passíveis de *enforcement* se não cumpridos os resultados e diretrizes, isso também estabelecido por lei.

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 antimonopólio a agir.

Este documento não pretende dar resposta a todos os desafios que se apresentam em matéria de governança dos conteúdos online, como no tema da "desinformação", mas consideramos que o conjunto de padrões aqui propostos em matéria de transparência, garantias processuais, limites aos termos de serviço, entre outros, terão um efeito positivo sobre esses problemas, ao restringir as condições favoráveis à sua difusão, estabelecer responsabilidades das grandes corporações no debate público e estabelecer um ambiente regulatório adequado para enfrentá-los de forma compatível com o direito à liberdade de expressão. Soluções mais específicas para outros temas serão objeto de documentos complementares.

Tampouco se incluem nesta proposta questões importantes como as políticas e 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 governança e moderação de conteúdos, 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, com a condição de que desenvolvam 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. Devido processo
5. Direito a defesa e apelação
6. Prestação de contas
7. Co-regulação e 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 (“plataformas de conteúdo”). Entre eles: redes sociais, mecanismos de busca e outras plataformas para troca de textos, imagens e vídeos. O documento excluiu expressamente os serviços de mensagem⁹.
- 1.2 O sistema deverá ser estruturado com base num modelo de co-regulação, onde as estruturas de autorregulação e regulação pública são complementadas¹⁰ para formular soluções legais, contratuais e técnicas que garantam a liberdade de expressão online, em equilíbrio com outros direitos fundamentais. Os instrumentos de regulação e co-regulação¹¹ deveriam ser o resultado de um processo de governança multissetorial que leve em consideração os contextos locais e regionais.
- 1.3 Todas 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 de terceiros serão guiadas pelos mesmos critérios que governam a proteção da expressão por qualquer meio¹². Entre esses princípios se incluem: transparência, prestação de contas, devido processo, necessidade, proporcionalidade, não discriminação e direito à defesa. Todas as plataformas devem também garantir o respeito total pelos direitos do consumidor.
- 1.4 As plataformas de conteúdos que sejam importante fonte de acesso à informação e opiniões de interesse público, de reconhecida influência no debate público ou autodefinidas como tais e/ou importante alcance (“grandes plataformas de conteúdos”) devem ter uma regulação assimétrica em relação a outros provedores do mesmo campo, tendo em vista a importância e o impacto que suas decisões empresariais têm sobre o debate público e o intercâmbio de informações, opiniões e bens culturais, bem como ao exercício da liberdade de expressão e o debate público, com efeitos políticos, sociais e culturais. A definição de reconhecida influência e de grande alcance deve ser feita pelo

⁹ É sabido que, em alguns casos, os serviços de mensagem funcionam como serviços de comunicação massiva, excedendo sua função original de comunicação interpessoal. Mesmo assim, por suas características e funções primordiais e pela ausência de moderação de conteúdo por parte das empresas, esses serviços não estão incluídos na presente proposta.

¹⁰ O desenho institucional e a divisão de responsabilidades são desenvolvidos no capítulo 7 desta proposta.

¹¹ Por co-regulação entendemos um esquema onde as diretrizes gerais e os resultados esperados nas políticas das plataformas - para que sejam compatíveis com os princípios de direitos humanos - são definidos mediante um instrumento legal formal, nutrido de aportes multissetoriais, que deve ser aplicado diretamente pelas plataformas levando em consideração os contextos locais e regionais. Um órgão idôneo, com garantias de independência e autonomia, faria a supervisão da aplicação dos parâmetros pelas empresas.

¹² Regulação do conteúdo online gerado por usuários, Relatoria Especial sobre a Promoção do Direito ao Liberdade de Opinião e Expressão das Nações Unidas, 2018.

órgão regulador legitimado com base na realidade específica de cada país e seu mercado de audiência¹³.

1.5 O tratamento assimétrico não supõe a violação do princípio de igualdade, já que todas as plataformas de conteúdos devem respeitar os princípios e padrões mínimos de direitos humanos, mas uma regulação inteligente é aquela que não coloca cargas excessivas a atores que, por seu desenvolvimento e características, não podem cumpri-las, considerando de maneira adequada e diferenciada as grandes plataformas de conteúdo em relação àquelas que sejam de menor porte ou estejam direcionadas a finalidades específicas¹⁴.

¹³ Versões anteriores do documento sugeriam a utilização de conceitos de concorrência econômica, como poder significativo de mercado, mercado relevante e substituíbilidade, mas preferimos uma perspectiva a respeito de sua importância no acesso à informação, o exercício da liberdade de expressão e sua influência no debate público, antes que concorrência. O objetivo central aqui é não impor cargas desproporcionais a atores com baixa capacidade econômica, start-ups ou ainda atores de serviços de interesses específicos com obrigações que poderiam converter-se em barreiras à sua entrada e permanência, afetando a diversidade de serviços ao alcance dos usuários de Internet. Posições monopolísticas ou oligopolísticas poderiam agravar as exigências, devido ao seu efeito ainda mais grave sobre o debate público. A definição de grande alcance pode considerar, por exemplo, o percentual de penetração em relação ao total de usuários ou quantidade de usuários, como faz a Alemanha na NetzDG.

¹⁴ Busca-se estabelecer igualdade dentro de condições similares e aplicar obrigações específicas somente a certo tipo de atores quando, por suas circunstâncias, tenham um papel preponderante no discurso público, bem como capacidade de responder a requerimentos específicos. Isso não significa negar a existência de pressupostos mínimos para garantir a proteção de direitos humanos para a totalidade das plataformas de conteúdo (não discriminação, transparência etc.), incluindo aquelas sem finalidades lucrativas, as que tenham finalidade científica ou educativa ou aquelas que reúnem grupos reduzidos ou fechados.

2. TERMOS E CONDIÇÕES DE SERVIÇO

- 2.1 Os termos de serviço (TOS) de todas as plataformas de conteúdo¹⁵, bem como outros documentos complementares (como guias ou diretrizes de aplicação de conteúdo), devem ser escritos de forma clara, precisa, inteligível e acessível¹⁶ para todos os usuários em seus idiomas nacionais. Nos serviços apropriados a crianças e adolescentes, sem prejuízo das limitações legais relacionadas à capacidade para contratá-los¹⁷, é necessário que os termos e condições do serviço adotem uma linguagem que seja compreensível para esse grupo de pessoas.
- 2.2 Todas as plataformas de conteúdo devem estabelecer e implementar TOS que sejam transparentes, claros, acessíveis e de acordo com as normas 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 usuários¹⁸. 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 alcance de suas expressões e conteúdos a partir de modificações unilaterais de algoritmos de curadoria¹⁹.
- 2.3 Nenhuma plataforma de conteúdo deve poder alterar unilateralmente os termos de serviço e normas comunitárias, ou aplicar novos termos, sem informar claramente ao usuário da justificativa e sem obter consentimento explícito, sob aviso razoável²⁰. Em qualquer situação, os termos de serviço não devem conter cláusulas abusivas, injustas ou desproporcionais.
- 2.4 Os usuários manterão os direitos autorais (tanto morais como patrimoniais) reconhecidos pelas leis de seu país de origem sobre o conteúdo publicado. As plataformas devem obter cessões de direitos patrimoniais de autor por parte de seus usuários individuais por meio de autorizações expressas, sem impor condições abusivas nem que se aproveitem das assimetrias entre as partes.
- 2.5 No processo de moderação de conteúdos, em todos os casos, as restrições derivadas da proteção de direitos autorais sobre conteúdos protegidos devem considerar as limitações e exceções reconhecidas em tratados internacionais e legislações nacionais, como usos

¹⁵ Idealmente, todas as grandes plataformas deveriam unificar o vocabulário que utilizam em seus TOS, com o objetivo de facilitar o trabalho dos usuários, organizações de direitos digitais e reguladores.

¹⁶ Uma boa referência de acessibilidade está nas "Web Content Accessibility Guidelines".

¹⁷ A suposição de uso das plataformas por parte de menores, sua capacidade ou incapacidade legal para compreender os termos de uso e as situações de controle parental efetivo na hora de aceitar as condições contratuais excedem o alcance desta proposta. De toda forma, a linguagem de redação dos termos de uso deve ter clareza suficiente para ser compreendida por crianças e adolescentes.

¹⁸ Liberdade de expressão e Internet, Relatoria Especial para a Liberdade de Expressão CIDH, 2013, par. 112.

¹⁹ Algoritmos que incidem sobre o processo de priorização e moderação de conteúdos.

²⁰ Acordo da União Europeia com Facebook, Google e Twitter em 2018: "*Better social media for European consumers*".

justos ("*fair use*") e uso de trechos curtos, especialmente para finalidade de crítica, denúncia social ou para fins educativos. Em particular, os filtros de *upload* por eventuais violações a direitos autorais são inconsistentes com a proibição de censura prévia estabelecida na Convenção Americana de Direitos Humanos. O usuário deve ser informado sobre quem fez a reclamação. O tratamento de conteúdo apontado como infrator deve seguir os trâmites de notificação e contra-notificação, como está previsto no item 4.8, sobretudo ao se considerar que as plataformas não devem ser responsabilizadas por conteúdos de terceiros (ver item 4.9).

2.6 Os termos de serviços não devem conceder poder ilimitado e discricionário às grandes plataformas para determinar a idoneidade do conteúdo gerado pelo usuário²¹. 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.

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

A. Tornar transparentes os critérios utilizados pelos algoritmos para a sua ordenação, priorização, redução de alcance ou direcionamento, explicitando os efeitos para o usuário²².

B. Não usar critérios discriminatórios nem que gerem diferenciação injusta²³ 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 dos usuários, para que eles sejam os que decidem o conteúdo que desejam priorizar e como desejam fazê-lo (por exemplo: ordem cronológica).

²¹ Última frase tomada do Acordo da União Europeia com Facebook, Google e Twitter em 2018 "*Better social media for European consumers*"

²² A obrigação de transparência sobre os algoritmos não deve implicar na violação de segredos comerciais nem de direitos de propriedade intelectual. No entanto, deve ser possível verificar-se os efeitos dos algoritmos para o processo de moderação com a finalidade de se avaliar sua compatibilidade com padrões de direitos humanos. Nesse sentido, pode-se atribuir a um grupo determinado de pessoas a função de auditar os algoritmos, incluindo aqueles que são elaborados por aprendizagem de máquinas e/ou inteligência artificial, garantindo aos responsáveis acesso total sob condição de confidencialidade que proteja o segredo comercial.

²³ O conceito de diferenciação injusta (*unfair differentiation*) foi discutido pelo Conselho Europeu em suas formulações sobre inteligência artificial e decisão tomada por algoritmo, no entanto não há um conceito claramente definido. Um dos desafios incluídos em tal formulação é uma consideração consistente sobre o significado de "justiça" ou "injustiça".

D. Respeitar o direito do usuário de conhecer e definir quais de seus dados pessoais são coletados e armazenados, e como são utilizados no direcionamento de conteúdo, respeitando o princípio da autodeterminação informativa.

2.8 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 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 sexual de menores ou a propaganda em favor da guerra e qualquer apologia ao 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 nacional²⁴.

B. Poderiam restringir, como medida cautelar não definitiva, conteúdos que, ainda que não reconhecidos como ilegais, causem um dano grave, iminente e irreparável ou de difícil reparação a outras pessoas, tais como: divulgação não autorizada de conteúdo sexual, violência de gênero e de orientação sexual ou atos de violência ou crueldade explícita e excessiva ou aberrante ou ainda danos iminentes ou irreparáveis à saúde pública ou individual. Nesses casos, a lista e as definições de conteúdo restrito devem ser incluídas nos Termos de Serviço de forma taxativa, clara e precisa e deve-se considerar, na análise da medida a ser tomada, o contexto da expressão publicada e que não seja utilizada no âmbito de expressões legítimas (conteúdo educacional, informativo, para finalidade de crítica, denúncia ou outras).

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

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 – “ofensivos”, “inapropriados”, “indecentes” e definições vagas ou amplas semelhantes que possam afetar ilegitimamente a liberdade de expressão, as grandes plataformas devem fornecer mecanismos e avisos para que sejam os outros usuários - voluntariamente e com base em suas preferências morais, religiosas, culturais, políticas ou outras - que decidam a que conteúdos querem ter acesso e a quais não²⁵. Esses conteúdos não devem ser proibidos, excluídos ou reduzidos em escopo por padrão, se não cumpram o teste tripartite de legalidade, necessidade e

²⁴ Convenção Americana de Direitos Humanos, art. 13, inc. 5. Com as interpretações realizadas pelos distintos órgãos do Sistema Interamericano, o alcance desta disposição inclui questões de gênero, orientação sexual e outros ("qualquer motivo").

²⁵ Deve-se proceder com cuidado em relação às restrições oferecidas ao usuário, de forma que conteúdos com pensamentos divergentes e outras expressões diversas sumam totalmente de seu campo de visão.

proporcionalidade, o que afetaria desproporcionalmente o direito à liberdade de expressão de seus usuários.

- 2.9 Os serviços que não sejam concebidos ou desenhados para o uso de crianças e adolescentes devem manter medidas efetivas para restringir seu uso por esse público, promovendo a divulgação ativa e massiva dessa condição para que as crianças e seus responsáveis tenham conhecimento.

3. TRANSPARÊNCIA

- 3.1 As plataformas devem publicar online suas políticas de restrição e priorização de conteúdo, em linguagem clara e em formatos acessíveis, mantê-las atualizadas à medida que evoluem e notificar os usuários sobre alterações, conforme apropriado²⁶. Os serviços que podem ser acessados por crianças e adolescentes devem considerar a utilização de linguagem adequada a esse público.
- 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 quê²⁷.
- 3.3 Nas ações de priorização de conteúdos online acessíveis ao usuário (*feeds, timelines, resultados de busca e outros*), a natureza comercial das comunicações, o conteúdo patrocinado, bem como a propaganda eleitoral ou política devem ser claramente identificados, incluindo a identificação da parte contratante e sem gerar dúvidas sobre o seu significado²⁸, além de ser transparente sobre os metadados do conteúdo (preços, etc.). As plataformas devem também ser transparentes sobre os vínculos que possam existir entre elas e as empresas que têm seus produtos ou conteúdos recomendados em seus serviços.
- 3.4 As grandes plataformas de conteúdo devem informar a seus usuários, de forma clara, explícita e acessível²⁹, pelo menos sobre:
- A. Que tipos de conteúdo e atividades estão 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; como o algoritmo de curadoria de conteúdo utilizado afeta a visibilidade de conteúdos³⁰.
 - C. Em que casos, quando e como a automação de remoção de conteúdo se aplica³¹.

²⁶ Princípios de Manila.

²⁷ Idem.

²⁸ Texto base: Acordo da União Europeia com Facebook, Google e Twitter em 2018: "*Better social media for European consumers*".

²⁹ 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)

³⁰ Por exemplo, com painéis de transparência em tempo real, que permitam aos usuários comparar os conteúdos a que foram expostos com o universo total de histórias publicadas durante um período determinado de tempo, com o objetivo de aumentar a compreensão pública sobre o funcionamento do algoritmo.

³¹ Regulação do conteúdo online gerado por usuários, Relatoria Especial sobre a Promoção do Direito à Liberdade de Opinião e Expressão das Nações Unidas, 2018.

D. Em que casos, quando e como a revisão humana de conteúdos se aplica. Em particular, que critérios foram utilizados para a tomada de decisões, levando em conta o contexto, a ampla variação de nuances idiomáticos e o significado e as peculiaridades linguísticas e culturais dos conteúdos sujeitos a uma possível restrição³².

E. Quantos moderadores tem, descrevendo em detalhes seu perfil profissional (experiência, especialização ou conhecimento), sua localização geográfica e sua distribuição de tarefas (em termos de temas, áreas geográficas etc.)³³.

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.

G. Como são utilizados e qual é o tratamento dos dados pessoais dos usuários, incluídos os pessoais e sensíveis, feito para tomar decisões algorítmicas que impactem os seus direitos³⁴.

3.5 Os governos e autoridades com poder de polícia devem ter o dever de informar sobre o modo como essas autoridades se relacionam com as empresas sobre temas de moderação de conteúdo, incluindo aspectos como:

- A. Quantos pedidos de dados de usuários foram feitos.
- B. Que motivos foram utilizados para justificar tais pedidos.
- C. Em que marcos legais as solicitações realizadas se basearam.
- D. Em que casos foram solicitadas medidas específicas de moderação de conteúdo tais como a remoção.
- E. Qual foi a resposta da empresa em cada caso.
- F. Quantas publicações foram removidas ou tiveram seu alcance limitado e quantas delas foram reabilitadas.

³² Idem.

³³ Sem prejuízo ao respeito ao direito à privacidade e ao anonimato dos moderadores.

³⁴ É um direito das pessoas ter acesso a esta informação a respeito do uso de seus dados pessoais, não apenas para ordenar ou priorizar conteúdo, como também para que seja possível conhecer como são tomadas as decisões que tenham impacto em sua qualidade de vida e em seus direitos fundamentais.

4. 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 ilegítima ou desproporcional o direito à liberdade de expressão³⁵, que deve respeitar os requisitos de busca de uma finalidade imperativa, bem como a necessidade, idoneidade e proporcionalidade da medida para atingir a finalidade pretendida³⁶.
- 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 nuances idiomáticos e o significado e as peculiaridades linguísticas e culturais dos conteúdos sujeitos a uma possível restrição³⁷.
- 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, seu alcance, a recorrência das violações e o impacto que tal restrição poderia ter na capacidade de Internet para garantir e promover a liberdade de expressão em relação aos benefícios que a restrição traria para a proteção de outros direitos³⁸.
- 4.4 Os usuários têm direito de que as decisões de restrição de conteúdo feitas pelas grandes plataformas sejam cumpridas com respeito ao devido processo³⁹, em particular quando se trata de medidas que possam afetar seu direito à liberdade de expressão. Como princípio geral, e salvo casos excepcionais devidamente justificados⁴⁰, as pessoas afetadas por uma medida de restrição ou interferência por parte das plataformas e, se for o caso, do público em geral, devem ser previamente notificadas⁴¹ das medidas restritivas que as afetam⁴², sendo oferecida a possibilidade de apresentar contra-argumentos ou retirar voluntariamente o conteúdo publicado antes da efetivação de uma medida adotada de forma unilateral pela plataforma.

³⁵ Liberdade de expressão e Internet, Relatoria Especial para a Liberdade de Expressão CIDH, 2013, par. 112.

³⁶ Liberdade de expressão e Internet, Relatoria Especial para a Liberdade de Expressão CIDH, 2013, par. 55

³⁷ Extraído do Relatório de 2018 de David Kaye (no original refere-se a recomendações de transparência para as plataformas).

³⁸ Liberdade de expressão e Internet, Relatoria Especial para a Liberdade de Expressão CIDH, 2013, par. 54.

³⁹ O devido processo implica, pelo menos, a garantia da igualdade de tratamento, que as decisões sejam motivadas, que seja possível o exercício de uma defesa efetiva por parte dos usuários afetados, a possibilidade de apelar das decisões e que os procedimentos tenham uma duração razoável.

⁴⁰ Ver 4.6.

⁴¹ Liberdade de expressão e Internet, Relatoria Especial para a Liberdade de Expressão CIDH, 2013, par. 115.

⁴² A documento da RELE (CIDH) citado na nota anterior expressa com clareza que a notificação prévia deve ser interpretada como um requisito absoluto e que deve ser aplicada "quando procedente". Sem prejuízo do anterior, consideramos que podem existir situações excepcionais que exijam agir com celeridade (moderar conteúdos, remover etc.) e logo fundamentar.

- 4.5 Atendendo aos princípios de necessidade e proporcionalidade acima mencionados, em caso de possíveis violações dos TOS, as plataformas devem adotar medidas menos onerosas do que a remoção ou outras de efeitos semelhantes, optando por mecanismos de aviso ou notificação, sinalização, vinculação com informações contraditórias ou outras.
- 4.6 As medidas unilaterais mais drásticas que forem adotadas sem cumprir com o estabelecido no item 4.4 - como remoção de contas ou perfis, remoção de conteúdo ou outras medidas que tenham um impacto semelhante de exclusão - devem ser tomadas pelas grandes plataformas apenas sob as seguintes condições:
- A. Quando se tratar de intervenções de gerenciamento técnico não arbitrárias nem discriminatórias (como *spam*, contas falsas⁴³, *bots* maliciosos ou similares),
 - B. No caso de duplicatas ou reiteraões 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 os padrões acima mencionadas,
 - C. Quando as seguintes situações foram identificadas:
 - a. Os fundamentos apresentados no item 2.8 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 graves, iminentes e irreparáveis ou de difícil reparação aos direitos de outras pessoas como nos casos listados no item 2.8 B e 2.8 C.

Em todos esses casos, exceto no caso de ordens de autoridades competentes⁴⁴, a plataforma deve proceder à imediata notificação subsequente ao usuário, com a possibilidade de recurso para uma possível revisão da medida nos termos do capítulo 5 deste documento. A decisão também deve ser notificada ao público em geral, alcançando ao menos os usuários que já tenham tido contato com esse conteúdo.

- 4.7 A filtragem e bloqueio de *uploads* (*upload filters*) por parte das plataformas só são legítimos e compatíveis com os padrões internacionais de direitos humanos quando se trata de proteção à infância e adolescência⁴⁵ ou nas duas primeiras situações descritas no item anterior. Caso contrário, deve ser considerado um ato de censura prévia, nos termos estabelecidos pela Convenção Americana de Direitos Humanos⁴⁶.

⁴³ Não são consideradas como contas falsas aquelas que sejam utilizadas para a paródia ou a sátira.

⁴⁴ Desde que não seja de responsabilidade da plataforma.

⁴⁵ Convenção Americana de Direitos Humanos, art. 13, inc. 4.

⁴⁶ Como já expressamos anteriormente, deve-se reconhecer que há, atualmente, outras obrigações legais de filtragem automática a que as plataformas estão obrigadas. No entanto, em nossa visão, essas são ilegítimas diante dos padrões internacionais e deveriam ser modificadas.

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 algo que afete direitos autorais), o conteúdo questionado deve ser mantido à disposição do público na plataforma até que uma decisão definitiva decorra de um devido processo onde, após a notificação ao usuário:

A. promova-se a retirada voluntária do conteúdo questionado o

B. garanta-se o exercício do direito de defesa⁴⁷ permitindo uma contra-notificação com a apresentação de contra-argumentos antes de se 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 conteúdos, nem se recuse a executar ordens judiciais ou de autoridades oficiais competentes e independentes, sempre e quando tais ordens cumpram com garantias adequadas correspondentes de devido processo legal e identifiquem o conteúdo a ser restringido e que definam de forma clara por que o conteúdo é ilegal.

4.10 As grandes plataformas de conteúdo devem ser consideradas responsáveis por ações próprias se censuram conteúdos de interesse público protegidos pelo direito à liberdade de expressão⁴⁸ e por suas ações na promoção ativa de expressões que podem afetar os direitos de terceiros, se afastarem-se dos padrões estabelecidos no ponto 2.8. Também devem ser consideradas responsáveis se não mantiverem a devida diligência em relação a conteúdo questionado judicialmente ou para evitar ou limitar ações maliciosas coordenadas.

⁴⁷ Ver seção 5.

⁴⁸ Entre estes, não se incluem peças comerciais publicitárias, que devem ser entendidas como ação econômica das plataformas.

5 DIREITO A DEFESA E APELAÇÃO

5.1 Todas as plataformas de conteúdo devem explicar claramente aos usuários por que seu conteúdo foi restringido, 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 deve incluir informações sobre como o conteúdo ou a conta foi detectada, avaliada e excluída ou restringida.

C. As pessoas devem ter informações claras sobre como recorrer da decisão⁴⁹
⁵⁰.

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 apelar⁵¹, para que possam exercer seu direito de defesa e evitar abusos. A esse respeito, as plataformas devem oferecer aos usuários a oportunidade de apelar das 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 e que não sejam parte interessada.

C. Os usuários devem ter o direito de propor novas evidências ou materiais a serem considerados em sua opinião.

D. As apelações devem resultar em rápida determinação e resposta ao usuário.

⁴⁹ Princípios de Santa Clara.

⁵⁰ 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 dados 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".

⁵¹ Acordo da União Europeia com Facebook, Google e Twitter em 2018 "*Better social media for European consumers*".

E. Qualquer exceção às garantias de dupla instância consagradas nos padrões de devido processo⁵² deve ser claramente divulgada e compatível com os demais princípios internacionais de direitos humanos⁵³.

5.3 Os usuários afetados por qualquer medida de restrição à 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 contestar tal decisão e mecanismos de reparação em relação à eventual violação de seus direitos⁵⁴.

5.4 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 Justiça⁵⁵ de maneira subsidiária ou paralela às reivindicações por meio dos mecanismos de apelação interna. Para esse efeito, o contrato celebrado entre o usuário com uma plataforma de conteúdo deve incluir expressamente que os litígios serão regidos pela lei e pela justiça do país em que o usuário tem a sua residência habitual e não pelo local onde estão localizados os escritórios⁵⁶ da empresa responsável pela plataforma em questão⁵⁷.

⁵² Convenção Americana de Direitos Humanos (São José da Costa Rica), artigo 8.2, inciso h.

⁵³ Princípios de Santa Clara.

⁵⁴ Liberdade de expressão e Internet, Relatoria Especial para a Liberdade de Expressão CIDH, 2013, par. 115.

⁵⁵ Acordo da União Europeia com Facebook, Google e Twitter em 2018: *“Better social media for European consumers”*.

⁵⁶ Acordo da União Europeia com Facebook, Google e Twitter em 2018: *“Better social media for European consumers”*.

⁵⁷ O tema da jurisdição é de grande complexidade e foi objeto de inúmeras considerações e sugestões na consulta pública, pela definição do alcance dos danos e dos remédios e pela sobreposição de normas conflitantes sobre distintos temas. Esse tema será objeto de análise e desenvolvimento mais aprofundado em documentos futuros.

6 PRESTAÇÃO DE CONTAS

6.1 As grandes plataformas de conteúdo devem publicar relatórios de transparência que forneçam informações específicas e desagregadas sobre todas as restrições de conteúdo adotadas pelo intermediário, incluindo ações tomadas mediante solicitações do governo, ordens judiciais, requisitos privados e sobre a implementação de suas políticas⁵⁸ de restrição de conteúdo⁵⁹.

6.2 As grandes 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 estão restritas (texto, foto ou vídeo; violência, nudez, violações de direitos autorais ou de direitos de crianças e adolescentes etc.), bem como o número de publicações que foram restringidas ou removidas em cada categoria, detalhadas por país⁶⁰.

B. Dados sobre quantas ações de moderação de conteúdo foram iniciadas por denúncia (*flag*) de um usuário, por um programa de informes 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 automático)⁶¹.

C. Dados sobre o número de decisões que foram efetivamente apeladas ou que foram consideradas como tomadas por engano⁶².

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 empresa⁶³.

E. Dados agregados que ilustram tendências no âmbito 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 concretas⁶⁴.

⁵⁸ Princípios de Manila.

⁵⁹ A informação divulgada deve ser competitivamente neutra e deve-se proceder com cautela a respeito dos segredos comerciais e procedimentos garantidos por propriedade intelectual.

⁶⁰ Princípios de Santa Clara.

⁶¹ Princípios de Santa Clara.

⁶² Princípios de Santa Clara.

⁶³ Princípios de Santa Clara.

⁶⁴ Regulação do conteúdo online gerado por usuários, Relatoria Especial sobre a Promoção do Direito ao Liberdade de Opinião e Expressão das Nações Unidas, 2018

6.3 Sem prejuízo ao relatório periódico de transparência, as grandes plataformas devem emitir alertas de casos específicos como interrupção de serviço, comportamentos anômalos de solicitação de remoção de conteúdo ou de contas.

7 CO-REGULAÇÃO E REGULAÇÃO

- 7.1 Considerando que se trata de medidas que podem afetar os direitos fundamentais, os aspectos substantivos da regulamentação surgida a partir desses padrões 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 nos órgãos de aplicação devem ser cuidadosamente estabelecidas por lei.
- 7.2 As plataformas de conteúdo não devem depender de autorizações prévias para sua operação em um determinado país, mas deve haver, ao menos, a 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.
- 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.
- 7.4 O funcionamento das plataformas de conteúdo deve enquadrar-se num ambiente de co-regulação adequado às características do ambiente digital, nos seguintes termos:
- A. Os princípios e padrões incluídos nesta proposta - aprovados mediante instrumento legal formal - devem ser incluídos pelas plataformas de conteúdo em seus termos de serviço e outros documentos complementares (como *guidelines*);
 - B. As plataformas devem aplicar esses princípios e padrões sem prévia intervenção de organismos estatais;
 - C. A implementação dessas políticas deve ser supervisionada por um órgão público especializado com perspectiva de defesa da liberdade de expressão - que opere com garantias suficientes de independência, autonomia técnica e decisória e imparcialidade e que tenha a capacidade para avaliar os direitos em jogo e oferecer as garantias necessárias ao usuário⁶⁵ - e tenha capacidade para indicar a adequação⁶⁶ das mesmas em relação ao cumprimento de A e B.
 - D. O órgão regulador pode ser criado especialmente para a função ou ser adaptado a partir de órgão já existente no país que atue em áreas afins, desde que cumpra com as garantias do ponto C. De qualquer forma, deverá ser estabelecido por lei com tramitação ordinária e ter recursos próprios para seu funcionamento. Deverá

⁶⁵ Nos termos expressados em Liberdade de Expressão e Internet, Relatoria Especial para a Liberdade de Expressão CIDH, 2013, par. 56.

⁶⁶ Não pressupõe a imposição de políticas ou padrões de tratamento de conteúdos específicos às plataformas.

contar com um conselho multissetorial de caráter consultivo, com representantes de todos os setores interessados, incluindo a sociedade civil.

E. Em caso de não cumprimento das obrigações de transparência, devido processo, direito à defesa e outros, o órgão deve ter capacidade de *enforcement* suficiente, podendo aplicar sanções proporcionais, se necessário. Entretanto, mesmo que possa avaliar casos individuais que podem ser emblemáticos para a análise das políticas das plataformas, não deve aplicar sanções ou tomar decisão vinculante sobre os mesmos, com exceção de autoridades independentes e especializadas que cumpram funções quase-judiciais, como os órgãos eleitorais.

F. O regulador deve ter o poder de exigir qualquer tipo de informação detalhada que seja necessária para cumprir com suas tarefas de supervisão e deve ter o poder de aplicar multas ou outras ações corretivas quando as plataformas não apresentam essa informação de forma oportuna.

G. Entre suas atribuições pode estar a realização de estudos nacionais e comparados, campanhas dirigidas à população sobre seus direitos, além de cooperação com autoridades regulatórias independentes de outros países e ações de cooperação com entidades de autorregulação das empresas.

H. Por regra, a jurisdição do regulador deve ser nacional, mas pode-se adotar soluções regionais, a partir da aprovação dos parlamentos envolvidos, se a legislação e os costumes da região são suficientemente consistentes e coerentes.

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

7.6 Deve-se promover a existência - ou o fortalecimento, caso existam - das defensorias do público, das audiências, *ombudsman* ou organismos similares de caráter não punitivo com o objetivo de defender e promover os direitos das pessoas diante das plataformas, com capacidade para receber e fazer denúncias em sua representação, diante de eventuais violações de direitos fundamentais tanto por plataformas quanto por organismos estatais, inclusive em relação a casos individuais.

7.7 Sem prejuízo do disposto nos itens anteriores, as plataformas devem dispor de mecanismos internos adequados e efetivos de apelação, assim como instâncias externas independentes às empresas para a revisão de casos e políticas adotadas.