

## **Decisão do caso 2020-004-IG-UA**

### **Resumo do caso**

O Conselho de Supervisão anulou a decisão do Facebook de remover uma publicação no Instagram. Após o Conselho ter selecionado este caso, o Facebook restaurou o conteúdo. Os sistemas automatizados do Facebook originalmente removeram a publicação por violar o Padrão da Comunidade sobre Nudez Adulta e Atividade Sexual da empresa. O Conselho determinou que a publicação era permitida de acordo com uma exceção da política para "conscientização sobre o câncer de mama" e que a moderação automática do Facebook, nesse caso, levanta questões importantes de direitos humanos.

### **Sobre o caso**

Em outubro de 2020, um usuário brasileiro denunciou uma foto no Instagram com título em português indicando que o objetivo era aumentar a conscientização sobre sinais do câncer de mama. A imagem era rosa, remetendo a "Outubro rosa", uma campanha internacional para aumentar a conscientização sobre a doença. Oito fotos da imagem mostravam sintomas do câncer de mama com as respectivas descrições. Cinco delas incluíam mamilos femininos visíveis e descobertos, enquanto as três fotografias restantes incluíam seios femininos, com os mamilos não enquadrados na imagem ou cobertos por uma mão. A publicação foi removida por um sistema automatizado, aplicando o Padrão da Comunidade sobre Nudez Adulta e Atividade Sexual do Facebook. Após o Conselho ter selecionado o caso, o Facebook determinou que foi um erro e restaurou a publicação.

### **Principais achados**

Em sua resposta, o Facebook declarou que o Conselho deveria negar o prosseguimento do caso. A empresa argumentou que, tendo restaurado a publicação, não há mais discordância entre o usuário e o Facebook de que o conteúdo deve permanecer no ar, tornando esse caso encerrado.

O Conselho rejeita o argumento do Facebook. A necessidade de discordância é aplicável somente no momento em que o usuário esgota o processo de recurso interno do Facebook. Como o usuário e o Facebook discordaram naquele momento, o Conselho poderá conhecer do caso.

A decisão do Facebook de restaurar o conteúdo também não torna esse caso encerrado, como a empresa afirma. Além de decisões de cumprimento obrigatório sobre a restauração de conteúdo, o Conselho também oferece aos usuários uma explicação completa do motivo pelo qual a publicação foi removida.

O uso de moderação automatizada do Facebook levanta algumas questões de direitos humanos. A detecção e remoção dessa publicação foi automatizada. Os sistemas automatizados do Facebook falharam em reconhecer as palavras "Câncer de Mama", que apareciam na imagem em português, e a publicação foi removida incorretamente. Como as regras do Facebook tratam mamilos femininos e masculinos de maneira diferente, usar automação imprecisa para aplicar essas regras afeta a liberdade de expressão das mulheres de forma desproporcional. A aplicação baseada apenas em moderação automatizada, sem análise humana, interfere na liberdade de expressão.

Nesse caso, o usuário foi informado de que a publicação violava as Diretrizes da Comunidade do Instagram, implicando que o compartilhamento de fotos de mamilos femininos descobertos para aumentar a conscientização sobre o câncer de mama não é permitido. No entanto, o Padrão da Comunidade sobre Nudez Adulta e Atividade Sexual do Facebook expressamente permite a nudez quando o usuário busca "aumentar a conscientização sobre uma causa ou por motivos educacionais ou médicos" e especificamente permite que mamilos femininos descobertos avancem na "conscientização sobre o câncer de mama". Como os Padrões da Comunidade do Facebook se aplicam ao Instagram, a publicação do usuário é coberta pela exceção acima. Portanto, a remoção do conteúdo pelo Facebook era inconsistente com os Padrões da Comunidade.

### **A decisão do Conselho de Supervisão**

A decisão do Conselho de Supervisão anula a decisão original do Facebook de remover o conteúdo e exige que a publicação seja restaurada. O Conselho observa que o Facebook já tomou medidas para esse efeito.

Em uma recomendação de políticas, o Conselho recomenda que o Facebook:

- Informe aos usuários quando a aplicação automatizada é usada para moderar o conteúdo, garanta que os usuários possam recorrer de decisões automatizadas a um humano em determinados casos e melhore a detecção automática de imagens com sobreposição de texto, para que as publicações que visam aumentar a conscientização sobre os sintomas do câncer de mama não sejam sinalizadas

incorretamente para revisão. O Facebook também deve aprimorar os relatórios de transparência sobre o uso de aplicação automatizada.

- Revise as Diretrizes da Comunidade do Instagram para especificar que mamilos femininos podem ser mostrados para aumentar a conscientização sobre o câncer de mama e esclarecer que há inconsistências entre as Diretrizes da Comunidade do Instagram e os Padrões da Comunidade do Facebook, o qual deve prevalecer.

\*Resumos de casos oferecem uma visão geral do caso e não têm valor de precedente.

## **Decisão completa do caso**

### **1. Resumo da decisão**

O Conselho de Supervisão anulou a decisão original do Facebook de remover o conteúdo, observando que o Facebook restaurou a publicação após a decisão do Conselho de conhecer do caso. A decisão do Facebook de restaurar o conteúdo não exclui a autoridade do Conselho de conhecer do caso.

O Conselho determinou que o conteúdo era permitido de acordo com uma exceção da política para "conscientização sobre o câncer de mama" no Padrão da Comunidade do Facebook sobre Nudez Adulta e Atividade Sexual.

O Conselho emitiu uma recomendação de políticas sobre a relação entre as políticas de conteúdo no Instagram e Facebook, além do uso de automação na moderação de conteúdo e a transparência dessas práticas.

### **2. Descrição do caso**

Em outubro de 2020, um usuário brasileiro postou uma foto no Instagram com título em português indicando que o objetivo era aumentar a conscientização sobre sinais do câncer de mama. A imagem era rosa, remetendo a "Outubro rosa", uma campanha popular internacional no Brasil para aumentar a conscientização sobre o câncer de mama. Oito fotografias em uma única imagem publicada mostravam sintomas do câncer de mama com a respectiva descrição como ondulações, nódulos e lesões na parte inferior. Cinco das fotografias incluíam mamilos femininos visíveis e descobertos. As três fotografias restantes incluíam seios femininos, com os mamilos

não enquadrados na imagem ou cobertos por uma mão. O usuário não compartilhou comentários adicionais com a publicação.

A publicação foi detectada e removida por um classificador de machine learning treinado para identificar nudez em fotos, aplicando os Padrões da Comunidade do Facebook sobre Nudez Adulta e Atividade Sexual, também aplicáveis ao Instagram.

O usuário recorreu desta decisão ao Facebook. Em declarações públicas, o Facebook afirmou anteriormente que nem sempre poderia oferecer aos usuários a opção de recorrer devido à redução temporária da capacidade de revisão como resultado da COVID-19. Além disso, o Facebook afirmou que nem todos os recursos serão revisados por humanos.

O usuário enviou uma solicitação de revisão ao Conselho e o Conselho decidiu aceitar o caso. Após o Conselho selecionar e atribuir o caso a um painel, o Facebook reverteu a decisão de remoção original e restaurou a publicação em dezembro de 2020. O Facebook alega que a decisão original de remover a publicação foi automatizada e subsequentemente identificada como um erro de aplicação. No entanto, o Facebook apenas tomou conhecimento do erro após ele ter sido destacado por meio dos processos do Conselho.

### **3. Autoridade e escopo**

O Conselho tem autoridade para revisar a decisão do Facebook nos termos do Artigo 2 (Autoridade para revisar) da [Carta do Conselho](#) e pode manter ou reverter essa decisão nos termos do Artigo 3, Seção 5 (Procedimentos para revisão: Resolução da Carta). O Facebook não apresentou motivos para que o conteúdo fosse excluído de acordo com o Artigo 2, Seção 1.2.1 (Conteúdo não disponível para revisão do Conselho) dos [Regulamentos do Conselho](#), nem indicou que considera o caso inelegível no âmbito do Artigo 2, Seção 1.2.2 (Obrigações legais) dos Regulamentos.

Embora o Facebook tenha publicamente aceitado a revisão do Conselho desse caso, o Facebook propôs que o Conselho negasse o prosseguimento do caso em seus arquivos perante o Conselho, pois o problema está agora encerrado.

O Facebook alega que, tendo restaurado o conteúdo, não há discordância de que ele deve permanecer no Instagram e esse é um requisito para que um caso tenha prosseguimento, de acordo com o Artigo 2, Seção 1 da Carta do Conselho:

*nos casos em que as pessoas discordam do resultado da decisão do Facebook e esgotaram os recursos, uma solicitação de revisão pode ser enviada ao Conselho.*

O Conselho discorda e interpreta que a Carta exige a discordância entre o usuário e o Facebook somente no momento em que o usuário esgota o processo interno do Facebook. Esse requisito foi atendido. O processo de revisão do Conselho é separado e não uma extensão do processo de recurso interno do Facebook. Pois para o Facebook corrigir erros evidenciados pelo Conselho e, *dessa forma, excluir casos da revisão*, o Conselho seria integrado de forma inapropriada ao processo interno do Facebook e prejudicaria a autonomia do Conselho.

Embora o Facebook tenha revertido sua decisão e restaurado o conteúdo, ainda houve danos irreversíveis nesse caso. A decisão do Facebook de restaurar o conteúdo no início de dezembro de 2020 não anula o fato de que a publicação do usuário foi removida durante toda a campanha do "mês rosa" em outubro de 2020.

Restaurar o conteúdo, nesse caso, não é o único propósito que a reparação do Conselho oferece. De acordo com o Artigo 4 (Implementação) da Carta do Conselho, e do Artigo 2, Seção 2.3.1 (Implementação das decisões do Conselho) dos Regulamentos, o Facebook tem o compromisso de tomar medidas em relação a "conteúdo idêntico em contexto paralelo". Dessa forma, o impacto do Conselho na tomada de decisões inclui muito mais que o conteúdo desse caso.

Além disso, uma decisão completa, mesmo quando o Facebook implementa o resultado de forma antecipada, é importante. O processo do Conselho oferece aos usuários uma oportunidade de serem ouvidos e de obter uma explicação completa do motivo pelo qual o conteúdo foi removido indevidamente. Quando a remoção do conteúdo é realizada inteiramente por meio de automação, as políticas de conteúdo estão essencialmente incluídas no código e podem ser consideradas inseparáveis dele e autoexecutáveis. O prosseguimento do caso permite que o Conselho emita recomendações de políticas sobre como as práticas de moderação de conteúdo do Facebook são aplicadas, incluindo com o uso de automação.

Por esses motivos, o Conselho considera que sua autoridade para revisar esse caso não é afetada pela decisão do Facebook de restaurar o conteúdo após o Conselho selecionar o caso. O Conselho prossegue com a revisão da decisão original de remover o conteúdo.

#### 4. Padrões relevantes

O Conselho considerou os seguintes padrões na decisão:

##### *I. Políticas de conteúdo do Facebook:*

A justificativa da política do Padrão da Comunidade sobre Nudez Adulta e Atividade Sexual declara que o Facebook visa restringir a exibição de nudez ou atividade sexual porque algumas pessoas "podem ser sensíveis a esse tipo de conteúdo" e "para evitar o compartilhamento de conteúdo sem o consentimento da pessoa ou de menores de idade". Os usuários não devem publicar imagens de adultos nus reais, onde nudez é definida como [...] mamilos femininos descobertos, exceto no contexto de [...] situações relacionadas à saúde (por exemplo, pós-mastectomia, conscientização do câncer de mama [...]).

As Diretrizes da Comunidade do Instagram definem uma remoção geral referente a mamilos femininos, especificando algumas exceções relacionadas à saúde, mas sem incluir especificamente a "conscientização sobre o câncer de mama". As Diretrizes da Comunidade estão vinculadas aos Padrões da Comunidade do Facebook.

##### *II. Valores do Facebook:*

Os valores do Facebook relevantes para esse caso estão descritos na introdução dos Padrões da Comunidade. O primeiro é "Voz", que é descrito como "primordial".

*O objetivo dos Padrões da Comunidade sempre foi criar um espaço para se expressar e permitir que as pessoas usem sua voz. [...] Queremos que as pessoas conversem abertamente sobre os problemas que são importantes para elas, mesmo se as opiniões forem conflitantes.*

O Facebook limita o valor "Voz" devido a quatro valores. O Conselho considera que dois destes valores são relevantes para essa decisão:

*Segurança: temos o compromisso de tornar o Facebook um local seguro. Uma expressão que ameaça as pessoas tem o potencial de intimidar, excluir ou silenciar outros, e essas práticas não são permitidas no Facebook.*

*Privacidade: temos o compromisso de proteger a privacidade e as informações pessoais. A privacidade oferece às pessoas a liberdade de poder ser quem são e escolher como e quando compartilhar informações no Facebook e se conectar com mais facilidade.*

### III. Padrões de Direitos Humanos Internacionais:

Os Princípios Orientadores da ONU sobre Empresas e Direitos Humanos (UNGPs), endossados pelo Conselho de Direitos Humanos da ONU de 2011, estabeleceram uma estrutura voluntária para a responsabilidade das empresas privadas em matéria de direitos humanos. A análise do Conselho nesse caso foi embasada nas provisões de tratados da ONU e na orientação oficial dos mecanismos de direitos humanos da ONU, incluindo os seguintes:

- O direito à liberdade de expressão: Pacto Internacional dos Direitos Civis e Políticos ([PIDCP](#)), Artigo 19, Comentário Geral nº 34, Comitê de Direitos Humanos (2011) ([Comentário Geral 34](#)); o Relator Especial da ONU sobre liberdade de opinião e expressão, relata: [A/HRC/38/35](#) (2018); [A/73/348](#) (2018); and [A/HRC/44/49](#) (2020).
- O direito à saúde: Pacto Internacional dos Direitos Econômicos, Sociais e Culturais ([PIDESC](#)), Artigo 12, Comentário Geral nº 14, o Comitê sobre os Direitos Econômicos, Sociais e Culturais, [E/C.12/2000/4](#) (2000).
- O direito a um recurso efetivo: PIDCP, Artigo 2, Comentário Geral nº 31, o Comitê de Direitos Humanos [CCPR/C/21/Rev.1/Add. 13](#) (2004).
- O direito à privacidade: PIDCP, Artigo 17.
- O direito a não discriminação: PIDCP, Artigo 2 (para. 1); Convenção sobre a eliminação de todas as formas de discriminação contra as mulheres (CEDAW), Artigo 1.
- Os direitos da criança: Convenção internacional sobre os direitos da criança (CRC), Artigo 6, Comentário Geral nº 13, o Comitê sobre os Direitos da Criança, [CRC/C/GC/13](#) (2011).

### 5. Declaração do usuário

O usuário relata que o conteúdo foi publicado como parte da campanha nacional "Outubro Rosa" de prevenção contra o câncer de mama. A publicação mostra alguns dos principais sinais do câncer de mama, que, segundo o usuário, são essenciais para a detecção precoce dessa doença e contribui para salvar vidas.

### 6. Explicação da decisão do Facebook

O Facebook esclareceu que sua decisão original de remover o conteúdo foi um erro. A empresa explicou ao Conselho que os Padrões da Comunidade se aplicam ao Instagram. Apesar de os Padrões da Comunidade normalmente proibirem mamilos femininos visíveis e expostos, eles são permitidos para "fins médicos e educacionais", incluindo para a conscientização do câncer de mama. O Facebook restaurou o conteúdo, pois ele se encaixava nessa exceção.

O Facebook afirma que a permissão desse conteúdo na plataforma tem uma contribuição importante para os valores "Voz" e "Segurança". A empresa declara que a identificação e a aplicação original desse conteúdo foi completamente automática. O processo automático não foi bem sucedido em determinar que se tratava de um conteúdo para "fins médicos e educacionais". O Facebook também afirma que entender se o conteúdo foi removido por meio de um processo automático ou por uma análise interna de um moderador humano não é relevante para o Conselho em consideração do caso. O Facebook gostaria que o Conselho se concentrasse nas consequências da aplicação em vez de no método.

## **7. Contribuições de terceiros**

O Conselho de Supervisão considerou 24 comentários públicos para esse caso: oito da Europa, cinco da América Latina e Caribe e 11 dos Estados Unidos e Canadá. Sete foram enviados em nome de uma organização. Um comentário foi enviado sem consentimento para publicação.

Os envios envolviam os seguintes temas: se a publicação cumpria os Padrões da Comunidade e os valores do Facebook, a importância da conscientização sobre o câncer de mama para um diagnóstico precoce, críticas sobre a sexualização e censura exacerbada dos mamilos femininos em comparação aos mamilos masculinos; a influência do Facebook na sociedade, a remoção excessiva devido à automação da moderação do conteúdo, assim como o feedback para a melhoria do processo de comentários públicos.

## **8. Análise do Conselho de Supervisão**

### **8.1 Conformidade com as políticas de conteúdo do Facebook**

A decisão do Facebook de remover a publicação do usuário no Instagram não estava em conformidade com a política de conteúdo da empresa.

De acordo com o Facebook, os Padrões da Comunidade operam em todos os produtos da empresa, incluindo o Instagram. Neste caso, o usuário foi informado que o conteúdo violava as Diretrizes da Comunidade do Instagram, que foram apresentadas ao usuário. A diferença entre essas regras merece análises separadas.

### *I. Diretrizes da Comunidade do Instagram*

A versão mais curta das Diretrizes da Comunidade resume as regras do Instagram como: "Respeite todos no Instagram, não envie spam ou publique conteúdo de nudez." Analisadas isoladamente, essas diretrizes implicam que a publicação do usuário viola as regras do Instagram.

A versão mais longa das Diretrizes da Comunidade menciona mais detalhes. Na seção "publique fotos e vídeos apropriados para um público variado", as diretrizes esclarecem:

*Por diversos motivos, não permitimos nudez no Instagram [...] Isso também inclui algumas fotos de mamilos femininos, mas fotos de cicatrizes resultantes de mastectomia e mulheres no momento de amamentação são permitidas.*

Essa explicação não permite explicitamente fotos de mamilos femininos expostos para incentivar a conscientização do câncer de mama. Apesar das Diretrizes da Comunidade do Instagram incluírem um hyperlink para o Padrão da Comunidade sobre Nudez Adulta e Atividade Sexual do Facebook, a relação entre essas duas regras, incluindo a precedente, não é explicada.

### *II. Padrões da Comunidade do Facebook*

O Padrão da Comunidade sobre Nudez Adulta e Atividade Sexual, na seção Conteúdo questionável, declaram que a exibição de nudez adulta, definida para incluir "mamilos femininos expostos", assim como atividade sexual, é normalmente restrita na plataforma. Duas razões são oferecidas para justificar essa posição: "algumas pessoas na comunidade podem ser sensíveis a esse tipo de conteúdo" e "para evitar o compartilhamento de conteúdo sem o consentimento da pessoa e de conteúdo impróprio para menores de idade".

O Padrão da Comunidade especifica que a nudez adulta consensual é permitida quando o usuário claramente indica que o conteúdo serve para "incentivar a

conscientização sobre uma causa para fins educacionais e médicos". A seção "Não publicar" do Padrão da Comunidade menciona "conscientização sobre o câncer de mama" como um exemplo de situação relacionada à saúde em que a publicação de mamilos femininos expostos é permitida.

O Conselho conclui que a publicação do usuário, apesar de ilustrar mamilos femininos expostos, enquadra-se na exceção relacionada à saúde para a conscientização do câncer de mama. Levando em consideração a explicação do Facebook de que os Padrões da Comunidade operam no Instagram, o Conselho decide que a publicação do usuário está em conformidade com eles.

Portanto, a decisão do Facebook de remover o conteúdo estava inconsistente com os Padrões da Comunidade. O Conselho reconhece que o Facebook concordou com essa conclusão.

## **8.2 Conformidade com os valores do Facebook**

Os valores do Facebook estão descritos na introdução dos Padrões da Comunidade, mas não estão diretamente referenciados nas Diretrizes da Comunidade do Instagram.

A decisão do Facebook de remover o conteúdo do usuário não estava em conformidade com os valores do Facebook. O valor "Voz" claramente inclui discussões relacionadas à saúde e é especialmente importante por incentivar a conscientização dos sintomas de câncer de mama. As imagens de sintomas precoces de câncer de mama são bastante importantes para tornar essa informação médica mais acessível. O compartilhamento dessa informação contribui para a "Segurança" de todas as pessoas vulneráveis a essa doença. Não há indícios de que as fotos contenham algum tipo de conteúdo não consensual. Portanto, o valor "Voz" não foi sobreposto por "Segurança" e "Privacidade", neste caso.

## **8.3 Conformidade com os padrões de direitos humanos internacionais**

### *I. Liberdade de expressão (Artigo 19 do PIDCP)*

A decisão do Facebook de remover a publicação também não estava em conformidade com os padrões de direitos humanos sobre liberdade de expressão (Artigo 19 do PIDCP). As informações relacionadas à saúde são particularmente importantes (A/HRC/44/49, para. 6), além de estarem protegidas como parte de um

direito à saúde (Artigo 12, IESCR; E/C.12/2000/4, para. 11). No Brasil, onde campanhas de conscientização são cruciais para incentivar o diagnóstico precoce do câncer de mama, o Conselho enfatiza a conexão entre estes dois direitos.

O direito à liberdade de expressão não é absoluto. Ao restringir a liberdade de expressão, o Facebook deve atender aos requisitos de legalidade, objetivo legítimo e necessidade e proporcionalidade. A remoção de conteúdo do Facebook não passou nas primeira e terceira partes deste teste.

#### a. Legalidade

As regras que restringem a liberdade de expressão devem ser claras, precisas e acessíveis ao público (Comentário Geral 34, para. 25). Os Padrões da Comunidade do Facebook permitem mamilos femininos no contexto da conscientização do câncer de mama, enquanto as Diretrizes da Comunidade do Instagram mencionam somente as cicatrizes pós-mastectomia. A prevalência dos Padrões da Comunidade do Facebook em relação às Diretrizes da Comunidade também não é informada aos usuários do Instagram. Essa inconsistência e falta de clareza são agravadas por avisos de remoção a usuários somente fazendo referência às Diretrizes da Comunidade. Portanto, as regras do Facebook nessa área não passam no teste de legalidade.

#### b. Objetivo legítimo

As restrições da liberdade de expressão devem ter objetivos legítimos, que estão listados no Artigo 19, para. 3 do PIDCP. O Facebook afirma que seu Padrão da Comunidade sobre Nudez Adulta e Atividade Sexual ajuda a evitar o compartilhamento de imagens de abuso infantil e imagens íntimas não consensuais no Facebook e no Instagram. O Conselho nota que as duas categorias de conteúdo são proibidas de acordo com Padrões da Comunidade separados e não estão sujeitas às exceções que se aplicam à nudez adulta consensual. Esses objetivos são consistentes com a restrição da liberdade de expressão com base no direito internacional dos humanos para proteger "os direitos de outros" (Artigo 19, para. 3, PIDCP). Eles incluem o direito à privacidade de vítimas do compartilhamento não consensual de imagens íntimas (Artigo 17, PIDCP) e os direitos da criança à vida e ao desenvolvimento (Artigo 6, CRC), que são ameaçados em casos de exploração sexual (CRC/C/GC/13, para. 62).

#### c. Necessidade e proporcionalidade

As restrições da liberdade de expressão "devem ser apropriadas para cumprir sua função protetiva; devem ser o instrumento menos intrusivo entre os que podem cumprir sua função protetiva; ao interesse a ser protegido" (Comentário Geral 34, para. 34).

O Conselho acredita que remoção, sem causa, de informações que servem a um interesse público não pode ser proporcional.

O Conselho preocupa-se que o conteúdo tenha sido removido erroneamente por um sistema automatizado de aplicação e possivelmente sem recurso ou análise humana. Isso reflete as limitações de tecnologias automatizadas para entender contextos e captar a complexidade da comunicação humana na moderação de conteúdo (Relatório especial da ONU sobre liberdade de expressão, A/73/348, para. 15). Nesse caso, essas tecnologias não conseguiram reconhecer as palavras "câncer de mama" que são exibidas na parte superior à esquerda da imagem em português. O Conselho reconhece que as tecnologias automatizadas são essenciais para a detecção de conteúdo possivelmente em violação. No entanto, uma aplicação que depende somente da automação, principalmente ao usar tecnologias com habilidade limitada de entender contextos, leva a restrições excessivas que interferem na expressão do usuário de forma desproporcional.

O Conselho reconhece que a aplicação automatizada pode ser necessária para remover rapidamente imagens íntimas não consensuais e imagens de abuso infantil, a fim de evitar prejuízos imediatos e irreparáveis. Entretanto, quando o conteúdo é removido para proteção contra esses danos, a ação deve ser baseada nas políticas aplicáveis sobre exploração sexual, e os usuários devem ser notificados de que o conteúdo foi removido por esses motivos. Ainda assim, as remoções automatizadas devem estar sujeitas a um procedimento de auditoria interna, explicado na seção 9.2 (l) e o recurso para análise humana deve ser oferecido (A/73/348, para. 70), permitindo que erros de aplicação possam ser corrigidos.

A moderação automatizada de conteúdo sem as proteções necessárias não é uma forma proporcional de o Facebook abordar formas de nudez adulta em violação.

#### d. Igualdade e não discriminação

As restrições da liberdade de expressão devem respeitar o princípio de igualdade e não discriminação (Comentário Geral 34, paras. 26 e 32). Vários comentários do

público argumentam que as políticas do Facebook sobre nudez adulta discriminam mulheres.

Dado que as regras do Facebook tratam mamilos masculinos e femininos de forma diferente, depender de uma automação imprecisa para aplicar essas regras provavelmente terá um impacto desproporcional sobre as mulheres, levando a preocupações com a discriminação (Artigo 1, CEDAW; Artigo 2, PIDCP). No Brasil e em muitos outros países, a conscientização dos sintomas do câncer de mama é uma questão de importância crítica. Assim, as ações do Facebook prejudicam não apenas o direito à liberdade de expressão das mulheres, como também seu direito à saúde.

## *II. Direito a um recurso efetivo (Artigo 2, PIDCP)*

O Conselho reconhece que o Facebook restaurou o conteúdo. Porém, os impactos negativos desse erro não podem ser totalmente revertidos. A publicação, referente ao mês de conscientização do câncer de mama, em Outubro, foi restaurada somente no início de dezembro. Restaurar o conteúdo não torna o assunto encerrado: como o Conselho havia selecionado esse caso, o usuário tinha o direito de ser ouvido e receber uma decisão totalmente fundamentada.

O relator especial da ONU sobre a liberdade de opinião e expressão identificou a responsabilidade de oferecer medidas como um dos aspectos mais relevantes de UNGPs, pois estão relacionadas a corporações que se envolvem na moderação de conteúdo (A/HRC/38/35, para. 11). A dependência excessiva de automação na moderação, sem a possibilidade de recurso, resulta em desrespeito ao direito do usuário a um recurso efetivo (Artigo 2, PIDCP; CCPR/C/21/Rev.1/Add. 13, para. 15) e na falha em atender às responsabilidades da empresa no âmbito dos Princípios Orientadores da ONU (Princípios 29 e 31). O Conselho preocupa-se principalmente que o Facebook não informe os usuários quando o conteúdo é aplicado por meio de automação e que o recurso para análise humana não esteja disponível em todos os casos. Isso reflete uma preocupação mais ampla com a falta de transparência do Facebook sobre a automação e as circunstâncias em que o recurso interno pode não estar disponível.

## **9. Decisão do Conselho de Supervisão**

### **9.1 Decisão de conteúdo**

O Conselho de Supervisão reverte a decisão inicial do Facebook de retirar o conteúdo do ar, exigindo que a publicação seja mantida no ar. O Conselho reconhece que o Facebook já tomou medidas para isso.

## **9.2 Declaração consultiva da política**

### *I. Automação na aplicação, transparência e o direito a um recurso efetivo*

O Conselho recomenda que o Facebook:

- Melhore a *detecção automatizada* de imagens com sobreposição de texto para garantir que as publicações de conscientização dos sintomas do câncer de mama não sejam sinalizadas erroneamente para análise.
- Garanta que os usuários sempre sejam notificados dos motivos da verificação de políticas de conteúdo em relação a eles, informando a regra específica do Padrão da Comunidade no qual o Facebook se baseou para tomar sua decisão.
- Informe os usuários quando a automação é aplicada em seu conteúdo, incluindo descrições acessíveis sobre o que ela significa.
- Garanta que os usuários possam abrir recursos sobre as decisões tomadas por sistemas automatizados para análise humana quando o conteúdo é visto como em violação do Padrão da Comunidade do Facebook sobre Nudez Adulta e Atividade Sexual. Como o Facebook está buscando impedir a exploração sexual infantil ou a disseminação de imagens íntimas não consensuais, deve agir de acordo com os Padrões da Comunidade sobre Nudez, Abuso e Exploração Sexual de Adultos e de Crianças, em vez de depender das políticas excessivas em relação à nudez adulta. Os recursos ainda devem estar disponíveis nesses casos para que as remoções incorretas de nudez adulta consensual e permitida possam ser revertidas.
- Implemente um procedimento de auditoria interna para analisar continuamente uma amostra representativa estatística das decisões automatizadas de remoção de conteúdo, para reverter e aprender com erros de aplicação.
- Expanda os relatórios de transparência para divulgar dados sobre o número de decisões automatizadas de remoção de acordo com o Padrão da Comunidade e a proporção das decisões revertidas após a análise humana.

Essas recomendações não devem ser implementadas de forma que prejudique o direito dos moderadores de conteúdo à saúde durante a pandemia da COVID-19.

### *II. A relação entre os Padrões da Comunidade e as Diretrizes da Comunidade:*

O Conselho recomenda que o Facebook:

- Revise a explicação na versão mais curta das Diretrizes da Comunidade do Instagram para esclarecer que a proibição sobre a nudez adulta não é absoluta;
- Revise a explicação na versão mais longa das Diretrizes da Comunidade do Instagram para esclarecer que mamilos femininos visíveis podem ser mostrados para aumentar a conscientização do câncer de mama;
- Esclareça que as Diretrizes da Comunidade do Instagram são interpretadas em alinhamento aos Padrões da Comunidade do Facebook e que os padrões prevalecem quando há inconsistências entre eles.

**\*Nota de procedimento:**

As decisões do Conselho de Supervisão são preparadas por comitês de cinco membros e devem ser aceitas pela maioria do Conselho. As decisões do Conselho não necessariamente representam as opiniões pessoais de todos os membros.