

# Contribución al Consejo Asesor de Contenido de Meta sobre la expansión del programa de notas de la comunidad fuera de Estados Unidos.

Diciembre, 2025.

Esta contribución es presentada por Derechos Digitales<sup>1</sup>, una organización no gubernamental latinoamericana que desde 2005 trabaja para la protección y garantía de los derechos humanos en entornos digitales.

## 1. Sobre el caso

El 7 de enero de 2025, Meta anunció el fin del programa global de verificación de hechos por terceros (*third-party fact-checking* en inglés), inicialmente sólo en Estados Unidos, reemplazándolo por un sistema de notas de la comunidad similar al utilizado por X (antes Twitter)<sup>2</sup>. A través de un video de cinco minutos<sup>3</sup>, Mark Zuckerberg confirmó el final -a casi 9 años- del programa de verificación de hechos por terceros el cual, desde 2016, delegaba a expertos independientes como organizaciones de medios y periodistas<sup>4</sup> la identificación, contextualización y profundización sobre contenidos potencialmente engañosos.

En contraste, el sistema de notas de la comunidad deja en manos de personas usuarias añadir notas de contexto a publicaciones que consideren engañosas, dejando la responsabilidad de combatir la desinformación en manos de lo que expertos en el tema han llamado la “sabiduría de las masas”<sup>5</sup>. Según Meta, las notas de la comunidad, como producto, se encuentran en una etapa temprana de desarrollo, aplicando por el momento sólo para Estados Unidos. Zuckerberg justificó el tránsito a este modelo de gestión de contenido en nombre de la “libertad de expresión” acusando a las personas “verificadoras” de la excesiva censura y el sesgo ideológico y político en sus plataformas<sup>6</sup>. No obstante, es clave recordar que los equipos de verificación de hechos por terceros no eliminaban contenido, ni sancionaban cuentas de personas usuarias; la decisión sobre qué contenidos eliminar o no siempre ha recaído en Meta con base en sus Normas Comunitarias<sup>7</sup>.

En respuesta a la solicitud de Meta al Consejo Asesor de Contenido sobre qué factores considerar al momento de incluir o no un país en la expansión de implementación de las notas de la comunidad, en esta contribución presentamos evidencia sobre, por un lado, las fallas y la ineficacia, ya documentadas a partir del caso de X, del sistema de notas de la comunidad para hacer frente a la desinformación y la diseminación de contenidos engañosos en redes sociales. Y por otro lado, argumentamos que en países del Sur Global, como los de América Latina, donde los esfuerzos estatales para abordar la desinformación han sido insuficientes, heterogéneos y desequilibrados ante los múltiples intereses en juego<sup>8</sup>; y donde en los últimos años ha persistido el silenciamiento de voces críticas junto con la restricción de la libre expresión de comunidades históricamente marginalizadas mediante marcos normativos ambiguos<sup>9</sup>, un sistema como el de las notas comunitarias de Meta incrementa el riesgo de un aumento significativo del abuso, la violencia, la desinformación y la manipulación en las redes sociales. Afirmamos que la expansión de las notas de la comunidad puede dejar desprotegidas

---

<sup>1</sup> Más información en: <https://www.derechosdigitales.org/>

<sup>2</sup> Derechos Digitales (2025). Meta elimina la verificación de datos en Estados Unidos: los peligros para el ecosistema digital. Disponible en: <https://www.derechosdigitales.org/noticias/meta-elimina-la-verificacion-de-datos-en-estados-unidos-los-peligros-para-el-ecosistema-digital/>

<sup>3</sup> Ver: <https://www.facebook.com/watch/?v=1525382954801931>

<sup>4</sup> Poynter (2025). Meta is ending its third-party fact-checking partnership with US partners. Here’s how that program works. Disponible en: <https://www.poynter.org/fact-checking/2025/meta-ends-fact-checking-community-notes-facebook/>

<sup>5</sup> Allen J, Arechar AA, Pennycook G, et al. (2021) Scaling up fact-checking using the wisdom of crowds. *Science Advances American Association for the Advancement of Science* 7(36). [10.1126/sciadv.abf4393](https://doi.org/10.1126/sciadv.abf4393)

<sup>6</sup> Incluso, llegó a afirmar la existencia de cortes secretas en América Latina que podían ordenar a las empresas silenciosamente la eliminación de posts: <https://www.facebook.com/watch/?v=1525382954801931>

<sup>7</sup> Ver: <https://transparency.meta.com/policies/community-standards/>

<sup>8</sup> Camacho, O.L. (2025). Meta y la desinformación: el fin de la verificación independiente. Disponible en: <https://razonpublica.com/meta-la-desinformacion-fin-la-verificacion-independiente/>

<sup>9</sup> Lara, J.C. (2025). La censura evoluciona: ¿estamos al día en la defensa de la libre expresión? Disponible en: <https://www.derechosdigitales.org/recursos/la-censura-evolucion-a-estamos-al-dia-en-la-defensa-de-la-libre-expresion/>

a millones de personas que se apoyan en la información en línea para participar en la vida democrática y social de sus comunidades<sup>10</sup>.

## 2. Antecedentes y contexto político del sistema de notas comunitarias de Meta.

En esta contribución, sostenemos que los mecanismos y las políticas de moderación de contenido, lejos de ser meros aspectos técnicos, se insertan, reflejan y son moldeados por el contexto social, político e institucional en el cual de adoptan; así lo han confirmado estudios sobre el tema<sup>11</sup>. El anuncio de Meta sobre el fin de la verificación independiente de hechos se dio en un contexto de creciente autoritarismo a nivel global y local<sup>12</sup>. En consonancia con ello, la empresa ha impulsado otras medidas como la eliminación de las políticas DEI (*Diversity, Equity and Inclusion*, en inglés)<sup>13</sup>, la reducción de restricciones en temas sensibles relacionados a migración e identidad de género, la reintroducción de contenido político en los “feeds” de las personas usuarias<sup>14</sup>, y la suma de nuevos miembros alineados con el conservadurismo político en la junta directiva de la compañía<sup>15</sup>.

Dichas decisiones coincidieron con una serie de directrices del ejecutivo estadounidense referentes a las políticas y los programas de diversidad, equidad e inclusión de las empresas<sup>16</sup>, el cese de los esfuerzos oficiales para frenar la desinformación<sup>17</sup>, y un tumulto de órdenes ejecutivas y pronunciamientos que amenazan los derechos humanos de las mujeres, las diversidades y personas migrantes, así como las acciones para combatir el cambio climático<sup>18</sup>. Lo anterior da cuenta, por un lado, del incumplimiento por parte de Meta de los Principios Rectores de las Naciones Unidas sobre las Empresas y los Derechos Humanos (UNGPs)<sup>19</sup>; cuyos pilares recaen en el respeto a los derechos humanos y a la diversidad en los entornos digitales. Por otro lado, evidencia la maleabilidad e inestabilidad de sus políticas y Normas Comunitarias, condicionadas por tensiones políticas e intereses económicos <sup>20</sup>.

Es importante destacar que no es la primera vez que Meta ajusta estratégicamente sus mecanismos de moderación de contenido según el contexto sociopolítico<sup>21</sup>, incluso adoptando posiciones diametralmente opuestas en distintos momentos. De hecho, la verificación de hechos por terceros surgió en 2016 tras las acusaciones de que el algoritmo de Facebook había influido en el Brexit y la primera elección presidencial de Donald Trump<sup>22</sup>, cuyas políticas eran criticadas por Zuckerberg en contraste con su relación de cooperación actual<sup>23</sup>. Para ese entonces, Meta estableció restricciones estrictas sobre el contenido político patrocinado y se posicionó de forma contundente, anunciando: “nos tomamos muy en serio la desinformación (...) Hemos logrado avances importantes, pero todavía queda mucho trabajo por hacer”.

---

<sup>10</sup> Derechos Digitales (2025). Meta elimina la verificación de datos en Estados Unidos: los peligros para el ecosistema digital. Disponible en: <https://www.derechosdigitales.org/noticias/meta-elimina-la-verificacion-de-datos-en-estados-unidos-los-peligros-para-el-ecosistema-digital/>

<sup>11</sup> Matamoros-Fernández, A., Jude, N. (2025). The importance of centering harm in data infrastructures for ‘soft moderation’: X’s Community Notes as a case study. *New Media & Society*, 27(4). <https://doi.org/10.1177/14614448251314399>

<sup>12</sup> Lara-Castro, P. (2025). El rol de las Big Tech en el auge del autoritarismo. Disponible en: <https://www.derechosdigitales.org/recursos/el-rol-de-las-big-tech-en-el-auge-del-autoritarismo/>

<sup>13</sup> The Guardian (2025). Meta terminates its DEI programs days before Trump inauguration. <https://www.theguardian.com/us-news/2025/jan/10/meta-ending-dei-program>

<sup>14</sup> Ver: BBC (2025). Politics content to be pushed on all Instagram and Threads users. Disponible en: [https://www-bbc-com.translate.goog/news/articles/clyjyd0297go?\\_x\\_tr\\_sl=en&\\_x\\_tr\\_tl=es&\\_x\\_tr\\_hl=es&\\_x\\_tr\\_pto=tc&\\_x\\_tr\\_hist=true](https://www-bbc-com.translate.goog/news/articles/clyjyd0297go?_x_tr_sl=en&_x_tr_tl=es&_x_tr_hl=es&_x_tr_pto=tc&_x_tr_hist=true)

<sup>15</sup> Derechos Digitales (2025). Meta elimina la verificación de datos en Estados Unidos: los peligros para el ecosistema digital. Disponible en: <https://www.derechosdigitales.org/noticias/meta-elimina-la-verificacion-de-datos-en-estados-unidos-los-peligros-para-el-ecosistema-digital/>

<sup>16</sup> Ver: <https://www.whitehouse.gov/presidential-actions/2025/01/ending-illegal-discrimination-and-restoring-merit-based-opportunity/>

<sup>17</sup> Ver: <https://www.whitehouse.gov/presidential-actions/2025/01/restoring-freedom-of-speech-and-ending-federal-censorship/>

<sup>18</sup> Ver a manera de ejemplo: New York Times (2025). Esto es lo que el presidente Trump ordenó cambiar en su primera semana. Disponible en: <https://www.nytimes.com/es/2025/01/25/espanol/estados-unidos/trump-medidas-primera-semana.html>

<sup>19</sup> Ver: [https://www.ohchr.org/sites/default/files/documents/publications/guidingprinciplesbusinesshr\\_en.pdf](https://www.ohchr.org/sites/default/files/documents/publications/guidingprinciplesbusinesshr_en.pdf)

<sup>20</sup> Lara-Castro, P. (2025). El rol de las Big Tech en el auge del autoritarismo. Disponible en: <https://www.derechosdigitales.org/recursos/el-rol-de-las-big-tech-en-el-auge-del-autoritarismo/>

<sup>21</sup> Marinangeli, J.M. (2025). Notas de la comunidad en Meta: una nota sobre su comunidad. Disponible en: <https://observatoriolegislativocele.com/notas-de-la-comunidad-en-meta-una-nota-sobre-su-comunidad-por-juan-martin-marinangeli/>

<sup>22</sup> Camacho, O.L. (2025). Meta y la desinformación: el fin de la verificación independiente. Disponible en: <https://razonpublica.com/meta-la-desinformacion-fin-la-verificacion-independiente/>

<sup>23</sup> El País (2025). La metamorfosis trumpista de Mark Zuckerberg. Disponible en: <https://elpais.com/tecnologia/2025-01-12/la-metamorfosis-trumpista-de-mark-zuckerberg.html>

En este escenario, la apuesta actual de Meta por las notas comunitarias combinada con la flexibilización de sus políticas de moderación sobre contenido relacionado con género, diversidad y migración habilita un entorno especialmente propicio para la proliferación de discursos de odio, anti derechos y narrativas discriminatorias. Estos, han sido históricamente utilizados para silenciar a determinados grupos, con consecuencias profundamente dañinas. Como resultado, comunidades como las mujeres en toda su diversidad, las personas migrantes y otras que han sido históricamente marginalizadas podrían enfrentar un aumento de ataques digitales que normalicen la violencia y perpetúen la discriminación en su contra<sup>24</sup>.

### 3. La ineficacia de las notas de la comunidad para combatir la desinformación, los aprendizajes del caso de X.

El modelo de notas de la comunidad de X (antes Twitter), cuyo algoritmo y principio de funcionamiento están en la base del mecanismo de notas comunitarias de Meta<sup>25</sup>, ha sido estudiado ampliamente en la medida que promete oportunidades para que las personas compartan información y participen en el discurso público. Sin embargo, es un modelo que también crea un entorno donde las personas usuarias pueden interactuar fácilmente con la desinformación masiva, lo que tiene impactos negativos en la sociedad y el debate democrático en su conjunto.

La evidencia disponible sobre este enfoque de moderación de contenido ha mostrado que las notas de la comunidad de X no han tenido un efecto significativo en reducir la interacción de las personas usuarias con los contenidos engañosos<sup>26</sup>, y han sido ineficaces para frenar la desinformación masiva. Por un lado, según el algoritmo de consenso detrás de las notas comunitarias<sup>27</sup>, aquellas notas “fiables” requieren evaluaciones suficientes desde diferentes perspectivas para ser publicadas, lo que significa que los contenidos con desinformación primero deben circular ampliamente, llamar la atención entre las personas usuarias y generalizarse para recibir suficientes valoraciones<sup>28</sup>. Por otro lado, el tiempo promedio para la creación y la publicación de una nota comunitaria, que suele ser mayor a dos días, resulta problemático en plataformas donde los contenidos tienen un impacto inmediato en las audiencias y sus reacciones, como Instagram o Threads<sup>29</sup>. Si bien la verificación de hechos por terceros tenía problemáticas como la conformación de equipos mono-demográficos con falta de inclusión de personas LGBTIQA+ del sur global, según estudios sobre el tema, constituía un mecanismo mucho más protector hacia los derechos humanos de las personas en tanto convocaba a expertos independientes familiarizados con, por ejemplo, la diversidad de culturas y los matices sociopolíticos locales, siendo más efectivo no sólo para combatir la desinformación, sino también para evaluar el potencial daño de ciertos discursos sobre las personas usuarias<sup>30</sup>.

Algunos estudios han señalado también la dificultad de que las notas comunitarias alcancen un consenso sobre temas con matices sociopolíticos propios del contexto de cada país, corriendo el riesgo de ser manipuladas por grupos coordinados, lo que perpetúa la desinformación en línea. Se ha encontrado, por ejemplo, que la afiliación política compartida es un fuerte indicador de si una persona usuaria califica un trino como engañoso o una nota comunitaria como útil<sup>31</sup>. Así, las publicaciones de los adversarios políticos se consideran más engañosas y sus notas comunitarias menos útiles en comparación con las de los miembros del mismo partido. Al ser el partidismo una parte importante de la identidad de las

---

<sup>24</sup> La Nota (2025). Zuckerberg y el mito de la libertad de expresión. Disponible en: <https://lanotatucuman.com/zuckerberg-y-el-mito-de-la-libertad-de-expresion/actualidad/13/01/2025/84454/>

<sup>25</sup> Ver: <https://www.oversightboard.com/wp-content/uploads/2025/11/Community-Notes-PAO-Request.pdf>

<sup>26</sup> Chuai, Y., Tian, H., Pröllochs, N., & Lenzini, G. (2024). Did the roll-out of community notes reduce engagement with misinformation on X/Twitter?. *Proceedings of the ACM on human-computer interaction*, 8(CSCW2), 1-52. <https://dl.acm.org/doi/pdf/10.1145/3686967>

<sup>27</sup> Bajo este algoritmo, si las personas usuarias que normalmente discrepan sobre cómo valoran las notas acaban poniéndose de acuerdo en una nota concreta como 'útil', esa nota es publicada. Ver más información en: <https://www.meta.com/es-la/technologies/community-notes/>

<sup>28</sup> Matamoros-Fernández, A., Jude, N. (2025). The importance of centering harm in data infrastructures for 'soft moderation': X's Community Notes as a case study. *New Media & Society*, 27(4). <https://doi.org/10.1177/14614448251314399>

<sup>29</sup> Chuai, Y., Tian, H., Pröllochs, N., & Lenzini, G. (2024). Did the roll-out of community notes reduce engagement with misinformation on X/Twitter?. *Proceedings of the ACM on human-computer interaction*, 8(CSCW2), 1-52. <https://dl.acm.org/doi/pdf/10.1145/3686967>

<sup>30</sup> Riedlinger, M., Montaña-Niño, S., Watt, N., García-Perdomo, V., & Joubert, M. (2024). Fact-Checking Role Performances and Problematic Covid-19 Vaccine Content in Latin America and Sub-Saharan Africa. *Media and Communication*, 12. [https://www.ssoar.info/ssoar/bitstream/handle/document/97319/ssoar-mediacomm-2024-riedlinger\\_et\\_al-Fact-Checking\\_Role\\_Performances\\_and\\_Problematic.pdf?sequence=1](https://www.ssoar.info/ssoar/bitstream/handle/document/97319/ssoar-mediacomm-2024-riedlinger_et_al-Fact-Checking_Role_Performances_and_Problematic.pdf?sequence=1)

<sup>31</sup> Allen, J., Martel, C., & Rand, D. G. (April). Birds of a feather don't fact-check each other: Partisanship and the evaluation of news in Twitter's Birdwatch crowdsourced fact-checking program. <https://dl.acm.org/doi/pdf/10.1145/3491102.3502040>

personas usuarias, y dado el creciente ambiente político en las plataformas de redes sociales, es posible que la participación en las comunidades de colaboradores de las notas comunitarias recaiga en la promoción de opiniones partidistas y en la obtención de beneficios políticos, por encima de una preocupación genuina por combatir la desinformación. Este hallazgo es preocupante en un contexto global donde la desinformación política, que se alimenta en parte del partidismo, no solo puede afectar discusiones electorales, sino la percepción pública de los derechos humanos y la democracia<sup>32</sup>. Al respecto, Meta no ha ofrecido claridades sobre los criterios de diversidad, inclusión y selección de las personas colaboradoras en su sistema de notas comunitarias, un silencio que posibilita la manipulación con corte político de la moderación de contenido en sus plataformas.

Además, ante la diversidad de actores, intereses y motivaciones que alimentan el intercambio de información engañosa en línea, Meta no ha implementado salvaguardas suficientes durante la fase piloto del mecanismo de notas comunitarias en Estados Unidos. Esta omisión dificulta determinar la utilidad general de las notas comunitarias en otros contextos geopolíticos para frenar la desinformación. En la medida que Meta sigue teniendo un control centralizado sobre lo que se comunica en sus plataformas, es crucial la implementación de medidas claras, en concordancia con los UNGPs, para prevenir el uso de las redes sociales como herramientas de amplificación del odio, la violencia y la desinformación. Esto incluye garantizar condiciones de transparencia sobre la forma en que se toman decisiones de cara las políticas de moderación de contenido<sup>33</sup>.

#### **4. El uso de IA en la redacción de notas comunitarias.**

Hasta el momento, Meta no ha ofrecido claridad alguna sobre el posible uso de inteligencia artificial (IA) en la elaboración de sus notas comunitarias<sup>34</sup>. En ese sentido es relevante señalar que X anunció que permitirá la elaboración de notas comunitarias mediante la utilización de grandes modelos de lenguaje, con el fin de aclarar o corregir publicaciones polémicas, incluso antes de que las personas usuarias las aprueben<sup>35</sup>. Según la plataforma, la IA tiene el potencial de "acelerar la velocidad y la escala" de las notas de la comunidad<sup>36</sup>. Este anuncio genera preocupaciones válidas respecto a la posibilidad de que Meta implemente medidas similares, considerando que dicho modelo fue la base del suyo, lo que refuerza la necesidad de que la plataforma considere salvaguardas ante el uso de esta tecnología emergente en el funcionamiento del sistema de notas comunitarias y en su expansión a nivel global.

#### **5. Los desafíos de las notas comunitarias en América Latina.**

En regiones como América Latina, el sistema de notas comunitarias de Meta se desplegaría en un escenario marcado por factores contextuales con la potencialidad de aumentar, de forma significativa, la amplificación de discursos de odio, violencia y discriminación, así como la desinformación y la manipulación del debate público en entornos como Instagram, Threads o Facebook. En primer lugar, encontramos el desarrollo desigual de políticas para regular las plataformas digitales y combatir la desinformación. Por ejemplo, se evidencia la falta de mecanismos unificados, con financiación y recursos humanos suficientes, para fortalecer el ecosistema de la información en los países de la región, invertir en educación informativa, garantizar el pluralismo informativo en las redes sociales y exigir a las grandes plataformas rendición de cuentas y responsabilidad por la difusión de contenidos engañosos. En segundo lugar, desde hace algunos años se vive

---

<sup>32</sup> Camacho, O.L. (2025). Meta y la desinformación: el fin de la verificación independiente. Disponible en: <https://razonpublica.com/meta-la-desinformacion-fin-la-verificacion-independiente/>

<sup>33</sup> La Nota (2025). Zuckerberg y el mito de la libertad de expresión. Disponible en: <https://lanotatucuman.com/zuckerberg-y-el-mito-de-la-libertad-de-expresion/actualidad/13/01/2025/84454/>

<sup>34</sup> La Relatora Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión de las Naciones Unidas señaló con preocupación los retrocesos de las plataformas digitales en su compromiso con la integridad de la información mediante la reversión de políticas clave, mientras que amplían su inversión en herramientas de IA para sus sistemas. Este argumento aplica también a los cambios en los mecanismos de moderación de contenido, como el de Meta. Para más información, ver: Khan, I. (2025). Freedom of expression and elections in the digital age. Disponible en: <https://www.ohchr.org/sites/default/files/documents/hrbodies/hrcouncil/sessions-regular/session59/advance-version/a-hrc-59-50-aev.pdf>

<sup>35</sup> The Guardian (2025). Fears AI factcheckers on X could increase promotion of conspiracy theories. Disponible en: <https://www.theguardian.com/technology/2025/jul/02/fears-ai-factcheckers-on-x-could-increase-promotion-of-conspiracy-theories>

<sup>36</sup> Infobae (2025). X quiere que la IA redacte las notas de la comunidad. Disponible en: <https://www.infobae.com/america/agencias/2025/07/02/x-quiere-que-la-ia-redacte-las-notas-de-la-comunidad/>

un cierre progresivo del espacio cívico combinado con el aumento de la censura y el silenciamiento de voces críticas en entornos digitales, así como la creciente implementación y normalización de tecnologías de vigilancia digital<sup>37</sup>.

En tercer lugar, en América Latina no existe rendición de cuentas ni transparencia en el funcionamiento de la moderación de contenido de las grandes plataformas tecnológicas, incluyendo a Meta. Como ejemplo, cuando Meta anunció el reemplazo de la verificación de hechos por terceros por las notas de la comunidad, el gobierno actual de Brasil solicitó una audiencia con la compañía exigiendo información sobre los cambios en sus políticas en un plazo de 72 horas<sup>38</sup>. Además, dicho gobierno señaló que los cambios en el mecanismo de moderación de contenido no estaban en línea con la legislación brasileña y eran insuficientes para proteger los derechos fundamentales de la ciudadanía. Esta petición no fue atendida por Meta, lo que refleja tanto el incumplimiento de legislaciones locales por parte de la compañía con su nuevo modelo de moderación de contenido, como la falta de importancia que le da a las políticas nacionales de cara al funcionamiento de las plataformas digitales en los territorios latinoamericanos.

En cuarto lugar, ante la alta susceptibilidad de la moderación comunitaria hacia la manipulación y la amplificación de la desinformación, preocupa su inserción en contextos donde la polarización política es elevada, como es el caso de muchos de los países latinoamericanos. Sumado a esto, los discursos de odio y la violencia en línea son un problema creciente, especialmente dirigido a grupos vulnerables como mujeres, personas LGBTIQ+ y comunidades étnicas. La moderación mediante notas de la comunidad podría dificultar la aplicación de criterios uniformes para la identificación del discurso de odio y el enfrentamiento de la violencia en línea, o incluso la permisividad hacia diversas formas de agresión y discriminación.

## Conclusiones.

Solicitamos al Consejo Asesor de Meta que recomiende a Meta:

- Priorizar y respetar sus obligaciones con los derechos humanos<sup>39</sup> sobre cualquier interés político o económico al momento de modificar sus políticas, comprometiéndose seriamente con la transparencia y la rendición de cuentas en sus prácticas de moderación de contenido, y adoptando medidas inclusivas y efectivas para proteger a todas las personas usuarias, en concordancia con estándares internacionales de derechos humanos.
- Adoptar medidas que aseguren un entorno habilitante para el ejercicio de la libertad de expresión de voces críticas y grupos históricamente en situación de vulnerabilidad en sus plataformas.
- Realizar evaluaciones de impacto en derechos humanos, de la mano de organizaciones locales especializadas, respecto al despliegue de sistemas de gestión de contenido como las notas de la comunidad.
- Establecer un proceso de evaluación continua respecto a la efectividad de las notas comunitarias para disminuir la amplificación de contenidos engañosos y la desinformación masiva. Este proceso, en línea con las recomendaciones de la Relatora Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión de las Naciones Unidas<sup>40</sup>, debe contemplar la publicación regular de informes que transparenten los resultados -tanto avances como limitaciones- y un compromiso explícito de no ampliar ni consolidar el sistema si la evidencia no demuestra impactos positivos verificables.
- Dar cumplimiento a las legislaciones nacionales sobre las obligaciones y responsabilidades en materia de derechos humanos de las plataformas digitales, atendiendo a los llamados de solicitud de información a través de audiencias o reuniones con los gobiernos sobre los impactos de los cambios en políticas de moderación de contenido en los contextos locales, así como en los derechos y libertades fundamentales de la ciudadanía.

---

<sup>37</sup> IACHR (2025). The Impact of Digital Surveillance on Freedom of Expression in the Americas. Special Rapporteurship for Freedom of Expression of the Inter-American Commission on Human Rights. Disponible en: <https://www.oas.org/en/iachr/expression/reports/VigilanciaRELECIDH.pdf>

<sup>38</sup> El País (2025). Brasil le da a Meta 72 horas para que explique el fin del programa de verificación. Disponible en: <https://elpais.com/america/2025-01-10/brasil-le-da-a-meta-72-horas-para-que-explique-el-fin-del-programa-de-verificacion.html>

<sup>39</sup> Meta (2021). Corporate Human Rights Policy. <https://about.fb.com/wp-content/uploads/2021/03/Facebooks-Corporate-Human-Rights-Policy.pdf>

<sup>40</sup> Khan, I. (2025). Freedom of expression and elections in the digital age. Disponible en: <https://www.ohchr.org/sites/default/files/documents/hrbodies/hrcouncil/sessions-regular/session59/advance-version/a-hrc-59-50-aev.pdf>