



---

# NOTA TÉCNICA

# PL 2.630/2020

**SOBRE A INCLUSÃO DE MECANISMOS  
DE TRANSPARÊNCIA ALGORÍTMICA  
NO PROJETO DE LEI**



**LAPIN**

**Realização:**

Laboratório de Políticas Públicas e Internet (LAPIN)

**Autoria:**

Felipe Rocha da Silva (felipe@lapin.org.br)

Isabela Maria Rosal Santos (isabela@lapin.org.br)

José Renato Laranjeira de Pereira (joserenato@lapin.org.br)

Mariane Andrade Moreira (mariane@lapin.org.br)

Paulo Henrique Atta Sarmiento (paulo@lapin.org.br)

Thiago Guimarães Moraes (thiago@lapin.org.br)

**Imagem de capa:**

Evgeni Tcherkasski, Unsplash



Este trabalho está licenciado com uma Licença Creative Commons  
Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)  
<https://creativecommons.org/licenses/by-sa/4.0/>

## Sobre esta nota técnica

---

O **Laboratório de Políticas Públicas e Internet - LAPIN** vem, com a presente nota, apresentar sua contribuição ao **Projeto de Lei (PL) 2.630/2020**, que tramita na Câmara dos Deputados. Este documento dá foco a um elemento importante que não se encontra presente no Projeto de Lei sobre Liberdade, Responsabilidade e Transparência na Internet: a **transparência algorítmica**.

A tomada de decisões automatizadas é utilizada por provedores de redes sociais para direcionamento e moderação de conteúdo, ocorrendo, na maioria das vezes, de forma invisível ao usuário. Desta forma, para que a lei possa garantir segurança, ampla liberdade de expressão, comunicação e manifestação do pensamento, a transparência de decisões automatizadas deve fazer parte das regras trazidas pelo Projeto de Lei.

No corpo do documento, trazemos aspectos teóricos a respeito da importância da transparência algorítmica e da explicabilidade para melhor compreensão do funcionamento de serviços prestados *online*. **Ao final, trazemos sugestões de redação** que julgamos oportunas para serem incorporadas pelo Projeto.

## Quem somos nós

---

O **Laboratório de Políticas Públicas e Internet (LAPIN)** é um *think tank* com sede na capital federal brasileira, de composição multidisciplinar e cujo objetivo é apoiar o desenvolvimento de políticas públicas voltadas para a regulação das tecnologias digitais por meio da pesquisa e da conscientização da sociedade.

# SUMÁRIO

---

<b>I - Introdução</b>	<b>5</b>
<b>II - Transparência na personalização de conteúdo</b>	<b>8</b>
<b>III - Transparência na moderação de conteúdo</b>	<b>14</b>
<b>IV - Sugestões para o texto legislativo</b>	<b>18</b>

# I - Introdução

---

Ao navegarmos na Internet, nossas interações deixam rastros no formato de dados. Cada clique que fazemos em uma rede social ou qualquer pesquisa que realizamos em um mecanismo de busca se tornam informações valiosas que, coletadas em ampla quantidade, permitem a criação de perfis sobre nós que serão utilizados para a veiculação de publicidade de produtos e serviços.

Com o aumento da capacidade de processamento de informações, **torna-se cada vez mais fácil extrair detalhes profundos a respeito da vida e personalidade de um indivíduo a partir de dados coletados.** De nossa atividade *online*, é possível identificar se uma mulher está grávida,<sup>1</sup> se um pai de família está desempregado,<sup>2</sup> ou qual partido político é mais provável que você seja filiado.<sup>3</sup> Isso se traduz em recomendações personalizadas em sites de compra, no direcionamento de anúncios publicitários em *sites*, ou no ranqueamento de conteúdo em redes sociais. Em casos extremos, essas ferramentas podem ser utilizadas para a veiculação de informações falsas.

Essa forma de uso de dados pessoais revela um modelo de negócio amplamente explorado, baseado na extração de previsões sobre quem somos e o que desejamos. **O problema é que isso frequentemente não é feito de forma transparente**, a partir de uma descrição adequada sobre quais dados são coletados, quais inferências são feitas deles e qual é o impacto que isso tem naquilo que consumimos.

---

<sup>1</sup> MAYER-SCHONBERGER, Viktor. **Big Data** : como extrair volume, variedade, velocidade e valor da avalanche de informação cotidiana. Tradução Paulo Polzonoff Junior. - 1. ed. - Rio de Janeiro : Elsevier, 2013, p. 40.

<sup>2</sup> *Idem*, p. 95

<sup>3</sup> HOWARD, Philip N. **Lie Machines**: *How to Save Democracy from Troll Armies, Deceitful Robots, Junk News Operations, and Political Operatives*. Yale University Press. New Haven, Londres. 2020.

Casos emblemáticos em que o tratamento de dados pessoais facilitou o direcionamento de conteúdo em campanhas desinformativas foram as operações realizadas pela **Cambridge Analytica** durante as campanhas do Brexit e das eleições estadunidenses de 2016,<sup>4</sup> a perseguição contra os Rohingya no **Mianmar**,<sup>5</sup> ou mesmo as **eleições brasileiras de 2018**. O ponto em comum foi o direcionamento de um amplo volume de informações falsas ou descontextualizadas de forma personalizada para perfis específicos em redes sociais, para favorecer determinado grupo político ou ideologia.<sup>6</sup>

O que vemos *online* é em grande medida fruto de decisões realizadas automaticamente por máquinas baseadas em inteligência artificial a partir de inferências feitas sobre nossa personalidade a partir de como atuamos no ambiente digital. Com base nesses dados, **o conteúdo é personalizado para cada indivíduo, que tem pouca ou nenhuma autonomia sobre esse processo de perfilhamento**. Participar do controle de como conteúdos nos são ofertados passa diretamente por saber quais dados estão sendo coletados e em quais perfis estamos sendo encaixados.

Neste sentido, **a transparência surge como conceito fundamental para que usuários saibam quais dados pessoais estão sendo coletados e quais inferências estão sendo feitas a partir deles**. Isso não tem só um aspecto ético, no sentido de que não sejamos feitos de meros receptores passivos de informação, mas também um aspecto jurídico, já que a transparência é um princípio positivado na Lei Geral de Proteção de Dados - **LGPD**,<sup>7</sup> que fundamenta

---

<sup>4</sup> THE GUARDIAN. **What is the Cambridge Analytica scandal?** - video explainer. 19 mar 2018.

<sup>5</sup> HAHM, Jasmine. **Facebook: Myanmar's Misinformation Megaphone**. 23 dez 2018. Disponível em: <https://bpr.berkeley.edu/2018/12/23/facebook-myanmars-misinformation-megaphone/>. Acesso em 11 ago 2020.

<sup>6</sup> EVANGELISTA, R. & BRUNO, F. **WhatsApp and political instability in Brazil**. Internet Policy Review, Vol 1, Issue 4. 31 Dez 2019. Disponível em <https://policyreview.info/articles/analysis/whatsapp-and-political-instability-brazil-targeted-messages-and-political>. Acesso em 11 ago 2020.

<sup>7</sup> Art. 6º As atividades de tratamento de dados pessoais deverão observar a boa-fé e os seguintes princípios:

VI - transparência: garantia, aos titulares, de informações claras, precisas e facilmente acessíveis sobre

o exercício de direitos tais como os de acesso, de correção e de eliminação dos dados pessoais.

Assim, **a transparência se funda como mecanismo essencial ao funcionamento da democracia**, ao garantir ferramentas para que um usuário ganhe controle sobre como decisões automatizadas afetam sua percepção da realidade por meio da personalização de conteúdo. Com isso em vista, esta nota técnica pretende analisar ferramentas para que redes sociais ofereçam maior **transparência algorítmica** e, portanto, garantam maior autonomia ao indivíduo para que direitos como liberdade de expressão, proteção de dados, privacidade e autodeterminação informativa lhe sejam assegurados.

---

a realização do tratamento e os respectivos agentes de tratamento, observados os segredos comercial e industrial;

## II - Transparência na personalização de conteúdo

---

Plataformas como redes sociais, mecanismos de busca e serviços por *streaming* aplicam sistemas de personalização de conteúdo para ofertar informações que julguem mais relevantes para cada usuário.

A partir da coleta de dados sobre o indivíduo e sobre suas interações com a plataforma, como pelo uso de palavras-chave específicas ou por cliques em um botão “curtir”, algoritmos são capazes de criar perfis de pessoas e, a partir deles, selecionar quais informações direcionar para um determinado indivíduo.<sup>8</sup> Com base nesses perfis, é possível ter uma visão ampla acerca das atitudes, preferências e comportamentos de um indivíduo.<sup>9</sup>

No entanto, ao veicular conteúdo que diz respeito somente aos interesses de uma pessoa, acaba-se por restringir a diversidade de ideias às quais ela tem acesso, criando as chamadas **câmaras de eco**, em que indivíduos têm acesso somente ao conteúdo formado por usuários que pensam de forma semelhante, retroalimentando um sistema que funciona sempre na base das mesmas visões de mundo.<sup>10</sup>

O objetivo dessas técnicas é frequentemente antecipar ou mesmo influenciar caminhos, comportamentos, preferências e ações de indivíduos.<sup>11</sup> Por isso, quando exploradas com fins políticos, podem afetar as bases de um

---

<sup>8</sup> BOZDAG, Engin. **Bias in algorithmic filtering and personalization**. *Ethics and Information Technology*, 15(3), p. 211. 23 Jun 2013. Disponível em <http://doi.org/10.1007/s10676-013-9321-6>. Acesso em 11 ago 2020.

<sup>9</sup> HOWARD, Philip N. **Lie Machines: How to Save Democracy from Troll Armies, Deceitful Robots, Junk News Operations, and Political Operatives**. Yale University Press. New Haven, Londres. 2020.

<sup>10</sup> MITTELSTADT, Brent. **Auditing for Transparency in Content Personalization Systems**. *International Journal of Communication* 10(2016), 4991-5002. Disponível em <https://www.ijoc.org/index.php/ijoc/article/view/6267>. Acesso em 11 ago 2020.

<sup>11</sup> EDPS Ethic Advisory Group. **Towards a Digital Ethics. Report 2018**. Disponível em [https://edps.europa.eu/sites/edp/files/publication/18-01-25\\_eaq\\_report\\_en.pdf](https://edps.europa.eu/sites/edp/files/publication/18-01-25_eaq_report_en.pdf).



sistema democrático, já que limitam o acesso do usuário a informações adicionais que venham a questionar crenças já consolidadas.

Como consequência, **a personalização de conteúdo acaba por frequentemente criar novas barreiras para a auto-reflexão crítica** sobre nossas próprias visões de mundo, ao impedir que tenhamos acesso a informações que questionem nossas opiniões.<sup>12</sup> O efeito disso tem sido a radicalização de populações em democracias mais ou menos consolidadas ao redor do mundo. Com base em dados pessoais obtidos através de plataformas de redes sociais, equipes de publicidade política conseguem criar conteúdo personalizado para cada público com base em sua suscetibilidade a ser influenciado pela informação.

**Isso abre margem, inclusive, para a veiculação de conteúdo desinformativo.** Durante as eleições de 2018 no Brasil, isso foi feito amplamente dentro do Whatsapp. Indivíduos eram segregados em diferentes grupos de apoiadores de determinado candidato ou ideologia de acordo com critérios como sua visão política, seu histórico de saúde, sua classe social ou sua localidade geográfica.

Em muitos desses grupos, mensagens com conteúdo falso eram disseminadas e compartilhados com muito maior frequência do que outro tipo de conteúdo.<sup>13</sup> Em alguns casos, inclusive, colaboradores de campanhas políticas se infiltravam em grupos que a princípio não tinham objetivo político, como grupos de bairro ou de pessoas que compartilhavam informações sobre determinada doença para veicular essas informações.<sup>14</sup>

---

<sup>12</sup> MITTELSTADT, Op Cit., p. 4994.

<sup>13</sup> RESENDE, G. et al. **(Mis)information Dissemination in WhatsApp: Gathering, Analysis and Countermeasures**. WWW '19: The World Wide Web Conference. Disponível em <https://dl.acm.org/doi/10.1145/3308558.3313688>. Acesso em 11 ago 2020.

<sup>14</sup> AUDI, A. & DIAS, T. **Video: Seu Número de Telefone Vale 9 Centavos no Zap dos Políticos**. The Intercept, 22 out 2018. Disponível em <https://theintercept.com/2018/10/22/whatsapp-politicos/>

Várias dessas mensagens possuíam conteúdo falso, sensacionalista, extremista, conspiracionista, profundamente enviesado ou comentário mascarado como notícia, elementos que o professor da Universidade de Oxford Philip Howard chama de *junk news*<sup>15</sup> (que pode ser traduzido como *notícias lixo*). O pior é que os participantes desses grupos muitas vezes não sabiam que seus dados pessoais estavam sendo tratados para veiculação desse tipo de conteúdo.

O mesmo acontece em outras plataformas, sejam elas Google, YouTube, Facebook, Twitter ou TikTok. No entanto, além de não serem sempre transparentes quanto aos perfis que utilizam para enquadrar seus usuários, sua opacidade também é muitas vezes refletida na falta de informações relevantes sobre o processo de tomada de decisão automatizada utilizada tanto no perfilamento quanto na recomendação de conteúdo.

**Por isso, é importante que esses provedores tornem mais transparente a funcionalidade de seus sistemas de recomendação de conteúdo por decisão automatizada.** Isso garantiria tanto ao usuário quanto a pesquisadores e reguladores uma melhor noção a respeito de como os algoritmos de personalização e microdirecionamento influenciam a veiculação de informações.

**A LGPD já pressupõe mecanismos de transparência ativa**, pela qual o titular de dados tem o direito de requerer “informações claras e adequadas a respeito dos critérios e dos procedimentos utilizados para a decisão automatizada” (art. 20, §1º). **No entanto, falta a previsão de uma transparência passiva**, pela qual o controlador de dados (como é o caso de uma plataforma) informa o usuário sobre a existência de decisões automatizadas e como elas ocorrem. Afinal, se a pessoa nunca é informada passivamente que decisões

---

<sup>15</sup> HOWARD, 2020, Op Cit.

automatizadas estão ocorrendo, como ela vai tomar a iniciativa de solicitar alguma revisão delas?

O Regulamento Geral de Proteção de Dados europeu - RGPD traz regras nesse sentido. Em seus artigos 13(2)(f) e 14(2)(g), a norma determina que as seguintes informações devem ser prestadas de forma espontânea pelo controlador de dados, independentemente de requisição:

1. A existência de decisões automatizadas;
2. Informações relevantes sobre a lógica das tomadas de decisão automatizadas;
3. Impacto e possíveis consequências dessas tomadas de decisão.

É importante também que, especialmente em redes sociais, seja tornado mais claro ao usuário quais as principais características que são identificadas de sua personalidade no processo de perfilização. Uma forma de apresentar esse tipo de informação é pela revelação das *tags* ou categorias que o sistema lhe encaixa para direcionamento de conteúdo.

Outros aspectos de grande importância para melhor compreensão tanto da lógica de um sistema quanto de uma decisão específica são:<sup>16</sup>

1. Quais os principais fatores (*inputs*) tomados em conta na tomada de decisão automatizada - como a localização de um sujeito, sua raça, o fato de participar de determinado grupo ou se interagiu recentemente com um conteúdo específico;
2. Qual foi o peso de cada um desses fatores na tomada de decisão;
3. Se algum desses elementos pode receber pesos distintos em decisões automatizadas semelhantes.

---

<sup>16</sup> DOSHI-VELEZ, Finale et al. **Accountability of AI Under the Law: The Role of Explanation**. Berkman Klein Center Working Group on Explanation and the Law, Berkman Klein Center for Internet & Society working paper, 2017, p. 3. Disponível em [nrs.harvard.edu/urn-3:HUL.InstRepos:34372584](https://nrs.harvard.edu/urn-3:HUL.InstRepos:34372584). Acesso em 11 ago 2020.

Essas questões são essenciais para **garantir que saibamos quais fatores foram determinantes para a tomada de uma decisão automatizada**, e consequentemente identificar os vieses de um sistema algorítmico de personalização de conteúdo, tanto para permitir a compreensão de por que determinada informação foi mostrada a um usuário quanto para compreender quais informações não estão sendo mostradas.

Vale ressaltar, no entanto, que **nem toda decisão deverá ser explicada para o usuário, sob pena de tornar o oferecimento do serviço inviável**. Devem ser explicadas as decisões que:

1. Tenham um impacto sobre os direitos do indivíduo;
2. Sirvam para comprovar a responsabilização da plataforma por determinado dano causado; e
3. Possuam fundados indícios de terem sido feitas de forma equivocada.<sup>17</sup>

**As informações aqui listadas devem ser apresentadas ao usuário individualmente**, já que envolvem o tratamento de dados pessoais que devem ser protegidos, em termos de simples compreensão, **e também em relatórios de transparência**.

É importante ainda que venham acompanhadas de um juízo de quais informações são mais relevantes para serem abertas de modo a alcançar o objetivo final desse mecanismo de transparência: compreender como nossas informações pessoais afetam aquilo que acessamos e como garantir autonomia do titular de dados sobre quais informações suas estão sendo tratadas.<sup>18</sup> Vale chamar atenção que isso não significa a abertura de segredos comerciais por uma organização.

---

<sup>17</sup> *Idem*, p. 4.

<sup>18</sup> *Idem*, p. 9.

Dada a complexidade do tema, e considerando a previsão, no Projeto de Lei n. 2.630/2020, de um **Conselho de Transparência** que terá dentre suas competências a de regulamentar o cumprimento de normas de eficácia limitada do Projeto, **recomenda-se que este órgão seja responsável por definir detalhes a respeito de quais decisões automatizadas deverão ser explicáveis, e quais métodos devem ser utilizados.** Afinal, considerando o rápido desenvolvimento tecnológico de nossa era, é essencial que regulamentações específicas a respeito do funcionamento de algoritmos sejam realizadas por instrumentos mais céleres do que uma lei, com a participação de especialistas no tema e **baseadas em normas de caráter principiológico.** Trazemos sugestões na Seção IV desta nota.

### III - Transparência na moderação de conteúdo

---

Redes sociais têm um desafio enorme na tarefa de garantir que usuários cumpram os princípios éticos previstos em seus termos de uso quanto à publicação de conteúdo em suas plataformas: como garantir que um controle eficiente possa ser realizado quando milhares de mensagens são publicadas a todo momento?

Em um relatório produzido pela *Cambridge Consultants*,<sup>19</sup> a pedido da autoridade reguladora britânica de comunicações, OfCom, foi demonstrado que **o monitoramento do conteúdo veiculado nas plataformas seria impossível sem o auxílio de tecnologias de inteligência artificial**, que realizam atividades de moderação *ex ante* (i.e. antes do conteúdo ser publicado na rede) e *ex post* (i.e. depois do conteúdo ser publicado na rede) para auxiliar os provedores de redes sociais.

A **pré-moderação** automatizada (ou semi-automatizada) é utilizada para a **remoção de conteúdo ilegal** ou que conflita diretamente com as políticas da plataforma, sendo comumente usada para detectar conteúdo "claramente" ilegal, como material de abuso infantil ou violência explícita. Esta atividade é muito importante para evitar que conteúdo impróprio seja removido antes mesmo que uma denúncia seja necessária, visto que, em muitos casos, o simples ato de publicização de uma imagem ou texto violento é capaz de causar danos a indivíduos. Este é um dos motivos pelo qual empresas como o Facebook investem em aprimoração de tecnologias que permitam a remoção de conteúdo impróprio antes de sua publicação.<sup>20</sup> Essas tecnologias podem realizar a

---

<sup>19</sup> CAMBRIDGE CONSULTANTS. **Use of AI in online content moderation**. Ofcom, 2018. Disponível em [https://www.ofcom.org.uk/\\_\\_\\_data/assets/pdf\\_file/0028/157249/cambridge-consultants-ai-content-moderation.pdf](https://www.ofcom.org.uk/___data/assets/pdf_file/0028/157249/cambridge-consultants-ai-content-moderation.pdf). Acesso em 11 ago 2020.

<sup>20</sup> FACEBOOK. **FB 2018: Como a tecnologia nos ajuda a remover conteúdo impróprio antes mesmo de ser denunciado**. 2018. Disponível em

remoção com ou sem auxílio de um humano, a depender da técnica implementada.

Já a **pós-moderação** é usualmente semi-automatizada, ou seja, necessita da intervenção de uma equipe humana para confirmar a remoção de conteúdo impróprio.<sup>21</sup> Geralmente, o conteúdo removido *ex post* é aquele que **depende da análise de elementos contextuais**, tais como pontos de vista políticos, crenças culturais, eventos históricos e leis locais, e dificilmente podem ser removidos automaticamente sem trazer riscos à liberdade de expressão.

Por exemplo, em 2016, o Facebook sofreu duras críticas por ter removido uma famosa foto em que aparece uma menina nua gritando e correndo após um ataque de napalm atingir o Vietnã em 1972.<sup>22</sup> Diante da pressão social, a empresa se retratou e tornou a disponibilizar a imagem. Embora riscos de falso positivo existam, tecnologias de inteligência artificial podem ser utilizadas para auxiliar operadores humanos, sinalizando publicações que possivelmente apresentam conteúdo impróprio, para que decidam quanto à necessidade de remoção da imagem ou texto sinalizado.

Embora tecnologias de automatização sejam importantes para a moderação de conteúdo, um dos grandes problemas é que, em muitos casos, não há informações precisas sobre essas atividades, isto é - **não se informa quais tipos de conteúdo são removidos, quantos conteúdos são removidos de forma automatizada ou quando as informações foram retiradas**. As informações prestadas aos usuários responsáveis pelo conteúdo também são frequentemente vagas e impedem uma compreensão mais detalhada do motivo por trás dessas remoções.

---

<https://about.fb.com/br/news/2018/05/f8-2018-como-a-tecnologia-nos-ajuda-a-remover-conteudo-improprio-antes-mesmo-de-ser-denunciado/>. Acesso em 11 ago 2020.

<sup>21</sup> CAMBRIDGE CONSULTANTS. Op cit.

<sup>22</sup> G1. **Facebook apaga foto da Guerra do Vietnã e premiê da Noruega republica**. 2016. Disponível em <http://g1.globo.com/tecnologia/noticia/2016/09/facebook-apaga-foto-da-guerra-do-vietna-e-premie-da-noruega-republica.html/>. Acesso em 11 ago 2020.

Um exemplo de prática de transparência, mas que tem espaço para ser aprimorado, são os relatórios de cumprimento das diretrizes de comunidade do Youtube, que informa a quantidade de vídeos removidos em um determinado período, e destaca os casos de remoções automatizadas.<sup>23</sup> O relatório mais recente, referente ao período de Janeiro a Março de 2020, revela que mais de 6.1 milhões de vídeos foram removidos, dos quais 5.7 milhões (93%) foram por detecção automatizada. Infelizmente, o relatório não faz a distinção entre remoções automatizadas e não-automatizadas na categoria de motivos, o que permitiria identificar quais tipos de conteúdo são mais comumente removidos de forma automática.

Deste modo, é fundamental que uma Lei de Transparência de plataformas de redes sociais reconheça a importância da **transparência de decisões automatizadas**, seja no momento de notificação do indivíduo, seja na elaboração de relatórios. Devem existir regras quanto à necessidade de que essas informações sejam reveladas em relatórios de transparência e usuários sejam notificados quando a remoção de conteúdo ocorrer de forma automatizada.

Da mesma forma que ressaltamos na Seção II, qualquer regulamentação legal do tema deve ser de ordem **principiológica e tecnologicamente neutra**, permitindo que cada provedor de rede social se adapte às regras tendo em vista as especificidades de seu funcionamento. É importante também que, nos relatórios de transparência previstos no escopo do Projeto, conste informações a respeito de como e em quais proporções decisões automatizadas atuaram nos processos de moderação de conteúdo das plataformas.

Por fim, ações específicas poderão vir a ser estabelecidas pelo Conselho de Transparência. Essa atuação foi amplamente discutida durante o Ciclo de

---

<sup>23</sup> GOOGLE. **Cumprimento das diretrizes da comunidade do YouTube**. 2020. Disponível em [https://transparencyreport.google.com/youtube-policy/removals?content\\_by\\_faq=period:Y2018Q3;exclude\\_%20automated:&lu=total\\_removed\\_videos&total\\_removed\\_videos=period:Y2020Q1;exclude\\_automated:all](https://transparencyreport.google.com/youtube-policy/removals?content_by_faq=period:Y2018Q3;exclude_%20automated:&lu=total_removed_videos&total_removed_videos=period:Y2020Q1;exclude_automated:all). Acesso em 11 ago 2020.



Debates Públicos para aprimorar o PL nº 2630/20, em especial no Painel 3, sobre Moderação de Conteúdos e Liberdade de Expressão. Na ocasião, os painelistas ressaltaram a importância do Conselho para definir regras de conduta, competência legitimada pela composição multissetorial do órgão.

## IV - Sugestões para o texto legislativo

Tendo em vista o exposto, apresentamos as seguintes sugestões de texto legislativo, tanto para aspectos relacionados ao direcionamento de conteúdo personalizado, quanto para mecanismos de moderação.

### a. Transparência no processo de personalização de conteúdo

Sugestão de redação	Justificativa
<p>Art. 14-A - Os provedores de redes sociais deverão informar o usuário a respeito da existência de decisões automatizadas, incluindo a definição de perfis, e garantir acesso a informações relativas ao funcionamento do sistema responsável por essas decisões, bem como o impacto e as consequências previstas de tal tratamento para o usuário.</p>	<p>O dispositivo prevê obrigação de uma transparência passiva, pela qual o controlador de dados (como é o caso de uma plataforma) informa o usuário sobre a existência de decisões automatizadas e como elas ocorrem.</p>
<p>§1º Serão disponibilizadas ao usuário informações individualizadas sobre os critérios e os dados pessoais utilizados para definição de seu perfil.</p>	<p>Esse dispositivo prevê mecanismos de explicabilidade por parte da plataforma a respeito dos critérios utilizados para a criação de perfis (<i>profiling</i>) e quais as características subsumidas sobre cada indivíduo.</p>
<p>§2º O provedor de redes sociais deverá disponibilizar ferramentas para que o</p>	<p>Muitas plataformas criam perfis de usuários a partir de dados pessoais que</p>

<p>usuário acesse as informações sobre sua personalidade que foram identificadas por sistemas de decisões automatizadas e para requerer correção e se opor ao tratamento de seus dados pessoais por esses sistemas.</p>	<p>fundamentarão a personalização de conteúdo, mas não lhes dão a possibilidade de corrigir ou atualizar os dados que sustentam essa perfilização. Esse dispositivo prevê que o usuário deverá ter acesso às informações que o sistema deduziu sobre sua personalidade, bem como a meios adequados a requerer modificações a respeito dessas informações.</p>
---	---

#### **b. Transparência na moderação de conteúdo**

<b>Texto original</b>	<b>Sugestão de redação</b>	<b>Justificativa</b>
N/A	<p>Art. 12, §7º Em caso de remoção de conteúdo por decisão automatizada, o usuário deverá ser notificado sobre a existência da automatização e o motivo da remoção.</p>	<p>O art. 12, sobre moderação de conteúdo, deve conter natureza principiológica e evitar estabelecer as categorias de conteúdo que devem ser removidas. O rol apresentado no §2º traz casos típicos de remoção automatizada. Por isso, sugere-se que ele seja substituído por previsão que estabeleça a obrigação</p>

		de informar o usuário que teve seu conteúdo removido por decisão automatizada sobre este fato e o motivo da remoção. Esta informação é necessária para garantir ao usuário exercer seu direito de revisão da decisão automatizada, nos termos do art. 20 da LGPD.
Art. 12, §3º Deve ser garantido pelo provedor o direito de o usuário recorrer da indisponibilização de conteúdo e contas.	Art. 12, §3º Deve ser garantido pelo provedor o direito de o usuário recorrer da indisponibilização de conteúdo e contas, <b>e de solicitar que o recurso seja revisado por operador humano.</b>	O usuário que recorre sobre remoção de conteúdo deve ter o direito de ter seu recurso avaliado por um operador humano, garantindo que a nova decisão a respeito do conteúdo removido não seja realizada por um sistema automatizado.
Art. 13, §1º Os relatórios devem conter, no mínimo: [...] II - número total de medidas de moderação de contas e conteúdos adotadas em razão do	Art. 13, §1º Os relatórios devem conter, no mínimo: [...] II - número total de medidas de moderação de contas e conteúdos adotadas em razão do	Os relatórios de transparência, no que diz respeito a informações sobre moderação de conteúdo, deve discriminar quando técnicas de moderação automatizada

<p>cumprimento dos termos de uso privados dos provedores de redes sociais, especificando sua motivação e metodologia utilizada na detecção da irregularidade e o tipo de medida adotada;</p>	<p>cumprimento dos termos de uso privados dos provedores de redes sociais, especificando sua motivação e metodologia utilizada na detecção da irregularidade e o tipo de medida adotada, <b>em particular, se mecanismos de decisão automatizada foram utilizadas e em que proporção;</b></p>	<p>foram utilizadas e a quantidade de conteúdo removido por decisões automatizadas.</p>
<p>III - número total de medidas de moderação de contas adotadas em razão do cumprimento desta Lei, especificando sua motivação e metodologia utilizada na detecção da irregularidade e o tipo de medida adotada;</p>	<p>III - número total de medidas de moderação de contas adotadas em razão do cumprimento desta Lei, especificando sua motivação e a metodologia utilizada na detecção da irregularidade e o tipo de medida adotada, <b>em particular, se mecanismos de decisão automatizada foram utilizadas e em que proporção;</b></p>	<p>Os relatórios de transparência, no que diz respeito a informações sobre moderação de conteúdo, deve discriminar quando técnicas de moderação automatizada foram utilizadas e a quantidade de conteúdo removido por decisões automatizadas.</p>