CARRERA DE DERECHO

Tema: La moderación de contenidos en Instagram y su incidencia en el derecho a la libertad de expresión de mujeres.

Proyecto de investigación previo a la obtención del título de Abogado

Línea de investigación:

ESTADO, DERECHO Y SOCIEDAD

Autor:

Karla Cristina Vera López

Director:

Abg. Santiago Javier Paliz Ibarra Mgs.

Esmeraldas - Ecuador

Septiembre 2026

DEDICATORIA

Culminar esta etapa representa mucho más que la finalización de un trabajo académico representa años de esfuerzos, aprendizajes y desafíos superados. Este trabajo no solo refleja todos los conocimientos adquiridos durante toda mi etapa universitaria. Sino la perseverancia necesaria para continuar incluso en los momentos en los que el camino parecía ser difícil.

Por ello, dedico esta investigación en primer lugar a Dios, porque es en él que pude encontrar la fortaleza y la confianza necesaria para poder continuar con mis estudios. Porque nunca me abandonó en los momentos más difíciles que atravesé durante mi etapa universitaria. En segundo lugar, dedico este logro a mis padres Carlos Vera y Diana López, quienes han caminado a mi lado durante todo este proceso y fueron parte importante para que yo culminara mi carrera universitaria. Sin su apoyo nada de esto hubiese sido posible, gracias por acompañarme con amor y paciencia. Gracias por motivarme y recordarme que todo esfuerzo tiene su recompensa. Su apoyo incondicional fue esencial para superar cada obstáculo y llegar a esta etapa tan importante de mi vida.

Finalmente, me dedico este trabajo a mí misma, por la constancia, la disciplina y la fortaleza necesarias para cumplir este objetivo. Por cada noche de estudio, cada dificultad superada y cada momento en el que decidí continuar. Esta investigación es una muestra de todo lo que puedo alcanzar con esfuerzo y determinación.

AGRADECIMIENTO

Expreso mi más sincero agradecimiento a todas las personas que hicieron posible la realización de esta investigación y que, de una u otra manera, contribuyeron a mi formación profesional y al cumplimiento de esta importante etapa académica.

En primer lugar, agradezco a Dios por brindarme fortaleza, sabiduría y perseverancia durante este proceso, permitiéndome superar cada desafío y continuar con determinación hasta alcanzar esta meta.

A mi familia, mi profundo agradecimiento por ser mi principal fuente de apoyo y motivación. Gracias por acompañarme durante este camino, por su paciencia, comprensión y confianza en cada momento. Su respaldo fue fundamental para mantenerme firme y continuar trabajando por este objetivo.

A mi abuelita Cristina Quiroz por ser un pilar importante en mi vida y por estar a mi lado siempre.

A Kevin por ser un apoyo incondicional durante todos estos años, por ser quien más confía en mí, incluso en los momentos en los que ni yo creo en mí. Gracias por recordarme lo valiosa que soy y por todo el amor que me das.

A mi mejor amigo Jorge Armando por recordarme lo inteligente que soy y de lo que soy capaz, por estar conmigo en mis momentos tristes y en mis logros.

A mi perrita Moet por acompañarme cada noche mientras realizaba mi proyecto de titulación.

A mi Camilita, gracias por ser un apoyo incondicional durante toda esta etapa, por acompañarme y alentarme siempre. Uno de los regalos más lindos que me dejó la universidad es tu amistad, te amo con todo mi corazón, gracias por hacer este camino más llevadero y divertido, por ser mi duplita y mi chicle

durante todos estos años, espero la vida no nos separe y conservar esta amistad tan linda porque para mí eres mi hermana, la mejor amiga que me pudo dar la universidad y quien espero que sea la tía consentidora de mis hijos, te amo.

A mis amigos de la Unidversidad, sin ustedes este camino no hubiese sido tan divertido, los amo con todo mi corazón, voy a extrañar verlos todos los días, reírnos e incluso discutir. Gracias por siempre estar, los llevo por siempre en mi corazón y espero la vida nos siga uniendo.

En definitiva, agradezco infinitamente a mis profesores Alex, Andrés, Santiago y Alejandro, quienes han sido mi inspiración en todo este trayecto y parte fundamental para mi formación profesional. Los quiero muchísimo y los llevo por siempre en mi corazón.

Gracias Andrés, por alentarme y no permitir que abandone la carrera infinidad de veces, por escucharme y aconsejarme, por ser un amigo más que un coordinador, te quiero un montón

Asimismo, agradezco a todas aquellas personas que, con sus palabras de apoyo, consejos y acompañamiento, hicieron que este proceso fuera más llevadero. Cada aporte, por pequeño que pareciera, tuvo un significado importante en la culminación de esta etapa. Finalmente, agradezco a todas las experiencias adquiridas durante mi formación universitaria, pues cada aprendizaje, desafío y esfuerzo realizado contribuyeron a la persona y profesional que aspiro ser. Este trabajo representa el cierre de una etapa llena de retos, pero también el inicio de nuevos caminos que asumiré con compromiso y responsabilidad.

RESUMEN

Esta investigación analiza la tensión jurídica que existe entre la facultad de Instagram para modera contenidos publicados por los usuarios y el derecho a la libertad de expresión de las mujeres en los entornos digitales en Ecuador. Este estudio parte del reconocimiento de que las redes sociales se han convertido en espacios esenciales para la participación ciudadana, el intercambio de ideas y el acceso a la información. Sin embargo, las decisiones de moderación adpotadas por estas plataformas pueden afectar el ejercicio del derecho a la libertad de expresión .

Aunque la moderación de contenidos constituye una herramienta fundamental para garantizar entornos digitales seguros, esta facultad no debe ejercerse de manera arbitraria, sino respetando los principios de legalidad, necesidad, proporcionalidad, transparencia y debido proceso. Finalmente se proponen lineamientos jurídicos que están orientados a fortalecer la protección de la libertad de expresión de las mujeres frente a las decisiones de moderación de contenidos de instagram.

Palabras Clave: moderación de contenidos, Instagram, libertad de expresión, mujeres, plataformas digitales, derechos fundamentales, Ecuador, libertad de empresa, lineamientos jurídicos.

ABSTRACT

This research analyzes the legal tension between Instagram's authority to moderate user-generated content and women's right to freedom of expression in digital environments in Ecuador. The study begins by acknowledging that social media platforms have become essential spaces for civic participation and the exchange and receipt of information. However, moderation decisions made by these platforms can impact the exercise of the right to freedom of expression.

Although content moderation is a fundamental tool for ensuring safe digital environments, this authority must not be exercised arbitrarily; rather, it must respect the principles of legality, necessity, proportionality, transparency, and due process. Finally, the study proposes legal guidelines aimed at strengthening the protection of women's freedom of expression regarding Instagram's content moderation decisions.

Keywords: content moderation, Instagram, freedom of expression, women, digital platforms, fundamental rights, Ecuador, freedom of enterprise, legal guidelines.

Pontificia Universidad Católica del Ecuador
Seréis mis testigos
ESMERALDAS

INDICE GENERAL DE CONTENIDOS

INTRODUCCIÓN 8

MARCO TEÓRICO 12

1. La moderación de contenidos en Instagram como mecanismo de gobernanza privada. 12

1.1. Plataformas digitales como espacios para el ejercicio de derechos fundamentales. 12

1.2. La moderación de contenidos: concepto, finalidad y evolución. 14

1.3. Mecanismos de moderación de contenidos en Instagram. 17

1.4. La facultad jurídica de las plataformas privadas para moderar contenidos. 22

2. La libertad de expresión de las mujeres en los entornos digitales 24

1.1. La libertad de expresión como derecho humano y fundamental. 24

1.2. Tipos de discursos protegidos por la libertad de expresión. 27

1.3. Restricciones legítimas al derecho a la libertad de expresión 29

1.4. La libertad de expresión de las mujeres en los entornos digitales. 33

3. La tensión entre la moderación de contenidos en Instagram y la libertad de expresión de las mujeres. 37

3.1. La colisión entre la facultad de moderación de Instagram y el derecho a la libertad de expresión. 37

3.4. La protección jurídica de la libertad de expresión de las mujeres frente a la moderación de contenidos en Ecuador. 40

4. Lineamientos jurídicos para fortalecer la protección de la libertad de expresión de las mujeres frente a la moderación de contenidos en Instagram 42

4.1. Principios y criterios jurídicos que deberían orientar la moderación de contenidos. 42

4.3. Propuesta de lineamientos jurídicos para el contexto ecuatoriano 45

CONCLUSIONES 47

RECOMENDACIONES 48

Referencias 49

INTRODUCCIÓN

La consolidación de las plataformas digitales ha transformado de manera profunda la forma en que las personas ejercen derechos fundamentales, en particular la libertad de expresión. Redes sociales como Instagram se han convertido en espacios de participación pública, intercambio de ideas, activismo y denuncia social, donde millones de usuarios difunden opiniones e información de forma inmediata. No obstante, el funcionamiento de estos entornos también depende de mecanismos privados de moderación de contenidos, a través de los cuales las plataformas establecen reglas para regular las publicaciones de sus usuarios.

En este escenario, resulta relevante considerar la situación de las mujeres. Históricamente, han enfrentado mayores obstáculos para participar en los espacios públicos debido a prácticas de discriminación, violencia y desigualdad estructural, realidades que también se han trasladado al entorno digital. En la actualidad, las redes sociales constituyen un medio clave para que las mujeres ejerzan su libertad de expresión, visibilicen situaciones de violencia, impulsen causas vinculadas a la igualdad de género y participen en debates de interés público. En este sentido, las decisiones de moderación adoptadas por las plataformas pueden generar efectos diferenciados cuando inciden sobre este tipo de expresiones.

En este contexto, la presente investigación analiza la tensión jurídica existente entre la libertad de empresa, que sustenta la facultad de Instagram para establecer y aplicar sus Normas Comunitarias, y el derecho a la libertad de expresión de las mujeres en los entornos digitales. Si bien las plataformas privadas cuentan con autonomía para gestionar sus servicios, dicha facultad no es absoluta, ya que debe ejercerse en armonía con los derechos fundamentales reconocidos en la Constitución de la República del Ecuador y en los instrumentos internacionales de derechos humanos.

El objetivo general de la presente investigación es analizar la tensión jurídica entre el derecho a la libertad de empresa, que fundamenta la facultad de Instagram para moderar contenidos, y el derecho a la libertad de expresión de las mujeres en los entornos digitales en Ecuador. Para ello, se han determinado dos objetivos específicos que guiarán la investigación, el primero es describir los mecanismos de moderación de contenidos aplicados por Instagram y su relación con el derecho a la libertad de expresión de las mujeres y como segundo objetivo

específico proponer lineamientos jurídicos orientados a la protección del derecho a la libertad de expresión de las mujeres frente a la moderación de contenidos en Instagram.

La presente investigación surge de la necesidad de analizar un problema jurídico que ha tomado gran relevancia debido al auge de las plataformas digitales en la moderación de los contenidos publicados por sus usuarios. Si bien es cierto que Instagram cuenta con la facultad de poder moderar contenidos a través de sus Normas Comunitarias, el ejercicio de esta facultad puede generar tensiones cuando las decisiones adoptadas recaen sobre contenidos que se encuentran protegidos por el derecho a la libertad de expresión, particularmente cuando son emitidas por mujeres.

La importancia de este estudio se basa en que, aunque la moderación de contenidos se realice con el fin de mantener espacios digitales que sean seguros, esto no exime de analizar los límites que deben observar las plataformas cuando sus decisiones pueden vulnerar el ejercicio de los derechos fundamentales de las personas, en especial, el derecho a la libertad de expresión. En Ecuador, este conflicto presenta grandes desafíos debido a la insuficiente regulación sobre la materia y al limitado desarrollo doctrinario respecto de la incidencia de la moderación de contenidos en la libertad de expresión dentro de los entornos digitales.

Desde un punto de vista académico, la investigación tiene la finalidad de aportar al estudio de un tema contemporáneo que constantemente se encuentra en evolución y que ha sido objeto de análisis, finalmente, la presente investigación posee utilidad práctica, pues, el análisis permitirá identificar criterios jurídicos que permitan contribuir a comprender el alcance de la facultad de moderación por parte de plataformas digitales y los límites que esta facultad debe respetar para que así sea garantice el ejercicio efectivo del derecho a la libertad de expresión en el contexto ecuatoriano.

La presente investigación se desarrollará en el contexto jurídico de la República del Ecuador, debido a que el estudio se orienta a examinar la incidencia de la moderación de contenidos de Instagram en el ejercicio del derecho a la libertad de expresión de las mujeres desde la perspectiva del ordenamiento jurídico ecuatoriano, particularmente a partir de las

disposiciones constitucionales, la legislación nacional y los estándares internacionales que forman parte de su bloque de constitucionalidad.

Asimismo, la investigación comprenderá el período 2021-2025, por cuanto durante estos años se ha evidenciado un incremento en el uso de las plataformas digitales como espacios para el ejercicio de la libertad de expresión, así como un mayor debate jurídico sobre la moderación de contenidos, impulsado por el desarrollo de sistemas automatizados de moderación, la evolución de las políticas de Meta y la consolidación de estándares nacionales e internacionales en materia de derechos humanos aplicables a los entornos digitales.

La presente investigación es de tipo descriptiva y propositiva. Es descriptiva porque permitió estudiar la moderación de contenidos en Instagram y su relación con el derecho a la libertad de expresión de las mujeres, mediante el análisis de la normativa, la doctrina y la jurisprudencia aplicable. Asimismo, es propositiva, ya que, con base en los resultados obtenidos durante la investigación, se plantean lineamientos jurídicos orientados a fortalecer la protección de este derecho frente a las decisiones de moderación de contenidos en la plataforma.

La investigación se desarrolló bajo un enfoque cualitativo, debido a que el estudio se centró en el análisis e interpretación de fuentes jurídicas, tales como la Constitución de la República del Ecuador, instrumentos internacionales de derechos humanos, jurisprudencia, doctrina y las políticas de moderación de contenidos de Instagram. Este enfoque permitió comprender la problemática desde una perspectiva jurídica y establecer la relación existente entre la libertad de empresa de la plataforma y la libertad de expresión de las mujeres.

Para el desarrollo de la investigación se utilizó el método jurídico-dogmático, por cuanto fue necesario analizar e interpretar el ordenamiento jurídico, la doctrina y la jurisprudencia relacionada con la moderación de contenidos y la libertad de expresión. De igual forma, se empleó el método analítico-sintético, el cual permitió estudiar por separado los elementos que conforman la problemática investigada y, posteriormente, integrarlos para comprender la tensión jurídica existente entre los derechos objeto de estudio. Finalmente, se recurrió al método inductivo, ya que, a partir del análisis de normas, criterios doctrinarios y

pronunciamientos jurisprudenciales, fue posible formular las conclusiones y recomendaciones que se presentan en esta investigación.

MARCO TEÓRICO

1. La moderación de contenidos en Instagram como mecanismo de gobernanza privada.

1.1. Plataformas digitales como espacios para el ejercicio de derechos fundamentales.

En 1969, las universidades de los Ángeles y de Stanford culminaron un proyecto denominado ARPANET. Este proyecto transformó la manera en la que las personas intercambiaban información, dado que consistía en la creación del primer enlace que permitió la interacción entre los usuarios por medio de 2 dispositivos informáticos remotos. ARPANET fue incentivado por la Defense Advanced Research Projects Agency (DARPA) y constituye el origen de lo que hoy en día conocemos como internet.

No obstante, el éxito del internet no se debió únicamente a la creación de ARPANET, sino también al conjunto de avances tecnológicos que tuvieron lugar durante las décadas anteriores. En otras palabras, su funcionamiento fue posible gracias a diversas invenciones creadas, como, por ejemplo, el teléfono móvil, las computadoras, entre otras. En este sentido, toda la combinación de factores derivados de estos avances permitió que se desarrolle lo que hoy conocemos como las nuevas tecnologías.

Como resultado de este proceso evolutivo de la tecnología surgieron las plataformas digitales. Este fenómeno ha tenido gran impacto en la sociedad, pues su evolución ha transformado la manera en que las personas se comunican, participan e intercambian información. De la misma manera, han facilitado la creación de espacios de interacción y participación en los que millones de personas pueden expresar sus opiniones, compartir información y acceder a la misma de manera inmediata.

Sobre este aspecto, Wilfredo Fabián García (2020) define a las plataformas digitales como "Infraestructuras digitales que posibilitan la interacción entre dos o más grupos. Por lo tanto, se posicionan como intermediarios que reúnen a diferentes usuarios, clientes, anunciantes, prestadores de servicios, productores, proveedores y hasta objetos físicos".

Es decir, las plataformas no solo facilitan la interacción entre las personas, sino que también organizan y administran los espacios donde esta interacción se desarrolla. Esta característica adquiere relevancia, porque permite comprender a las plataformas digitales no solo como

espacios tecnológicos o empresariales, sino también como un fenómeno social que evolucionó hasta el desarrollo de las redes sociales.

En ese sentido, Francisco Valiente establece que los datos de uso de las redes no admiten discusión:

> "A principios de 2021, Facebook contaba con 2.740 millones de usuarios, YouTube con 2.291 millones, Whattsap con 2.000, Instagram con 1.221, TikTok con 689, Telegram con 500, Snapchat con 498, Pinterest con 442 y Twitter con 353. Todo ello sin citar redes profesionales como LinkedIn, de contactos como Tinder o Bumble, u otras de carácter regional, como la rusa VK. Cada usuario pasa conectado un promedio de dos horas y veinte minutos al día" (2023, pág. 169).

A partir de esta evolución, las redes sociales dejan de ser simples plataformas de interacción para convertirse en espacios donde las personas ejercen su derecho a la libertad de expresión, acceden a la información y participan activamente. En consecuencia, las plataformas digitales deben considerar el respeto de los derechos fundamentales de sus usuarios y, además, deben apegarse a los estándares establecidos por el derecho internacional, con el fin de crear un espacio digital seguro.

De acuerdo con datos emitidos por la UNESCO, en el año 2023, el 60% de la población mundial utilizó las redes sociales para expresarse e informarse. Este dato evidencia el crecimiento y el impacto que las redes sociales han generado en las personas dando lugar a sociedades cada vez más organizadas en redes. Para la UNESCO las redes sociales se han convertido en "El ámbito digital, espacio de libertad y nuevo foro de expresión y debate, que entrelaza nuestras relaciones sociales, identidades y vidas" (2023).

Lo anterior permite afirmar que las redes sociales tienen una dimensión pública, pues se trata de un ambiente que reúne a millones de personas, lo que permite que los contenidos publicados sean observados por terceros y que se expongan a la opinión pública. Debido a esto, las redes sociales adquieren relevancia jurídica, porque esta exposición da la apertura para que las plataformas digitales se vean en la necesidad de tomar decisiones que pueden tener como consecuencia la vulneración de los derechos de los usuarios que forman parte de los entornos digitales.

De este modo, podemos decir que las redes sociales son espacios de interacción que requieren el estudio y análisis jurídico de la facultad que tienen las redes sociales para moderar el contenido. Puesto que, a partir de aquellas decisiones dirigidas a mantener los espacios digitales seguros, se podrían afectar los derechos de las personas.

1.2. La moderación de contenidos: concepto, finalidad y evolución.

Los internautas no son solo receptores de información, como ocurre en medios tradicionales como la prensa, la televisión o la radio, sino que también actúan como creadores y difusores de contenido. Cada vez que una persona accede a internet, se pone en marcha una serie de actores que hacen posible su funcionamiento, entre los cuáles podemos encontrar a los proveedores del servicio de internet. Cada uno de ellos cumple con una función específica, lo que facilita la interacción social, la creación y la difusión del contenido.

Sobre este aspecto, Francisco Valiente (2023, pág. 170) señala que los roles que cumple cada actor son los siguientes:

> "En primer lugar, encontramos a los proveedores de acceso, encargados de suministrar el hardware básico (los routers, el sistema de cableado, el wifi…) y de su instalación. En segundo lugar, están los servidores de tránsito, también llamados conductores, que permiten la interacción entre el terminal de cada usuario y el punto de alojamiento de la información consultada. En tercer lugar, encontramos los servidores de alojamiento que, pese a ser desconocidos para el gran público, realizan la crucial labor de albergar todos los contenidos. En cuarto lugar, intervienen los proveedores de servicios en línea, mucho más conocidos, pues son los que permiten que realicemos actividades cotidianas como disponer de correo electrónico, leer la prensa, conectarnos a las redes sociales, etcétera (es decir, Facebook, Outlook, YouTube…). Por último, los llamados proveedores de búsqueda facilitan los listados de enlaces url que nos llevan directamente a la información solicitada" (2023, pág. 170).

Regular estos procesos es complejo, sobre todo en un espacio donde participan millones de personas que permanecen activas en las redes sociales desde cualquier parte del mundo. El crecimiento y el impacto de las redes sociales ha sido tan significativo que estos espacios también requieren mecanismos de regulación.

Cuando hablamos de moderación de contenidos, nos referimos a todos aquellos procesos o mecanismos destinados a revisar, filtrar, suprimir o eliminar los contenidos publicados por los usuarios cuando incumplen sus políticas y normas comunitarias.

En este sentido, Francisco Valiente (2023, pág. 182) sostiene que la moderación de contenidos "Engloba las prácticas dirigidas a la evaluación, categorización y posterior ratificación o eliminación de materiales tenidos por inadecuados de acuerdo con los criterios establecidos por los administradores de un portal de Internet".

Toda red social cuenta con sus propias políticas y normas comunitarias, sin embargo, estas difieren en su alcance y forma de aplicación. Puede suceder que una red social sea más permisiva que otra, no obstante, en esencia cada normativa se desarrolla respetando los derechos fundamentales.

Para la UNESCO (2023): "Los sistemas de moderación y curación de contenidos deben ser fiables y eficaces y a una escala que se ajuste al volumen de contenidos que se moderan, en todas las jurisdicciones en las que opera la plataforma". Es obligación de los espacios digitales contar con mecanismos de moderación que realmente cumplan con su función y que garanticen que no se vulneren derechos como el derecho a la libertad de expresión.

Las políticas de moderación de contenidos son el resultado de un proceso evolutivo influido por debates públicos, recomendaciones de organismos internacionales, demandas de la sociedad civil y evidencia académica sobre los riesgos del entorno digital. Diversos estudios y reportes combinados de UNICEF, ECPAT e INTERPOL (2022) han documentado cómo, desde mediados de la década de 2010, el crecimiento de las redes sociales generó preocupaciones globales relacionadas con la exposición de menores a contenido sexual, la explotación sexual en línea, la trata de personas, el grooming y la difusión de imágenes íntimas sin consentimiento, lo que llevó a exigir a las plataformas tecnológicas mayores estándares de responsabilidad en la moderación de contenidos.

Asimismo, la evolución de las políticas de moderación también ha estado influenciada por la necesidad de responder a otras problemáticas que se encuentran presentes en las redes sociales, entre ellas aquellas que afectan de manera particular a las mujeres. Actualmente, las redes sociales constituyen un espacio donde las mujeres ejercen su derecho a la libertad de expresión mediante la difusión de opiniones, manifestaciones artísticas, campañas de

sensibilización y denuncias relacionadas con violencia de género, igualdad o la salud sexual o reproductiva.

En este contexto, las normas comunitarias se han ido ajustando progresivamente para responder a estas preocupaciones sociales y perseguir fines legítimos como: Proteger a menores de edad, mujeres y la comunidad en general frente a contenido sexual explícito, prevenir la explotación sexual, trata y contenido íntimo no consentido, garantizar un entorno digital seguro y adecuado para una comunidad global diversa e impedir la circulación de contenidos que faciliten, fomenten o coordinen encuentros o servicios comerciales entre adultos.

Estos fines reflejan una respuesta institucional a problemáticas ampliamente reconocidas en el ecosistema digital contemporáneo y buscan regular conductas específicas relacionadas con el tipo de contenido publicado o el uso que se hace de la plataforma.

La UNESCO (2023, pág. 44) ha manifestado que:

> "Las plataformas digitales deben garantizar una acción rápida y decisiva para eliminar todo contenido sobre abusos sexuales a menores de que se tenga conocimiento o la retransmisión en directo de actos de terror, respetando los derechos de todas las personas, incluidos los grupos en situación de vulnerabilidad y marginación" (2023, pág. 44).

A partir de lo anterior, la moderación de contenidos surge como un mecanismo de regulación y control cuya finalidad es garantizar un entorno digital seguro para todos aquellos usuarios que se encuentran inmersos dentro una red social. Las redes sociales constituyen un fenómeno de gran alcance, pues en estos espacios podemos encontrar incluso a menores de edad quienes se pueden estar expuestos a diversas situaciones y contenidos no deseados.

En esta misma línea, la UNESCO (2023, pág. 40) también dice que:

> "En cualquier tipo de regulación, las plataformas digitales deben poder demostrar los sistemas o procesos que han implementado para garantizar que se aplica la debida diligencia en materia de derechos humanos de forma continua, incluidas las evaluaciones de impacto en los derechos humanos y de géneros, así como las medidas de mitigación de riesgos. Estos sistemas deben revisarse periódicamente y la revisión debe hacerse pública" (2023, pág. 40).

De lo anterior se desprende que la moderación de contenidos no consiste en una facultad absoluta de las plataformas digitales, sino una actividad que tiene que ejercerse bajo criterios de transparencia, diligencia y respeto por los derechos humanos. En consecuencia, las decisiones adoptadas en el marco

de estos procesos deben sustentarse en mecanismos claros, revisables y compatibles con los estándares internacionales de protección de derechos, a fin de evitar que la aplicación de las normas comunitarias genere restricciones al ejercicio del derecho a la libertad de expresión.

1.3. Mecanismos de moderación de contenidos en Instagram.

Instagram es una red social estadounidense que tiene como finalidad compartir fotografías, videos, enviar mensajes, compartir opiniones, entre otras funciones. Esta aplicación inició con un público joven, sin embargo, a medida que ha ido evolucionando y creciendo, ha logrado captar la atención de un público de todas las edades y de todas las profesiones. Instagram constituye una de las redes sociales más utilizada a nivel mundial como menciona Raquel Tarullo (2022) se ha consolidado como uno los principales espacios digitales para el ejercicio del derecho a la libertad de expresión. El servicio de Instagram es uno de los productos de Meta que proporciona Meta Platforms, Inc. (pág. 86)

Al contar con millones de usuarios en todo el mundo, resulta necesario establecer mecanismos que permitan controlar el tipo de contenido que se publica en la plataforma. Este control se realiza mediante la moderación de contenidos, entendida como una práctica que tiene el objetivo de revisar los contenidos publicados por los usuarios en redes sociales y plataformas digitales, es fundamental para el funcionamiento las plataformas en internet.

De acuerdo Pedro José Peña Jiménez (2021) “Para expertos como Tarleton Gillespie, una plataforma no solo sería poco funcional sin la moderación, sino que la moderación es un elemento indispensable para que exista una plataforma”. En tal sentido, la moderación de contenidos no constituye una actividad accesoria, sino un mecanismo necesario para garantizar el funcionamiento de las plataformas digitales y la convivencia entre usuarios.

Todas las plataformas moderan el contenido, unas intentan ser menos notarias, sin embargo, la moderación es inevitable. Entre las principales razones que justifican la moderación de contenidos pueden mencionarse las siguientes:

• La imagen Corporativa: Si bien muchas plataformas justifican la moderación de contenidos como una medida de responsabilidad social, en otros casos responde también al interés de preservar su imagen corporativa.

• Por razones económicas: Los incentivos económicos son quizás una de las principales razones que explica el actuar de las plataformas digitales, pues su objetivo es atender a un modelo de ganancia que se genera con la presencia de los usuarios en las redes.

Para que exista verdaderamente una moderación de contenidos efectiva, es fundamental establecer las reglas y principios orientadores que le permitan al usuario conocer cuáles son las normas comunitarias. Es entonces que la moderación requiere de reglas claras, precisas y públicas para los usuarios que participan en Instagram. La publicidad y claridad son características fundamentales de la moderación de contenidos para que los intermediaros puedan actuar dentro de un margen transparente.

Para las redes sociales como Instagram que permiten la publicación de contenido de terceros existen dos textos que son esenciales, estos son los términos y condiciones y las normas comunitarias. Estos documentos establecen las reglas que el usuario debe cumplir al utilizar la plataforma, así como las facultades que tiene la plataforma para administrar y moderar los contenidos publicados.

Desde una perspectiva jurídica, los términos y condiciones constituyen un contrato de adhesión mediante el cual el usuario acepta las condiciones establecidas unilateralmente por la plataforma para acceder al servicio. En virtud de este contrato, Instagram adquiere la facultad de aplicar medidas de moderación cuando considere que una publicación infringe sus políticas internas.

De hecho, en el apartado de servicio de ayuda en las condiciones de uso de Instagram, se da la bienvenida a la red social y se establece que “El Servicio de Instagram es uno de los Productos de Meta que te proporciona Meta Platforms, Inc., por lo que estas Condiciones de uso constituyen un acuerdo entre tú y Meta Platforms, Inc. Si no estás de acuerdo con ellas, no accedas a Instagram ni lo utilices” (Instagram, 2026). Esto reafirma el contrato que existe entre la plataforma digital y el usuario.

Por su parte las normas comunitarias son documentos dirigidos a los usuarios, son escritos en un lenguaje claro de tal manera que puedan ser entendidas por el usuario. En ellas se

establecen las conductas esperadas dentro de la red social, así como las actitudes o acciones que no se permiten, se establece la visión y valores de la red social.

De acuerdo con Meta:

> "Estas normas se basan en los comentarios que recibimos de las personas y en los consejos de especialistas en campos como tecnología, seguridad pública y derechos humanos. Para garantizar que se valore la opinión de todos, diseñamos con mucho cuidado normas que incluyen diferentes perspectivas y creencias, en particular las de personas y comunidades que, de otro modo, podrían ignorarse o marginarse.
> Cada día, las personas usan Facebook, Instagram, Messenger y Threads para compartir sus experiencias, conectarse con amigos y familiares, y crear comunidades. Nuestros servicios permiten que miles de millones de personas se expresen libremente en diversos países y culturas y en decenas de idiomas.
> Meta sabe lo importante que es que las personas se sientan seguras al comunicarse en Facebook, Instagram, Messenger y Threads, por lo que nos tomamos muy en serio nuestra responsabilidad de mantener los abusos fuera de nuestro servicio. Por eso, desarrollamos normas de lo que está y no está permitido en estos servicios" (Meta, 2026).

De esta manera, Meta presenta las Normas Comunitarias como un instrumento destinado a promover un entorno seguro para los usuarios. Sin embargo, dichas normas también constituyen el fundamento jurídico interno que legitima las decisiones de moderación adoptadas por la empresa respecto de los contenidos publicados en la plataforma.

De acuerdo Vladimir Alexei Chorny Elizalde, Luis Fernando García Muñoz y Grecia Elizabeth Macias Llanas (2022, pág. 41) en los procedimientos de moderación de contenidos pueden distinguirse dos momentos:

Tabla número 1: Procedimientos de moderación de contenidos

Tipo de moderación	Característica
Moderación ex ante	Trata de la moderación humana o sistematizada que sea realiza antes de que el contenido sea publicado.
Moderación ex post	Consiste en la revisión de contenidos después que han sido publicados.

Nota. Elaborada por Karla Vera López (2026).

Tabla número 2: La moderación de contenidos ex post

Tipo de moderación	Característica
Moderación Reactiva	Este es uno de los métodos más utilizado por las plataformas y consiste en que el moderador revisa la información de manera pasiva o hasta que esto sea solicitado por un tercero.
Moderación Activa	Este método es el más usado frente a discursos que no se encuentran protegidos por la libertad de expresión. Los moderadores constantemente de manera proactiva se encuentran buscando contenido que incumpla con las normas y términos y condiciones.

Nota. Elaborada por Karla Vera López (2026).

Tanto la moderación reactiva como la activa se puede realizar por dos tipos de moderadores:

Humanos: Son personas a quienes se le da un entrenamiento para moderar cierto tipo de contenido, esto va a depender de la red social en la que se encuentre.

La (UNESCO, 2023) establece lo siguiente:

> "Las personas moderadoras de contenido ya sean empleadas por las plataformas directamente o contratadas como contratistas externos a través de funciones externalizadas, deben recibir una formación adecuada, dominar el idioma o idiomas utilizados en las plataformas y estar familiarizadas con los contextos lingüísticos y culturales locales, ser evaluadas, autorizadas y recibir apoyo psicológico. Además, las plataformas deben establecer programas de apoyo, bien financiados y dotados de suficiente personal, para los moderadores de contenidos con la finalidad de minimizar el daño causado por su exposición recurrente a contenido violento o perturbador mientras realizan su trabajo. El número de personas moderadoras empleadas debe ajustarse a la complejidad y al volumen de contenidos que deben tratar" (UNESCO, 2023).

Es decir, la moderación de contenidos realizada por personas no puede limitarse a la aplicación mecánica de las normas comunitarias. La persona que revise el contenido debe

hacer una valoración de las circunstancias según lo requiera cada caso, ya que de su decisión va a depender la eliminación o no del contenido y con ello el pleno ejercicio de la libertad de expresión.

Algoritmos: Se trata de sistemas automatizados conocidos como machine learning, usan aprendizaje de máquina y son entrenados para buscar cierto tipo de contenido y eliminarlo si no cumple con los términos y condiciones.

Es importante señalar que los sistemas de moderación automatizados tienen como desventaja que tienen limitaciones para interpretar el verdadero significado de una publicación. Esto se debe a que no puede interpretar el contexto de una publicación de la misma manera que lo haría una persona.

En la red social de Instagram existen tanto la moderación de contenidos automatizada como sistematizada y se realiza siempre y cuando se contravengan los términos y condiciones y las normas comunitarias.

Si se viola una de estas normas el contenido será moderado por una de las formas ya antes mencionadas. Sin embargo, es importante tomar en cuenta que de considerar que la moderación realizada vulneró sus derechos como usuario de Instagram y realmente se siente inconforme con la decisión, la presente red social también cuenta con mecanismo de apelación a fin de que se garantice los derechos de los usuarios y el debido proceso.

El proceso es sencillo, en el Centro de Ayuda de Instagram existe un apartado llamado "Apelar una decisión de Instagram relacionada con el contenido ante el Consejo asesor de contenido" (Instagram, 2024), el usuario puede solicitar una revisión interna a Instagram. Si la decisión se mantiene tras dos revisiones humanas, puede acudir al Consejo Asesor de Contenido, que conoce tanto casos sobre contenido propio como de terceros reportados. En caso de que no se pueda acceder a la opción de apelar la persona puede usar el formulario de cuenta inhabilitada en el Centro de ayuda.

El Consejo Asesor está integrado por expertos de diversos países, incluyendo académicos, juristas, periodistas y especialistas en derechos humanos. Se trata de un órgano independiente

cuyas decisiones son vinculantes para Meta. Este órgano actúa bajo estándares de derechos humanos y permite a los usuarios apelar, de esta forma se garantiza los mecanismos de defensa y transparencia.

En definitiva, Instagram cuenta con sus propias normas comunitarias, políticas y formas de moderación del contenido, así como mecanismos de apelación a fin de que se garantice los derechos de los usuarios y la transparencia. Además de ello, cuenta también con un órgano independiente que es el Consejo Asesor, lo que da seguridad al usuario de que se va a tomar decisiones de manera imparcial y respetando sus derechos fundamentales.

1.4. La facultad jurídica de las plataformas privadas para moderar contenidos.

La posibilidad de que las plataformas digitales moderen los contenidos publicados por los usuarios genera controversia. Ello se debe a que, mediante el ejercicio de esta facultad, pueden verse afectados derechos fundamentales como la libertad de expresión. En este contexto, resulta primordial el análisis jurídico que permita fundamentar la libertad de las empresas como Instagram para establecer reglas de comportamiento y determinar qué medidas tomar en caso de que esto se incumpla.

Según José Anzures (2019, pág. 87) la libertad de empresa es "Un derecho propio de un Estado constitucional y democrático de derecho que, en términos generales, puede definirse como la facultad de toda persona para constituir una empresa y decidir sobre distintas opciones relacionadas con la actividad económica que desempeña".

En el Estado Ecuatoriano, esta facultad denominada libertad de empresa, se encuentra fundamentada en el artículo 66 numeral 15 de la Constitución, que garantiza el derecho a desarrollar actividades económicas y, por tanto, la facultad de las empresas de organizar sus servicios, definir reglas de funcionamiento y adoptar mecanismos de moderación de contenidos.

En este sentido el Relator Especial de Naciones Unidas sobre la libertad de expresión, David Kaye, señaló en su informe al Consejo de Derechos Humanos que las plataformas digitales desarrollan normas comunitarias como mecanismos de gobernanza privada para gestionar riesgos asociados a la seguridad de los usuarios y la integridad de los servicios digitales. Por

lo anterior, la facultad de Meta Platforms, Inc. de establecer y aplicar normas comunitarias constituye una manifestación legítima de la libertad de empresa.

No obstante, la facultad de las plataformas digitales para moderar contenidos no constituye una potestad ilimitada. Aunque las empresas privadas gozan de autonomía para organizar el funcionamiento de sus servicios, dicha autonomía encuentra límites cuando las decisiones adoptadas pueden afectar el ejercicio de derechos fundamentales de los usuarios. En consecuencia, la libertad de empresa debe ejercerse de manera compatible con otros derechos igualmente protegidos por el ordenamiento jurídico.

Es cierto que las plataformas digitales actúan bajo un ámbito privado, sin embargo, tienen un alcance global y no pueden actuar de manera aislada al derecho. Por esta razón, los organismos internacionales han manifestado que los procesos de moderación se deben aplicar en virtud de los estándares establecidos por los derechos humanos.

En virtud de lo establecido por las Naciones Unidas (2011), existen derechos y obligaciones que tienen las empresas, estas son:

- Reconocimiento del deber de respetar derechos.
- Prevención de riesgos (debida diligencia)
- Instancias de revisión externa.
- Transparencia y rendición de cuentas.

Debido a esto, las empresas privadas tienen la posibilidad de poder determinar que contenido se pública o no en sus redes sociales. Sin embargo, esta moderación del contenido no debe cruzar los límites establecidos por los derechos humanos, más bien debe existir respeto por los mismos. De esta manera se garantiza que los usuarios gocen efectivamente de sus derechos.

2. La libertad de expresión de las mujeres en los entornos digitales

2.1. La libertad de expresión como derecho humano y fundamental.

La libertad de expresión es un derecho fundamental dentro de un Estado democrático, pues permite que todas las personas puedan opinar, recibir, buscar o difundir información. Reconocer este derecho permite el desarrollo de sociedades más participativas y libres al facilitar el debate público y el intercambio de ideas. De la misma manera, es un pilar fundamental para el pleno ejercicio de otros derechos.

El derecho a la libertad de expresión se encuentra reconocido en el artículo 18 de la Constitución del Ecuador, el cual garantiza a todas las personas, de manera individual o colectiva, la facultad de buscar, recibir, producir y difundir información, así como acceder libremente a la información pública. De igual forma, este derecho se encuentra reconocido en el artículo 13 de la Convención Americana de Derechos Humanos ("CADH"), 14 y el artículo 19 del Pacto Internacional de Derechos Civiles y Políticos ("PIDCP"). Es decir, la libertad de expresión es un derecho que tanto en el ámbito internacional como en el ámbito nacional se encuentra reconocido como un derecho fundamental propio de todas las personas.

La libertad expresión como un derecho fundamental posee un contenido y alcance jurídico definido. Su protección no se reduce únicamente a la facultad de manifestar ideas, opiniones o pensamientos de manera libre, sino que comprende diversas facultades que garantizan el intercambio de información y fortalecer el debate público en una sociedad democrática. La amplitud de este derecho responde a la importancia que tiene la libertad de expresión para el funcionamiento de una sociedad democrática.

El artículo 13 de la CADH en sus numerales 1 ,2 y 3 establece lo siguiente:

> "1. Toda persona tiene derecho a la libertad de pensamiento y de expresión. Este derecho comprende la libertad de buscar, recibir y difundir informaciones e ideas de toda índole, sin consideración de fronteras, ya sea oralmente, por escrito o en forma impresa o artística, o por cualquier otro procedimiento de su elección.
>
> 2. El ejercicio del derecho previsto en el inciso precedente no puede estar sujeto a previa censura sino a responsabilidades ulteriores, las que deben estar expresamente fijadas por la ley y ser necesarias para asegurar: a. el respeto a los derechos o a la reputación de los demás, o b. la protección de la seguridad nacional, el orden público o la salud o la moral públicas.

> 3. No se puede restringir el derecho de expresión por vías o medios indirectos, tales como el abuso de controles oficiales o particulares de papel para periódicos, de frecuencias radioeléctricas, o de enseres y aparatos usados en la difusión de información o por cualesquiera otros medios encaminados a impedir la comunicación y la circulación de ideas y opiniones".

Es decir, se reconoce la libertad de expresión como un derecho amplio que protege la facultad de no solo de expresar ideas, sino también de buscar y difundir información. Además, prohíbe la censura previa y cualquier limitación que se realice a la libertad de expresión, será posible siempre y cuando se establezca en una ley y tenga fines legítimos.

Tabla número 3: Doble dimensión de la libertad de expresión

Dimensiones de la libertad de expresión	
Dimensión Individual	Dimensión Colectiva.
Consiste en la posibilidad de compartir y contrastar sus opiniones, pensamientos o ideas, así como a la posibilidad de tener acceso y difundir informaciones que son importantes para la sociedad. Esta dimensión permite que al emisor compartir la información por el medio que crea conveniente, cualquier impedimento hacía esta posibilidad restringe el derecho a la libertad de expresión	Trata del intercambio de ideas e información de manera masiva, es decir, consiste en una sociedad donde se cuenta con una diversidad de voces que participan de manera activa en una esfera pública y accesible. Implica el derecho que tienen todas las personas de conocer opiniones y noticias

Nota. Elaborada por Karla Vera López (2026).

Estas dos dimensiones se deben garantizar de manera simultánea para el pleno ejercicio del derecho a la libertad de expresión.

Las dimensiones individual y colectiva de la libertad de expresión deben garantizarse respecto de todas las personas sin discriminación. No obstante, el Derecho Internacional de los Derechos Humanos ha reconocido que las mujeres han enfrentado históricamente

obstáculos estructurales que limitan el ejercicio pleno de sus derechos, razón por la cual ha desarrollado instrumentos específicos destinados a garantizar la igualdad sustantiva y eliminar las distintas formas de discriminación que las afectan. En consecuencia, el análisis de la libertad de expresión de las mujeres no puede limitarse a las disposiciones generales contenidas en la Constitución, la Convención Americana sobre Derechos Humanos o el Pacto Internacional de Derechos Civiles y Políticos, sino que debe complementarse con los estándares internacionales especializados en materia de derechos de las mujeres.

En este sentido, la Convención sobre la Eliminación de Todas las Formas de Discriminación contra la Mujer (CEDAW) constituye el principal tratado internacional orientado a garantizar la igualdad entre hombres y mujeres. Aunque la Convención no reconoce expresamente la libertad de expresión como un derecho autónomo, sí impone a los Estados la obligación de adoptar medidas apropiadas para eliminar toda forma de discriminación contra las mujeres y asegurar que puedan ejercer plenamente sus derechos civiles, políticos, económicos, sociales y culturales en condiciones de igualdad. Desde esta perspectiva, el ejercicio efectivo de la libertad de expresión constituye un presupuesto indispensable para que las mujeres participen activamente en la vida pública y puedan hacer visibles sus opiniones, necesidades e intereses.

En desarrollo de la CEDAW, el Comité para la Eliminación de la Discriminación contra la Mujer emitió la Recomendación General n.º 23, relativa a la participación de las mujeres en la vida pública y política. En ella se reconoce que la participación efectiva de las mujeres exige que puedan expresar libremente sus opiniones e intervenir en los asuntos públicos sin discriminación ni obstáculos que limiten su voz.

Aunque esta recomendación fue adoptada en un contexto previo a la consolidación de las redes sociales, sus principios resultan plenamente aplicables a los entornos digitales, los cuales constituyen hoy uno de los principales espacios para el debate público y el intercambio de ideas.

Por otra parte, la Recomendación General n.º 35, que actualiza la Recomendación General n.º 19, reconoce que la violencia por razón de género contra las mujeres puede manifestarse mediante el uso de las tecnologías de la información y la comunicación. El Comité advierte

que estas formas de violencia tienen la capacidad de desalentar la participación de las mujeres en los espacios públicos y limitar el ejercicio de diversos derechos, entre ellos la libertad de expresión.

Este reconocimiento pone de manifiesto que los entornos digitales deben garantizar condiciones que permitan a las mujeres participar libremente sin enfrentar obstáculos derivados de prácticas discriminatorias o de medidas que restrinjan injustificadamente el ejercicio de sus derechos.

Finalmente, la Recomendación General n.º 40 reafirma la obligación de los Estados de adoptar medidas destinadas a garantizar la participación plena, igualitaria y efectiva de las mujeres en todos los ámbitos de la vida pública y en los procesos de toma de decisiones. Si bien esta recomendación se dirige principalmente a fortalecer la representación de las mujeres en la esfera pública, sus lineamientos también resultan relevantes para los entornos digitales, dado que estos se han consolidado como espacios de participación ciudadana, deliberación democrática y ejercicio de la libertad de expresión.

En consecuencia, el estudio de la moderación de contenidos en plataformas como Instagram adquiere especial relevancia, pues permite analizar si las decisiones adoptadas por estas empresas pueden incidir en el ejercicio efectivo del derecho a la libertad de expresión de las mujeres.

2.2.Tipos de discursos protegidos por la libertad de expresión.

La libertad de expresión protege ciertos discursos y prohíbe otros, esto se hace en razón de la calidad del sujeto que emite el discurso, es preciso aclarar que existen determinadas categorías de personas cuyo discurso merece una protección reforzada, entre ellos, servidores públicos, periodistas, los grupos en situación de vulnerabilidad y las víctimas de violencias basadas en género.

Respecto del tipo de discurso, existen diferentes tipos/categorías de contenidos y que dependiendo de la categoría hay una correlación con su protección a través de la libertad de expresión:

Tabla número 4: Tipos de discursos.

Tipos de discursos	
Discursos prohibidos	Es decir, aquellos que no se encuentran protegidos por el derecho a la libertad de expresión. Entre estos se incluyen la propaganda en favor de la guerra, la apología del odio nacional, racial o religioso que constituya incitación a la discriminación, la hostilidad o la violencia contra una persona o grupo, la pornografía infantil y la incitación directa y pública a cometer genocidio. (Corte Constitucional de Colombia, 2025)
Discursos especialmente protegidos.	En principio todas las formas de discursos se encuentran protegidos por la libertad de expresión independientemente de sus contenidos. No obstante, existen discursos que reciben una especial protección por la relevancia que tienen para el ejercicio de los derechos y para preservar la democracia. En esta categoría se encuentran los discursos como, por ejemplo, el discurso político, el debate sobre asuntos de interés público, y aquellos constituyen un ejercicio directo e inmediato de derechos fundamentales adicionales que se vinculan necesariamente a la libertad de expresión. (Corte Constitucional de Colombia, 2025)

Discursos que no están especialmente protegidos, pero tampoco prohibidos	En la tercera categoría se encuentran aquellos que no están especialmente protegidos, pero tampoco prohibidos. Por lo tanto, la eliminación de contenidos y cuentas en esta categoría deberá sujetarse al test de proporcionalidad. (Corte Constitucional de Colombia, 2025)

Nota. Elaborada por Karla Vera López (2026).

2.2. Restricciones legítimas al derecho a la libertad de expresión

En el punto anterior hemos dejado claro que ciertos discursos tienen protección bajo la libertad de expresión. La libertad de expresión recibe una protección especial que es fundamental dentro de una sociedad democrática para que las personas puedan acceder, recibir y transmitir la información. De esta manera se garantiza el debate público, que las personas puedan construir sus propios criterios y que puedan participar de manera activa en la sociedad. Sin embargo, pese a que existe esta protección para la libertad de expresión este derecho no es absoluto y existen situaciones en las cuáles se puede limitar la libertad de expresión.

De hecho, la Corte Constitucional del Ecuador ha señalado, en la sentencia 785-20-JP/22, que efectivamente este derecho no es absoluto y reconoció que puede estar sujeto a restricciones legítimas cuando estas se encuentren expresamente previstas en la ley, persigan un fin legítimo y resulten idóneas, necesarias y estrictamente proporcionales para la consecución de dicho fin. Estas restricciones tienen que provenir de una ley que sea clara, precisa y deben perseguir fines legítimos y necesarios dentro de una sociedad, estos fines han sido establecidos en la Convención Americana de Derechos Humanos.

Según Botero Marino Catalina, Guzmán Federico, Jaramillo Sofia y Gómez Salomé (2017, pág. 96) :

> "Una limitación de la libertad de expresión puede provenir de cualquiera de las autoridades estatales en cualquiera de las ramas del poder o demás segmentos de la estructura del Estado; también puede provenir eventualmente de particulares alentados por el Estado o por su propia iniciativa". (Botero Et al., 2017).

Lo dicho anteriormente resulta relevante para la presente investigación porque nos indica que la limitación a la libertad de expresión no solo puede ser limitada por la entidad estatal, sino que también por particulares. Actualmente las plataformas digitales como Instagram pueden limitar la libertad de expresión a través de la moderación de contenidos, es por esta razón que entra en debate hasta qué punto las decisiones tomadas por este tipo de plataformas, pueden vulnerar derechos fundamentales.

La Corte Interamericana de Derechos Humanos ha establecido el denominando test tripartito que consta de tres requisitos para determinar si una restricción de la libertad es compatible con la Convención Americana de derechos humanos. A su vez, la Corte Constitucional, en la sentencia 785-20-JP/22 como ya se mencionó anteriormente, ha señalado que la libertad de expresión no es un derecho absoluto y que puede ser limitado. Para que exista esta limitación, la medida debe superar el test de proporcionalidad que incluye el análisis de la legalidad de la medida, fin legítimo, idoneidad, necesidad y estricta proporcionalidad.

2.3. Test de proporcionalidad referente a libertad de expresión en redes sociales.

El test de proporcionalidad constituye una herramienta de interpretación jurídica para analizar si una medida que limita el ejercicio de un derecho fundamental resulta constitucionalmente válida.

Para Elena Córdoba Azcárate y Dolores Martín Villalba (2019, pág. 53) el test de proporcionalidad es:

> "Es un test tripartito aplicado por las instituciones europeas en el marco de la construcción del mercado interior desde los años 70 y que comenzó, más concretamente en la jurisprudencia del Tribunal europeo sobre libre circulación de mercancías, extendiéndose más adelante al resto de las liberta-des fundamentales" (2019, pág. 53).

En el ámbito de la libertad de expresión en redes sociales, este examen permite determinar se la restricción aplicada por redes sociales como Instagram, encuentra una justificación suficiente frente a la afectación al derecho a expresarse libremente. Este test comprende las fases de legalidad de la medida, fin legítimo, idoneidad, necesidad y estricta proporcionalidad.

2.3.1. Legalidad de la medida.

Este primer requisito del test exige que toda restricción a la libertad de expresión se fundamente en normas claras, previas y accesibles, esto se encuentra reconocido en el artículo 82 de la Constitución de la República del Ecuador. El hecho de que las limitaciones a la libertad de expresión se fundamenten en normas claras, previas y accesibles, permite que los usuarios puedan prever qué conductas están prohibidas y cuáles son las consecuencias de incumplirlas.

En el caso de Instagram, todos los usuarios conocen previamente lo que está permitido o no, pues al crear y utilizar su cuenta en esta red social acepta expresamente las normas de la plataforma. Dichas normas son accesibles y forman parte de un sistema único y centralizado de reglas comunitarias, disponibles desde el Centro de Ayuda y las Condiciones de Uso de la plataforma Instagram, lo cual facilita su conocimiento por parte de todos los usuarios.

Es importante tomar en consideración que existe una prohibición a la censura previa, la libertad de expresión no puede ser limitada antes que una persona exprese sus ideas. Si en el ejercicio de su derecho a la libertad de expresión infringe una ley o se vulnera los derechos de otros, la responsabilidad solo se puede imponer después de que el derecho a la libertad de expresión haya sido ejercido, de esta manera se garantiza que no exista la censura previa y se respete los estándares establecidos por la Convención Americana de Derechos humanos.

2.3.2. Fin legítimo.

El segundo requisito establece que la limitación debe perseguir un fin legítimo, no es suficiente con que una limitación se encuentre recogida en la ley, sino que además debe buscar proteger un interés muy importante. Las limitaciones a la libertad de expresión deben perseguir los fines establecidos en el artículo 13.2 de la Convención Americana de derechos humanos, estos son: la protección de los derechos de los demás, la protección de la seguridad nacional, del orden público, de la salud pública o de la moral pública.

Es en este sentido que, las normas comunitarias de Meta Platforms y en consecuencia de Instagram persiguen fines legítimos como: Proteger a menores de edad y la comunidad en general frente a contenido sexual explícito, prevenir la explotación sexual, trata y contenido

íntimo no consentido, garantizar un entorno digital seguro y adecuado para una comunidad global diversa, impedir la circulación de contenidos que faciliten, fomenten o coordinen encuentros o servicios comerciales entre adultos.

2.3.3. Idoneidad.

El tercer requisito señala que la limitación tiene que ser necesaria y proporcional dentro de una sociedad, aunque se encuentre la limitación recogida en una norma y aunque tenga un fin legítimo, la limitación será válida únicamente si no existe otra medida menos restrictiva, si el beneficio es mayor que la afectación al derecho y si realmente es idónea. En la sentencia 11-18-CN/19 párrafo 110 de la Corte Constitucional Ecuatoriana se establece que el requisito de idoneidad exige verificar si la medida adoptada es apta para contribuir al logro del fin legítimo perseguido.

2.3.4. Necesidad.

El análisis de necesidad de acuerdo la sentencia 21- 18-IN/24, 11 de julio de 2024, párr. 116 exige verificar que la medida adoptada sea la menos restrictiva para el ejercicio del derecho y que siga siendo igualmente idónea para alcanzar el fin constitucionalmente válido Ello implica determinar si al momento en que se realiza una limitación a la libertad de expresión se recurrió a medidas graduales antes de adoptar la sanción más intensa.

La Corte Interamericana de Derechos Humanos por su parte ha manifestado que la limitación a la libertad de expresión debe responder a las necesidades legítimas de la sociedad. Sin embargo, cuando se habla de una medida necesaria no se refiere a que sea útil, sino que en efecto como lo menciona la Corte Constitucional Ecuatoriana, implica que necesariamente se debe aplicar dicha limitación porque no existe una medida menos restrictiva para lograr el objetivo pretendido.

En este sentido es necesario determinar si la medida a adoptar constituye la medida menos gravosa, se debe analizar si existen otros mecanismos alternativos que permitan alcanzar estos objetivos con una menor afectación al derecho a la libertad de expresión. Si la respuesta es que, si se aplicaron las medidas menos lesivas y no existen otros mecanismos, entonces la medida resulta necesaria.

2.3.5. Estricta proporcionalidad

La estricta proporcionalidad de acuerdo el artículo 2.3 de LOGJCC, busca que exista un debido equilibrio entre la protección del derecho y la restricción realizada. En sentencia 18-18-IN/24, fecha 11 de enero de 2024, párr. 51 se establece que este último paso del test exige evaluar si los beneficios derivados de la medida superan la afectación que esta produce sobre los derechos de la persona afectada.

La Corte IDH, caso Kimel vs. Argentina. Fondo, Reparaciones y Costas. Sentencia de 2 de mayo de 2008, serie C núm. 177, párr. 84. Ha establecido ciertas pautas para la apreciación de la estricta proporcionalidad:

> "Deben evaluarse tres factores: i) el grado de afectación del derecho a la libertad de expresión grave, intermedia, moderada–, ii) la importancia de satisfacer el derecho humano que se busca proteger mediante la limitación de la libertad de expresión, y iii) si la satisfacción de dicho derecho justifica en consecuencia la afectación de la libertad de expresión" (CIDH, 2008).

Es decir, el requisito que se refiere a la estricta proporcionalidad consiste en un ejercicio valorativo y ponderativo entre los derechos e intereses en juego que solo podrá hacerse en ciertos casos por un juez y tomando en consideración las cuestiones particulares de cada caso. De esta manera se lograría determinar si una restricción a la libertad de expresión resulta admisible o no y que no resulte desproporcional o arbitraria.

En consecuencia, al ponderar los intereses en juego y los beneficios derivados de la medida se puede concluir si superan la afectación limitada que experimenta la persona que exige sus derechos. Por ello, las medidas que se adopten deben resultar razonables y proporcionales dentro del funcionamiento de los servicios digitales contemporáneos, y además se deben alinear con la evolución del marco regulatorio ecuatoriano.

2.4. La libertad de expresión de las mujeres en los entornos digitales.

La llegada de la era digital provocó que las redes sociales se convirtieran en un espacio importante para libertad de expresión. La pandemia del COVID-19 en el año 2020 aceleró la digitalización de múltiples actividades sociales. Esta emergencia sanitaria obligó a todas las personas del mundo a permanecer en cuarentena, una cuarentena que a medida que pasaba el tiempo se fue alargando. Tomando en consideración este contexto donde las personas

permanecían encerradas en sus hogares y de que sus actividades cotidianas cambiaron, el uso de las tecnologías tomó un rol importante que revolucionó al mundo.

Las personas empezaron a utilizar las herramientas tecnológicas en su día a día como un medio para cumplir con sus obligaciones como el trabajo o estudio. De la misma manera, el mundo empezó a estar mucho más conectado a través de las redes sociales como Instagram. A raíz de esta interacción digital se ampliaron las posibilidades de que todos, incluidas las mujeres tuvieran mayor participación en redes sociales. Las redes sociales consolidaron su papel en espacios de participación pública, permitiendo que un mayor número de mujeres difundieran información, compartieran experiencias, promovieran causas sociales y participaran activamente en debates de interés público.

No obstante, el incremento de la participación de las mujeres en los entornos digitales también ha evidenciado diversos desafíos para el ejercicio efectivo de su libertad de expresión. Entre ellos se encuentran la violencia digital, el acoso en línea, las campañas de hostigamiento y otras prácticas que pueden desalentar su participación en el debate público. La participación de las mujeres en espacios digitales también encuentra fundamento en el derecho internacional, específicamente el artículo 7 de la Convención sobre la Eliminación de Todas las Formas de Discriminación contra la Mujer (CEDAW), aquí se reconoce la obligación que tienen los Estados de garantizar la participación de las mujeres en la vida política y pública, de la misma manera esta disposición debería aplicarse en los entornos digitales.

De acuerdo con lo dicho por el informe de las Naciones Unidas (2018):

> “En 2018 la Relatora Especial sobre la violencia contra las mujeres y las niñas, sus causas y consecuencias definió la violencia en línea contra las mujeres para abarcar todo acto de violencia por razón de género contra las mujeres cometido, con la asistencia, en parte o en su totalidad, del uso de las tecnologías de la información y las comunicaciones (TIC), o agravado por este, como los teléfonos móviles y los teléfonos inteligentes, Internet, plataformas de medios sociales o correo electrónico, dirigida contra una mujer porque es mujer o que la afecta en forma desproporcionada” (CEDAW, 2018).

La violencia digital genera consecuencias que van más allá de las afectaciones individuales. En muchos casos produce un efecto inhibidor sobre el ejercicio de la libertad de expresión, ya que el temor a recibir ataques, amenazas o campañas de hostigamiento puede llevar a que

las mujeres decidan limitar su participación en los espacios digitales, eliminar publicaciones o abstenerse de pronunciarse sobre determinados temas de interés público. En consecuencia, el entorno digital deja de ser un espacio en el que todas las personas ejercen este derecho en condiciones de igualdad.

De hecho, Pooran Chandra Pande y Dr. KB Asthana (2025) afirman que:

> "Esta libertad de expresión se ve constantemente obstaculizada por barreras institucionalizadas como el acoso en línea, la violencia cibernética contra las mujeres, las normas sociales y el analfabetismo digital. A pesar de la creciente interconexión digital, la mayoría de las mujeres sufren acoso de género, estigma cultural y una alfabetización digital limitada que dificulta su participación en línea. Mediante una revisión de los desafíos existentes, la evaluación de programas de sensibilización y el análisis de intervenciones educativas, este estudio propone medidas para crear entornos en línea más seguros e inclusivos donde las mujeres puedan expresarse libremente" (Pooran & Asthana, 2025).

Sin embargo, los desafíos que enfrentan las mujeres en los entornos digitales no provienen únicamente de las intenciones de los usuarios. El ejercicio de la libertad de expresión también puede verse afectado por decisiones adoptadas por las propias plataformas digitales en el marco de sus políticas de moderación de contenidos. Si bien estas políticas persiguen objetivos legítimos, como la protección de la seguridad de los usuarios o la prevención de contenidos ilícitos, su aplicación puede incidir en la permanencia, difusión o eliminación de determinadas expresiones relacionadas con cuestiones de género, derechos sexuales y reproductivos, salud femenina o activismo social.

Justamente porque las redes sociales se han consolidado como espacios para el ejercicio de los derechos, las decisiones que sean adoptadas por las diversas plataformas respecto del contenido que se publica o no adquiere relevancia jurídica. Esto significa que cuando una plataforma realice la moderación de contenidos debe analizar la situación respecto de lo que se está publicando, tomando en consideración los estándares internacionales que protegen la libertad de expresión.

La libertad de expresión está garantizada como un derecho fundamental en el artículo 19 de la Declaración Universal de Derechos Humanos, en dicho artículo se establece que todos tenemos derecho a "buscar, recibir y difundir información e ideas por cualquier medio de expresión y sin consideración de fronteras". Es decir, las redes sociales deberían ser el espacio

que garantice lo dicho por el artículo ya antes mencionado, pues se trata de un espacio donde cualquier persona sin excepciones puede participar en igualdad de condiciones, compartiendo sus fotos, donde los ciudadanos puedan expresar sus opiniones, debatan o reclamen sus derechos

Desde un punto de vista de género, el derecho a la libertad de expresión está estrechamente vinculado al derecho de las mujeres a participar activamente en la vida política y pública en igualdad de condiciones. De hecho, el artículo 7 de la Convención sobre todas las formas de discriminación de las mujeres, establece que se debe garantizar a las mujeres el ejercicio de sus derechos sin discriminación alguna. De la misma manera, la Relatoría Especial para la Libertad de Expresión de la Comisión Interamericana de Derechos Humanos establece que los medios de comunicación (lo que abarcaría también Instagram), tienen un papel fundamental en la visibilización de las mujeres en el debate público.

No obstante, el problema de la libertad de expresión de mujeres en redes sociales no se trata únicamente de una cuestión digital, sino también de una cuestión de derechos humanos que tienen repercusiones en la vida real. La libertad de expresión femenina se ve censurada comúnmente, es mucho más probable que una mujer sea amenazada con violencia física o sexual que un hombre. El obstáculo más evidente que viven las mujeres en el entorno digital es el acoso, que se manifiesta de desde mensajes explícitos hasta difusión de imágenes no deseadas.

Por ello, resulta necesario analizar si dichas decisiones respetan los estándares nacionales e internacionales que protegen la libertad de expresión. En este sentido, la moderación de contenidos constituye un elemento que también puede incidir en el ejercicio de la libertad de expresión de las mujeres, razón por la cual su aplicación debe desarrollarse de manera compatible con los estándares internacionales de derechos humanos y con el principio de igualdad y no discriminación.

3. La tensión entre la moderación de contenidos en Instagram y la libertad de expresión de las mujeres.

3.1. La colisión entre la facultad de moderación de Instagram y el derecho a la libertad de expresión.

La libertad de expresión es un derecho que tenemos todas las personas sin excepción alguna y que ha sido reconocido por el derecho internacional, de la misma manera se debe garantizar el pleno ejercicio de este, tanto el ámbito público como privado, este derecho es ejercido por todos los usuarios de Instagram al momento que publican una foto, idea, opiniones, difunden o reciben información. Sin embargo, cada una de estas publicaciones o formas de expresarse están sujetas a un sistema de moderación establecido por Meta platforms Inc.

Como ya lo mencionamos en capítulos anteriores de la presente investigación la moderación de contenidos es una facultad que se le otorga a Instagram debido al derecho de libertad de empresa reconocido en el artículo 66.15 de la Constitución del Ecuador. Este derecho da la posibilidad de que empresas como Instagram puedan decidir qué contenido se publica o no en sus redes sociales. Este control se hace a través de sistemas de moderación siempre y cuando no se cumpla con los términos y condiciones y las normas comunitarias de la red social.

Cuando estas decisiones derivadas de la moderación de contenidos recaen sobre expresiones protegidas por la libertad de expresión se genera tensión entre la moderación de contenidos que implica su derecho a la libertad de empresa y el derecho a la libertad de expresión. Cuando ambos intereses se encuentran en conflicto es que se genera el presente problema jurídico, sobre todo cuando se trata de mujeres ejerciendo su libertad de expresión, pues, son las más propensas a que su contenido sea moderado. Como podemos darnos cuenta, el conflicto no se genera porque uno de los derechos sea ilegítimo, sino que la problemática radica en que ambos derechos poseen reconocimiento jurídico y en ciertas circunstancias pueden entrar en contradicción.

La moderación de contenidos es realizada por empresas como Instagram con el fin de garantizar un entorno digital seguro para todos sus usuarios, es por ello por lo que se crean

las llamadas normas comunitarias y los términos y condiciones, con el objetivo de que los usuarios conozcan que está permitido publicar en la red social, cuáles son los valores que persiguen como plataforma digital y que conozcan lo que está totalmente prohibido porque no va a fin con los principios de la plataforma. Pero la verdadera pregunta es ¿Cómo se manifiesta esta colisión en Instagram?

Instagram es una plataforma digital que ha tenido gran alcance a nivel mundial, debiendo entonces moderar el contenido publicado de millones de personas alrededor del mundo, esta moderación como ya lo mencionamos se puede realizar de manera sistematizada o humanizada. Primero se realiza la moderación de contenidos sistematizada y como resultado, las decisiones tomadas son automatizadas porque son realizadas por el algoritmo, el principal problema de este sistema de moderación es que el algoritmo a veces no logra interpretar de manera adecuada el contexto de las publicaciones y es por esta razón que se puede afectar a publicaciones realizadas por mujeres.

No obstante, es importante tomar en consideración que no toda moderación de contenidos se considera una limitación a la libertad de expresión, esto se debe a que existen discursos que no se encuentran protegidos por la libertad de expresión, un ejemplo de ello es la publicación de contenidos relacionadas con terrorismo o explotación sexual infantil. El problema nace cuando se limitan contenidos que, si se encuentran protegidos por la libertad de expresión como el discurso político, el debate sobre asuntos de interés público, y aquellos constituyen un ejercicio directo e inmediato de derechos fundamentales adicionales que se vinculan necesariamente a la libertad de expresión.

Si la moderación del contenido se realiza a una expresión protegida por el derecho a la libertad de expresión, la decisión de la moderación se debe analizar desde otra perspectiva, dado que, ya no se trata únicamente del ejercicio de su derecho a la libertad de empresa, sino que ya implica una posible vulneración a un derecho fundamental que es inherente de cada persona y que debería ser garantizado dentro de plataformas como Instagram. Sin embargo, la tensión no se genera únicamente cuando se moderan discursos especialmente protegidos por la libertad de expresión.

Es decir, si se moderan contenidos que se encuentran prohibidos como la propaganda en favor de la guerra, la apología del odio nacional, racial o religioso que constituya incitación a la discriminación, la hostilidad o la violencia contra una persona o grupo, la pornografía infantil y la incitación directa y pública a cometer genocidio. No existiría una tensión entre ambos derechos porque estos discursos no están protegidos por la libertad de expresión y se justificarían las razones para su eliminación.

Ahora, si Instagram moderara contenidos especialmente protegidos por la libertad de expresión como el discurso político, el debate sobre asuntos de interés público, y aquellos constituyen un ejercicio directo e inmediato de derechos fundamentales adicionales que se vinculan necesariamente a la libertad de expresión, se vulneraría el ejercicio de este derecho porque se tratan de discursos protegidos y no existen razones que justifiquen la eliminación de estos, configurándose como una vulneración al derecho a la libertad de expresión porque no existe un verdadero conflicto entre dos derechos.

Por otro lado, si Instagram modera contenidos que no se encuentran protegidos por la libertad de expresión, pero tampoco prohibidos se genera un conflicto. Esto se debe a que el hecho de que se traten de discursos que no están protegidos por la libertan de expresión abre la posibilidad de que Instagram decida que contenido se publica o no en sus redes sociales. De la misma manera el hecho de que tampoco se traten de discursos protegidos permite que las mujeres publiquen su contenido y ejerzan su libertad de expresión. Como podemos darnos cuenta se genera tensión entre el derecho a la libertad de empresa y la libertad de expresión de las mujeres porque se trata de dos derechos reconocidos jurídicamente.

La tensión que se genera no se resuelve con la prevalencia de uno de estos derechos sobre otro, dado que, como ya lo analizamos anteriormente el derecho a la libertad de empresa posee reconocimiento jurídico, así como el derecho a la libertad de expresión. En consecuencia, cuando estos derechos colisionan se deben aplicar los estándares analizados en el capítulo II, es decir, se debe aplicar el test de proporcionalidad que implica el análisis de la legalidad de la medida, de los fines legítimos, de la idoneidad, necesidad y la estricta proporcionalidad. Aplicando este test se va a poder lograr ponderar ambos derechos y determinar si efectivamente existió o no una limitación al derecho a la libertad de expresión.

La aplicación de este test es fundamental en el contexto de la moderación de contenidos porque va a poder permitirnos identificar si la moderación del contenido se realizó en uso legítimo del derecho a la libertad de empresa que le da Instagram la facultad de administración de la plataforma y decidir qué contenido representa un expresión desproporcionada o restrictiva de la libertad de expresión. Dicho en otras palabras, el test de proporcionalidad representa una herramienta jurídica que busca armonizar ambos derechos.

3.4. La protección jurídica de la libertad de expresión de las mujeres frente a la moderación de contenidos en Ecuador.

Proteger la libertad de expresión de las mujeres frente a los mecanismos de moderación representa uno de los mayores desafíos del derecho contemporáneo. Aunque las decisiones de moderación son una facultad otorgada a empresas como Instagram por gozar del derecho a la libertad de empresa, sus decisiones de moderación podrían afectar derechos fundamentales de las mujeres, especialmente si recaen sobre contenidos relacionados con el cuerpo femenino, la salud sexual y reproductiva, la igualdad de género o la denuncia de hechos de violencia. Es por es por esta razón que resulta imprescindible determinar cuáles son los mecanismos jurídicos otorgados por el sistema jurídico ecuatoriano para proteger a las mujeres frente a restricciones arbitrarias de su libertad de expresión.

En el Ecuador, la protección jurídica de la libertad de expresión encuentra su fundamento en la Constitución. El artículo 66 numeral 6 reconoce y garantizará a las personas el derecho a opinar y expresar su pensamiento libremente y en todas sus formas y manifestaciones. El presente artículo no hace distinción entre espacios físicos y digitales, por lo tanto, este artículo también es aplicable para expresiones que se realizan a través de plataformas digitales como Instagram. En consecuencia, de esto, las mujeres que usan Instagram para compartir sus ideas, pensamientos, opiniones, que difunden y reciben información, se encuentran amparadas por esta protección.

Sin embargo, su protección jurídica no se limita únicamente a este artículo, el artículo 11.2 de la Constitución reconoce el principio de igualdad y no discriminación. En este artículo se prohíbe toda forma de discriminación incluida la basada en género, este principio adquiere gran relevancia en el contexto del mundo digital porque obliga a que plataformas como

Instagram tengan sistemas de moderación que eviten un trato diferenciado o desproporcional hacia las mujeres. Dicho en otras palabras, los sistemas de moderación no deben reforzar estereotipos de género y así evitar la vulneración de derechos como la libertad de expresión de las mujeres.

A todo esto, se le suma la aplicación directa de los instrumentos de derechos humanos que forman parte del ordenamiento ecuatoriano. De acuerdo los artículos 417 y 416 de la Constitución, los tratados internacionales constituyen parámetros de interpretación para las autoridades judiciales. En consecuencia, de esto, tratados como la Convención Internacional de derechos humanos, la CEDAW y la Convención de Belém do Pará complementan esta protección constitucional del derecho a la libre expresión de mujeres.

Pese a que Instagram es una empresa privada, eso no significa que sus actuaciones deban ser ajenas al derecho. Esto se debe al gran impacto que ha tenido Instagram en la sociedad y al alcance que ha logrado constituyéndose como un espacio para el ejercicio de derechos fundamentales como lo son el derecho a la libertad de expresión. Debido a esto su sistema de moderación y sus decisiones de moderación deben ser compatibles como los estándares internacionales y nacionales de protección de derechos.

En el ordenamiento jurídico ecuatoriano no existe una ley específica que regule la moderación de contenidos en redes sociales y menos uno que regule la moderación de contenidos frente a la libertad de expresión de mujeres. Es por ello, que la protección de la libertad de expresión de mujeres frente a los sistemas de moderación de Instagram debe realizarse a partir de la interpretación en conjunta de la Constitución y de los Instrumentos Internacionales de Derechos Humanos. De esta manera se podría tomar una decisión adecuada y lograr determinar cuándo una decisión constituye un ejercicio legítimo del derecho a la libertad de empresa y cuándo, por el contrario, representa una decisión restrictiva del derecho a la libertad de expresión.

4. Lineamientos jurídicos para fortalecer la protección de la libertad de expresión de las mujeres frente a la moderación de contenidos en Instagram.

Los capítulos anteriores permitieron analizar la moderación de contenidos como una facultad de las plataformas digitales otorgada por el derecho a la libertad de empresa, el alcance de la libertad de expresión de mujeres y la tensión que puede surgir entre ambos derechos cuando una decisión de moderación afecta contenidos publicados por mujeres en Instagram. A parir de dicho análisis se logró evidenciar que, aunque la moderación de contenidos constituye un medio para lograr alcanzar un ambiente digital seguro, su aplicación podría generar restricciones a la libertad de expresión cuando no se aplican los estándares correspondientes.

En este sentido, el presente capítulo tiene como objetivo proponer los lineamientos jurídicos que contribuyan a fortalecer la protección del derecho a la libertad de expresión de las mujeres frente a la moderación de contenidos en Instagram. La propuesta no pretende restringir la facultad de las plataformas para administrar sus servicios, más bien busca establecer parámetros que permitan la armonización entre el derecho a la libertad de empresa de Instagram y el derecho a la libertad de expresión de mujeres, especialmente cuando las decisiones de moderación restrinjan el ejercicio de la libre expresión.

4.1. Principios y criterios jurídicos que deberían orientar la moderación de contenidos.

La presencia de sistemas de moderación de contenidos cumple una función esencial dentro de las plataformas digitales, esto se debe a que gracias a dichos sistemas se puede garantizar un entorno digital seguro. No obstante, tomando en consideración el alcance que tiene Instagram en el ejercicio a la libertad de expresión, las decisiones de moderación que tome no pueden responder únicamente a criterios empresariales o tecnológicos. Debido al impacto que pueden tener las decisiones de moderación en el ejercicio de los derechos fundamentales como la libertad de expresión.

Es por ello, que resulta imprescindible que los sistemas de moderación se desarrollen conformes a principios jurídicos que permitan garantizar el equilibrio entre el derecho a la libertad de empresa de Instagram y el derecho a la libertad de sus usuarios, particularmente de las mujeres que son quienes están más propensas a enfrentar riesgos específicos en los

entornos digitales. En esta línea de ideas, los lineamientos que se proponen a continuación toman como fundamento la Constitución de la República del Ecuador, los tratados internacionales de derechos humanos, la jurisprudencia de la Corte Interamericana de Derechos Humanos, la Corte Constitucional del Ecuador y los Principios Rectores sobre las Empresas y los Derechos Humanos de las Naciones Unidas.

Estos Instrumentos permiten identificar los principios que se deben orientar las decisiones de moderación cuando exista la posibilidad de que se afecte el derecho a la libertad de expresión, estos son:

4.1.1. Debida diligencia

Instagram en su sistema de moderación debe demostrar que en los procesos implementados se aplica la debida diligencia en materia de derechos humanos de manera continua. Estos sistemas deben ser revisados periódicamente y así mismo debe hacer publica dicha revisión. En razón de los Principios Rectores sobre las Empresas y los Derechos Humanos de las Naciones Unidas y la normativa Internacional, las plataformas como Instagram deben realizar evaluaciones periódicas para identificar y abordar cualquier tipo de impacto sobre los derechos fundamentales.

Las evaluaciones periódicas deben ser un elemento esencial para Instagram y servir de fundamento para el enfoque y funcionamiento de sus servicios. De acuerdo con el informe desarrollado por la (UNESCO, 2023):

> "Como mínimo, deben realizarse evaluaciones de riesgos y en materia de derechos humanos:
> a. Antes de realizar cualquier cambio significativo de diseño, tomar decisiones políticas importantes (incluidas las relacionadas con el sistema de publicidad, si procede), realizar cambios en el modo de funcionamiento o iniciar nuevas actividades o relaciones/asociaciones.
> b. Periódicamente, para proteger los derechos de todos los grupos en situación de vulnerabilidad y marginación, así como mujeres y niñas, periodistas, artistas, personas defensoras de los derechos humanos y defensoras del medio ambiente.
> c. Antes de cualquier proceso electoral para proteger su integridad.
> d. En respuesta a emergencias, crisis, conflictos o cambios significativos en el entorno operativo" (UNESCO, 2023).

Durante el proceso de debida diligencia Instagram debe garantizar un compromiso significativo con las partes involucradas, para así identificar riesgos específicos para ciertos grupos vulnerables como las mujeres.

4.1.2. Adherir a la normativa internacional en materia de derechos humanos.

Las plataformas digitales como Instagram deben garantizar que todas las consideraciones sobre los derechos humanos y el debido proceso se encuentren integrados en cada proceso, políticas y normas de moderación. Instagram debe garantizar igualdad en todos los procesos de moderación y la no discriminación, esto implica abordar los perjuicios, los estereotipos o las prácticas de moderación de contenidos que podrían afectar los derechos de las mujeres.

4.1.3. La transparencia de las plataformas.

Instagram debe informar periódicamente sobre cómo se están cumpliendo los principios de transparencia y explicabilidad. Esto implica comunicar como actúan en relación con sus condiciones de servicios y las normas comunitarias, así mismo debe dar respuestas a las solicitudes de información o eliminación de contenidos. La información debe ser lo más clara, precisa y lo más detallada posible. De acuerdo la UNESCO “La efectividad de los mecanismos de transparencia de las plataformas digitales debe evaluarse de manera independiente con respecto de la normativa internacional, a través de evaluaciones cualitativas y cuantitativas empíricas, para determinar si la información proporcionada cumplió su propósito”.

4.1.4. Información y herramientas a disposición de las personas usuarias.

Instagram debe tener sus condiciones de servicios completas y disponibles en el idioma oficial y en los idiomas de los países en los que opera, en el presente caso como nos encontramos en Ecuador el idioma debe ser en español. Esto asegura que la plataforma pueda dar respuesta a las usuarias en su idioma y tramitar sus reclamos por igual. Así mismo, Instagram debe garantizar que los informes y procesos de reclamación se encuentren en el idioma de las usuarias para garantizar un buen servicio.

4.2. Propuesta de lineamientos jurídicos para el contexto ecuatoriano

Las plataformas digitales si tienen derecho a moderar los contenidos que publican sus usuarios, ya que esta facultad forma parte de su libertad de empresa y de su autonomía para establecer reglas dentro de sus servicios. Sin embargo, ese poder no puede ejercerse de manera ilimitada, porque las decisiones que toman también pueden afectar derechos fundamentales especialmente la libertad de expresión de las mujeres.

En Ecuador no existe una norma específica que regule cómo deben moderarse los contenidos en las plataformas digitales. A pesar de ello si existen principios y reglas que sirven de guía, los cuáles se encuentran en la constitución en los tratados internacionales de derechos humanos y en la jurisprudencia nacional e interamericana con base en estos criterios, considero que la moderación debería orientarse por varios lineamientos.

En primer lugar, las plataformas tendrían que tomar sus decisiones respetando los derechos humanos y no únicamente sus políticas internas, además es importante que incorporen una perspectiva de género, ya que las mujeres enfrentan formas particulares de violencia y discriminación en los entornos digitales que deben ser consideradas antes de restringir una publicación.

Además, toda medida de moderación debería orientarse a ser proporcional. Es decir, antes de eliminar un contenido o suspender una cuenta deberían analizarse otras alternativas menos restrictivas cuando estas sean suficientes para cumplir el objetivo buscado.

Otro aspecto fundamental es la transparencia, las plataformas deberían explicar de forma clara por qué adoptan una determinada decisión indicando qué norma aplicaron y cómo se realizó la revisión del contenido. Del mismo modo es importante que existan mecanismos de apelación eficaces para que los usuarios puedan solicitar la revisión de decisiones que consideran injustas especialmente cuando esté en juego el ejercicio de derechos fundamentales.

La implementación de estos lineamientos no pretende limitar la autonomía de las plataformas digitales ni imponerles obligaciones desproporcionadas por el contrario busca que el ejercicio de sus facultades se desarrolle de manera compatible con el respeto y la protección de los

derechos fundamentales de las personas usuarias. En un contexto en el que gran parte del que transcienden el ámbito privado de las plataformas y pueden influir directamente en el ejercicio de la libertad de expresión, el acceso a la información y la participación en asuntos de interés público.

Asimismo, la adopción de estos criterios contribuiría a generar mayor seguridad jurídica tanto para las plataformas como para las mujeres usuarias al establecer parámetros claros que orienten las decisiones de moderación, esto permitiría reducir el riesgo de actuaciones arbitrarias o discriminatorias y fortalecer la confianza en los mecanismos de revisión y protección de derechos dentro de los entornos digitales.

Finalmente es importante reconocer que la moderación de contenidos constituye un desafío permanente debido al constante avance de la tecnología y a la aparición de nuevas formas de comunicación en internet, por esa razón los criterios propuestos deben entenderse como una base susceptible de actualización conforme evolucionen los estándares internacionales de derechos humanos la jurisprudencia y las herramientas tecnológicas utilizadas por las plataformas digitales y solo mediante a una revisión continua será posible alcanzar un equilibrio entre la libertad empresarial, la innovación tecnológica y la protección efectiva de los derechos fundamentales en especial de la libertad de expresión de las mujeres.

CONCLUSIONES

- La presente investigación permitió determinar que la moderación de contenidos constituye una herramienta esencial dentro del funcionamiento de Instagram, esto se debe a que permite establecer límites frente a publicaciones que pueden afectar a terceros, generar violencia o vulnerar las normas de convivencia dentro del entorno digital. Sin embargo, estas decisiones basadas en su derecho a la libertad de empresa pueden generar tensiones con el derecho a la libertad de expresión de mujeres.
- Se concluye que los mecanismos de moderación utilizados por Instagram, tienen una incidencia directa en la forma en que las mujeres participan y expresan sus ideas en espacios digitales. Aunque estos mecanismos persiguen finalidades legítimas, su aplicación puede producir afectaciones cuando se ejecutan sin considerar el contexto del contenido publicado, especialmente en expresiones relacionadas con igualdad de género, denuncia de violencia o reivindicación de derechos.
- Del estudio de la relación entre la libertad de empresa de Instagram y la libertad de expresión de las mujeres se concluye que la facultad de moderar contenidos no puede ejercerse de manera ilimitada. Al involucrar el ejercicio de un derecho fundamental, las decisiones adoptadas por las plataformas digitales deben responder a criterios de proporcionalidad, de modo que las restricciones aplicadas sean realmente necesarias y guarden equilibrio entre la protección de otros derechos y la garantía de la expresión dentro del entorno digital.
- Finalmente, se concluye que es necesario promover lineamientos jurídicos que permitan establecer un equilibrio entre la autonomía de las plataformas digitales y la protección de la libertad de expresión de las mujeres. Para ello, resulta fundamental fortalecer la transparencia en los procesos de moderación, garantizar mecanismos adecuados de reclamación y revisión de decisiones, e incorporar un enfoque de género que permita valorar el contexto de los contenidos antes de aplicar restricciones que puedan limitar la participación de las mujeres en los espacios digitales.

RECOMENDACIONES

- Se recomienda que las plataformas digitales, especialmente Instagram, fortalezcan sus políticas de moderación de contenidos mediante criterios más claros, transparentes y accesibles para los usuarios. Si bien las plataformas cuentan con la facultad de establecer reglas internas para garantizar un entorno seguro, estas deben encontrarse debidamente fundamentadas y permitir que los usuarios conozcan las razones por las cuales un contenido es restringido o eliminado.
- También, que los procesos de moderación de contenidos incorporen un análisis contextual de las publicaciones, especialmente cuando se trate de expresiones relacionadas con los derechos de las mujeres, igualdad de género o denuncia de situaciones de violencia. La aplicación automática de restricciones puede generar interpretaciones limitadas del contenido y afectar expresiones que requieren una valoración más amplia considerando las circunstancias en las que fueron realizadas.
- Además, que la actuación de las plataformas digitales frente a la restricción de contenidos que involucren el ejercicio de la libertad de expresión se base en criterios de proporcionalidad. En este sentido, antes de aplicar medidas como la eliminación de publicaciones o la suspensión de cuentas, deberían considerarse alternativas menos restrictivas que permitan alcanzar el objetivo de protección sin afectar de manera innecesaria la participación de las mujeres en los espacios digitales.
- Se recomienda promover desde el ámbito jurídico ecuatoriano un mayor desarrollo normativo y académico sobre la responsabilidad de las plataformas digitales en la protección de derechos fundamentales dentro de internet. Esto permitiría establecer parámetros orientados a garantizar la transparencia, la existencia de mecanismos efectivos de apelación y la incorporación de una perspectiva de género en los procesos de moderación de contenidos, buscando un equilibrio entre la libertad empresarial de las plataformas y la protección de la libertad de expresión.

REFERENCIAS BIBLIOGRÁFICAS

Galarza, M. S. (2022). GARANTÍAS FRENTE A LA MODERACIÓN DE CONTENIDOS EN LA PROPUESTA DE REGLAMENTO ÚNICO DE SERVICIOS DIGITALES. *Revista CESCO*, 160-179.

Anzures Gurría , J. J. (2019). La libertad de empresa como derecho fundamental en México. *Revista Jurídica Digital UANDES* , 80-93.

Barata Mir, J. (2022). LIBERTAD DE EXPRESIÓN, REGULACIÓN Y MODERACIÓN PRIVADA DE CONTENIDOS. *Revista De Pensamiento jurídico*, 88-107.

Botero Marino, C., Guzmán Duque, F., Jaramillo Otoya, S., & Gómez Upegui, S. (2017). *El derecho a la libertad de expresión* . Bogatá.

Chorny Elizalde, V. A., García Muñoz , L. F., & Macias Llanas, G. E. (2022). *LA MODERACIÓN DE CONTENIDOS DESDE UNA PERSPECTIVA INTERAMERICANA.* Ciudad de México: Gibrán Aquino Pineda.

Cordova Azcárate , E., & Martín Villalba , D. (2019). LA APLICACIÓN DEL TEST DE PROPORCIONALIDAD . *REVISTA ESPAÑOLA DE EDUCACIÓN FÍSICA Y DEPORTES* .

García , W. F. (2020). *Plataformas Digitales.* Ciudad de México: ISEF EMPRESA LIDER .

ONU. (2011). *PRINCIPIOS RECTORES SOBRE LAS EMPRESAS Y LOS DERECHOS HUMANOS.* New york.

ONU. (2018). *Informe de la Relatora Especial sobre la violencia contra la mujer, sus causas y consecuencias acerca de la violencia en línea contra las mujeres y las niñas desde la perspectiva de los derechos humanos.*

Ordoñez, A., Espinoza , F., & Vizuete , C. (2023). Libertad de Expresión, Discurso de Odio y Violencia Política de Género en Procesos Electorales. *Revista Enfoques de la Comunicación*.

Peña Jimenéz, P. J. (2021). ENTRE ANALOGÍAS Y METÁFORAS: EL DEBATE SOBRE LA MODERACIÓN DE CONTENIDOS EN LAS REDES SOCIALES. *Revista de las Cortes Generales*, 265-311.

Pooran , P., & Asthana, K. (2025). Libertad de expresión de las mujeres en las redes sociales. *REVISTA DE INVESTIGACIÓN CIENTÍFICA Y TECNOLOGÍA*, 64-69.

Relatoría Especial para Libertad de expresión Comisión Interoamericana de Derechos Humanos. (2025). *Marco jurídico.*

Tarullo, R. (2022). *Jóvenes, medios y redes sociales.* Ciudad Autónoma de Buenos Aires: CEDi Centro de Edición y Diseño. UNNOBA.

UNESCO. (2023). *Directrices para la gobernanza de las plataformas digitales: salvaguardar la libertad de expresión y el acceso a la información con un enfoque de múltiples partes interesadas.*

UNICEF; ECPAT; INTERPOL. (2022). *Disrupting Harm in Southeast Asia: Evidence on Online Child Sexual.*

Vaca Villarreal , P. (2025). *Informe Anual de la Comisio Interamericana de Derechos Humanos.*

Valiente Martínez , F. (2023). *La libertad de expresión y las redes sociales.*