

ARTÍCULO DE INVESTIGACIÓN

Inteligencia artificial y creación de pornografía deepfake no consensuada: derechos humanos y perspectiva ético jurídico*

Artificial intelligence and the creation of non-consensual deepfake pornography: Human rights and an ethical-legal perspective

Jireth Melissa Becerra Rangel¹  [CvLAC](#), Enuar Andrés Ortiz Quintero¹  [CvLAC](#), Jean Carlos Vergara Llanos²  [CvLAC](#)

* Artículo de investigación derivado del proyecto de investigación titulado: El impacto de la inteligencia artificial generativa en la creación de pornografía deepfake frente a la afectación de los derechos de imagen, intimidad personal y el buen nombre.

¹ Estudiante de octavo semestre de derecho de la Universidad Francisco de Paula Santander Ocaña. Estudiante del semillero de derecho público (Opus Status). Filiación institucional: Universidad Francisco de Paula Santander Ocaña.

² Especialista en Contratación Estatal de la Universidad Libre. Abogado. Docente catedrático de la Universidad Francisco de Paula Santander Ocaña. Coordinador del semillero de investigación en Derecho Público Opus Status UFPSO. Filiación institucional: Universidad Francisco de Paula Santander Ocaña.

Forma de citar: Becerra- Rangel, Jireth Melissa; Ortíz- Quintero, Enuar Andrés; Vergara- Llanos, Jean Carlos. “Inteligencia artificial y creación de pornografía deepfake no consensuada: derechos humanos y perspectiva ético jurídico”. En: *Revista CES Derecho*. Vol. 16. No. 2, mayo-agosto de 2025. pp. 38-60. <https://doi.org/10.21615/cesder.7910>

Resumen

Este artículo aborda el impacto de la pornografía deepfake no consensuada en los derechos de imagen, intimidad personal y buen nombre, planteando un dilema ético-jurídico sobre la necesidad de adaptar la legislación a los nuevos desafíos tecnológicos. El objetivo principal es analizar cómo las tecnologías emergentes, particularmente la inteligencia artificial generativa (IAG), afectan los derechos humanos en estos contextos. La metodología empleada es dogmática jurídica, utilizando un enfoque cualitativo y hermenéutica jurídica a través de un método exploratorio y descriptivo, con análisis de casos de pornografía deepfake. Los resultados indican que la IA afecta negativamente estos derechos, especialmente en mujeres, y revela que la legislación actual no está suficientemente actualizada para enfrentar estos nuevos desafíos. Se concluye que es urgente adecuar las normas jurídicas para garantizar la protección efectiva de los derechos humanos, ofreciendo un marco legal adecuado para la resolución de estos casos en la práctica judicial.

Palabras clave: afectación; derechos humanos; inteligencia artificial; impacto; pornografía deepfake.

Abstract

This article addresses the impact of non-consensual deepfake pornography on image rights, personal privacy, and good name, raising an ethical and legal dilemma regarding the need to adapt legislation to new technological challenges. The main objective is to analyze how emerging technologies, particularly generative artificial intelligence (GAI), affect human rights in these contexts. The methodology employed is legal dogmatic, utilizing a qualitative approach and legal hermeneutics through an exploratory and descriptive method, with analysis of deepfake pornography cases. The results indicate that AI negatively affects these rights, especially those of women, and reveal that current legislation is not sufficiently updated to address these new challenges. It is concluded that it is urgent to adapt legal norms to guarantee the effective protection of human rights, offering an adequate legal framework for the resolution of these cases in judicial practice.

Keywords: Impact; human rights; artificial intelligence; deepfake pornography.

Introducción

La presente investigación plantea un análisis sobre la evolución de la inteligencia artificial (IA), específicamente en la creación de pornografía deepfake no consensuada y el impacto que esto genera en los derechos fundamentales como el derecho a la imagen, intimidad personal y el buen nombre, entendida como una afectación sistemática de derechos humanos, el cual implica retos y desafíos desde la legislación vigente en Colombia. En ese sentido, la IA según autores como Kaelbling y Moore indicaron que ésta se entiende “como la capacidad de las máquinas para adaptarse a nuevas situaciones, resolver problemas, diseñar planes y realizar otras funciones que requieren un cierto nivel de inteligencia” (Como se citó en Martínez et. al., 2023); es decir, presenta un enfoque en el desarrollo de programas o máquinas que se centran en la solución de problemas que podrían imitar el comportamiento humano e incluso superar su capacidad, razón por la cual, la inteligencia artificial es una tecnología que está transformando la era y la forma en cómo se interactúa con el mundo digital; sin embargo, su gran desarrollo no está exento de desafíos y controversias que amenazan los derechos humanos.

Al hablar acerca de la inteligencia artificial, es importante indicar que esta se puede organizar de distintas formas, una de ellas, es la inteligencia artificial generativa (IAG) la cual “se centra en la creación de datos, imágenes, texto y otros contenidos de forma autónoma” (ISDI, 2023). Ahora bien, es necesario indicar que esta tecnología usa una serie de algoritmos articulados con redes neuronales con la capacidad de elaborar textos e inclusive imágenes para, posteriormente generar un contenido nuevo con características únicas, siendo uno de los grandes avances que se han generado con el uso de la IAG y se siga produciendo en el futuro. Sin embargo, debido a esto se observa un crecimiento exponencial en la creación y difusión de imágenes y vídeos falsificados, lo que ha provocado controversias significativas en la sociedad ya que este fenómeno conocido como “*deepfake*”, implica la manipulación de contenido audiovisual de manera que parezca auténtico, aunque en realidad ha sido alterado digitalmente. Los deepfakes pueden engañar a los espectadores al modificar caras, voces, movimientos de labios e incluso posturas corporales, todo mediante tecnología avanzada de procesamiento audiovisual (Boté y Váñez, 2021).

Gran parte de los deepfakes generados en internet, son de contenido sexual explícito de personas ajenas a la industria del porno, las cuales en su mayoría son mujeres; en ese sentido, realizando una aproximación conceptual sobre la pornografía deepfake, esta hace referencia a “los vídeos e imágenes porno modificados utilizando tecnología de intercambio de caras a través de inteligencia artificial, por lo que el rostro del protagonista se reemplaza por la de otra persona”. (BBC News Mundo, 2018); es decir, las personas son propensas a ser retratadas en deepfakes pornográficos sin dar consentimiento alguno y los usuarios de internet sin tener la capacitación técnica o profesional pueden crear material de este tipo, incluyendo menores de edad.

La inteligencia artificial en su esencia es una herramienta con potencial beneficioso para la humanidad, pero no siempre es ocupada para tal fin. Su mal uso transgrede los límites morales y ético generando como consecuencia la vulneración derechos fundamentales.

El propósito de esta investigación es analizar la creación de pornografía deepfake no consensuada y su repercusión en los derechos de imagen, intimidad personal y el buen nombre; razón por la cual, se busca plantear cómo el empleo de inteligencia artificial puede generar situaciones que pongan en riesgo estos derechos, provocando así un debate sobre las implicaciones éticas y legales del contenido multimedia producido mediante esta tecnología.

Teniendo en cuenta la problemática expuesta es necesario que surja el siguiente interrogante, ¿de qué manera puede el marco jurídico adaptarse para garantizar la protección efectiva de los derechos de imagen, intimidad personal y el buen nombre frente a los riesgos generados por la pornografía deepfake no consensuada?

Metodología

La investigación principal se desarrolló mediante un enfoque cualitativo, con apoyo de la hermenéutica jurídica

la cual se entiende como el arte y/o el proceso de interpretación, siendo de gran importancia para el derecho, toda vez que permite darle sentido a la norma y la jurisprudencia en cada caso en concreto (Blanco, 2018. P. 39); con énfasis en el análisis de casos concretos sobre pornografía deepfake y vulneración de derechos humanos, apoyado en el estudio de la literatura académica y jurídica relevante (Marshall y Rossman, 2016).

Así mismo, el presente estudio se enmarca dentro de la ciencia jurídica, siendo una investigación dogmática jurídica, en la cual el propósito del derecho en estas investigaciones, es resolver el estudio a través de las fuentes formales del derecho (Giraldo, 2012).

Debe indicarse que se utilizaron fuentes secundarias o indirectas; es decir, la revisión bibliográfica, consistente en el análisis de información pertinente desde un análisis internacional y nacional respectivamente (Creswell y Creswell, 2018).

En cuanto al nivel de profundidad, se trata de una investigación exploratoria y descriptiva. Lo anterior, teniendo en cuenta que el estudio de la IA y la creación de pornografía Deepfake no consensuada ha tenido un gran auge en los últimos años, y ha sido poco estudiado, entendido como un espacio de análisis constante, lo cual implica para los Estados la obligación de adecuar su legislación vigente a los cambios que trae consigo el surgimiento de las nuevas tecnologías.

Finalmente, se expone que mediante esta investigación se pretende analizar los diferentes casos registrados de pornografía Deepfake y su impacto en los derechos de imagen, intimidad y buen nombre; en segunda medida, un análisis de legislación comparada que pretende observar la normativa vigente para cada país frente al tema objeto de estudio y, por último, plantear recomendaciones frente a la IA y su impacto en relación con los derechos humanos.

Fundamento y aproximación conceptual sobre inteligencia artificial

Para realizar una aproximación conceptual sobre el uso de la inteligencia artificial, debe indicarse que hace referencia a sistemas de carácter informáticos que tienen como propósito “imitar la función cognitiva humana a través de máquinas, procesadores y softwares con el objetivo de realizar tareas de procesamiento y análisis de datos. En términos sencillos, se trata de máquinas diseñadas para razonar, aprender, realizar acciones y resolver problemas”. Lo anterior permite entender que la IA está integrada por un sistema de programación con la capacidad de almacenar información sobre un asunto en particular para convertirla en nuevo conocimiento al servicio de las actividades humanas (Ferroval, s.f).

Por consiguiente, “el avance en el desarrollo de la inteligencia artificial (IA) logró la posibilidad de crear imágenes falsas indistinguibles de la realidad. La existencia de esta tecnología derivó en la reconstrucción de perfiles de personas ya existentes para recrear acciones que ellas nunca hicieron” (Urquiza, 2022). Ahora bien, es necesario entender el concepto sobre deepfake, el cual es una “técnica con la que se consiguen imágenes o vídeos de objetos o rostros de personas aparentemente reales a través de algoritmos de aprendizaje de IA”. (Urquiza, 2022), que puede ser utilizado para distintas actividades, pero también puede ser usada por cibercriminales para suplantar identidades de las personas.

Es por ello que, en la actualidad la IA ha ido generando incertidumbre a escala global, toda vez que su impacto y alcance es cada vez mayor y su usabilidad se ha venido presentando en la vida cotidiana con más frecuencia; siendo necesario estar atentos frente a cualquier situación que por sus características pueda suponer una desventaja en su implementación (Murillo y Ramirez, s.f).

Finalmente, estos avances conllevan también importantes retos éticos desde una perspectiva jurídica precisando que “la ética jurídica aplicada a la inteligencia artificial (IA) diverge de la ética general, ya que se orienta en las implicaciones legales y en los derechos esenciales que podrían verse afectados por el uso de estas tecnologías” (Pérez, et. al., 2025, p. 9); razón por la cual es necesario indicar que entre los retos desde una perspectiva ética

jurídica se tienen la “transparencia y la responsabilidad en el uso de la tecnología, el riesgo de contenidos engañosos o sesgados, la protección de la privacidad y de los derechos de autor, y la necesidad de evitar que el uso de la IA genere discriminación o acentúe desigualdades. Son cuestiones clave que se deben abordar para asegurar que la tecnología se utilice de manera responsable y se respeten los valores fundamentales de la sociedad” (Franganillo, J. 2023).

Inteligencia artificial generativa y pornografía deepfake

La pornografía deepfake se popularizó en el año 2017, en reddit, una red social que sirvió como instrumento para que se desarrollara el primer brote de deepfake, el cual se propagó y convirtió a reddit en el punto de intercambio de pornografía adulterada. Todo empezó cuando el usuario denominado “deepfakes” (de ahí proviene el término) compartió varios videos e imágenes manipuladas de actrices y cantantes famosas realizando actividades sexuales o eróticas. Los clips fueron manipulados gracias a la inteligencia artificial generativa, la cual utiliza un sistema de redes generativas antagónicas (GAN) para crear imágenes o vídeos hiperrealistas con base en las indicaciones proporcionadas por el usuario (Rodrigo, . 2022).

GAN utiliza dos redes neuronales de IA: la primera, generativa, se encarga de crear una variabilidad aleatoria, a partir de datos que ya conoce, y la segunda, discriminativa, se encarga de evaluar e identificar si los datos creados por la red generativa son falsos, o si por el contrario son parte de un conjunto de datos ya existentes. El proceso implica entrenar ambas redes a través de un juego competitivo donde se enfrentan los dos modelos. El generador crea los datos de tal manera que el discriminador no sea capaz de identificar entre los datos reales o falsos. Mientras más se enfrentan mejores serán los resultados, tanto de generador, para crear los datos más realistas posibles con el objetivo de engañar a discriminador, y discriminador para identificar hasta la más mínima diferencia entre aquello que es real y falso (Moncada, 2020).

Esta tecnología permite la creación de contenido multimedia falso, especialmente vídeos, conocidos como deepfakes. GAN genera imágenes, audios y videos de manera hiperrealista. Actualmente los deepfakes han sido orientados para difundir información falsa, lo que abre un campo amplio de desconfianza sobre lo que antes era considerado confiable dado a su origen: imágenes y videos. Asimismo, debido a la hiperrealidad del deepfake resulta bastante fácil que las personas compartan este contenido falso creyendo que es real.

Los deepfake varían según su contenido y propósito, mientras que algunos se emplean con fines humorísticos en redes sociales o entre amigos, otros en cambio, difunden videos o imágenes falsos de una persona realizando actividades sexuales, esto último es conocido como pornografía deepfake. Según un informe realizado por Home Security Heroes:

La pornografía deepfake representa el 98% de los deepfakes online[...]. De acuerdo a su estudio, “entre 2022 y 2023 la cantidad de pornografía deepfake creada aumentó un 464%, en 2022, Internet tenía alrededor de 3725 vídeos de pornografía deepfake. Sin embargo, en 2023, esta cifra aumentó a 21.019” [...]. Además, el informe logró demostrar que la aplicación de estas prácticas tiene un enfoque de género, “sus hallazgos revelan que el 99% de la pornografía deepfake presenta a mujeres como sujetos principales, mientras que solo el 1% del contenido presenta a hombres” (Home Security Heroes, 2023).

Otro estudio llevado a cabo por Internet Watch Foundation, revela de forma preocupante que los niños también son víctimas de la pornografía deepfake. El estudio realizado en Inglaterra presentó los siguientes resultados:

Durante un mes se monitorizó la actividad de la web y se encontraron 11.108 imágenes sospechosas de infringir algún delito relacionado con el abuso sexual infantil. De estas, cerca de 3.000 eran ilegales, según la ley del Reino Unido sobre el abuso infantil, y una de cada cinco, 564, se clasifican como categoría A, que es el tipo más grave (Bigas, 2023).

Estos videos se notan reales debido a que se usan algoritmos específicos para el deepfake, donde aprenden a

partir de la repetición, y con esta se explora y experimenta en varias posibilidades como: la creación de rostros nuevos no preexistentes, suplantación de identidad, manipulación de atributos (color de piel, ojos, boca, nariz, color de ojos, etc.), cambios de expresión, sincronización del movimiento de los labios al momento de articular palabras y/o reproducción de movimientos corporales (Morales, 2023).

La creciente accesibilidad a las plataformas y aplicaciones móviles como fakeapp permite a los usuarios sin experiencia técnica crear deepfakes, tan solo subiendo unas cuentas imágenes de su objetivo. La aplicación deepnude es otra de las herramientas que evidencia el desarrollo de las nuevas tecnologías del deepfake ya que, por ejemplo, permite “poner al desnudo a una mujer a partir de un retrato, esta analiza el rostro a partir de base de datos de 10,000 fotos de mujeres desnudas extraídas de las plataformas digitales, añade ese cuerpo desnudo al rostro del retrato” (Morales, 2023).

La popularización de estas aplicaciones, tanto gratuitas como de pago, es tanta que hoy en día existen aproximadamente 42 herramientas de este tipo:

[...]. En conjunto, estas herramientas generan la asombrosa cifra de 10 millones de búsquedas mensuales, lo que subraya el interés generalizado y la demanda de soluciones deepfake fáciles de usar. [...] el 40% son aplicaciones descargables y el 60% restante son plataformas basadas en web” (Home Security Heroes, 2023).

El auge del deepfake en redes sociales ha llevado a crear métodos para detectar videos e imágenes deepfake. Los investigadores crearon un método de detección del parpadeo en los videos deepfake, que consisten en evaluar los movimientos fisiológicos de los párpados en los videos falsos sintetizados, los cuales no logran representar de manera perfecta dicho movimiento, el método ha demostrado ser efectivo a la hora de detectar los videos generados por inteligencia artificial (IEEE Biometrics Compendium, 2018), al parecer la IA es el problema pero también es la solución; sin embargo su avance crece vertiginosamente, lo que implica que la creación de contenido multimedia deepfake, que tarde o temprano serán indistinguibles, superando la capacidad humana e inclusive la misma inteligencia artificial para detectar contenidos que han sido manipulados.

Pornografía deepfake no consensuada y su impacto en derecho a la imagen, intimidad personal y buen nombre en Colombia

En Colombia, la carta política de 1991 trajo consigo un hito importante: el establecimiento de los derechos fundamentales, entre ellos el derecho al buen nombre, intimidad personal y el derecho de imagen. El crecimiento exponencial de los entornos digitales y la inteligencia artificial ha sembrado ciertas dudas en cuanto a la protección de estos derechos dado al uso indebido de la IA, que en muchos casos ha sido utilizada para la creación de pornografía deepfake.

El artículo 15 de la Constitución Política de Colombia garantiza el derecho al buen nombre consagrado de la siguiente manera: “Todas las personas tienen derecho a su intimidad personal y familiar y a su buen nombre, y el Estado debe respetarlos y hacerlos respetar (...)” (Asamblea Nacional Constituyente, 1991, art. 15).

La teleología de este derecho busca salvaguardar la reputación que tienen los demás de alguien, ante comentarios o expresiones ofensivos e injuriosos, así como frente a la difusión de información falsa o errónea que pueda distorsionar esa percepción.

La Corte Constitucional de Colombia ha indicado “el derecho al buen nombre es esencialmente un derecho de valor porque se construye por el merecimiento de la aceptación social, esto es, gira alrededor de la conducta que observe la persona en su desempeño dentro de la sociedad” (Corte Constitucional, Sentencia SU 056 de 1995, párr. 2).

En esa línea, la jurisprudencia ha determinado el punto de afectación de tal derecho, según la Corte Constitucional de Colombia “es vulnerado cuando se difunde información falsa o inexacta, o que se tiene derecho a mantener en reserva, con la intención de causar desdoro contra el prestigio público de una persona” (Corte Constitucional de Colombia, , Sentencia T-949 de 2011, párr. 1).

Al igual que el buen nombre, la Constitución Política de Colombia establece en su artículo 15 la garantía fundamental a la intimidad personal, el cual lo define de la siguiente manera: “todas las personas tienen derecho a su intimidad personal y familiar y a su buen nombre, y el Estado debe respetarlos y hacerlos respetar” (Asamblea Nacional Constituyente, 1991).

El núcleo esencial del derecho a la intimidad personal es que cada persona tenga su propia esfera privada, protegida del control del Estado o de intrusiones arbitrarias de la sociedad, para que pueda desarrollar plenamente su vida personal, espiritual y cultural, dicho por la Corte Constitucional de Colombia: “está orientado a garantizar a las personas una esfera de privacidad en su vida personal y familiar, al margen de las intervenciones arbitrarias del Estado o de terceros. Comprende de manera particular la protección frente a la divulgación no autorizada de los asuntos que conciernen a ese ámbito de privacidad” (Corte Constitucional, Sentencia C 489 de 2002, párr. 1).

La misma corporación judicial mencionada ha indicado las formas de vulneración a este derecho, resumidas en las siguientes:

“(i) La intromisión en la intimidad de la persona que sucede con el simple hecho de ingresar en el campo que ella se ha reservado. Es un aspecto meramente material, físico, objetivo, independiente de que lo encontrado en dicho interior sea publicado o de los efectos que tal intrusión acarree. Cabe en este análisis la forma en que el agente violador se introduce en la intimidad del titular del derecho y no tanto el éxito obtenido en la operación o el producto de la misma, que se encuentran en el terreno de la segunda forma de vulneración antes señalada. (ii) En la divulgación de hechos privados, en la cual incurre quien presenta al público una información cierta, veraz, pero no susceptible de ser compartida, es decir, perteneciente al círculo íntimo de cada quien, siempre y cuando no se cuente con autorización para hacerlo bien de su titular, bien de autoridad competente. (iii) Finalmente, la presentación falsa de aparentes hechos íntimos no corresponde con la realidad” (Corte Constitucional de Colombia, Sentencia C-881 de 2014, párr. 5).

Por otro lado, la Carta política de Colombia no establece taxativamente el derecho a la imagen; sin embargo, la Corte Constitucional ha reiterado en diferentes providencias que este mismo se deriva de la personalidad jurídica, garantía contenida en el artículo 14 de la Constitución, el cual se define en los siguientes términos “toda persona tiene derecho al reconocimiento de su personalidad jurídica” (Asamblea Nacional Constituyente, 1991, art. 14).

Desde la perspectiva jurisprudencial es importante indicar que el derecho a la imagen ha sido interpretado como un derecho de carácter autónomo, el cual se deriva de la dignidad humana, íntimamente relacionado con el derecho al libre desarrollo de la personalidad y la personalidad jurídica, el cual según la Corte Constitucional “comprende la facultad de disponer de su apariencia y de su privacidad, autorizando o no la captación y difusión de ella” (Corte Constitucional de Colombia, Sentencia T-094 de 2000).

El órgano constitucional ha resaltado que el derecho de imagen comprende un núcleo duro referido como “aquellas expresiones que dan cuenta con claridad del aspecto físico, en general, y de los rasgos del rostro que permiten identificar a las personas. Por ende, es evidente el uso de la imagen a través de las fotografías, esculturas, videos y demás soportes que permitan identificar con precisión al individuo (...)” (Corte Constitucional de Colombia, Sentencia T-007 de 2020).

El derecho de imagen comprende que toda persona tiene la disposición de manejar su propia imagen, en ese

sentido el manejo de la imagen implica la facultad que tiene la persona para otorgar el consentimiento o autorización para que un tercero la utilice o explote, no obstante, ello no significa la renuncia definitiva de tal derecho. Por lo tanto, el uso no autorizado de la imagen por parte de terceros constituye sin más la vulneración a esta garantía constitucional.

El ente constitucional estableció determinantes para tener en cuenta si existe o no una vulneración al derecho de imagen:

1. La autorización para el uso de la propia imagen no puede implicar la renuncia definitiva del mismo.
2. La autorización comprende el consentimiento informado no solo acerca del uso de la propia imagen sino sobre las finalidades de ésta.
3. La autorización de uso de la propia imagen no puede constituir un límite absoluto al carácter necesariamente dinámico y cambiante de la autodeterminación de las personas o a su libre desarrollo de la personalidad.
4. La autorización de uso de la propia imagen, como expresión de un acuerdo de voluntades y de la libertad contractual en general, encuentra un límite constitucional en el respeto a los derechos fundamentales. (Corte Constitucional de Colombia, Sentencia T-634 de 2013).

La Corte Constitucional de Colombia ha señalado que el derecho a la imagen es independiente, aunque puede ser afectado junto con los derechos a la intimidad, a la honra y al buen nombre, estando estrechamente ligado a la dignidad y libertad del sujeto (Corte Constitucional de Colombia, Sentencia T-634 de 2013).

El derecho al buen nombre, la intimidad personal y a la propia imagen son derechos fundamentales que van ligados pero que también son autónomos, estos hacen parte de la personalidad de cada persona, por lo tanto, son inalienables, imprescriptibles e inembargables. Aunque cada uno de estos derechos tiene su propia protección legal, es importante destacar que no es necesario que todos coincidan en una misma situación para recibir protección por parte del Estado. Cada uno de estos derechos es crucial en la esfera general de una persona, simplemente por el hecho de existir.

Por otro lado, con el auge de la tecnología ha conllevado a la creación de herramientas de inteligencia artificial generativas, que, si bien en principio no suena amenazante puesto que su buen uso facilita las actividades cotidianas de las personas y empresas, recientemente su uso se ha desviado, afectando los derechos anteriormente mencionados.

Discusión y análisis de resultados

A continuación, se relacionan los resultados derivados de la investigación, en el cual se expondrán los análisis de las diferentes categorías conceptuales estudiadas en el presente artículo científico.

Análisis de casos relevantes sobre pornografía deepfake: retos y consecuencias

La inteligencia artificial está transformando la manera de ver el mundo, cambiando la forma de vivir, trabajar y percibir la realidad digital. Una herramienta que genera enormes beneficios, ofreciendo ventajas excepcionales en la sociedad y economía, pero su avance presenta una serie de desafíos éticos que genera incertidumbre sobre la confiabilidad de la inteligencia artificial, por ejemplo, los casos de pornografía deepfake que se presentan en redes sociales son aún más evidentes, exhibiendo a menores de edad y mujeres como principales víctimas. La situación se agrava dado que los responsables no hacen distinción alguna basada en el sexo, la edad, la cultura o la nacionalidad, gracias a la accesibilidad y utilización sencilla de la IA. Esto se evidencia en los casos siguientes:

El caso de Taylor Swift y la difusión de pornografía deepfake en la plataforma de la red social X

A finales de enero de 2024, la reconocida cantante de pop estadounidense Taylor Swift se vió envuelta en un escándalo desencadenado por la difusión de una imagen de pornografía deepfake en la plataforma de la red social X que luego se propagó en otras redes sociales. Esta publicación logró acumular más de 40 millones de visitas antes de que fuera retirada. Swift expresó su preocupación por la seguridad de las mujeres y anunció su intención de emprender acciones legales contra las plataformas que permitieron la difusión de dicho contenido. Su denuncia provocó un debate público sobre la necesidad de establecer regulaciones específicas en Estados Unidos y en las redes sociales para abordar este tipo de situaciones (Duque, 2024)

Es aquí donde se evidencia el alto riesgo que presentan las mujeres reconocidas por su popularidad, puesto que suelen ser vulnerables a este tipo de prácticas debido a que son más visibles en las plataformas digitales, por lo tanto, sus imágenes están expuestas todo el tiempo siendo susceptibles a ser utilizadas en herramientas de deepfake que puedan realizar todo tipo de contenido, en este caso de tipo sexual y erótico. Este incidente expone el peligro que corre la sociedad frente a las consecuencias generadas por la creación de contenido de pornografía deepfake en cuanto a la privacidad y seguridad de cualquier persona. Además, destaca la necesidad de tomar medidas legales y abordar por un cambio normativo, desde un enfoque proactivo tanto a nivel individual como institucional para abordar los desafíos emergentes en el ámbito de la seguridad digital y la integridad de la información en línea.

Desde otra perspectiva, el caso de esta figura pública internacional, presenta un impacto directo en el derecho de imagen, ya que cómo persona tiene derecho a decidir sobre la utilización de ésta. En Colombia, el derecho de imagen está amparado constitucionalmente, la Corte Constitucional de Colombia ha señalado que, en el contexto colombiano, el alcance de este derecho comprende que es necesario el consentimiento para su utilización, así mismo, reitera que se garantiza protección constitucional, por lo tanto la imagen “no debe ser objeto de libre e injustificada disposición y manipulación de terceros” (Corte Constitucional de Colombia, Sentencia T-634 de 2013) elementos que no concurren en esta situación, ya que no existe un consentimiento para la utilización del contenido expuesto y aun así fue dispuesto y manipulado por un tercero.

El derecho a la intimidad de Swift se ve considerablemente afectado, ya que, como titular del mismo, tiene la capacidad de permitir o promover la revelación de aspectos privados de su vida. Sin embargo, en esta controversia, no hay un consentimiento explícito por parte de ella, lo que podría generar una percepción negativa por parte de sus seguidores, la industria del entretenimiento y el público en general, por lo tanto, podría influir tanto en su vida personal como profesional ya que su ámbito cómo ser humano es de carácter reservado. La difusión de contenido sexual falso menoscaba el control que Swift ejerce sobre la divulgación de información íntima, incluidos temas de naturaleza sexual, para preservar su privacidad personal. Finalmente, se encuentra inmiscuida la vulneración directa del buen nombre, el cual se ha entendido a través de la jurisprudencia y doctrina cómo la *reputación personal* y que se configuró su afectación a través de la propagación que por medio de informaciones falsas o tendenciosas generó una inadecuada valoración en el caso de Taylor con los demás miembros en sociedad, por la difusión de información errónea sobre opiniones tendenciosas que impactan negativamente y ocasionando daño moral a ella como sujeto de derechos.

Estudiantes de una escuela en Quito emplearon imágenes de sus compañeras para generar material de naturaleza sexual utilizando inteligencia artificial

Dos estudiantes de primero de bachillerato de un colegio en Ecuador utilizaron las fotografías de 24 compañeras de su curso y de otros cursos, para recrear imágenes y videos de tipo sexual a través de herramientas de inteligencia artificial, las cuales, se difundieron entre 700 videos e imágenes. Las víctimas se enteraron de este

suceso a través de la difusión por vía de mensajes textos que circulaban en el centro educativo, las menores de edad afectadas por el suceso decidieron desvincularse de la institución debido a este tipo de violencia digital. Los adolescentes autores de la agresión también desertaron voluntariamente del colegio luego de ser expuestos por las víctimas. La Fiscalía General del Estado de Ecuador comunicó que abriría una investigación por el presunto delito de pornografía infantil por la difusión del material audiovisual alterado mediante inteligencia artificial (Redacción Vistazo, 2023).

Es claro que en este tipo de actividades no sólo involucra adultos, sino también a menores de edad, en el presente caso, por ejemplo, se evidencian dos extremos: el primero implica a los responsables del contenido sexual, quienes son niños; mientras que el segundo abarca a las víctimas, también menores de edad. Este hecho es realmente preocupante, ya que demuestra el fácil acceso a estas tecnologías y la facilidad con la que pueden ser empleadas no solo por adultos, sino también por menores de edad que consideran este tipo de producciones como algo divertido.

Cada individuo tiene derecho a su intimidad personal y así mismo controlar la divulgación de su propia imagen, con el fin de no ser objeto de intromisión en su vida privada, cosa que se vió ligada en esta problemática, debido a que se involucró estudiantes de Quito en un contenido sexual falso pero que resulta aparentemente real, sin embargo lo impactante es que estas son niñas, por lo tanto se generó mayor impacto social puesto que estas tienen una protección especial sólo por el hecho de ser menores de edad, además desde el otro extremo, en este caso, otra menor de edad (quién realizó el contenido) desconoce sobre las implicaciones de sus actos debido a la poca madurez de edad, generando actos que provocaron una repercusión psicológica y legal sus compañeras, es por ello que el uso no consensuado de contenido para fines sexuales, generan una violación clara del ámbito personal y sexual.

El streamer Atrioc y las fotos desnudas de sus compañeras generadas por IA; el escándalo

El 30 de enero de 2023 explotó un escándalo entre reconocidos streamers estadounidenses, en una transmisión en vivo, el streamer Brandon Ewing, también conocido como “Atrioc” en twitch, mediante propio descuido fue sorprendido mirando páginas web dedicadas a la creación de imágenes falsas de contenido sexual de mujeres, algunas colegas, amigas y conocidas de él, como QTCindrella, Sweet Anita y Maya Higa. Todas manifestaron su rechazo en contra de las imágenes sexualizadas propagadas en las redes sociales, QTCindrella escribió a través de la red social X: “Ser vista “desnuda” en contra de tu voluntad no debería ser parte de este trabajo”, por su parte Maya Higa también pronunció: “He sido usada por miles de hombres para su gratificación sin mi consentimiento” (Carranza, M. 2023). Finalmente, debe indicarse que las imágenes fueron bajadas de las redes al poco tiempo y el streamer Atrioc pidió disculpas por lo ocurrido en un video, sin embargo, el daño ya estaba hecho.

En este caso, las streamers QTCindrella, Sweet Anita y Maya Higa presentan una clara vulneración al derecho de imagen, de acuerdo con la sentencia T-634 de 2013, la cual establece los lineamientos para determinar tal afectación, se logra observar que (i) no existe previa autorización para el uso de la imagen de cada una de ellas; y (ii) evidentemente no existe consentimiento informado sobre las finalidades de utilización del derecho de imagen de las víctimas, lo que constituye la violación al derecho de imagen puesto a que ellas no tenían el consentimiento ni conocimiento de la utilización a su propia imagen. Además, en las imágenes sexuales se logra diferenciar que efectivamente los cuerpos desnudos presentan los rasgos faciales de las víctimas, que, aunque son falsos parecen reales.

La intimidad personal es otro derecho afectado, dado que implica el quebrantamiento de la esfera privada de las víctimas. Publicar fotos de las streamers desnudas implica (i) la intromisión a un campo reservado por las streamers: su intimidad sexual y (ii) la difusión del material sexual sintetizado los cuales no corresponden con la realidad, lo cual provoca que las streamers pierdan el carácter reservado, puesto que la publicación de

información sexual íntima destruye su círculo privado; dicho por las víctimas “verte desnuda en contra de tu voluntad y difundida en internet es como sentirse violada”. A pesar de que las imágenes son manipuladas constituye un peligro real para las afectadas, interferir arbitrariamente en la vida íntima publicando imágenes sexuales de una persona, implica sin más, una clara vulneración al derecho de intimidad personal.

El derecho al buen nombre también se ve comprometido, es evidente que las streamers sufrieron una afectación en su buen nombre considerando que al difundir información falsa podrían verse perjudicadas, ya que tiene por objetivo causar deshonra contra el prestigio de las streamers, lo que podría generar juicios y estigmas sociales injustos hacia ellas, lo que provoca opiniones dañinas sobre su reputación.

Pornografía deepfake creada por la inteligencia artificial generativa involucra a estudiante llamada Ellis en un colegio de Estados Unidos

Ellis, una adolescente de 14 años de Texas, se despierta una mañana de octubre y se da cuenta que tiene varias llamadas perdidas y mensajes de textos los cuales contenían fotografías de contenido explícito que era divulgado a través de redes sociales con su cara real y el cuerpo expuesto a la sociedad, las cuales fueron enviadas a sus compañeros de clases, quienes no se fijaron que eran creadas con IA. Sin embargo, no solo ella fue usada para este tipo de prácticas, sino que también otras amigas de clase, donde sus fotos fueron extraídas de Instagram y compartidas por Snapchat. Quien realizó este tipo de contenido fue una compañera del mismo nivel de clase, quien posteriormente es expulsada del colegio (Montevideo Portal, 2023).

Las autoridades han intentado regular este tipo de prácticas con el fin de prohibir su uso, pero hasta el momento no han logrado implementar un plan efectivo que pueda limitar la utilización de estas tecnologías. Dichas herramientas permiten una significativa vulneración de derechos fundamentales. Entre ellos, se destaca el derecho al buen nombre, ya que la difusión de este contenido puede afectar gravemente la percepción y reputación de las víctimas. Como consecuencia, la persona afectada puede ver restringido su estilo de vida debido al temor al juicio y señalamiento por parte de la sociedad. Analizando el caso desde la perspectiva jurídica colombiana, la Corte Constitucional de Colombia en su sentencia T-610 de 2019 estableció que se afecta este derecho cuándo “se propagan entre el público -bien sea de forma directa o personal, o a través de los medios de comunicación de masas-informaciones falsas o erróneas que distorsionan el concepto público que se tiene del individuo” (Corte Constitucional de Colombia, Sentencia T 610 de 2019).

Además, estas prácticas también violan el derecho a la intimidad personal, ya que implican una exposición no autorizada del entorno privado que contribuye a conformar la identidad y la personalidad del individuo, puesto que sus fotos son extraídas con el fin de realizar materia deepfake, así establece la Ley 1098 del años 2006, esto es el código de infancia y adolescencia de Colombia, establece en su artículo 33 “los niños, las niñas y los adolescentes tienen derecho a la intimidad personal, mediante la protección contra toda injerencia arbitraria o ilegal en su vida privada, la de su familia, domicilio y correspondencia” (Congreso de la República de Colombia, 2006). La intromisión de la privacidad provocó consecuencias profundas y perjudiciales para las víctimas, afectando su bienestar emocional y su capacidad para desenvolverse libremente en su entorno (France 24, 2024).

Difunden un video pornográfico deepfake de la primera ministra de Italia

Giorgia Meloni, primera ministra de Italia, expresa que fue víctima de un vídeo pornográfico deepfake dónde su cara es puesta en el cuerpo de otra persona. Las investigaciones evidencian que se tratan de dos hombres anónimos, uno de 40 años y otro de 73, este contenido multimedia fue publicado en línea y fueron observados por millones de personas en un sitio web de pornografía realizado en Estados Unidos, esto confirmado por la agencia de noticias italianas ANSA. Mediante este hecho, el abogado de Giorgia manifiesta que los daños y perjuicios morales están catalogados en un monto de 100.000 euros o unos 108.500 dólares y que, en caso de tener éxito, es decir se demuestre a favor los hechos, este monto se destinará a apoyar a las mujeres víctimas

producto del abuso sexual. Melodi a través de la denuncia quiere enviar un mensaje de apoyo a todas las víctimas deepfake, debido a este tipo de herramienta que genera un abuso de poder, así mismo, incentiva a no tener miedo para presentar cargos (Forbes, 2024).

Un aspecto crucial del presente caso es cómo esta mujer anima a las víctimas deepfake a seguir los procedimientos adecuados para iniciar las investigaciones correspondientes ante eventos como este. Además, el dinero que obtenga en caso de ganar el caso será utilizado para ayudar a las mujeres víctimas de abuso digital. Melodi experimentó una afectación a su derecho de imagen, sin lugar a dudas la ministra de Italia en ningún momento consintió el uso de su imagen para la creación de pornografía adulterada, en consecuencia, considerando que la imagen fue publicada sin la autorización de su titular configura el quebrantamiento del derecho a la propia imagen.

Además, no sólo se encuentra inmiscuido el anterior derecho, sino que también se presenta una vulneración en el derecho a la intimidad personal, ya que se trata de información privada o reservada que involucra aspectos cómo la información íntima o sexual de Giorgia y no existe un consentimiento de ella para poder divulgar la información (que resulta siendo fake). Por lo tanto, este acto es una intromisión a la esfera privada de la víctima que integra la autonomía y dignidad humana, por ello este derecho permite controlar el acceso a la información de la vida privada.

El derecho al buen nombre se ve comprometido como consecuencia de la difusión y manipulación de este contenido, ya que ha generado un gran revuelo a nivel internacional. La perspectiva de que este contenido sea real puede perjudicar la reputación y la percepción social de Melodi, especialmente en aquellos que no están familiarizados con este tipo de prácticas. Esto puede afectar la forma en cómo es vista por otros, así como su reputación y posición en la sociedad.

Alexandria Ocasio-Cortéz, congresista y víctima de contenido pornográfico con IA

Alexandria Ocasio-Cortez, la congresista progresista estadounidense, compartió su experiencia al descubrir un deepfake explícito y sexual de sí misma durante un viaje en coche en febrero. En una entrevista con Rolling Stone, describió cómo fue impactada por esta manipulación generada por inteligencia artificial, destacando que una vez que eres víctima de tal cosa y te ves expuesta a millones de extraños, es algo que nunca puedes olvidar, incluso si la imagen en sí es falsa. Según Ocasio-Cortez, ver ese contenido es algo irreversible: “Una vez que lo has visto, lo has visto. Comparó la experiencia con la intención de una violación física y agresión sexual, enfatizando que se trata de poder, dominación y humillación. Para ella, los deepfakes representan una forma digital de humillación violenta hacia otras personas”. En realidad, Ocasio-Cortez, de 34 años, respalda una ley en el Congreso que busca permitir a las víctimas de pornografía deepfake demandar a los responsables de crear, distribuir y consumir representaciones digitales sexualmente explícitas no consentidas. (Magas, 2024). El uso del deepfake en el caso descrito ocasiona un detrimento en el derecho de imagen, intimidad personal y el buen nombre. Pero, ¿De qué manera se encuentran involucrados los anteriores derechos?

En primer lugar, se debe precisar que el derecho de imagen implica el control y la autorización sobre el uso de la imagen de una persona. En este caso, se observa que la imagen de Alexandria fue utilizada con el fin de crear, producir y difundir imágenes de contenido sexual sin su consentimiento distorsionando su identidad y generando que sea involucrada con actividades eróticas que no tienen relación con su vida real ni con su consentimiento.

Por otro lado, la intimidad personal protege la esfera privada de la persona en su dimensión de sujeto individual y en su entorno familiar, es por ello que Alexandria al ser expuesta de manera íntima y comprometedora en un contexto sexual falso, experimentó una intrusión en su esfera privada que puede causar daño emocional y psicológico significativo, además de múltiples señalamientos y fuertes críticas.

Por último, es necesario destacar que “el buen nombre es un derecho íntimamente relacionado con la dignidad humana de toda persona, valor porque se construye por el merecimiento de la aceptación social, esto significa que gira alrededor de la conducta que observe la persona en su desempeño dentro de la sociedad” (Montenegro, 2006). Es por ello que esto compromete su buen nombre al asociarla falsamente con comportamientos sexualmente explícitos, lo cual puede afectar su credibilidad y la percepción pública de ella, al ser esta una mujer política de alto rango.

En Corea del Sur un hombre fue detenido por usar inteligencia artificial para la creación de imágenes sexuales de menores de edad

En septiembre de 2023, un hombre de 40 años en Corea del Sur, cuya identidad no se conoce, fue condenado por el Tribunal de Distrito de Busan y la Fiscalía del distrito a dos años y medio de prisión. Fue declarado culpable de producir 360 imágenes generadas por inteligencia artificial, las cuales fueron confiscadas por la policía y no se distribuyeron. La sentencia resaltó que estas imágenes, creadas con tecnología avanzada, representaban de manera tan realista a niños y menores que se consideraron como contenido sexualmente abusivo (CNN Mundo, 2023).

Los niños también son foco de atención para los ciberdelincuentes, lo que agrava aún más la situación. El obtener imágenes de los rostros de las personas en internet es muy sencillo, basta con entrar a la red social de la persona o algún familiar de ella para extraer el retrato de la persona objetivo, sin importar si es un niño, es así como sin el consentimiento de su representante legal toma sin autorización la imagen y la manipula a través de AIG, materializando la violación al derecho fundamental de la propia imagen.

A demás, al hacerlo se entromete en la burbuja privada del menor, puesto que desnudarlo, poniéndolo en poses eróticas o cometiendo actos sexuales ya es sobrepasar la línea roja de su intimidad sexual, ya que se habla de un menor y este no puede (aunque así lo quisiera) conceder autorización alguna para la realización de este tipo de atrocidades, por lo que constituye sin más, la vulneración al derecho de su intimidad personal.

Respecto al buen nombre, de acuerdo a la sentencia SU 420 de 2019 “Su desconocimiento se presenta cuando se difunde información falsa o errónea, o se afecta la reputación o el concepto de una persona como consecuencia de expresiones ofensivas o injuriosas, lo que conlleva adicionalmente la transgresión de su dignidad humana” (Corte Constitucional de Colombia, Sentencia SU-420 de 2019), por lo tanto, al no publicar las imágenes no se altera o perjudica la percepción pública de los menores de edad, en consecuencia no vulnera el derecho fundamental al buen nombre, pues no genera un detrimento de la imagen personal de las víctimas.

Estos desafíos exigen una normatividad ejemplar para evitar el impacto en la reputación de las personas víctimas de este contenido, simplemente eliminar fotos o vídeos de las plataformas no es suficiente para abordar la explotación digital. Aunque elimina una parte del problema, los internautas pueden descargar el contenido, grabarlo, capturarlo en pantalla y compartirlo con otros, lo que significa que el contenido puede seguir circulando y permanecer en la memoria de los usuarios.

Uso de la inteligencia artificial en la creación de contenido sobre abuso sexual infantil y su tráfico ilegal

Unos cuantos pedófilos usan la inteligencia artificial generativa, para crear y vender material que muestra situaciones reales de abuso sexual infantil. Los creadores de imágenes de abuso sexual infantil están utilizando un software de inteligencia artificial llamado Stable Diffusion, originalmente diseñado para generar imágenes destinadas al arte y el diseño gráfico. Este software permite a los usuarios describir la imagen que buscan, a través de palabras clave. Esta es usada para crear imágenes de abuso sexual, donde se incluía la violación de bebés y menores.

Las imágenes de abuso son compartidas a través de un proceso de tres etapas:

1. Generan las imágenes con software de IA.
2. Promueven las fotos en plataformas como Pixiv, un sitio japonés de intercambio de imágenes.
3. Estas cuentas tienen vínculos que dirigen a los clientes a imágenes más explícitas, para las que pueden pagar para mirar en cuentas en sitios como Patreon (BBC New Mundo, 2023).

El programa de IA Stable Diffusion, diseñado para generar arte y gráficos, ha sido explotado por delincuentes para crear imágenes de abuso sexual infantil. Una versión de código abierto permite a los usuarios eliminar filtros, facilitando la creación de contenido ilegal. Stability AI condena su uso indebido y apoya a las autoridades para frenar esta práctica. Las autoridades advierten sobre el impacto de estas imágenes "sintéticas" en la identificación de víctimas reales y los riesgos futuros de la tecnología si no se regula adecuadamente (BBC New Mundo, 2023).

Es aquí donde el derecho de imagen se encuentra vulnerado porque se utilizan representaciones de menores, sin consentimiento y autorización previa, lo que afecta su identidad visual. El derecho a la intimidad personal se vulnera debido a que se crean o distribuyen imágenes que involucran situaciones sexuales, afectando la privacidad de los menores. El buen nombre se ve perjudicado porque estas imágenes asocian a las víctimas (reales o ficticias) con actividades ilegales y moralmente reprochables, afectando su reputación y dignidad en la sociedad.

Revisión de avance legislativo internacional sobre regulación de pornografía deepfake

A nivel internacional se ha reconocido el peligro potencial que genera la AIG, en respuesta a esa amenaza múltiples países han tratado de reforzar sus legislaciones para proteger a su nación de posibles violaciones masivas de derechos fundamentales, manipulación de realidades y destrucción de reputaciones. Por esta razón, países como España, Estados Unidos, China, entre otros países han implementado acciones significativas para combatir los riesgos de la inteligencia artificial generativa.

Análisis de propuesta normativa en España

Para abordar este desafío, España ha tomado una medida significativa al presentar el 13 de octubre de 2023 una Proposición de Ley Orgánica 122/000011 destinada a regular las simulaciones de imágenes y voces creadas mediante inteligencia artificial, su fin es garantizar los derechos de los menores en el campo digital, centrándose en la privacidad, la protección de datos personales y el acceso a contenidos adecuados a su edad. Dentro de esta propuesta se tiene como objetivo reformar el código penal y tipificar como delito de injuria de carácter grave la creación y difusión de imágenes corporales o audio generado por inteligencia artificial, sin el debido consentimiento de la persona retratada, con la finalidad de perjudicar o afectar negativamente el honor, la reputación, la dignidad o la propia estimación de una persona esto con el fin de reforzar la integridad de los menores de edad, dentro de esta reforma se incluye órdenes de alejamiento virtual para aquellos que cometan delitos contra menores online, incluye el aumento de la edad para que los niños puedan autorizar el tratamiento de información personal mediante redes sociales, pasando de 14 a 16 años de edad respectivamente. También se contempla la participación de pediatras en la identificación de adicciones y situaciones de riesgo en internet que afecten a los menores. (Congreso de los diputados, 2023, Ley 122/000011) La difusión de tales simulaciones en redes sociales se considerarán una injuria de carácter público, catalogadas como graves, de acuerdo con el artículo 209 de la ley 10/1995 (Código penal) de España la conducta se castigará con hasta catorce (14) meses de prisión (BOE, 1995).

Análisis de propuesta normativa en México

En el senado de la Republica de México analizan la propuesta de ampliar la Ley Olimpia la cual reconoce y castiga

el acoso digital reflejado en la viralización de contenidos sexuales sin consentimiento de la víctima; esto con el fin de reconocer como una forma de acoso digital la creación y difusión de imágenes sexuales creadas mediante inteligencia artificial. El partido político Morena en cabeza de Alejandro Armenta es quien impulsa esta idea, la cual propone imponer penas de hasta seis años (6) de prisión a aquellos que divulguen, compartan, distribuyan o publiquen imágenes, videos o grabaciones de contenido sexual íntimo de personas adultas creadas utilizando inteligencia artificial.

Armenta expresó que “esta iniciativa busca tipificar la violación a la intimidad sexual a través del contenido generado por inteligencia artificial. Con inteligencia artificial se produce este tipo de intimidación sexual y de violación que, desde luego, genera afectación a las mujeres, a las jóvenes principalmente” (El economista, 2024). Este proyecto de ampliación propone modificar el artículo 20 Quarter de la Ley General de Acceso de las Mujeres a una Vida Libre de Violencia y el artículo 199 Octies del Código Penal Federal. De acuerdo con estas modificaciones, se consideraría como delito la violación a la intimidad sexual, la divulgación, compartición, distribución o publicación de imágenes, videos o audios de contenido íntimo sexual de personas adultas sin su consentimiento, aprobación o autorización, utilizando Tecnologías de la Información y Comunicación, incluyendo Inteligencia Artificial. Según el Código Penal Federal, esta infracción con llevaría a una pena de tres (3) a seis (6) años de prisión y multas que oscilarán entre 54,000 y 109,000 pesos mexicanos (El economista, 2024).

Análisis de propuesta normativa en Estados Unidos

En la Secretaría del Congreso de Estados Unidos circula un proyecto de ley que pretende combatir ferozmente la pornografía deepfake no consensuada desde otros ángulos, un grupo bipartidista de senadores encabezados por el senador Ted Cruz presentaron el proyecto de ley: TAKE IT DOWN.

Este proyecto de ley propone la criminalización a nivel federal de la pornografía deepfake no consensuada, así como la publicación de imágenes íntimas sin el consentimiento de la persona afectada. La iniciativa "TAKE IT DOWN" no solo busca responsabilizar a quienes crean y/o publican material sexual no autorizado, sino también a los sitios web que permiten su difusión. Esta ley garantiza a las víctimas que las redes sociales y los portales web prioricen sus denuncias, exigiendo a las empresas eliminar el contenido denunciado y sus duplicados dentro de las 48 horas siguientes a la recepción de la notificación de la denuncia. En caso de incumplimiento, la ley facultará a la Comisión Federal de Comercio para tomar medidas enérgicas contra las redes sociales o páginas web que no acaten estas disposiciones. Las sanciones establecidas en el proyecto de ley disponen una diferenciación entre delitos cometidos contra adultos y menores de edad, este último siendo más grave. Cuando la víctima es un adulto dispone como sanción multas económicas ajustadas a la gravedad del delito y/o una pena máxima de hasta 2 años de prisión; del mismo modo, cuando la víctima es menor de edad la condicionalidad de las multas económicas son las mismas, es decir, relativo a la gravedad del delito y/o una pena máxima de hasta 3 años de prisión. (Congreso de los Estados Unidos, 2023).

Caso del Estado de California, Estados Unidos

El Estado de California reconociendo los riesgos del deepfake dio un paso el 03 de octubre de 2019 al aprobar el proyecto de ley N.º 602, que agregó al artículo 1708.86 del código civil de California algunas disposiciones para determinar la responsabilidad de una persona que cree o divulgue intencionalmente material sexual de otra, a sabiendas que la persona retratada no dió su consentimiento para la creación o divulgación del material. No obstante, se menciona que al existir consentimiento no existiría ningún tipo de responsabilidad, pero el consentimiento debe ser expreso en escrito, firmado consciente y voluntariamente, especificando la descripción del material sexual en general y de la herramienta audiovisual a utilizar, además el consentimiento se puede retirar mediante notificación escrita dentro de los 3 días hábiles siguientes a partir de la fecha en la que se dió el consentimiento. Al hallarse responsabilidad en el actor, el demandante vencedor quien sufre daños por la divulgación o creación del material sexual podrá solicitar una indemnización que varía de acuerdo a las circunstancias, pero que, si la conducta se cometió con intenciones de causar daño, la reparación podría llegar hasta los 150.000 dólares. Por último, la ley establece que la acción judicial debe iniciarse a más tardar en 3 años,

contados desde la fecha en la que se descubrió o debería descubrirse mediante diligencia razonable, la creación, desarrollo o divulgación del material no consensuado. (Assembly Bill No. 602, 2019).

Caso del Estado de Texas, Estados Unidos

El Estado federal de Texas ha optado por criminalizar la pornografía deepfake no consensuada, convirtiéndola como delito dentro de su legislación, así lo establece el artículo 21.165 del Código Penal de Texas, titulado: "Producción o distribución ilegal de determinados vídeos de contenido sexual explícito", artículo el cual fue agregado por la Ley N°1361, la cual entró en vigor el 1 de septiembre de 2023.

De acuerdo con sus disposiciones, la persona comete un delito siempre que, (i) sin el consentimiento efectivo de la persona que aparece retratada (ii) produce o distribuye mediante medios electrónicos (iii) un vídeo deepfake que parece mostrar a la persona con sus partes íntimas expuestas o participando en una, conducta sexual.

La ley categoriza a este delito como un delito menor de clase A, por lo que en este nivel la condena puede resultar en una multa de hasta 4.000 dólares y/o hasta un año en la cárcel del condado.

Esta ley forma parte del bloque de defensa legislativa del Estado de Texas, para contraatacar la creciente preocupación sobre el uso de inteligencia artificial en la creación de contenido sexualmente explícito sin consentimiento.

Caso del Estado de Nueva York

El Estado de New York ha aprobado la Ley del Senado S5959D que aborda el uso de la tecnología para generar reproducciones digitales de contenido sexual explícito y establece regulaciones en relación con su utilización. El legislador estableció la responsabilidad contra la persona que produce, publique, difunda o divulgue el material sexual explícito de la persona retratada que no dió el debido consentimiento, esto es firma consciente y voluntaria manifestado en escrito, que incluya la descripción general del desarrollo del material sexual a la cual da autorización y la obra audiovisual que se integrará (imágenes, videos, audios), con la salvedad de que la persona retratada puede retirar su consentimiento mediante aviso escrito dentro de los 3 días hábiles contados desde la fecha en la que se dió la autorización. De no existir el consentimiento la persona retratada puede emprender acciones legales con la finalidad de proteger su reputación, a través de la compensación económica impuesta por un juez que varía de acuerdo con el caso. La reclamación podrá realizarse en un máximo de tres años desde la publicación del material sexual, pero que puede reducirse el término a un año desde que la víctima tuvo el conocimiento o por razones lógicas debió haber descubierto la difusión del material, por último, la ley hace énfasis que no se podrá perseguir alguna acción penal en contra del responsable (The New York State Senate, 2020).

Caso del Estado de Georgia

Georgia, otro Estado de EE.UU que decidió tomar medidas legislativas frente al fenómeno de los deepfakes pornográficos. El legislador modificó la Parte 3 del Artículo 3, Capítulo 11 del Título 16 del Código Oficial. Esta sección castiga a quienes transmitan o publiquen electrónicamente material sexual explícito, ya sea imágenes o videos, tanto reales como falsificados. Las sanciones bajo esta normativa se clasifican como delitos menores graves, con penas que oscilan entre uno (1) y cinco (5) años de cárcel. Además, se pueden imponer multas de hasta \$100,000, y en ciertos casos, ambas sanciones pueden aplicarse simultáneamente (Asamblea General de Georgia, 2019).

Caso del Estado de Virginia

De igual forma, en el Estado de Virginia incluyó en su legislación penal, específicamente en los artículos 18.2-386.1 y 18.2-386.2, sanciones contra quienes, produzcan, comercialicen o distribuyan, de manera intencional y consciente, material sexual audiovisual o imagen fija de otra persona sin su consentimiento, las disposiciones

incluyen representaciones generadas por inteligencia artificial. Estas conductas se categorizan como un delito de clase 1, con una sanción de hasta doce (12) meses de prisión y multas de hasta 2.500 dólares. Sin embargo, cuando el material involucra a menores de edad (menores de 18 años), el delito se clasifica como más grave, siendo un delito de Clase 6 en Virginia, lo que puede resultar en penas de uno (1) a cinco (5) años de prisión. (Asamblea General de Virginia, 2019).

Análisis de propuesta normativa en Reino Unido

El Ministerio de Justicia del Reino Unido ha introducido propuestas legislativas que se centran en los deepfakes de contenido sexual explícito creados sin el consentimiento de la persona afectada. Esta nueva legislación cambia su enfoque hacia la producción de dicho material, convirtiendo en una actividad ilegal crear contenido sin el consentimiento del sujeto, sin importar la intención de distribuirlo. Este proyecto de ley incluye no solo imágenes pornográficas, sino cualquier contenido que presente desnudez, independientemente de las acciones del individuo representado.

Laura Farris, ministra de Víctimas y Protección del Reino Unido, junto con el Ministerio de Justicia, ha impulsado el proyecto de ley de justicia penal, que establece que la creación de deepfakes sexualmente explícitos constituye un delito, incluso si no hay intención de compartir dicho contenido. La medida abarca imágenes falsas de contenido pornográfico, así como desnudos parciales o totales, independientemente de si la persona representada está actuando de manera erótica o no. Los infractores enfrentarán sanciones que incluyen multas ilimitadas y la posibilidad de tener antecedentes penales. Aquellos que decidan distribuir este tipo de contenido podrían enfrentar hasta dos (2) años de prisión como consecuencia de sus acciones.

Esta iniciativa se enfoca en abordar la explotación sexual a través de deepfakes, con un énfasis especial en la protección de las mujeres como las principales víctimas. El objetivo es salvaguardar derechos fundamentales, como la intimidad personal, mediante la imposición de sanciones tanto para quienes creen como para quienes difundan contenido sexual generado por esta tecnología. Además de penalizar estas acciones, la legislación busca crear conciencia sobre los riesgos asociados con la manipulación digital de imágenes para evitar la explotación y el daño a la dignidad de las personas afectadas (20 minutos, 2024).

Análisis de propuesta normativa en Australia

El gobierno australiano presentó un proyecto de ley para penalizar la creación de pornografía deepfake no consensuada, dicha legislación propone castigos penales a todo aquel que distribuya contenido pornográfico deepfake sin consentimiento del retratado. La legislación propone una pena máxima de seis años de prisión para quienes compartan este tipo de material sin consentimiento, y penas de hasta siete años para reincidentes y creadores de contenido, cabe resaltar que el proyecto de ley se enfoca en material sexual protagonizado por adultos, mientras que el material de abuso infantil será tratado bajo cargos diferentes. (Parlamento de la Commonwealth de Australia, 2024).

Análisis de propuesta normativa en Corea del Sur

Corea del Sur dentro de su legislación regula las conductas relacionadas con la pornografía deepfake. En la Ley de Prevención de la Violencia Sexual (ley 19743 de 2023) en su artículo 14-2, contempla sanciones para quienes distribuyan, editen o sinteticen material con el propósito de causar deseo sexual o vergüenza, en contra de la voluntad del sujeto retratado, recibirá un castigo de hasta 5 años de prisión o una multa de hasta 50 millones de wones, sin embargo si utiliza redes de información o comunicaciones para compartir la información buscando obtener beneficio económico, la pena se incrementa a 7 años de prisión. En caso de reincidencia, la sanción podría aumentar hasta la mitad de la pena inicial. (Ministerio de Justicia de Corea del Sur, 2023).

Análisis de propuesta normativa en Argentina

El Diputado Nacional Juan Brügge ha presentado en el Congreso de la Nación de Argentina el Proyecto de Ley

3955-D-2024, que tiene como objetivo la inclusión de nuevas disposiciones en el Código Penal para sancionar la creación y difusión de contenido deepfake sin consentimiento. Este proyecto de ley surge en respuesta a la creciente preocupación por el uso indebido de tecnologías de Inteligencia Artificial Generativa, las cuales permiten la manipulación de videos, audios e imágenes para hacerlos parecer auténticos, y que pueden tener consecuencias graves para la integridad personal y profesional de las personas afectadas.

El proyecto de ley introduce una serie de modificaciones a los artículos 139 ter, 129 bis y 157 ter del Código Penal Argentino, el cual se pretende contemplar penas de prisión de entre 4 y 8 años para aquellos que, sin el consentimiento de las personas involucradas, generen o alteren contenido digital mediante el uso de inteligencia artificial para hacerlo pasar como real. Este tipo de delito será punible en tres escenarios específicos:

1. En perjuicio de actividades personales y profesionales;
2. Cuando genere un daño o perjuicio actual o potencial a la imagen o reputación;
3. En el marco de un proceso electoral, la pena se incrementa si las víctimas son menores, personas con discapacidad o en estado de vulnerabilidad;
4. Cuando se trate de contenido sexual. (Código Penal Argentino, 1921).

Además, el proyecto establece que las penas serán más severas si las víctimas son menores de edad, personas con discapacidad o aquellas en estado de vulnerabilidad, así como en casos en los que el contenido manipulado tenga un carácter sexual. En estas circunstancias, las penas podrían elevarse de 6 a 10 años de prisión. Este tipo de delito, cuando involucra a menores, podría también ser procesado junto con el delito de tenencia de pornografía infantil, dada la gravedad del impacto y las posibles consecuencias para las víctimas jóvenes.

La propuesta de ley de Brügge refleja una creciente necesidad de abordar los peligros que plantea el uso de tecnologías avanzadas como la inteligencia artificial en la creación de contenido digital, reconociendo que las implicaciones de estos actos pueden ser profundamente perjudiciales tanto a nivel personal como social (Perfil, 2024).

Análisis de propuesta normativa en la República Popular de China

Aunque China no cuenta con una regulación especial que trate sobre el fenómeno de los deepfakes pornográficos, si tiene algunas regulaciones que de cierta manera limitan este tipo de conductas, destacando el Reglamento para la Gestión Profunda de los Servicios de Información en Internet N°12. De acuerdo a la normativa, se requiere autorización expresa antes de compartir el contenido generado por inteligencia artificial. Además, establece la obligatoriedad de incluir marcas de agua que los identifiquen dicho contenido como producto creado por inteligencia artificial con la finalidad de prevenir confusiones públicas o identificación errónea. Asimismo, los proveedores de servicios de generación de contenido deben evitar procesar información personal y seguir normas estrictas como la autenticación de usuarios, evaluación de algoritmos y transparencia en la creación de videos.

Finalmente, los transgresores pueden acarrear sanciones penales privativas de la libertad, cuya pena varía de acuerdo con la gravedad de la conducta y el daño causado. De igual forma, la ley faculta a las víctimas para emprender acciones civiles con la finalidad de recibir indemnizaciones por los perjuicios causados (Ministerio de Industria y Tecnología de la Información de la República Popular China, 2022).

A nivel internacional, los países están adoptando diferentes enfoques para abordar el uso inadecuado de la inteligencia artificial generativa. En general, la legislación se está moviendo hacia la creación de marcos legales que no solo penalicen la producción y distribución de contenido deepfake, sino que también impongan responsabilidades a las plataformas digitales que facilitan su difusión. Estas leyes buscan enfrentar tanto la creación como la propagación de material no consensuado, estableciendo sanciones penales y, en algunos casos, medidas preventivas y de protección para las víctimas.

Las legislaciones varían en cuanto a su alcance y gravedad, es así cómo algunos países están implementando sanciones rigurosas y detalladas, en cambio, otros están en fases preliminares de desarrollo o abordan el problema a través de regulaciones más generales relacionadas con el contenido digital. La tendencia general es hacia un enfoque más proactivo y específico, que reconoce la necesidad de abordar la problemática de manera integral.

Ausencia de legislación normativa en Colombia en relación con la pornografía deepfake

Dada la poca actuación del Estado Colombiano para combatir la pornografía deepfake no consensuada, es necesario analizar los tipos penales existentes dentro del ordenamiento jurídico colombiano, que podrían abarcar algunos aspectos de esta conducta, esto con la finalidad de mitigar y castigar la violación masiva de los derechos de imagen, intimidad personal y el buen nombre con el objetivo constituir un mecanismo de protección con la legislación vigente.

Uno de ellos es la injuria, tipificada en el artículo 220 de la ley 599 del 2000 (Código Penal) que establece “El que haga a otra persona imputaciones deshonrosas incurrirá en prisión de dieciséis (16) a cincuenta y cuatro (54) meses y multa de trece puntos treinta y tres (13.33) a mil quinientos (1.500) salarios mínimos legales mensuales vigentes.” (Congreso de la República de Colombia, Ley 599 de 2000). La injuria busca proteger el honor (respeto, autoestima) y buen nombre (reputación) de las personas, derechos que consagran parte de la integridad moral de las mismas, fundamentados en su dignidad.

La Sala de Casación Penal de la Corte Suprema de Justicia en auto de 29 de septiembre de 1983 enumeró los requisitos que permiten estructurar de manera satisfactoria el punible de injuria, siendo los siguientes: “1) Que una persona impute a otra conocida o determinable un hecho deshonroso; 2) Que el imputado tenga conocimiento del carácter deshonroso del hecho, 3) Que el carácter deshonroso del hecho imputado dañe o menoscabe la honra de aquella persona; 4) Que el imputador tenga conciencia de que el hecho atribuido tiene esa capacidad de dañar o menoscabar la honra de la persona”. (Como se citó en la Corte Constitucional de Colombia, 2011).

Del mismo modo, la Sala de Casación Penal es enfática en destacar que la legislación colombiana establece como conducta sancionable para la injuria el comportamiento doloso (*animus injuriandi*); es decir, cuando el ofensor tiene pleno conocimiento que la imputación es deshonrosa para el agraviado y el daño que esta que fuere a producir en su patrimonio moral, y pese a ello decide hacerla. (Corte Constitucional de Colombia, Sentencia C-442 de 2011).

Conforme a estos preceptos jurisprudenciales, es evidente que al publicar el material sexual falso creado por inteligencia artificial se está realizando una imputación deshonrosa para la persona retratada, dado que; (i) implica la atribución de la realización de un acto sexual que no cometió una persona claramente identificable o reconocible, hecho considerado de carácter deshonroso; (ii) quien crea o distribuye a conciencia el deepfake pornográfico tiene pleno conocimiento de que el material es falso y de carácter deshonroso y humillante para la sociedad en general; (iii) la difusión o la mera existencia del deepfake puede menoscabar la integridad moral de la persona cuyo rostro ha sido retratado; y (iv) el deepfake no es un error inocente, quien lo crea o distribuye a conciencia lo hace con la intención de mostrar a la persona en una posición degradante y humillante, con la capacidad de menoscabar la honra y el buen nombre de la víctima. En términos generales, la conducta de quienes participan en la creación o difusión a conciencia de la pornografía deepfake no consensuada cumple con los requisitos para la estructuración del delito de injuria en Colombia.

Además, el artículo 222 del Código Penal Colombiano habla de la injuria y calumnia indirectas: “A las penas previstas en los artículos anteriores quedará sometido quien publicare, reproducire, repitiere injuria o calumnia imputada por otro, o quien haga la imputación de modo impersonal o con las expresiones se dice, se asegura u

otra semejante”. (Congreso de la República de Colombia, Ley 599 de 2000) Por consiguiente, cualquier persona que publique, comparta, reproduzca el deepfake pornográfico en plataformas digitales estarían incurriendo en la conducta descrita en el artículo anterior, aunque no hayan participado en la creación del contenido, ni a su difusión inicial incurrirá en las mismas penas previstas para el delito de injuria. Sin embargo, surge el siguiente interrogante ¿existe realmente el animus injuriandi en quienes comparten el contenido de deepfake pornográfico creado sin consentimiento, cuando no son capaces de diferenciar entre los videos reales y los creados por inteligencia artificial?

Por otro lado, algunos de los puntos negativos de la injuria es que hace parte de los delitos querellables en Colombia, lo que significa que son desistibles o conciliables. La querrela debe ser presentada dentro de los 6 meses siguientes a la realización de la conducta punible o desde que se tuvo conocimiento de la misma, un plazo relativamente corto considerando que en otros países se pretende establecer plazos de 3 años o más. Aunado a esto, la pena es moderadamente pequeña comparada con las penas previstas en otras legislaciones para una conducta que puede generar afectaciones graves para la víctima, como lo es la pornografía deepfake.

Aparte, el Código Penal Colombiano prevé la figura jurídica de la retractación en su artículo 225 que dispone:

No habrá lugar a responsabilidad si el autor o partícipe de cualquiera de las conductas previstas en este título, se retractare voluntariamente antes de proferirse sentencia de primera o única instancia, siempre que la publicación de la retractación se haga a costa del responsable, se cumpla en el mismo medio y con las mismas características en que se difundió la imputación o en el que señale el funcionario judicial, en los demás casos. No se podrá iniciar acción penal, si la retractación o rectificación se hace pública antes de que el ofendido formule la respectiva denuncia. (Congreso de la República de Colombia, 2000).

Esto significa que, si el creador o difusor a conciencia del contenido de pornografía deepfake admite voluntariamente que el video es falso y deshonesto, de forma pública y bajo las mismas condiciones en las que se difundió el deepfake no habrá lugar a la pena; en el caso de un deepfake viral la retractación deberá subirse por todos los medios digitales en donde fue difundido con la finalidad de evitar la responsabilidad penal que lo atañe. Pero, ¿es suficiente una simple retractación para resarcir el daño ocasionado por un deepfake pornográfico que podría seguir circulando por internet?

Si bien, la retractación es un paso importante para mitigar el impacto que este pueda generar en la víctima, no es suficiente para resarcir el daño completamente, dado que la retractación puede no ser tan viral o no tenga el mismo alcance de difusión como el material sexual creado sin consentimiento por inteligencia artificial; además, es posible que la persona retratada sufra afectaciones emocionales o psicológicas a causa de la difusión del material, y por si fuera poco, la retractación no garantiza la plena y completa eliminación del deepfake pornográfico en internet, especialmente si ha sido compartido por millones de personas.

Otro tipo penal que podría aplicarse en los deepfakes pornográficos no consensuados es el delito de falsedad personal tipificado en el artículo 296 del Código Penal Colombiano, el cual establece: “El que con el fin de obtener un provecho para sí o para otro, o causar daño, sustituya o suplante a una persona o se atribuya nombre, edad, estado civil, o calidad que pueda tener efectos jurídicos, incurrirá en multa, siempre que la conducta no constituya otro delito”. (Congreso de la República de Colombia, 2000) Este punible podría configurarse en casos específicos; para que ello suceda la persona debe hacerse pasar activamente por la víctima, por ejemplo, creando un perfil falso en su nombre para publicar deepfakes pornográficos no consensuados de la víctima, obteniendo beneficios económicos. En este tipo de situaciones se cumple, en su sentido más estricto, los elementos tradicionales de la falsedad personal o suplantación de identidad, puesto que adopta el papel de la víctima de manera directa, bien sea para obtener un beneficio económico o causarle daño a la persona retratada.

La pena establecida para este delito es una multa, por lo tanto, no dispone de una pena privativa de la libertad, lo que es insuficiente, considerando los daños ocasionados a la víctima; además, la aplicabilidad del delito es

considerablemente limitada, teniendo en cuenta que el artículo señala que solo se ejecutará la pena prevista si la conducta no constituye otro delito más grave. En la mayoría de los casos de pornografía deepfake podrían estar presentes otros delitos más graves como la injuria que podrían tener más prioridad sobre la falsedad personal.

El Estado colombiano como máximo garante de los derechos fundamentales a la imagen, intimidad personal y el buen nombre, tiene el deber de emprender acciones para protegerlos, en tal sentido, combatir la pornografía deepfake, esto debido al exponencial riesgo que representa este nuevo tipo de violencia digital para los derechos fundamentales, la cual ha estado acrecentando conforme avanza la inteligencia artificial. Por ello, el tema central en este contexto es si el Estado colombiano cuenta con las herramientas legales, tecnológicas y operativas necesarias para prevenir, mitigar y sancionar estos.

Como parte del trabajo investigativo del presente estudio, se radicó diferentes solicitudes a la fiscalía general de la Nación y a la Policía Nacional de Colombia con la finalidad obtener información detallada acerca de los casos de pornografía deepfake en Colombia y que tan preparadas están estas instituciones para afrontar la reciente y peligrosa problemática.

En las respuestas obtenidas de la fiscalía general de la Nación y la Policía Nacional de Colombia evidencian que no existen registros o datos oficiales sobre casos de pornografía deepfake. Ambas instituciones confirmaron que no han identificado denuncias o querellas relacionadas con este fenómeno. Por lo que la ausencia de estos datos limita el desarrollo de políticas públicas efectivas para abordar los deepfake pornográficos.

Además, en el marco de las respuestas, la Fiscalía mencionó que su Grupo de Informática Forense solo dispone de software que clasifica imágenes explícitas, pero carece de sistemas avanzados para identificar ediciones generadas por inteligencia artificial. De manera similar, el Grupo de Fotografía y Video del Cuerpo Técnico de Investigación (CTI) “no posee herramientas que permitan verificar la autenticidad de contenido deepfake”. Por su parte, la Policía Nacional señaló que no cuenta con software especializado para la detección de contenido deepfake en material de abuso sexual infantil ni en otros delitos informáticos. Además, reconoció que enfrenta desafíos en la capacitación y fortalecimiento de sus equipos de investigación forense, lo que limita su capacidad para analizar y judicializar estos casos.

Finalmente, ambas hicieron énfasis en la insuficiencia del marco legal en Colombia que impide una respuesta efectiva y contundente por parte de las autoridades judiciales, estableciendo la necesidad de una respuesta legal que permita a la formulación de directrices para el uso y desarrollo responsable de la IA en Colombia, asimismo una que sancione la producción y difusión de este contenido.

Conclusiones

La investigación tuvo como objetivo central analizar la creación de pornografía Deepfake no consensuada mediante IA en relación con un análisis de legislación comparada para estudiar cómo esto vulnera el derecho de imagen, intimidad personal y buen nombre, a partir de un estudio de casos. Para ello, el presente estudio expuso como problema de investigación el siguiente interrogante: ¿de qué manera puede el marco jurídico adaptarse para garantizar la protección efectiva de los derechos de imagen, intimidad personal y el buen nombre frente a los riesgos generados por la pornografía deepfake no consensuada?

Para dar respuesta al mismo, es importante mencionar que se realizó un análisis de casos desde una perspectiva internacional con el propósito de estudiar los elementos de cada uno frente a la vulneración de derechos humanos; en segunda medida, se realiza un análisis de legislación comparada, para identificar las normas de algunos países sobre el tema objeto de estudio y; finalmente, se llevó a cabo un estudio del caso del Estado colombiano, en el cual se estudia la normativa vigente, los diferentes pronunciamientos jurisprudenciales existentes y la postura actual del Estado colombiano en relación con la IA y su impacto en los derechos humanos.

Para dar respuesta al interrogante, debe indicarse que la creación de pornografía Deepfake no consensuada genera una vulneración sistemática de derechos humanos, y según se desarrolló anteriormente, impacta en el derecho a la imagen, intimidad personal y el buen nombre, los cuales deben ser protegidos por la legislación vigente de los Estados, como una obligación internacional, que proteja a toda persona en casos de trasgresión de sus derechos. Así mismo, la investigación logró evidenciar que el uso irresponsable de la IA afecta a toda persona sin distinción alguna, inclusive se observan casos registrados de menores de edad, que han sido víctimas de casos de pornografía Deepfake, lo que genera un dilema frente a la escasa normativa existente para garantizar una protección integral en estos casos. Aunado a ello, se logró evidenciar que a nivel internacional se han presentado diferentes propuestas legislativas para reglamentar el uso de la IA y su impacto en los derechos humanos, pero han sido pocos los avances en ese sentido. En tercera medida, debe indicarse que, no se trata de prohibir el uso de la IA, sino por el contrario, que sean los Estados los encargados de plantear iniciativas que proponen un debate ético jurídico que garantice los derechos de toda persona en estos casos, pero, que sea proporcional en relación con los casos de trasgresión particular.

Es importante mencionar que el impacto de la IA generativa en los derechos humanos como el buen nombre, intimidad y derecho de imagen, requiere la actualización normativa de cada Estado para que las diferentes instituciones puedan contar con las herramientas normativas y procesales para actuar en estos casos y así proteger los derechos de la población en cada caso en concreto, ya que el auge de las tecnologías emergentes implica la obligación de los Estados de generar las condiciones necesarias para una mayor seguridad digital de los usuarios de la IA, ya que cada vez se hacen mas accesibles a toda persona y ello requiere el estudio de la legislación para que los operadores judiciales puedan tener las herramientas jurídicas para resolver casos concretos amparados por la Constitución política, los instrumentos y pactos internacionales, leyes nacionales y demás normas concordantes que promuevan una seguridad jurídica en estos casos y dilucidar el dilema ético y legal de la creación de contenido pornográfico Deepfake no solo en Colombia sino en los diferentes países de la región.

Referencias

- 20minutos (2024) El Gobierno de Reino Unido penalizará con cárcel la creación de deepfakes porno. <https://www.20minutos.es/tecnologia/inteligencia-artificial/gobierno-reino-unido-carcel-deepfake-porno-5237101/>
- Asamblea General de Georgia. (2019) Code of Georgia (Estados Unidos). Recuperado de <https://law.justia.com/codes/georgia/2019/>
- Asamblea General de Virginia. (2019) Code of Virginia (Estados Unidos) <https://law.lis.virginia.gov/vacode/title18.2/chapter8/section18.2-386.1/>
- Asamblea Nacional Constituyente. (1991). Constitución Política de Colombia de 1991. Gaceta Constitucional 116 de 20 de julio de 1991. http://www.secretariassenado.gov.co/senado/basedoc/constitucion_politica_1991.html
- Assembly Bill No. 602 (2019) An act to add Section 1708.86 to the Civil Code, relating to privacy. https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=20190200AB602
- Balboa, M. (2022, 08 de septiembre). ¿Rostros aparentemente reales creados por computadora?, Sí, se hacen con deepfake y el ciberdelito ya los está usando. El comercio. https://elcomercio.pe/tecnologia/actualidad/deepfakes-que-es-tipos-riesgos-y-amenazas-inteligencia-artificial-como-lo-esta-usando-el-ciberdelito-noticia/#google_vignette
- BBC News Mundo (2018). Qué son los "deepfakes" y por qué se están convirtiendo en el nuevo "porno de la venganza". <https://www.bbc.com/mundo/noticias-43975322>
- BBC News Mundo (2023) Investigación BBC: el comercio ilegal de imágenes de abuso sexual infantil generadas por Inteligencia Artificial. <https://www.bbc.com/mundo/articles/c1vzyevl0nr0>
- Bigas, N. (2023). 'Deepfakes' pornográficos: Cuando la IA desnuda tu intimidad y vulnera tus derechos. UOC. <https://www.uoc.edu/es/news/2023/265-deepfakes-pornograficos-cuando-la-ia-desnuda-tu-intimidad-vulnera-tus-derechos>
- BOE (1995). Ley Orgánica 10/1995, de 23 de noviembre, del Código Penal. <https://www.boe.es/buscar/pdf/1995/BOE-A-1995-25444-consolidado.pdf>
- Boté Vericad J. J. y Vállez M. (2022). Aplicaciones de deepfakes. Manipulación de contenido audiovisual y riesgos para los usuarios basados en las políticas de privacidad. *Documentación de las Ciencias de la Información*, 45(1), 25-32.

- <https://doi.org/10.5209/dcin.77256>
- Carranza, M. (2023) El escándalo del streamer Atrio y las fotos de desnudos generadas por IA. Sopitas.com. <https://www.sopitas.com/geek/streamer-atric-que-paso-deepfake-fotos-desnudos-ia-escandalo/>
- CNN Mundo (2023) Corea del Sur encierra a un hombre por utilizar inteligencia artificial para crear imágenes sexuales de niños, en un hecho sin precedentes en el país.
- Congreso de la República de Colombia. (2000). Ley 599 de 2000. Por la cual se expide el Código Penal. Diario Oficial No. 44.097 de 24 de julio de 2000. <https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=6388>
- Congreso de los diputados (2023) Ley 122/000011. Proposición de Ley Orgánica de regulación de las simulaciones de imágenes y voces de personas generadas por medio de la inteligencia artificial. https://www.congreso.es/public_oficiales/L15/CO/NG/BOCG/B/BOCG-15-B-23-1.PDF
- Congreso de los Estados Unidos (2023). To require covered platforms to remove non consensual intimate visual depictions, and for other purposes. <https://files.constantcontact.com/1849eea4801/b6c3c645-ac61-4a95-9088-0d0e932fe633.pdf>
- Corte Constitucional de Colombia. La Sala Novena de Revisión de la Corte Constitucional. (2000). Sentencia T-094. Magistrado Ponente Dr. Álvaro Tafur Galvis. <https://www.corteconstitucional.gov.co/relatoria/2000/T-094-00.htm>
- Corte Constitucional de Colombia. La Sala Octava de Revisión de tutelas. (2020). Sentencia T-007. Magistrado Ponente José Fernando Reyes Cuatras. <https://www.corteconstitucional.gov.co/relatoria/2020/T-007-20.htm>
- Corte Constitucional de Colombia. La Sala Octava de Revisión de Tutelas. (2019). Sentencia T-610. Magistrado Ponente José Fernando Reyes Cuatras. <https://www.corteconstitucional.gov.co/relatoria/2019/T-610-19.htm>
- Corte Constitucional de Colombia. La Sala Plena Constitucional. (2019). Sentencia SU420. Magistrado Ponente José Fernando Reyes Cuatras. <https://www.corteconstitucional.gov.co/relatoria/2019/SU420-19.htm>
- Corte Constitucional de Colombia. La Sala Primera de Revisión de la Corte Constitucional. (2013). Sentencia T-634. Magistrado Ponente María Victoria Calle Correa. <https://www.corteconstitucional.gov.co/relatoria/2013/T-634-13.htm>
- Corte Constitucional de Colombia. La Sala Sexta de Revisión de la Corte Constitucional. (2011). Sentencia T-949. Magistrado Ponente Nilson Pinilla Pinilla. <https://www.corteconstitucional.gov.co/relatoria/2011/T-949-11.htm#:~:text=T%2D949%2D11%20Corte%20Constitucional%20de%20Colombia&text=Esta%20Corte%20ha%20se%C3%B1alado%20que,activa%20la%20protecci%C3%B3n%20del%20derecho>
- Corte Constitucional de Colombia. Sala Plena de la Corte Constitucional. (2011). Sentencia C-881. Magistrado Ponente Jorge Ignacio Pretelt Chaljub. <https://www.corteconstitucional.gov.co/relatoria/2014/C-881-14.htm#:~:text=El%20n%C3%BAcleo%20esencial%20del%20derecho,vida%20personal%2C%20espiritual%20y%20cultural>
- Corte Constitucional de Colombia. Sala Plena. (2002) Sentencia C 489. Magistrado Ponente Rodrigo Escobar Gil. <https://www.corteconstitucional.gov.co/relatoria/2002/C-489-02.htm>
- Corte Constitucional de Colombia. Sala Plena. (2011). Sentencia C-442. Magistrado Ponente Humberto Antonio Sierra Porto. <https://www.corteconstitucional.gov.co/relatoria/2011/C-442-11.htm>
- Corte Constitucional de Colombia. Salas Civil y Penal del Tribunal Superior de Antioquia. (1995). Sentencia SU 056. Magistrado Ponente Antonio Barrera Carbonell. <https://www.corteconstitucional.gov.co/RELATORIA/1995/SU056-95.htm>
- Creswell J. y Creswell J. (2018) Research Design (7ª ed.). https://spada.uns.ac.id/pluginfile.php/510378/mod_resource/content/1/creswell.pdf
- Duque, L. (2019) Taylor Swift, los deepfakes y la violencia de género. Podcast Universidad de los Andes. <https://cerosetenta.uniandes.edu.co/taylor-swift-deepfakes-violencia-genero/>
- El economista (2024) Deepfakes: Sancionarán con cárcel la difusión de imágenes porno creadas con IA. <https://www.eleconomista.com.mx/politica/Deepfakes-Sancionaran-con-carcel-la-difusion-de-imagenes-porno-creadas-con-IA-20240216-0094.html>
- France 24 (2023) Acoso en colegios de EEUU mediante pornografía falsa creada con IA. <https://www.france24.com/es/minuto-a-minuto/20231124-acoso-en-colegios-de-eeuu-mediante-pornograf%C3%ADa-falsa-creada-con-ia>
- Franganillo, J. (2023). La inteligencia artificial generativa y su impacto en la creación de contenidos mediáticos. Methadots Revista de Ciencias Sociales, 11(2), m231102a10. <http://dx.doi.org/10.17502/mrcs.v11i2.710>
- Home Security Heroes (2023). State Of Deepfakes. <https://www.homesecurityheroes.com/state-of-deepfakes/#key-findings>

- ISDI. (2023, 28 de Septiembre). Inteligencia artificial generativa: ¿Qué es y cómo funciona? Digital Skills. Recuperado el 23 de marzo de 2024. Obtenido de <https://www.isdi.education/es/blog/inteligencia-artificial-generativa-que-es>
- Magas (2024) La congresista Alexandria Ocasio-Cortez, víctima de un 'deepfake' pornográfico: "Es como una violación". https://www.elespanol.com/mujer/actualidad/20240409/congresista-alexandria-ocasio-cortez-victima-deepfake-pornografico-violacion/846415865_0.html
- Marshall y Rossman (2016). Designing Qualitative Research (6th ed.). Thousand Oaks, CA: Sage. <https://www.scirp.org/reference/referencespapers?referenceid=2914926>
- Martínez-Comesaña, M., Rigueira-Díaz, X., Larrañaga-Janeiro, A., Martínez-Torres, J., Ocaranza-Prado, I., & Kreibel, D. (2023). Impacto de la inteligencia artificial en los métodos de evaluación en la educación primaria y secundaria: revisión sistemática de la literatura. *Revista de psicodidáctica*.
- Ministerio de Industria y Tecnología de la Información de la República Popular China (2022) Disposiciones sobre la gestión de síntesis en profundidad de los servicios de información de Internet. https://www.gov.cn/zhengce/zhengceku/2022-12/12/content_5731431.htm
- Ministerio de Justicia de Corea del Sur (2023). Ley Especial de Sanción de Delitos de Violencia Sexual. <https://www.law.go.kr/%EB%B2%95%EB%A0%B9/%EC%84%B1%ED%8F%AD%EB%A0%A5%EB%B2%94%EC%A3%84%EC%9D%98%EC%B2%98%EB%B2%8C%EB%93%B1%EC%97%90%EA%B4%80%ED%95%9C%ED%8A%B9%EB%A1%80%EB%B2%95>
- Moncada, A. (2020). Ventajas y riesgos de las redes generativas antagónicas. Vertigo Político. <https://www.vertigopolitico.com/todo-menos-politica/conocimiento/ventajas-y-riesgos-de-las-redes-generativas-antagonicas>
- Montevideo Portal (2023) Acoso en centros educativos de EE. UU. mediante pornografía infantil falsa creada con IA. <https://www.montevideo.com.uy/Ciencia-y-Tecnologia/Acoso-en-centros-educativos-de-EE-UU-mediante-pornografia-infantil-falsa-creada-con-IA-uc872261>
- Morales, K. (2023). Pornografía Deepfake y Violencia de Género. Centro de Autonomía Digital. <https://autonomia.digital/2023/11/22/deepfakes-and-gender-violence.html>
- Murillo y Ramírez (s.f) Implicaciones de la Inteligencia Artificial. Universidad Libre, Pereira, Colombia. Recuperado de: <https://hdl.handle.net/10901/29284>.
- OBS Business School (s.f) ¿Qué es la Inteligencia Artificial Generativa?. <https://www.obsbusiness.school/blog/que-es-la-inteligencia-artificial-generativa>
- OCDE (2019) Cuarenta y dos países adoptan los Principios de la OCDE sobre Inteligencia Artificial. <https://www.oecd.org/espanol/noticias/cuarentaydospaisesadoptanlosprincipiosdelaocdesobreinteligenciaartificial.htm>
- Parlamento de la Commonwealth de Australia (2024). CRIMINAL CODE AMENDMENT (DEEPFAKE SEXUAL MATERIAL) BILL 2024. <https://www.diarioconstitucional.cl/wp-content/uploads/2024/06/Vea-proyecto-de-ley-1.pdf>
- Pérez Llain, Y. Y. ., Durán Chinchilla, C. M. ., & Casadiegos Santana, M. H. . (2025). La inteligencia artificial en la educación superior: un examen ético-jurídico de su perspectiva transformadora. *Academia & Derecho*, 18(30). <https://revistas.unilibre.edu.co/index.php/academia/article/view/12707>
- Perfil (2024) Proyecto de Ley de Deepfakes en Argentina. <https://www.perfil.com/noticias/opinion/proyecto-de-ley-de-deepfakes-en-argentina.phtml>
- Redacción Vistazo (2023) Alumnos de un colegio de Quito usaron inteligencia artificial para crear videos sexuales con los rostros de sus compañeras. Vistazo. <https://www.vistazo.com/actualidad/nacional/alumnos-de-un-colegio-de-quito-usaron-inteligencia-artificial-para-crear-videos-sexuales-con-los-rostros-de-sus-companeras-ME6107394>
- Rockcontent (2021) Tipos de Inteligencia Artificial: conoce cuáles existen y cómo usarlos. <https://rockcontent.com/es/blog/tipos-de-inteligencia-artificial/>
- Rodrigo, J (2022) Realismo Trágico: ¿Cómo regular los deepfakes en Colombia? *Latín American Law Review*, 1(8), 125-145. Universidad de los Andes. <https://revistas.uniandes.edu.co/index.php/lar/article/view/4715>
- Swissinfo.ch (2023) La India prepara una ley para regular y penar el uso de vídeos ultrafalsos o “deepfake”. <https://www.swissinfo.ch/spa/la-india-prepara-una-ley-para-regular-y-penar-el-uso-de-v%C3%ADdeos-ultrafalsos-o-deepfake/49002178>
- The New York State Senate (2020). Senate Bill S5959D. <https://www.nysenate.gov/legislation/bills/2019/S5959>