

## **Opinión sobre el proyecto de Ley Modelo Interamericana para Prevenir, sancionar y erradicar la violencia digital contra las mujeres por razones de género**

ABRIL 2025

OBSERVACOM reconoce el valor de esta iniciativa por cuanto recepta una preocupación que ha sido expresada por distintos organismos internacionales como la UNESCO, entre otros, y por el propio Mecanismo de Seguimiento de la Convención de Belém do Pará (MESECVI), impulso de esta iniciativa, en distintos pronunciamientos.

En efecto, la violencia digital contra las mujeres por motivos de género es una problemática facilitada por las tecnologías y la comunicación a través de plataformas digitales. Esta violencia tiene serios impactos en la vigencia de distintos derechos humanos, incluida la libertad de expresión que se ve cercenada por la autocensura y el miedo que esta modalidad de violencia impone.

Ahora bien, en relación con el proyecto de Ley Modelo sujeto a consulta, consideramos pertinente formular una serie de observaciones que están basadas en los estándares interamericanos e internacionales de libertad de expresión y, en particular- los referidos a la regulación democrática de las grandes plataformas digitales.

Antes de avanzar, en términos generales entendemos útil señalar que sería relevante, a los fines de una mayor comprensión y fundamentación del proyecto, que éste contara —por ejemplo, a través de un sistema de notas al pie— con la indicación de las fuentes en las que abrevia ya se trate de normativa internacional de derechos humanos (tanto de libertad de expresión como de protección de las mujeres contra la violencia), normativa internacional de gobernanza de plataformas, estándares en estas materias, documentos de relatorías especiales o cualquier otra referencia pertinente.

A continuación, detallaremos nuestros comentarios a la Ley Modelo según orden de tratamiento en el proyecto.

## ***Ámbito de aplicación y definiciones***

Consideramos relevante definir qué se entiende por *comunidad* en el inciso b. del artículo 3 toda vez que si bien el concepto surge de la Convención de Belem Do Pará (artículo 2.b), en el contexto de la Ley Modelo, y en caso de que refiera al conjunto de personas que habita un territorio común, los incisos 3a y 3c se subsumen al b. Es importante además detallar este concepto por cuanto, en los entornos digitales, es habitual referir a comunidad o comunidades, contexto que a la fecha de la adopción de la Convención era diferente al actual.

Por otra parte, la definición de *moderación de contenidos* (inciso c) omite dimensiones de las prácticas y políticas de las plataformas que es preciso considerar, como es la “*curación de contenidos*”.

La curación de contenidos es el proceso mediante el cual las plataformas digitales seleccionan, organizan y presentan contenidos a una audiencia, de acuerdo a criterios desconocidos para sus usuarios. La curación se diferencia de la moderación de contenidos, ya que esta última está más enfocada en la aplicación de políticas y normativas para filtrar o eliminar contenido que se considera inapropiado o que infringe las reglas de la plataforma. Por su parte, la curación determina qué contenidos tendrán mayor visibilidad y cuáles quedan relegados en los feeds, los resultados de búsqueda y las recomendaciones personalizadas de los usuarios de la plataforma.

La Relatoría Especial para la Libertad de Expresión de la Comisión Interamericana de Derechos Humanos considera a la curación de contenidos “como las decisiones automatizadas sobre el alcance, clasificación, promoción o visibilidad de los contenidos. Las plataformas suelen curar los contenidos con base en las recomendaciones personalizadas para los perfiles de las personas usuarias”. En la medida en la que se privilegian ciertos contenidos, la curación puede terminar por amplificar o reducir el alcance de determinados discursos, advierte”.

Al respecto, debe destacarse que al respecto la Ley Modelo toma como referencia la Digital Services Act de la Unión Europea. Ahora bien, en el art. 3 s. de esa misma norma define los sistemas de recomendación, que también pueden tener impacto en la violencia digital (en particular en su escala). La definición es la siguiente: “un sistema total o parcialmente automatizado y utilizado por una plataforma en línea para proponer en su interfaz en línea información específica para los destinatarios del servicio o priorizar dicha información, también como consecuencia de una búsqueda iniciada por el destinatario del servicio, o que determine de otro modo el orden relativo o la relevancia de la información presentada”.

Por último, la definición de *desinformación* (inciso e) se alinea con la adoptada por la Relatoría Especial de ONU (2021) y retomada en el Informe de la Alta Comisionada (2022). Ahora bien, es preciso abreviar a otros estándares para distinguir la desinformación de lo que, así definido como aparece en el texto propuesto, es una injuria, y no necesariamente forma parte de un esquema de desinformación.

En este marco, distintos informes de las Naciones Unidas ponen el foco en que “lo que sí es nuevo es la manera en que la tecnología digital ha hecho posible que diversos actores creen, difundan y amplifiquen información falsa o manipulada por motivos políticos, ideológicos o comerciales a una escala, a una velocidad y con un alcance sin precedentes”. Y las Directrices UNESCO (retomando el informe “Justicia de género y libertad de expresión” de la Relatoría ONU de 2021) toman el concepto desinformación de género como un “subtipo de violencia de género, [que] tiene algunas características particulares, ya que entraña el uso de relatos falsos o engañosos basados en el género y el sexo contra las mujeres, a menudo con cierto grado de coordinación, a fin de disuadir a las mujeres de participar en la esfera pública. La desinformación de género combina tres características definitorias de la desinformación en línea: la falsedad, la intención maliciosa y la coordinación”.

Entendemos que resultaría adecuado que la definición de la Ley Modelo diera cuenta de estos rasgos de la desinformación identificados por los estándares mencionados.

### ***Sujetos alcanzados y regulación asimétrica para la imposición de obligaciones a las plataformas digitales***

La definición prevista en la Ley Modelo entiende como proveedores —y por tanto obligados por Capítulo III— tanto a intermediarios como a quienes diseñen, fabriquen o comercialicen productos tecnológicos. Los estándares internacionales ponen en manos de los intermediarios obligaciones de debida diligencia en materia de derechos humanos, lo cual incluye la protección de derechos humanos, incluidos los de mujeres, adolescentes, niñas y niños. Esto implica que es preciso revisar la extensión del alcance de la norma a los proveedores tal como los considera la Ley Modelo o la definición de quién es proveedor de servicios.

De acuerdo con RELE (2024: 43) “Los intermediarios de internet van desde los proveedores del servicio hasta los motores de búsqueda, comprendiendo también las plataformas de redes sociales, las plataformas de comercio electrónico y los servidores web”. Las Directrices para la gobernanza de las plataformas digitales adoptadas por UNESCO en 2023 ponen en cabeza de las plataformas digitales las

obligaciones de debida diligencia en materia de derechos humanos, en particular la implementación de protecciones especiales (en sus sistemas de moderación, enfoques algorítmicos, y medidas de protección contra la violencia derivada de las tecnologías) para mujeres y niñas, entre otras personas y grupos en situación de vulnerabilidad.

En este sentido, consideramos necesario revisar la definición de *proveedores de servicio*, en tanto extiende el alcance de la norma a sujetos a los que no pueden imponerse similares obligaciones y responsabilidades. La experiencia y los estándares internacionales recomiendan -y suscribimos- la necesidad de un abordaje asimétrico de la regulación.

De acuerdo con los estándares en la materia, al momento de imponer obligaciones a las plataformas (por ejemplo, las previstas en los artículos 19, 20, 21 o 26 de la Ley Modelo) es preciso considerar las asimetrías entre los actores del entorno. Así, no debieran ser similares las obligaciones que pesen sobre Meta o Google que sobre una plataforma *startup* educativa pequeña de un país de América Latina, para señalar dos actores claramente distintos en escala e impacto de sus operaciones.

En el caso de las plataformas digitales de contenidos que sean importante fuente de acceso a información y opiniones de interés público, de reconocida influencia en el debate público o autodefinidas como tales y/o importante penetración (“grandes plataformas de contenidos”) deberían tener una regulación asimétrica respecto a otros proveedores del mismo campo, en atención a la importancia y el impacto que sus decisiones empresariales tienen sobre el debate público y el intercambio de información, opiniones y bienes culturales, así como al ejercicio de la libertad de expresión y otros derechos humanos.

El trato asimétrico no supone violar el principio de igualdad, ya que todas las plataformas de contenidos deben observar los principios y estándares mínimos de derechos humanos, pero una regulación inteligente es aquella que no pone cargas excesivas a actores que, por su desarrollo y características no pueden cumplirlas, considerando de manera adecuada y diferenciada a las grandes plataformas de contenidos, respecto de aquellas que sean de menor porte o estén direccionadas a finalidades específicas.

## ***Manifestaciones de violencia digital contra mujeres políticas***

OBSERVACOM toma nota de la relevancia que la situación particular de violencia desplegada contra las mujeres con participación política y con voz pública en el

entorno digital conlleva (artículo 8 del proyecto de Ley Modelo), y de las recomendaciones de los órganos internacionales en este sentido.

Ahora bien, teniendo en cuenta que las conductas definidas tanto en este artículo como en el anterior (el número 7) serán aquellas perseguibles por el Estado así como reclaman acciones por parte de las empresas —constituyan o no delitos—, es relevante advertir el riesgo del uso de categorías vagas como el “uso de lenguaje machista” (inciso b) y de la divulgación de “información personal sin consentimiento” como una conducta pasible de ser perseguida por las plataformas —e incluso de ser considerada delito en función del artículo 34— y al mismo nivel que la “incitación a la violencia”.

En términos de compatibilidad con la jurisprudencia del sistema interamericano, es preciso que se advierta sobre el necesario resguardo de la preeminencia del interés público que dicha información pudiera revestir. En este sentido, OBSERVACOM recuerda que, de acuerdo con los estándares del sistema interamericano de derechos humanos, los funcionarios y las personas cuya actividad reviste interés público, no deben tener una protección especial en función del principio 10 de la Declaración de Principios sobre Libertad de Expresión de la Comisión Interamericana de Derechos Humanos (CIDH) de 2000 y que la aplicación de responsabilidades tiene que cumplir con umbrales específicos señalados por el test tripartito.

Ahora bien, en caso de que el inciso d. del artículo 8 haya pretendido aludir a información de carácter “íntimo”, “sexual” o referida a “datos sensibles”, recomendamos expresamente referir a dicho tipo de información. De lo contrario, la mención sin especificaciones a “información personal” puede involucrar la persecución de “discursos protegidos” por los estándares internacionales protectorios de la libertad de expresión, cuando, claramente, la intención de la Ley Modelo no abrevia en tal sentido.

### ***Medidas de sanción***

Se advierte que el artículo 12, no obstante lo indicado en su título, no refiere a medidas sancionatorias. A su vez, el artículo 42 incluye en el mismo listado, medidas reparatorias (por ejemplo, la publicación en medios de una decisión administrativa prevista en el inciso b) y medidas tendientes a generar garantías de no repetición (por ejemplo, la obligación de reportar periódicamente medidas adoptadas). A su vez, la revocación de permisos para operar (inciso e) carece de sentido en el caso de los proveedores que no requieren de los permisos que sí requieren otros servicios.

También estimamos pertinente visitar las facultades de los jueces intervinientes en causas civiles, en la medida en que se les asignan atribuciones propias de las autoridades regulatorias como, por ejemplo, las decisiones sobre la vigencia de licencias.

### ***Responsabilidad del representante legal del proveedor de servicios***

La consideración del representante legal como responsable por el incumplimiento de las obligaciones previstas en el artículo 15 de la Ley Modelo, está tomada de la *Digital Services Act* (artículo 13.3). Ahora bien, las obligaciones previstas en esa norma europea, de acuerdo con el primer párrafo del mencionado artículo son “garantizar la cooperación eficiente y en tiempo oportuno con las autoridades competentes del Estado para la aplicación de la presente ley”.

En relación con las obligaciones previstas en la Ley Modelo, entendemos problemático y riesgoso en términos de libertad de expresión, además, considerar responsable legal al representante por el eventual incumplimiento del extenso catálogo de obligaciones que establece esta norma, cuando en muchos casos no tiene dominio del hecho de las cuestiones que la ley exige.

### ***La obligación de notificar sospechas o probabilidades de comisión de violencia digital***

El artículo 18 de la Ley Modelo pone en manos de las plataformas digitales la obligación de notificar a las autoridades estatales, las “sospechas” de ejercicio de violencia digital cuando tenga “indicios razonables” de que esta “se ha cometido, se está cometiendo o es probable que se cometa”.

Entendemos oportuno señalar la inconsistencia de esta previsión con los estándares en la materia que admiten la intervención de las plataformas para la remoción de contenidos con carácter excepcional y frente a la existencia efectiva y concreta de contenidos ilegales que podrían causar daños en los derechos de las personas.

Así, la obligación de comunicación a las autoridades cuando “es probable que se cometa”, pone en manos privadas un instrumento que puede representar un riesgo tanto en términos de protección efectiva de las mujeres contra la violencia como en relación con la garantía de la libertad de expresión. Así establecida, esta obligación es mayor que la de los funcionarios públicos en ejercicio de su cargo, e

incluso podría entenderse que excede las obligaciones de debida diligencia que los estándares, las directrices de UNESCO o en el derecho comparado se exigen a las plataformas.

### ***Los riesgos de las obligaciones de actuación de las plataformas***

A través de distintos artículos, la Ley Modelo pone en cabeza de las plataformas digitales la obligación de adoptar medidas en el marco de lo que identifica como “moderación de contenidos”. Estas medidas pueden involucrar, entre otras, la eliminación de contenidos y la suspensión de servicios. La Ley Modelo obliga además a que estas decisiones de moderación se concreten ante la sola sospecha de la posible comisión de violencia digital.

Si bien corresponde destacar que el artículo 23 establece una serie de pautas y criterios que deben orientar la actuación de un equipo de evaluación interna de la empresa, corresponde señalar el riesgo de dejar en manos de sistemas automatizados que operan mediante algoritmos de inteligencia artificial -que son los que utilizan las plataformas digitales para manejar un enorme volumen de contenidos- la identificación de los contenidos problemáticos. Tal como muestran los estudios sobre la materia, la moderación se encuentra cada vez en mayor medida en manos de estos sistemas que no interpretan, más allá de su creciente sofisticación, los marcos y contextos de los discursos (por ejemplo, el humor, la sátira o los usos coloquiales).

Así, en el marco de las obligaciones impuestas por la Ley Modelo, el filtrado automático, masivo y preventivo, y el monitoreo activo y sistemático, pueden legitimar y legalizar el poder de policía de contenidos de las grandes plataformas digitales. Además, pone en cabeza de las plataformas no solamente la valoración sobre la posible comisión de delitos, y la adopción de medidas al respecto, sino también la obligación de denunciarlos a las autoridades competentes.

Las obligaciones de vigilancia y el margen amplio de acción frente a los contenidos, incluso frente a la mera sospecha de violencia, pueden representar un serio riesgo para el ejercicio de la libertad de expresión, el acceso a la información y el debate público en general, además de otras garantías de derechos humanos.

Los estándares internacionales refieren, en esta materia, al necesario deber de cuidado de las plataformas e imponen análisis de riesgos sistémicos de su actividad.

## ***La responsabilidad de las plataformas por omisión***

Entendemos que corresponde realizar un señalamiento en relación con la responsabilidad por omisión que es cabría a las plataformas digitales.

El alcance de la responsabilidad de las plataformas en distintos campos es un debate central que continúa delineándose en regulaciones y decisiones de los poderes judiciales en distintos ámbitos.

Puntualmente en relación con la propuesta de Ley Modelo, entendemos pertinente señalar que se establecería una responsabilidad por omisión de adoptar medidas que, establecida de manera transversal respecto de todas las obligaciones impuestas en la Ley Modelo, no resulta adecuada a los estándares ni a los criterios utilizados por la jurisprudencia para delimitar la responsabilidad de estos actores.

Surgen múltiples interrogantes sobre los criterios que determinan cuándo una plataforma puede ser considerada responsable. ¿Es suficiente con la mera publicación de un contenido que contenga violencia digital para que se configure su responsabilidad? ¿O esta responsabilidad se activa cuando la plataforma amplifica y monetiza esos contenidos, beneficiándose de su circulación? ¿Resultan equivalentes estos escenarios? ¿Cuándo se asume que la plataforma ha sido notificada de la hipotética infracción? ¿Cuándo se entiende qué debe bajar los contenidos? ¿Cómo opera en ese caso la obligación de debido proceso que plantean los estándares?

También se plantea si el incumplimiento, total o parcial, de algunas de las obligaciones previstas en la Ley Modelo -por ejemplo, la implementación del marcado automático de contenidos-, es suficiente para atribuir responsabilidad. Del mismo modo, cabe preguntarse si la sola notificación, por cualquier medio, de la comisión de violencia digital basta para generar una obligación inmediata de actuar y, en caso de inacción, una responsabilidad por omisión.

Además, persisten dudas clave sobre cómo debe interpretarse la notificación: ¿cuándo se considera que la plataforma ha sido efectivamente notificada de una hipotética infracción? ¿debería actuar ante una simple denuncia de cualquier usuario? ¿En qué momento se espera que actúe y bajo qué condiciones se entiende que debe retirar un contenido? Estos interrogantes remiten a los estándares internacionales sobre libertad de expresión, en particular al principio de debido proceso, que exige garantías mínimas antes de que se adopten decisiones restrictivas sobre la circulación de contenidos en línea.

## ***Los mecanismos de denuncia y la necesaria incorporación de garantías de debido proceso***

Corresponde destacar el acierto y la importancia de que la Ley Modelo establezca obligaciones de disponer de mecanismos internos de denuncia tanto frente a situaciones de violencia digital como frente a decisiones propias de las plataformas enmarcadas en la moderación de contenidos (art. 20, 22, 24 y 25). Del mismo modo, respecto a las pautas de información sobre la investigación interna y la obligación de fundar las decisiones de moderación adoptadas.

No obstante, cabe señalar que no se incluyen garantías de debido proceso respecto a las personas cuyos contenidos son objeto de moderación y reciben acciones que eliminan sus expresiones o suspenden sus cuentas (defensa, revisión, presentación de prueba, etc.), si entienden que se trata de sanciones injustas o ilegítimas.

## ***Excepcionalidad de las medidas de eliminación de contenidos, suspensión de cuentas y similares***

Los estándares internacionales de libertad de expresión admiten que las plataformas digitales adopten decisiones de moderación de contenidos, incluida la eliminación o remoción. Frente a circunstancias específicas resulta indispensable que las plataformas tengan esta actuación y desplieguen estas herramientas en función de la garantía de derechos fundamentales.

Ahora bien, también de acuerdo con estos estándares, “...la remoción de contenidos sin orden judicial podría ser considerada admisible en circunstancias excepcionales, en las que el contenido sea ilegal y su circulación podría causar daño inmediato y grave en los derechos de las personas. Esto abarcaría las expresiones prohibidas por el derecho internacional de los derechos humanos, a saber: casos de propaganda de guerra, apología del odio que constituye incitación a la violencia, incitación al genocidio y pornografía infantil. En resumen, en circunstancias donde hay un peligro claro y presente de daño significativo, las restricciones podrían ser impuestas de manera preventiva. Sin embargo, aún en estos casos, el principio de proporcionalidad es clave y la acción de los intermediarios debe ser la menos perjudicial posible para la libertad de expresión.” (RELE, 2024: 52).

A continuación, refiere a la violencia señalando que: “La Relatoría resalta que el discurso de odio en línea afecta en particular a personas de grupos en situación de vulnerabilidad, reforzando la discriminación y la violencia, debido a obstáculos

estructurales. La violencia en internet contra estos grupos de personas puede derivar en violencia física y acentuar su exclusión en el debate público. En este mismo contexto, las plataformas digitales son cruciales para la promoción del empoderamiento y autonomía de estos grupos.”

Es decir que, en circunstancias específicas, en las cuales el contenido puede acarrear un daño grave, inminente, que no pueda subsanarse por otra vía (por ejemplo, frente a contenidos de abuso sexual contra personas adultas o niños, niñas u adolescentes, o frente a la transmisión de una masacre en tiempo real) los estándares habilitan la actuación de las plataformas digitales frente a esos contenidos. En este marco, la violencia digital contra las mujeres por motivos de género puede efectivamente justificar este tipo de intervenciones (porque, además, en muchos casos coincide con un discurso de incitación a la violencia), pero consideramos imprescindible que la Ley Modelo determine con claridad estos supuestos a fin de evitar la legalización de medidas de control indiscriminadas sobre los intercambios y contenidos que circulan en el entorno digital.

-