



R3D

Red en Defensa
de los Derechos Digitales

Ciudad de México, 08 de abril de 2025
Asunto: comentarios a ley modelo sobre violencia digital

Dirección General de Vinculación con las Organizaciones de la Sociedad Civil - Secretaría de Relaciones Exteriores

Comité de Expertas del Mecanismo de Seguimiento de la Convención de Belém do Pará - Organización de Estados Americanos

P R E S E N T E S

Introducción

R3D, Red en Defensa de los Derechos Digitales es una organización no gubernamental dedicada a la defensa y promoción de los derechos humanos en el entorno digital en México.

Desde R3D, Red en Defensa de los Derechos Digitales, consideramos que los intereses regulatorios para erradicar la violencia contra las mujeres en el ámbito digital deben armonizarse con el marco interamericano y los principios establecidos en él, además de tomar en cuenta las particularidades de la arquitectura de Internet para así desarrollar propuestas coherentes que impulsen legislaciones compatibles con la Convención Americana sobre Derechos Humanos (CADH) y con los estándares que han sido desarrollados a partir del Sistema Interamericano de Derechos Humanos (SIDH).

A partir de este enfoque, presentamos algunas preocupaciones que surgen de advertir esas contradicciones en el borrador de propuesta de ley modelo sobre violencia digital, tanto con los estándares del derecho a la libertad de expresión en Internet como con la arquitectura de Internet misma.

Además, presentamos una serie de recomendaciones con las que esperamos contribuir para mejorar la propuesta de ley que se nos presenta.

Observaciones a la propuesta de ley modelo

1. La definición de “Violencia digital” del artículo 2o es incompatible con el derecho a la libertad de expresión

El artículo 2o define la violencia digital como:

“**Cualquier acción, conducta u omisión** contra las mujeres, basada en su género, que cause muerte, **daño o sufrimiento** físico, sexual, psicológico, económico o simbólico, en cualquier ámbito de su vida, que sea cometida, instigada o agravada, en parte o en su totalidad, con el uso de las tecnologías”.
[Énfasis añadido]

La definición es en extremo amplia, y resulta vaga y ambigua, sin establecer el alcance de las posibles conductas y acciones. Además, establece como parámetro el concepto de “sufrimiento”, que también es vago y subjetivo. La amplitud de esta definición sin referir a la necesidad de que las conductas constituyan un hecho ilícito hace que, al menos en principio, puedan considerarse como violencia digital distintas expresiones protegidas por el derecho a la libertad de expresión. En este sentido, incorporar explícitamente la referencia a conductas y acciones ilícitas vuelve el concepto más concreto evitando la incertidumbre que una definición tan amplia puede generar.

El derecho a la libertad de expresión da una protección *ab initio* a todas las expresiones, incluso aquellas que pueden ser chocantes y ofensivas, pero además da una protección **reforzada** a expresiones especialmente protegidas como las de la crítica política y la crítica a personas funcionarias públicas o a figuras públicas que también estaría en tensión con esta definición. Para que una restricción a este derecho sea legítima **debe estar establecida de manera clara y precisa, y además ser necesaria, idónea y proporcional**.

La posibilidad de que una definición tan amplia genere incertidumbre sobre el tipo de expresiones que podrían configurarse como violencia digital puede generar también un efecto disuasorio (*chilling effect*) que lleve a la autocensura de expresiones protegidas por el derecho a la libertad de expresión.

2. Algunas de las manifestaciones de violencia digital contra mujeres con voz pública que establece el artículo 8o están en tensión con la libertad de expresión

Los incisos a, b, y c, de la ley modelo catalogan como violencia digital contra mujeres políticas, con voz pública o figuras públicas, acciones tales como: “b. **Usar lenguaje misógino, machista, racista** o de incitación a la violencia o conductas similares contra mujeres políticas o públicas en plataformas digitales; c. **Publicar y promover contenidos falsos o maliciosos y calumnias que afecten la imagen, reputación** o integridad de una mujer política o con voz pública”. [Énfasis añadido]

Estas disposiciones están en clara tensión con el discurso especialmente protegido que protege la crítica política, a personas funcionarias públicas y figuras públicas de manera reforzada en el SIDH. Dicha protección permite que discursos ofensivos e incluso perturbadores puedan ser dirigidos hacia personas que tengan esa posición, puesto que se considera que debe existir un ámbito mucho más amplio a la crítica y que el ámbito privado se reduce en dichos casos. La desinformación y la afectación al honor (inciso c) y el lenguaje misógino o machista (inciso b) no son discursos prohibidos por el derecho a la libertad de expresión y, en casos del ámbito público y de crítica política, son parte de las expresiones protegidas por el derecho a la libertad de expresión salvo excepciones muy concretas (como la de la figura de la **real malicia**).¹

El derecho a la libertad de expresión contempla la posibilidad de que existan expresiones que configuren tres tipos de discursos discriminatorios: i) discursos discriminatorios análogos a la figura de incitación a la violencia (regulables por vía penal); ii) discursos discriminatorios no análogos a la iniciación a la violencia (regulables por vía civil o administrativa); y iii) discursos discriminatorios que

¹ El Principio 10 de la Declaración de Principios sobre Libertad de Expresión de la OEA establece que puede sancionarse en casos de real malicia, donde “debe probarse que en la difusión de las noticias el comunicador tuvo la intención de infligir daño o pleno conocimiento de que se estaba difundiendo noticias falsas o se condujo con manifiesta negligencia en la búsqueda de la verdad o falsedad de las mismas”; ONU Plan de Acción de Rabat sobre la prohibición de la apología del odio nacional, racial o religioso, 2012; CIDH. Marco Interamericano sobre la libertad de expresión; CIDH. Libertad de expresión e Internet.

aunque puedan ser ofensivos no son ilegales (no regulados por vía de sanción en ningún sentido). El último tipo de discurso discriminatorio suele ser ignorado pero su existencia en democracia por fuera de la regulación legal es necesaria como parte del propio ejercicio de la libertad de expresión, la libertad de pensamiento y la diversidad y tolerancia en sociedades plurales.

En este caso, los incisos mencionados de la ley modelo no hacen esa distinción y catalogan como violencia a discursos que, aunque deben ser problematizados y combatidos desde una óptica democrática, no deben tener ese carácter si se quiere mantener el alcance del derecho a la libertad de expresión. Asimilar como "violencia digital" distintos tipos de expresiones implica el problema de catalogarlos como límites a discursos protegidos por este derecho, así como establecer criterios de decoro, gusto, estética u otros que son incompatibles con la exigencia de contar con un debate público robusto.

En términos democráticos, es necesario dividir las distintas expresiones en tres categorías: 1) discursos discriminatorios que pueden constituir incitación a la violencia (regulables por vía penal), 2) discursos discriminatorios que no constituyen incitación a la violencia pero pueden generar algún daño (regulables por vía civil o administrativa), y 3) discursos discriminatorios que no deben ser regulables por el derecho sino combatidos por otros medios como la educación y la cultura.

En este caso, las disposiciones del artículo 8 borran estas distinciones y establecen acciones y discursos que son distintos dentro de la misma categoría de violencia contra mujeres en el ámbito público. Sin embargo, el inciso b puede quedar comprendido dentro de la tercera categoría de discursos no regulables ni sancionables por la vía jurídica, y el inciso c equipara el problema de la desinformación a la violencia, en casos donde el margen para la crítica y las expresiones es más amplio y debe existir una tolerancia mucho mayor, por lo que establece un criterio contradictorio que puede generar un efecto disuasorio o incluso sanciones que son incompatibles con el derecho a la libertad de expresión (al considerar estas conductas como violencia).

3. El artículo 11.c y 11.e, y los artículos 30 y 31 de la propuesta de ley modelo son incompatibles en algunas disposiciones con el artículo 13.2 de la CADH

La CADH establece una prohibición expresa para llevar a cabo medidas de censura previa, estableciendo la regla de responsabilidades ulteriores para los casos en que el ejercicio de la libertad de expresión genere algún daño o pueda ser limitada -siempre ex post la emisión de las expresiones-. Como única excepción a la censura previa está reconocida la difusión de contenidos que constituyen abuso sexual infantil (también conocido como “pornografía infantil”, en el artículo 13.4) y, de manera posterior, los discursos prohibidos de propaganda de la guerra y la incitación a la violencia o al genocidio.

Sin embargo, los incisos c y e del artículo 11 de la ley modelo autorizan a utilizar medidas cautelares como medios de protección contra expresiones que puedan constituir violencia política por razones de género. Las medidas cautelares sobre expresiones en el ámbito público y sobre funcionarias públicas o figuras públicas no sólo está prohibido porque se trata de un discurso especialmente protegido, sino porque la naturaleza jurídica de las medidas cautelares implica bloquear un contenido en Internet sin que éste haya sido considerado como contenido ilegal por una autoridad judicial competente; es decir, que las medidas cautelares en Internet son una forma de control de la información que constituye censura previa en todos los casos, por lo que debe ser descartada como una medida legítima para limitar expresiones en Internet. Toda responsabilidad y posible expresión que constituya un daño y un exceso ilegítimo del derecho a la libertad de expresión en Internet debe, entonces, ser sancionada de acuerdo a la regla de responsabilidades ulteriores, una vez que la autoridad judicial competente ha determinado este hecho y siempre con sanciones que cumplan el test tripartito del SIDH, y sean necesarias, idóneas y proporcionales.

El SIDH tiene una larga doctrina y jurisprudencia en la que ha desarrollado la prueba de legitimidad de las restricciones a la libertad de expresión, denominado

“test tripartito”.² Este test debería estar presente en cualquier regulación que pretenda dirigirse al entorno digital, e incorporar lo que la Relatoría Especial para la Libertad de Expresión ha denominado “la perspectiva sistémica digital”, que sirve como un paso complementario (una especie de “cuarto paso”) en el caso de evaluar restricciones a la libertad de expresión en Internet. Este requisito implica que cualquier medida regulatoria debe de tomar en cuenta la arquitectura de Internet y su funcionamiento libre y abierto antes de tomar medidas que puedan restringirlo o modificarlo.

En el mismo sentido, los artículos 30 y 31 establecen la posibilidad de llevar a cabo medidas cautelares sobre la libertad de expresión y autorizan el bloqueo de contenidos sin una determinación final sobre la ilegalidad de los contenidos, afectando no sólo la prohibición de la censura previa sino las garantías mínimas reconocidas en los artículos 8 y 25 de la CADH.

Existe un consenso internacional de que el bloqueo de contenidos en Internet es tan grave como lo sería el cierre de una televisión o a la censura de un periódico. El hecho de que se trate de expresiones en el entorno digital no cambia la naturaleza de censura previa y el carácter desproporcionado del uso de las medidas cautelares para regular el discurso político. Así, La Declaración conjunta sobre libertad de expresión e Internet de los relatores especiales sobre libertad de expresión reconoce explícitamente que:

3. Filtrado y bloqueo

- a. El bloqueo obligatorio de sitios web enteros, direcciones IP, puertos, protocolos de red **o ciertos tipos de usos (como las redes sociales) constituye una medida extrema—análoga a la prohibición de un periódico o una emisora de radio o televisión— que solo podría estar justificada conforme a estándares**

² CorteIDH. Casos: Kimel v. Argentina, párr. 177; Palamara Iribarne v. Chile, párr. 135; Herrera Ulloa v. Costa Rica, párrs. 121-122; Tristán Donoso v. Panamá, párr. 119. CIDH. Marco Jurídico Interamericano del Derecho a la Libertad de Expresión, párrs. 85-87; CIDH. Estándares para una Internet Libre, Abierta e Incluyente, párr. 74; CiDH. Informe Anual 2013. Relatoría Especial para la Libertad de Expresión, párr. 65.

internacionales, por ejemplo, cuando sea necesaria para proteger a menores del abuso sexual.

b. Los sistemas de filtrado de contenidos impuestos por gobiernos o proveedores de servicios comerciales que no sean controlados por el usuario final **constituyen una forma de censura previa y no representan una restricción justificada a la libertad de expresión.**³
[Énfasis añadidos]

Por el contrario, se ha demostrado empíricamente que este tipo de recursos se utilizan por partidos políticos y personas servidoras públicas para clausurar la crítica política, perseguir medios de comunicación y periodistas, y manipular el debate público, sin que estas medidas tengan un impacto real en la lucha contra la violencia de género y generando un efecto disuasivo que es incompatible con el derecho a la libertad de expresión en el SIDH.⁴

Las disposiciones señaladas en este punto, entonces, son contrarias no sólo a la prohibición de censura previa, sino que se constituyen como medidas que no son idóneas y son claramente desproporcionales en su forma de afectar la libertad de expresión. Pero además, que tampoco cumplen con la perspectiva sistémica digital al afectar la forma libre y abierta de Internet, tal como explicamos en el punto siguiente al referir al principio de no responsabilidad de intermediarios.

³ Declaración conjunta sobre libertad de expresión e Internet. ONU. Relator Especial de las Naciones Unidas sobre la Promoción y Protección del derecho a la Libertad de Opinión y de Expresión, Representante para la Libertad de los Medios de Comunicación de la Organización para la Seguridad y la Cooperación en Europa, Relatora Especial de la Organización de Estados Americanos para la Libertad de Expresión, y Relatora Especial sobre Libertad de Expresión y Acceso a la Información de la Comisión Africana de Derechos Humanos y de los Pueblos. 1 de junio de 2011.

⁴ Al respecto ver nuestra investigación: Chorny, Vladimir. "Guía para interpretar restricciones a la libertad de expresión en Internet", *Red en Defensa de los Derechos Digitales*, Octubre de 2024. Disponible en www.r3d.mx. En el mismo sentido ver nuestra investigación: www.censuraelectoral.mx.

4. Los artículos 19, 20, 23 y 40 son incompatibles en diversas disposiciones con los artículos 8, 13.2 y 25 de la CADH, al establecer un sistema de censura previa de carácter privado y violar el principio de no responsabilidad de intermediarios en Internet

El artículo 19 y el artículo 20 de la propuesta de ley modelo establecen una obligación a las plataformas de llevar a cabo una vigilancia proactiva que lleve a eliminar contenidos de las personas usuarias en sus plataformas, sin que para ello medie una determinación por parte de una autoridad judicial competente, habilitando un mecanismo de "notificación y retirada" por parte de los usuarios en dichas plataformas y con ello eliminando las garantías mínimas de debido proceso que deben existir para que se respeten los artículos 8 y 25 de la CADH.

El inciso a del artículo 19 y el artículo 20 establecen un mecanismo de *notice and take down* que ya ha sido considerado como un procedimiento extrajudicial que es incompatible con la libertad de expresión en Internet y que viola también las garantías mínimas de debido proceso, estableciendo una obligación de control de contenidos por las plataformas de servicios digitales y convirtiéndolos en una especie de "jueces privados" que realizan medidas de censura previa sin que los contenidos en cuestión hayan sido previamente declarados como ilegales. Así, los relatores especiales sobre libertad de expresión determinaron explícitamente que:

2. Responsabilidad de intermediarios
 - a. Ninguna persona que ofrezca únicamente servicios técnicos de Internet como acceso, búsquedas o conservación de información en la memoria caché deberá ser responsable por contenidos generados por terceros y que se difundan a través de estos servicios, siempre que no intervenga específicamente en dichos contenidos ni se niegue a cumplir una orden judicial que exija su eliminación cuando esté en condiciones de hacerlo ("principio de mera transmisión").
 - b. Debe considerarse la posibilidad de proteger completamente a otros intermediarios, incluidos los mencionados en el preámbulo,

respecto de cualquier responsabilidad por los contenidos generados por terceros en las mismas condiciones establecidas en el párrafo 2(a). Como mínimo, **no se debería exigir a los intermediarios que controlen el contenido generado por usuarios y no deberían estar sujetos a normas extrajudiciales sobre cancelación de contenidos que no ofrezcan suficiente protección para la libertad de expresión (como sucede con muchas de las normas sobre "notificación y retirada" que se aplican actualmente).**⁵ [Énfasis añadido]

El inciso b del artículo 19 tiene los mismos problemas señalados en este punto y el punto anterior. Al habilitar una facultad amplia de "Solicitudes gubernamentales" para el bloqueo y la eliminación de contenidos, se reconoce la posibilidad de establecer un mecanismo de censura previa gubernamental que elimine las garantías procesales de los artículos 8 y 25, por un lado, y que cree un sistema incompatible con el derecho a la libertad de expresión en Internet por el que las medidas más extremas de restricción a la libertad de expresión estén habilitadas sin los controles judiciales mínimos exigidos por el SIDH.

Es fundamental subrayar que las medidas de bloqueos de contenidos en Internet deben considerarse como desproporcionales y excepcionalísimas sólo para los casos más graves reconocidos de discursos prohibidos por el marco interamericano, mientras que todas aquellas otras expresiones que puedan generar un daño y que afecten derechos de terceros deberían quedar en dado caso disponibles para establecer las responsabilidades ulteriores correspondientes a las personas, **respetando en todo momento el principio de no responsabilidad de intermediarios por contenidos de terceros en los que no intervengan como creadores o editores.**

El inciso c sobre "marcado automático" debería especificar que éste no debe traducirse en bloqueo de contenidos ni eliminación de los mismos, para evitar generar incentivos negativos de censura previa por parte de los sujetos particulares. En este sentido, el artículo 23 contempla como una de las medidas legítimas el bloqueo de contenidos y el bloqueo de personas usuarias de las plataformas sin que exista ninguna determinación de la ilegalidad de los

⁵ Declaración conjunta sobre libertad de expresión e Internet. *Op. cit.*, nota 1, artículo 2.a y 2.b.

contenidos y la necesidad extrema de llevar a cabo esas medidas, por lo que establece un sistema desproporcional de control previo de información que es incompatible con la prohibición de censura previa del artículo 13.2 de la CAHD.

Finalmente, el artículo 40 formaliza la eliminación del principio de no responsabilidad de intermediarios, al establecer sanciones para los mismos cuando no lleven a cabo las medidas de censura privada establecidas en la propuesta de ley modelo. Al hacerlo, se establece un sistema que les impone obligaciones de vigilancia proactiva que son incompatibles con la prohibición de censura previa del artículo 13.2, y afectan de manera categórica uno de los principios fundamentales para la arquitectura libre y abierta de Internet. Al revisar este artículo 40 a la luz de los artículos 41, 42 y 43 de la propuesta de ley modelo, queda claro que ésta invierte la lógica de responsabilidades y pone a los intermediarios como sujetos responsables por las expresiones de terceros en sus plataformas. Esto atenta contra cualquier noción de un ecosistema digital libre y abierto, y pone en riesgo la libertad de expresión en Internet contraviniendo al SIDH.

Recomendaciones:

Por las razones anteriormente expuestas, consideramos fundamental que:

1. La propuesta de ley modelo sea discutida de manera más profunda, en particular tomando en cuenta las preocupaciones aquí señaladas desde la óptica del derecho a la libertad de expresión en Internet. Dicha discusión debería ser lo más incluyente posible, incorporando a sectores de la sociedad civil y la academia que tengan una trayectoria y *expertise* demostrados en materia de libertad de expresión, para hacer los ajustes que permitan armonizar estos intereses con los intereses de combatir la violencia digital en contra de las mujeres.
2. Las consideraciones hechas en los cuatro puntos de los comentarios a la ley modelo sean tomadas en cuenta e incorporadas sustantivamente a la propuesta de ley modelo sobre violencia digital en contra de las mujeres.
3. Se incorporen explícitamente a cualquier propuesta regulatoria los principios fundamentales para la libertad de expresión en Internet desde el enfoque interamericano del SIDH, tales como: i) la prohibición de censura previa y la regla de responsabilidades ulteriores; ii) el principio de no responsabilidad de intermediarios en Internet; iii) la perspectiva sistémica digital; y iv) las obligaciones de transparencia (y transparencia algorítmica), rendición de cuentas y remedio adecuado que deben cumplir las empresas que presten servicios digitales.