



Desafíos y dimensiones de la desinformación en ALAC: deepfakes y la urgencia de proteger los derechos de las mujeres*

Challenges and dimensions of disinformation in ALAC: deepfakes and the
urgency to protect women's rights

Constanza González-Véliz**

contacto@constanzagonzalezveliz.com

Investigadora independiente. Talca, Chile.

ORCID: 0009-0000-4409-9741

Ximena Cuzcano-Chavez***

contacto@ximenacuzcanochavez.com

Investigadora independiente. Lima, Perú.

ORCID: 0009-0002-9854-0234

Palabras clave: *desinformación; deepfakes; ciberviolencia de género; derechos de las mujeres; regulación tecnológica*

Keywords: *disinformation; deepfakes; gender cyber-violence; women's rights; technological regulation*

El presente trabajo es un análisis que aborda los desafíos y dimensiones de la desinformación, centrándose en la proliferación de *deepfakes* y su impacto desproporcionado en las mujeres. Desde la irrupción de las *fake news* hasta el surgimiento de tecnologías *deepfake*, se destaca una transformación en la interacción social con la información. La propagación acelerada de desinformación, manifestada en formas como *deepfakes*, presenta riesgos significativos para la percepción individual y la participación democrática. Los *deepfakes*, videos manipulados mediante inteligencia artificial, representan un fenómeno más avanzado y desafiante que las *fake news*, al recrear hiperrealistamente situaciones inexistentes. Se examina cómo estas tecnologías, inicialmente desarrolladas para propósitos creativos, se han convertido en armas digitales dirigidas principalmente a las mujeres. La ciberviolencia de género, manifestada en *deepfakes* pornográficos y acosos en línea, compromete la privacidad, reputación y salud mental de las mujeres, socavando la igualdad de género en el entorno digital. La ausencia de marcos regulatorios específicos y la insuficiencia de métodos de detección confiables plantean desafíos

* Esta investigación forma parte del proyecto VERITAS ALAC: Vulnerabilidades, Estrategias y Reflexiones sobre la Innovación Tecnológica y Asimetrías de Género en América Latina y el Caribe. <https://veritas.situada.online>.

** Ingeniera Forestal, Especialista en Políticas Públicas y Justicia de Género, Máster en Estudios Feministas e Intervención para la Igualdad.

*** Ingeniera de Sistemas, Master en Ciberseguridad.



DÍA INTERNACIONAL DE LA MUJER

éticos y de derechos humanos. Se destaca la urgencia de adoptar enfoques de género en políticas de ciberseguridad, promover la alfabetización en información y pensamiento crítico, y cerrar la brecha de género en el acceso y uso de tecnologías digitales. La conclusión resalta la necesidad imperativa de proteger los derechos de las mujeres en línea mediante investigaciones continuas, marcos regulatorios específicos y estrategias efectivas para prevenir la violencia digital.

Introducción

La sociedad actual se encuentra inmersa en una transformación significativa en la manera en que accede y procesa la información. Este cambio ha dado lugar a la propagación acelerada de desinformación, que va más allá de las conocidas *fake news* y se manifiesta de manera más avanzada en los *deepfakes*. Estos videos manipulados mediante inteligencia artificial representan un desafío único, al recrear situaciones de manera hiperrealista, desafiando la percepción y la veracidad de la información.

La evolución de las *fake news* a los *deepfakes* ha llevado a la sociedad a enfrentarse a un fenómeno complejo que afecta tanto a nivel individual como social. La verificación tardía de hechos en el caso de las *fake news* ha permitido la propagación de información errónea, dificultando la corrección de percepciones equivocadas. Con la llegada de los *deepfakes*, esta problemática se intensifica, ya que estos videos avanzados pueden ser utilizados para manipular elecciones, exacerbar divisiones sociales, influir en la confianza en instituciones y autoridades, y promover un consumo sesgado de información partidista.

Este fenómeno no solo amenaza la integridad de la información, sino que también impacta de manera desproporcionada a las mujeres. Los *deepfakes*, más allá de su capacidad para crear narrativas engañosas, se han convertido en herramientas de acoso digital que violan los derechos fundamentales de las mujeres, incluyendo la privacidad, la reputación y la salud mental. La ciberviolencia de género, exacerbada por las brechas digitales y los sesgos tecnológicos, se ha normalizado, silenciando a las víctimas y legitimando este problema. En este contexto, se observa una falta de enfoque de género en la protección del derecho a la privacidad en línea, dejando desatendidas las experiencias de las mujeres y colocándolas en una posición vulnerable. La regulación para salvaguardar a las mujeres de la violencia en línea se presenta como crucial, y aunque se han adoptado enfoques de género en políticas de ciberseguridad en América Latina y el Caribe (ALAC), persisten desafíos éticos y de derechos humanos.

Este panorama revela la urgencia de abordar la desinformación y la ciberviolencia de género de manera integral. En este sentido, es necesario comprender el impacto de los *deepfakes*, desarrollar métodos de detección más efectivos y promover una regulación inclusiva que proteja los derechos de las mujeres en línea. Las perspectivas de investigación cruciales incluyen el cierre de las brechas de género en el acceso, uso y apropiación de tecnologías digitales, así como la implementación de estrategias para prevenir la violencia digital y promover un entorno en línea seguro y equitativo para todas las personas. Estas acciones son esenciales para avanzar hacia una sociedad más justa, inclusiva y libre de violencia de género en el ámbito digital.



DÍA INTERNACIONAL DE LA MUJER

MOVIMIENTO

El análisis exhaustivo sobre las dimensiones de la desinformación en ALAC realizado en este trabajo surge como respuesta a la creciente amenaza que representan los *deepfakes* en el panorama de la información y la ciberviolencia de género. El propósito fundamental del trabajo es comprender y destacar la complejidad de este fenómeno, que va más allá de la mera manipulación de información y alcanza dimensiones éticas, sociales y de género. El fin último, es contribuir a la comprensión y concienciación de la sociedad sobre los profundos efectos de los *deepfakes*, especialmente en el contexto de la ciberviolencia de género. Se busca destacar la necesidad imperativa de desarrollar métodos de detección más efectivos y promover regulaciones inclusivas que protejan los derechos de las mujeres en línea.

Perspectivas actuales sobre la desinformación: de las *fake news* a los *deepfakes*

La sociedad ha experimentado un cambio significativo en la forma de informarse, lo que ha facilitado las operaciones de desinformación (Fernández, 2020). No solo afecta a nivel individual, sino que también tiene consecuencias a nivel social. Dejando a las personas vulnerables de caer en un ciclo de engaños, afectando su capacidad de interpretar la realidad y volviendo a la sociedad susceptible a intereses fraudulentos (Zúñiga-Carrasco & Millar de Jesús, 2022). Es en este contexto donde surgen las *fake news* que se difunden principalmente por medios tradicionales, redes sociales y páginas especializadas en sátira. La efectividad en su propagación a menudo se atribuye a la verificación de hechos compartida con un retraso significativo después de la divulgación de la información falsa original. Por lo que una vez arraigada la percepción errónea, corregirla resulta extremadamente complicado. También, su propagación se vincula estrechamente al comportamiento y criterio de las personas, su atracción por lo impactante lleva a compartir esta información falsa sin importar su veracidad (Caiza et al., 2020).

El término *fake news* se entrelaza con *información engañosa*, *información errónea* y *desinformación*. Algunos planteamientos distinguen entre la *información errónea*, compartida sin intención dañina, y la *desinformación*, creada para engañar. Sin embargo, no existe un acuerdo definitivo en su utilización, más bien, son criterios de orientación. De todas maneras, al hablar de *fake news* es más propio hablar de *desinformación*, ya que imitan contenido mediático con la intención de engañar o causar daño, a menudo persiguiendo algún tipo de beneficio. Se trata de artículos o mensajes noticiosos que, independiente de sus medios y motivos, contienen información falsa y son difundidos a través de los medios de comunicación. Buscan influir en la opinión pública, generando inseguridad y desestabilización con diversos propósitos, ya sea apoyar o desacreditar alguna situación o persona, o simplemente por diversión (Caiza et al., 2020). De hecho, el término se consolidó en el discurso público a partir de finales de 2016, cuando agentes de inteligencia rusos buscaron influir en las elecciones presidenciales estadounidenses. Desde entonces, han surgido como una preocupación política prominente y un término ampliamente utilizado durante las candidaturas, destacando su relevancia, especialmente en el contexto de figuras como Donald Trump. Su efectividad radica en su capacidad para conectarse emocionalmente con el público (Cerdán & Padilla-Castillo, 2019).

Aunque la proliferación de *fake news* no es nueva, la magnitud actual es alarmante, pues tienen una probabilidad muy alta de ser compartidas y creídas en comparación con noticias

164

2024

DERECHOS

2017



DÍA INTERNACIONAL DE LA MUJER

MOVIMIENTO

verídicas. Además, abarcan diversas formas de desinformación, desde teorías de conspiración hasta propaganda (Zúñiga-Carrasco & Millar de Jesús, 2022). Es por esto que se ha generado una creciente preocupación global. En el campo académico, su estudio surgió después de las elecciones presidenciales de 2016 en EE. UU. y del proceso del Brexit en el Reino Unido, pues es crucial comprender su impacto en sociedades y personas que las generan, difunden y consumen (Caiza et al., 2020). Asimismo, a nivel individual, algunas personas han desarrollado sus propias defensas, tomando medidas para protegerse, como la verificación de la información con fuentes alternativas y evitando difundir noticias dudosas (Fernández, 2020). También, se destacan los esfuerzos para contener la propagación de *fake news*, con diversas técnicas de detección automática, incluyendo inteligencia artificial. No obstante, aún se carece de métodos altamente confiables para garantizar la veracidad de la información. La estrategia más efectiva combina detección computarizada con alfabetización en información y fomento del pensamiento crítico de la ciudadanía. Empoderar a la comunidad para identificar noticias mediante validación es esencial. Además, el proceso de verificación de hechos debe ser transparente y las plataformas de noticias deben proporcionar un entorno neutral para el debate, promoviendo así una información más precisa y confiable (Caiza et al., 2020).

La desinformación ha extendido su influencia al ámbito visual, especialmente a través de los *deepfakes*, videos manipulados mediante inteligencia artificial que simulan situaciones creíbles (Zúñiga-Carrasco & Millar de Jesús, 2022). Es una fase avanzada de las *fake news*, que utiliza técnicas de alteración de fotos y videos con la facilidad de un software de edición convencional para modificarlos (Fernández, 2020) y crear contenido falso (Caiza et al., 2020), mostrando a personas realizando acciones o expresando palabras inexistentes en la realidad (Mustak et al., 2023). Este contenido falso es hiperrealista que finalmente se percibe como auténtico (Bañuelos, 2021). Debido a la reducción de costos y la disminución de requisitos de entrenamiento se hace más fácil acceder a esta tecnología, mientras que cada iteración mejora la calidad, dificultando la detección de componentes sintéticos a simple vista (Helmus, 2022).

Los *deepfakes* alcanzan la perfección al recrear la realidad, desafiando la percepción convencional para discernir su autenticidad. Gracias a técnicas de inteligencia artificial, como el aprendizaje automático, se pueden integrar decenas de imágenes en un algoritmo que genera máscaras humanas altamente convincentes que reemplazan rostros en videos, utilizando datos similares y permitiendo que el software se perfeccione con el tiempo (Cerdán & Padilla-Castillo, 2019). Por lo que la evidencia audiovisual ya no garantiza la verdad, convirtiendo los *deepfakes* en contenidos aparentemente veraces, complicando la verificación de la autenticidad. Pues, la multiplicidad de herramientas digitales permite diversas versiones de la verdad, incluso mediante imágenes descontextualizadas (Zúñiga-Carrasco & Millar de Jesús, 2022), marcando un desafío significativo en la confiabilidad de la información mediática. Además, aprovechan la capacidad de confirmar prejuicios sin necesidad de realismo y tienen el potencial de afectar tanto a la sociedad como a la seguridad nacional. En relación a este último, hay cuatro formas cruciales en las que pueden ser utilizados: para manipular elecciones, para exacerbar divisiones sociales, para influir en la confianza en instituciones y autoridades, y para desencadenar un consumo sesgado de información partidista (Helmus, 2022). Todo con la intención de crear creencias erróneas. Y, dada su naturaleza emergente y compleja, la comprensión actual de sus implicaciones es limitada,

165

2024

DERECHOS

2017



DÍA INTERNACIONAL DE LA MUJER

aunque desde ya es posible notar que su escala, volumen y velocidad de distribución, junto con la omnipresencia digital, tendrán consecuencias significativas en diversos ámbitos de la sociedad, delineando un panorama complejo y dinámico (Mustak et al., 2023).

El surgimiento del deepfake: acoso digital que afecta a las mujeres

El término *deepfake* surgió en 2017, combinando aprendizaje profundo y falso. A diferencia de formas tradicionales de engaño que ocultan información o presentan datos falsos, las tecnologías *deepfake*, pueden generar narrativas humanas usando procesamiento de lenguaje natural (Mustak et al., 2023). Su proliferación ofrece nuevas oportunidades creativas, pero también plantea amenazas. Si bien la sátira y la fantasía tienen su espacio, la capacidad de retocar videos amplía el riesgo de distorsionar imágenes para humillar o sexualizar mujeres (Tripathi, 2019). Desde su surgimiento ha abarcado diversas categorías, como la pornografía no consentida, la extorsión pornográfica, la denuncia, actuaciones virtuales, retratos vivientes, entre otras (Bañuelos, 2021).

El precursor de esta práctica, fue un usuario de Reddit autodenominado Deepfake, quien empleó los rostros de Gal Gadot, Maisie Williams y Taylor Swift, fusionándolos con cuerpos de actrices de pornografía para crear videos convincentes que simulaban la participación de las celebridades. Utilizando su computadora personal y un algoritmo de *machine learning* accesible en línea, este usuario compartió los videos en Reddit, generando gran interés y llevando a la creación de un subreddit dedicado a este tema. En solo dos meses, la comunidad alcanzó 15.000 suscriptores, popularizando el término *deepfake* para describir videos creados con inteligencia artificial. Asimismo, otro usuario, Deepfakeapp, desarrolló la aplicación FakeApp en Reddit, permitiendo a cualquier persona, incluso sin experiencia en informática, crear sus propios videos mediante inteligencia artificial. Aunque la aplicación es versátil, ha ganado notoriedad principalmente por la creación de pornografía falsa (Cerdán & Padilla-Castillo, 2019). Además, actualmente, diversas páginas web brindan acceso a servicios *deepfake*, representando una amenaza tangible y un costo desproporcionado para mujeres que son víctimas de sitios de pornografía generada por inteligencia artificial. Un ejemplo de esto es DeepNude, que permite la carga de fotos generando imágenes en las que la persona aparenta estar desnuda (Helmus, 2022).

Una abrumadora mayoría de videos *deepfakes* en Internet se emplean en pornografía no consentida, revelando imágenes íntimas sin el consentimiento de las personas (Tripathi, 2019). En ellos se superponen caras de celebridades en actrices pornográficas, se difundieron rápidamente en plataformas como Pornhub, afectando no solo a famosas, sino también a mujeres anónimas cuyos ex parejas usan la tecnología para vengarse y humillarlas. Esto se vincula al problema del porno de venganza, siendo devastador para las mujeres al parecer auténtico. Esta tecnología, presente y futura en la desinformación, se ha convertido en una peligrosa herramienta para atentar contra la imagen de las mujeres, construyendo un imaginario falsificado y perverso que perpetúa la degradación y el daño (Cerdán & Padilla-Castillo, 2019).

La violencia digital dirigida a las mujeres se manifiesta, principalmente a través del ciberacoso, el hostigamiento sexual en línea y la difusión no consentida de imágenes íntimas. Las redes sociales, especialmente Facebook y X, son los principales entornos virtuales donde ocurren estos



DÍA INTERNACIONAL DE LA MUJER

MOVIMIENTO

actos de agresión y violencia de género (González-Véliz & de Andrés, 2023a). Estas acciones constituyen tendencias alarmantes que violan los derechos humanos, incluyendo privacidad, reputación y salud mental de las mujeres. Al socavar su dignidad, también infringen el principio fundamental de no discriminación e igualdad, abusando de múltiples derechos (Tripathi, 2019). Es así como el *deepfake* se presenta como una herramienta alarmante y perjudicial que socava la imagen de las mujeres al construir narrativas falsas y perversas. A través de escenas eróticas y pornográficas ficticias, las mujeres son cosificadas y despojadas de su identidad. Este contraste se evidencia cuando los hombres, en situaciones humorísticas o políticas, mantienen su integridad personal y protagonismo positivo. Las mujeres son víctimas de la despersonalización, presentadas en contextos íntimos, mientras los hombres mantienen su esencia. Por lo que es crucial explorar todas las formas de esta amenaza en Internet, considerando incluso formatos más breves como boomerangs, gifs y memes, a pesar de las restricciones en plataformas como Facebook, YouTube e Instagram (Cerdán & Padilla-Castillo, 2019).

Es así como el ámbito digital, inicialmente concebido como un espacio de liberación para las mujeres, se ha transformado en un arma de doble filo. Aunque empoderó a mujeres globalmente para expresarse, también otorgó libertades similares a quienes perpetran acoso en línea, que trasciende lo digital afectando el mundo físico y siendo más fácil cuando las mujeres son el blanco. Al cuestionar su intención, agencia, intelecto y carácter, los acosadores desencadenan consecuencias graves, aunque los perpetradores raramente enfrentan consecuencias, dificultando establecer vínculos causales. La falta de aplicación de la perspectiva de género en la protección del derecho a la privacidad en línea deja desatendidas las experiencias de mujeres en plataformas de redes sociales. La violencia de género en línea, al amenazar e intimidar a mujeres, coarta su libertad de expresión, acceso a información y participación en procesos políticos, infringiendo gravemente sus derechos y exponiéndolas a riesgos que atentan contra su derecho a vivir libres de violencia (Tripathi, 2019). Esta situación se relaciona estrechamente con las brechas digitales y el diseño tecnológico con sesgos de género que contribuyen a perpetuar la violencia digital contra mujeres y niñas, colocándolas en una posición altamente vulnerable (González-Véliz & de Andrés, 2023b).

Ciberviolencia: una amenaza para las mujeres en ALAC y la urgencia de regulaciones inclusivas

Es crucial reconocer que el derecho a una vida libre de violencia y discriminación es fundamental para lograr la igualdad de género y preservar la autonomía de las mujeres (CEPAL, 2023). Este principio abarca también la violencia en línea, destacando la necesidad de abordar y prevenir formas digitales de agresión que amenazan la integridad y la igualdad de oportunidades para mujeres en entornos virtuales. Porque la violencia de género constituye una forma discriminatoria y violatoria de los derechos humanos, impactando de manera desproporcionada a las mujeres. Adicionalmente, la brecha digital y los sesgos tecnológicos la acentúan (Vaca-Trigo & Valenzuela, 2022). Es un fenómeno complejo y en constante cambio que, sin embargo, debido a la falta de datos, impide conocer la magnitud del problema y su impacto real en las víctimas pues dificulta su seguimiento, tendencias y los impactos en la vida de las mujeres y niñas (OEA,

167

DERECHOS

2017

2024



DÍA INTERNACIONAL DE LA MUJER

MOVIMIENTO

2021a). Adicionalmente, mediante campañas de desinformación y difusión de información falsa ha aumentado la violencia política de género en línea (Vaca-Trigo & Valenzuela, 2022). Esta es una problemática que se ha intensificado en los últimos años, afectando a aquellas mujeres que tienen un perfil público, especialmente en la política, periodismo y defensa de derechos humanos. Este tipo de acoso y ataques tienen como objetivo desacreditarlas, socavar su reputación y credibilidad. Estas prácticas fomentan la cosificación, explotación y subordinación de las mujeres con fines económicos, sociales o de dominación, lo que perpetúa la desigualdad y violencia de género. De igual manera, obstaculizan su inclusión plena en los ecosistemas digitales, lo que las obliga a autocensurarse y reducir su interacción en línea, debilitando su autonomía (CEPAL, 2023).

Las mujeres son víctimas desproporcionadas de ciberviolencia en comparación con los hombres. Un ejemplo de esto lo expuso la Comisión de Banda Ancha de las Naciones Unidas que en 2015 identificó que 73% de las mujeres había experimentado alguna forma de violencia en línea. También, la Relatora Especial sobre Violencia de la ONU informó en 2018 que el 23% de las mujeres y niñas había sufrido abuso o acoso en línea al menos una vez en su vida. Además, un estudio de Plan Internacional de 2020 reveló que casi 60% de las niñas y jóvenes en todo el mundo habían sido víctimas de diferentes formas de ciberacoso, enfrentándose a esta forma de violencia desde los 8 años. Según Amnistía Internacional, 88% de las mujeres que utilizan X para hablar sobre feminismo y temas relacionados con la igualdad de género sufren acoso. Asimismo, 90% de las víctimas de la distribución no consentida de imágenes íntimas son mujeres y 80% de las víctimas de la trata de personas son mujeres y niñas (OEA, 2021b).

Durante los últimos años, en ALAC, ha quedado en evidencia esta situación, pues el aumento del uso de tecnologías digitales ha generado una exposición más amplia a los riesgos en línea, donde el cibercrimen, el abuso y la violencia digital existían previamente. Como resultado de la falta de conocimientos y herramientas de protección y seguridad digital ha generado un ambiente propicio para estas agresiones (Vaca-Trigo & Valenzuela, 2022). Sin embargo, la falta de registros estadísticos y estudios oficiales sobre la violencia de género en línea contra las mujeres ha llevado a diversas organizaciones de la sociedad civil a documentar el fenómeno. Estos indican que la violencia digital contra las mujeres se ha incrementado (OEA, 2021b). Ejemplo de esto es un estudio realizado en Chile donde se queda en evidencia la falta de información sobre la violencia digital, lo que sugiere que este tema no recibe la atención necesaria. Según el estudio de la Fundación Datos Protegidos en 2018, las formas principales de violencia digital experimentadas fueron la violencia verbal, como burlas, insultos y humillaciones, en un 88,1%, el acoso y hostigamiento en un 66,1%, y el envío de imágenes o videos sexualmente agresivos sin consentimiento en un 40,6%. Además, también se reveló que un 30,5% sufrió la exposición de información falsa sobre su vida personal, un 22% recibió amenazas explícitas y un 18,6% perdió el acceso a sus redes sociales o correo electrónico debido a hackeos. Estas formas de violencia digital también afectan a menores de edad, con un 32% de las niñas y un 1% de los niños recibiendo peticiones sexuales o solicitudes de fotos íntimas a través de internet, por su parte, ocurre lo mismo con 47% de las mujeres adolescentes entre 15 y 18 años (Matus et al., 2018).

En toda la región, se sigue normalizando la violencia de género en línea por parte de los medios de comunicación, plataformas de internet, autoridades y comunidades, lo que silencia a las víctimas y legitima el problema. Los medios presentan la violencia en línea contra las

DERECHOS

2017

2024



DÍA INTERNACIONAL DE LA MUJER

MOVIMIENTO

mujeres como un fenómeno neutral o una consecuencia del comportamiento inapropiado de las víctimas, y un cuarto de ellas son culpabilizadas por la violencia sufrida. Además, se culpa a mujeres jóvenes y niñas por prácticas de sexting y la distribución no consensuada de sus imágenes íntimas, y se exculpa constantemente a los perpetradores, minimizando y marginando su conducta agresiva. Los atacantes en línea son infantilizados y se responsabiliza a las víctimas de resolver el problema. El ciberacoso es calificado como "bromas" y se recomienda a las mujeres retirarse de Internet. Perpetradores, plataformas de internet y algunas organizaciones utilizan argumentos sobre libertad de expresión y censura, pero no se aborda la perspectiva de género y de derechos humanos sobre los daños sufridos por las víctimas de la violencia digital (OEA, 2021b).

Es vital establecer regulaciones que salvaguarden a las mujeres de la violencia en línea, en el caso de ALAC se han estado adoptando enfoques de género en políticas de ciberseguridad, abordando ciberacoso y pornovenganza. La seguridad digital, como derecho humano, no solo garantiza el acceso de las mujeres a Internet, sino también su capacidad para ejercer sus derechos (Vaca-Trigo & Valenzuela, 2022). Para garantizar la seguridad digital y contribuir a crear un entorno digital más seguro e inclusivo, se deben crear políticas con enfoque de género y estrategias basadas en derechos humanos digitales. Asimismo, es clave fortalecer la gestión para prevenir delitos digitales, apoyar a víctimas y, se deben implementar sistemas de monitoreo y campañas de concientización para promover espacios seguros. Es fundamental capacitar a docentes y familias sobre comportamiento en línea y protección contra violencia digital (CEPAL, 2023).

169

Perspectiva regional: *deepfakes* en ALAC y su impacto desproporcionado en la violencia de género

Desde el año 2020, se ha observado el surgimiento de una práctica preocupante en distintos países de Latinoamérica y el Caribe: la proliferación de videos *deepfakes*. En este contexto, resulta necesario investigar y comprender los casos de *deepfakes* que han impactado la región, a través de una exhaustiva revisión de los medios de comunicación en países de ALAC. La revisión mediática no solo destaca la actualidad del fenómeno, sino que también enfatiza la importancia de entender lo que se está discutiendo en relación con los *deepfakes* en la región. Estas manipulaciones digitales que sobreponen rostros de figuras públicas, como Lionel Messi (Balmaceda, 2020), tienen propósitos variados, desde propósitos humorísticos y de entrenamiento en redes sociales hasta situaciones alarmantes en el ámbito político argentino. Cerca de las elecciones presidenciales, un *deepfake* viral mostró a un candidato consumiendo drogas, generando preocupación (Urgente24, 2023; Galarza, 2023). Este video resultó ser falso, pero su impacto llevó a que la contienda electoral del 2023 se convirtiera en un campo de batalla digital lleno de imágenes *deepfakes*, utilizadas para respaldar, desprestigiar y sembrar desinformación entre los votantes (Nicas & Cholakian Herrera, 2023). Incluso, el actual presidente de Argentina difundió una imagen de su contrincante político retratado como un líder comunista chino, alcanzando 3 millones de vistas (Milei, 2023).

Brasil lidera en casos de *deepfakes* en la región, representando la mitad de los incidentes reportados, especialmente en industrias como criptomonedas, tecnología financiera y juegos en

DERECHOS

2017

2024



DÍA INTERNACIONAL DE LA MUJER

MOVIMIENTO

línea. Argentina y Colombia también registran casos significativos, aunque en menor proporción en comparación con Brasil (Sumsb, 2023). La incidencia de *deepfakes* no se limita a sectores económicos, ya que, durante las elecciones de 2022 en Brasil, se emplearon estas técnicas para distorsionar la información en el influyente medio Jornal Nacional. Esto implicó la difusión de videos falsos donde periodistas informaban incorrectamente sobre encuestas electorales y difamaban a candidatas y candidatos (Jornal Nacional, 2022). En 2023, Perú también enfrentó esta problemática cuando dos periodistas televisivos fueron víctimas de *deepfakes* con motivaciones económicas: una promocionando criptomonedas y la otra vendiendo supuestos productos curativos (Ramírez Mendoza, 2023). Un caso aún más perturbador involucró a estudiantes que manipularon imágenes de compañeras de 15 años, insertando sus rostros en situaciones sexuales ficticias para luego comercializarlas entre sus compañeros (Saavedra, 2023). Que la institución educativa afirmara que el incidente ocurrió en el entorno virtual, generó protestas, ya que se consideraba un delito de pornografía infantil. La respuesta tanto de la institución como la legislación peruana se limitó a trasladar a los agresores a otro salón, dejando a las víctimas vulnerables durante clases, y expuestas a los agresores en el centro educativo (Rojas Berríos, 2023). Además, la ley de Inteligencia Artificial en Perú no contempla disposiciones para sancionar el uso inapropiado de esta tecnología, como en los casos de *deepfakes* (Morán & Lavanda, 2023).

En México también se han reportado casos actuales de *deepfakes* para crear contenido pornográfico de mujeres, muchas de ellas menores. Uno de los casos más impactantes involucró a un estudiante universitario que alteró 166 mil fotos de sus compañeras para convertirlas en material pornográfico mediante el uso de *deepfake*, y luego venderlas en grupos de Telegram. Aunque las víctimas denunciaron el delito, la Fiscalía optó inicialmente por no procesarlo, a pesar de las pruebas en su contra (Rodríguez, 2023). En México existe la Ley Olimpia desde el 2020, que penaliza la difusión de contenido sexual sin consentimiento, sin embargo, esta ley resulta insuficiente para abordar delitos de *deepfake*, dado que sus estándares y sanciones difieren en identificación y aplicación con respecto a estos casos (Piña, 2023). Actualmente, la mayoría de videos *deepfakes* en Internet son de naturaleza pornográfica, principalmente protagonizados por mujeres cuyas imágenes son manipuladas sin su consentimiento. La creciente normalización de esta disparidad con respecto a otros tipos de *deepfakes*, como los políticos o económicos, es preocupante. Es crucial entender que la abrumadora mayoría de estos *deepfakes* pornográficos afectan a mujeres, resaltando de manera contundente la violencia de género en el entorno digital (Lucas, 2022).

En la web, diversas aplicaciones, mayormente desarrolladas por hombres, utilizan la tecnología de *deepfakes* para convertir imágenes de mujeres vestidas en desnudos realistas sin su consentimiento, considerándose una forma de *deepfake*. Estas aplicaciones no funcionan de la misma manera con imágenes de hombres (Cole, 2019a). En ciertos casos, cuando estas aplicaciones han sido retiradas, sus creadores han tratado de eximirse de responsabilidad al afirmar que nunca anticiparon la verdadera magnitud del impacto de su creación, argumentando que solo buscaban proporcionar entretenimiento. Sin embargo, tanto expertos como usuarios coinciden en que esta práctica carece de justificación y tiene consecuencias perjudiciales para la intimidad y dignidad de las personas afectadas (Cole, 2019b).

170

DERECHOS

2017

2024



DÍA INTERNACIONAL DE LA MUJER

Los videos no consensuados pueden tener consecuencias devastadoras para las víctimas, incluyendo daños psicológicos y reputacionales (Stevens, 2020). Los casos identificados revelan que las víctimas enfrentan constante humillación y desorientación al ser reducidas como objetos sexuales (Chesney & Keats Citron, 2019). Quienes consumen este tipo de contenido no distinguen entre la realidad y la manipulación, lo que resulta en suposiciones erróneas sobre la participación de la víctima, causando graves perjuicios en sus relaciones personales y profesionales (Uribarri, 2022).

Desafíos y propuestas frente al *deepfake*: democratización digital, violencia de género y necesidades legislativas

La promesa de la revolución digital era la democratización del acceso y la libre expresión en Internet y, aunque amplió el acceso, careció de protección para quienes cuentan con acceso limitado al poder. Si bien proporcionó un espacio para grupos históricamente excluidos, también permitió a la mayoría dominar, aprovechando su poder e influencia para silenciar voces disidentes, desvirtuando la igualdad de participación. La democratización inicial de la comunicación digital se desvaneció cuando aquellos con poder buscaron suprimir perspectivas divergentes (Tripathi, 2019). Por lo que es crucial tener en cuenta que Internet no es meramente un espacio donde se refleja la cultura, sino que también es un producto cultural que surge de la interacción social, lo que implica más que simplemente explorar la red; requiere conocimiento y apropiación del espacio, así como interacción activa con el entorno digital (González-Véliz & de Andrés, 2023a). En esta misma línea, existen propuestas que buscan abordar los desafíos emergentes de los *deepfakes*. Entre ellas, se destaca la necesidad de que plataformas de redes sociales colaboren en la detección, así como que se incorporen *deepfakes* notorios como *datos radiactivos* que revelen cualquier uso hostil. También, las tecnologías como modos seguros en cámaras y metadatos para garantizar la autenticidad de las imágenes. Además, la investigación constante y mejoras en la caracterización informan y educan al público, democratizando las herramientas accesibles y su respaldo gubernamental son esenciales para construir confianza en la era de *deepfakes* (Helmus, 2022).

La prohibición total de publicar *deepfakes* en plataformas de redes sociales está en marcha. TikTok está actualizando sus directrices para prohibir contenido sintético o manipulado que distorsione la verdad y cause daño. Reddit ha actualizado su política contra la suplantación de identidad, no permitiendo contenido que se haga pasar por personas de manera engañosa. YouTube prohíbe medios manipulados que fabriquen eventos con riesgo de daño atroz. A pesar de estas medidas, es importante reconocer que aún pueden existir lagunas en la detección. Por lo tanto, las empresas tienen la oportunidad de invertir en técnicas de detección computacional para abordar estos desafíos, ya que existen experimentaciones que reflejan la viabilidad de detectar los peligros de Internet utilizando técnicas de inteligencia artificial, como máquinas de vectores de soporte, bosques aleatorios y regresión logística (Cuzcano & Ayma, 2020). También, el aprendizaje profundo está siendo explorado, dado que los *deepfakes* pueden volverse increíblemente sutiles con el tiempo y es crucial contar con herramientas que presenten una alta tasa de precisión en la detección (James, 2023). Además, las plataformas de redes sociales deben ca-



DÍA INTERNACIONAL DE LA MUJER

MOVIMIENTO

pacitar a su personal en sensibilidad de género, aplicando estándares internacionales de manera informada y evitando el relativismo cultural, en ese sentido, la consulta con grupos de derechos humanos, mujeres y libertad de expresión es esencial. Asimismo, instituir paneles internos para recibir comentarios y colaborar con partes interesadas externas, desarrollando políticas robustas para negar acceso a quienes persisten en intimidar o acosar a mujeres o grupos vulnerables. Las empresas que cuentan con recursos deben patrocinar investigaciones sobre los riesgos para la salud mental de moderadores y víctimas, además de apoyar regulaciones gubernamentales y ampliar la verificación de datos para combatir la desinformación. Es imperativo garantizar la seguridad en línea para toda la población (Tripathi, 2019).

El fenómeno del *deepfake*, aún en sus primeras etapas de desarrollo, carece de legislación específica. No obstante, debido a sus objetivos, podría ser sancionado como un delito contra el derecho a la propia imagen, una forma de injuria o incluso un delito en sí mismo (Cerdán y Padilla-Castillo, 2019). La falta de marcos legales claros destaca la necesidad urgente de abordar estos vacíos para enfrentar los desafíos emergentes asociados con la manipulación audiovisual, protegiendo así los derechos individuales y preservando la integridad de la información y la imagen personal (Fernández, 2020). Sin duda, lo que no hay que perder de vista es que la tecnología desempeña un papel social crucial y es esencial transformarla aprovechando su potencial sin precedentes y la amplia gama de opciones tecnológicas, organizativas y de relaciones sociales disponibles. Por lo tanto, es fundamental comprender que no se limita a ser un conjunto de herramientas, sino que constituye un componente integral y dinámico de la sociedad (González-Véliz & de Andrés, 2023a).

Conclusión

En el análisis de los desafíos y dimensiones de la desinformación, particularmente en el contexto de los *deepfakes*, se evidencia una realidad compleja y preocupante que impacta de manera significativa a las mujeres. La transformación en la forma en que la sociedad interactúa con la información, desde las *fake news* hasta los *deepfakes*, ha planteado retos sustanciales para la veracidad de la información y la protección de los derechos fundamentales.

La propagación acelerada de desinformación, más allá de representar un riesgo para la percepción individual, se ha convertido en una herramienta de acoso digital dirigida principalmente a las mujeres. Los *deepfakes*, al recrear situaciones hiperrealistas, no solo desafían la veracidad de la información, sino que también violan la privacidad, reputación y salud mental de las mujeres. La ciberviolencia de género, exacerbada por la falta de enfoque de género en la protección en línea, ha generado un ambiente propicio para la normalización de esta problemática, silenciando a las víctimas y perpetuando la violencia digital.

La falta de regulaciones específicas y la insuficiencia de métodos de detección efectivos plantean desafíos éticos y de derechos humanos que requieren una atención inmediata. La urgencia de proteger los derechos de las mujeres en línea se refleja en la necesidad de adoptar enfoques de género en políticas de ciberseguridad, promover la alfabetización en información y pensamiento crítico, y cerrar las brechas de género en el acceso, uso y apropiación de tecnologías digitales. En el camino hacia una sociedad más justa e inclusiva, es imperativo abordar estos

DERECHOS

2017

2024



desafíos mediante investigaciones continuas sobre detección de *deepfakes*, marcos regulatorios específicos con perspectiva de género y estrategias efectivas para prevenir la violencia digital. La seguridad digital, como derecho humano, debe garantizar no solo el acceso equitativo a la tecnología, sino también la capacidad de las mujeres para ejercer sus derechos en un entorno en línea seguro y libre de violencia de género. La conclusión es clara: el compromiso colectivo en la protección de los derechos de las mujeres en el ámbito digital es esencial para construir una sociedad digital más equitativa y respetuosa.

Referencias

- Balmaceda, T.** (2020). De Messi a Francella: cómo se hacen los *deepfakes* argentinos. *LA NACION*. <https://www.lanacion.com.ar/tecnologia/de-lionel-messi-guillermo-francella-como-se-nid2438022/>
- Bañuelos Capistrán, J.** (2022). Evolución del Deepfake: campos semánticos y géneros discursivos 2017-2021. *Revista ICONO 14. Revista Científica De Comunicación Y Tecnologías Emergentes*, 20(1). <https://doi.org/10.7195/ri14.v20i1.1773>
- Caiza, E., Fernandez, A., & Torres, D.** (2020). Noticias falsas; en busca de la vacuna. *Revista Colombiana De Computación*, 21(2), 92–101. <https://revistas.unab.edu.co/index.php/rcc/article/view/4037>
- Cerdán Martínez V., & Padilla Castillo G.** (2019). Historia del "fake" audiovisual: "deepfake" y la mujer en un imaginario falsificado y perverso. *Historia y Comunicación Social*, 24(2), 505-520. <https://doi.org/10.5209/hics.66293>
- Chesney, R., & Keats Citron, D.** (2019). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. *107 California Law Review* 1753, 1753-1773.
- Cole, S.** (2019a). This Horrifying App Undresses a Photo of Any Woman With a Single Click. *Vice*. <https://www.vice.com/en/article/kzm59x/deepnude-app-creates-fake-nudes-of-any-woman>
- Cole, S.** (2019b). Creator of DeepNude, App That Undresses Photos of Women, Takes It Offline. *Vice*. <https://www.vice.com/en/article/qv7agw/deepnude-app-that-undresses-photos-of-women-takes-it-offline>
- Comisión Económica para América Latina y el Caribe (CEPAL).** (2023). La igualdad de género y la autonomía de las mujeres y las niñas en la era digital: aportes de la educación y la transformación digital en América Latina y el Caribe. Santiago, Chile.
- Cuzcano, X. M., & Ayma, V. H.** (2020). A Comparison of Classification Models to Detect Cyberbullying in the Peruvian Spanish Language on Twitter. *International Journal of Advanced Computer Science and Applications*, 132-138. doi:10.14569/IJACSA.2020.0111018
- Fernández, C. B.** (2020). La democracia amenazada por ciberpolítica. Los riesgos más allá de facebook. *Diálogo político*. <https://dialogopolitico.org/edicion-especial-2020-seguridad/la-democracia-amenazada-por-la-ciberpolitica/>



DÍA INTERNACIONAL DE LA MUJER

MOVIMIENTO

- Galarza, F.** (2023). Deepfake en la campaña electoral: el detrás de la inteligencia artificial que genera videos falsos de Sergio Massa. *El Destape*. <https://www.eldestapeweb.com/sociedad/elecciones-2023/deepfake-en-la-campana-el-detras-de-la-ia-que-genera-videos-falsos-de-massa-202311140554>
- González-Véliz, C., & de Andrés, S.** (2023a). Corrientes del ciberfeminismo y transición digital feminista en América Latina y el Caribe: hacia un futuro digital más equitativo. *Revista Punto Género*, (N° 20), pp, 31-61. <https://doi.org/10.5354/2735-7473.2023.73459>
- González-Véliz, C., & de Andrés, S.** (2023b). Comunes digitales, ciberfeminismo y ciencia abierta: estrategias para la igualdad de género en el entorno digital. *Hipertext.net*, Vol. 27, (N°7), pp. 67-79. <https://doi.org/10.31009/hipertext.net.2023.i27.07>
- Helmus, T. C.** (2022). Artificial Intelligence, deepfakes, and Disinformation: A Primer. *RAND Corporation*. <https://doi.org/10.7249/PEA1043-1>
- James, C.** (2023). Intel's deepfake detector tested on real and fake videos. BBC. Obtenido de <https://www.bbc.com/news/technology-66267961>
- Jornal Nacional** (2022). Deepfake: conteúdo do Jornal Nacional é adulterado para desinformar os eleitores. *Jornal Nacional*. <https://g1.globo.com/jornal-nacional/noticia/2022/09/19/deepfake-conteudo-do-jornal-nacional-e-adulterado-para-desinformar-os-eleitores.ghtml>
- Lucas, K.** (2022). Deepfakes and Domestic Violence: Perpetrating Intimate Partner Abuse Using Video Technology. *Victims & Offenders*, 647-659.
- Matus, J., Rayman, D., & Vargas, R.** (2018). Informe violencia de género en Internet en Chile. Datos Protegidos. Santiago, Chile.
- Morán, A., & Lavanda, M.** (2023). Reflexiones Legales sobre la IA en el Perú: Consideraciones desde Asilomar. *Pólemos*. <https://polemos.pe/reflexiones-legales-sobre-la-ia-en-el-peru-consideraciones-desde-asilomar/>
- Milei, J.** [@JMilei]. (2023) EL MÉTODO CASTA-MASSA. X. <https://x.com/JMilei/status/1719479197440127330>
- Mustak, M., Salmien, J., Mäntymäki, M., Rahman, A., & Dwivedi, Y. K.** (2023). Deepfakes. Deceptions, mitigations, and opportunities. *Journal of Business Research*. Vol. 154. <https://doi.org/10.1016/j.jbusres.2022.113368>
- Nicas, J., & Cholakian Herrera, L.** (2023). La IA hace campaña en las elecciones de Argentina. *The New York Times*. <https://www.nytimes.com/es/2023/11/15/espanol/elecciones-argentina-imagenes-inteligencia-artificial.html>
- Organización de los Estados Americanos (OEA).** (2021a). La violencia de género en línea contra las mujeres en línea. <https://www.oas.org/es/sms/cicte/docs/Guia-conceptos-basicos-La-violencia-de-genero-en-linea-contras-las-mujeres-y-ninas.pdf>
- Organización de los Estados Americanos (OEA).** (2021b). Ciberviolencia y ciberacoso contra las mujeres y las niñas en el marco de la convención Belém do Pará. <https://www.oas.org/es/mesecvi/docs/MESECVI-Ciberviolencia-ES.pdf>
- Piña, G.** (2023). Cinco casos de "DeepFake porn" en México. La ley es insuficiente: Olimpia Coral. <https://m-x.com.mx/entrevistas/cinco-casos-de-deepfake-porn-en-mexico-la-ley-es-insuficiente-olimpia-coral/>

174

2024

DERECHOS

2017



DÍA INTERNACIONAL DE LA MUJER

MOVIMIENTO

- Ramírez** Mendoza, S. (2023). Clonan rostros de conductoras peruanas de TV para usarlos en publicidad engañosa: nuevo caso de deepfake y qué medidas de prevención tomar. *El Comercio*. <https://elcomercio.pe/lima/clonan-los-rostros-de-conductoras-de-television-con-ia-para-utilizarlos-en-publicidad-enganosa-la-tecnica-del-deepfake-y-que-medidas-de-prevencion-tomar-maite-vizcarra-mavila-huertas-inteligencia-artificial-estafa-conductoras-pe>
- Rodríguez**, A. (2023). Acusado un universitario de alterar con inteligencia artificial miles de imágenes de alumnas para venderlas como pornografía. *El País*. <https://elpais.com/mexico/2023-10-14/acusado-un-universitario-de-alterar-con-inteligencia-artificial-miles-de-imagenes-de-alumnas-para-venderlas-como-pornografia.html>
- Rojas** Berríos, M. (2023). Chorrillos: Escolares que alteraron fotos de compañeras con IA y las comercializaron no fueron expulsados. *Infobae*. <https://www.infobae.com/peru/2023/08/29/chorrillos-escolares-que-alteraron-fotos-de-sus-companeras-con-ia-para-venderlas-no-fueron-expulsados/>
- Saavedra**, N. (2023). Chorrillos: denuncian a alumnos de colegio por editar fotos de sus compañeras para venderlas. *La República*. <https://larepublica.pe/sociedad/2023/08/28/chorrillos-denuncian-a-alumnos-de-colegio-por-editar-fotos-de-sus-companeras-para-venderlas-1338680>
- Stevens**, D. (2020). Regulating Deepfake Technology. Legislative possibilities in the Netherlands to obstruct the use of deepfake technology for the creation of non-consensual pornography: <https://arno.uvt.nl/show.cgi?fid=152071>
- Sumsub** (2023). *Sumsub Identity Fraud Report*. https://sumsub.com/files/sumsub_identity_fraud_report_2023.pdf
- Tripathi**, S. (2019). Women and online harassment. (N°6), pp. 133-158. *Feminist Dissent*. <https://doi.org/10.31273/fd.n6.2022.1265>
- Urgente24** (2023). Deepfake de Sergio Massa: No saben el daño que hacen. *Urgente24*. <https://urgente24.com/actualidad/deepfake-sergio-massa-no-saben-el-dano-que-hacen-n563440>
- Uribarri**, F. (2022). Aplicaciones de 'deepfake' colocan a mujeres en videos porno. *La Voz de Galicia*. <https://www.lavozdeg Galicia.es/xlsemanal/a-fondo/mujeres-porno-deepfake-montaje-caras-videos-pornografia-ilegal.html>
- Vaca** Trigo, I., & Valenzuela, M. E. (2022). Digitalización de las mujeres en América Latina y el Caribe: acción urgente para una recuperación transformadora y con igualdad. Comisión Económica para América Latina y el Caribe (CEPAL). Santiago, Chile.
- Zúñiga-Carrasco**, I., & Millar de Jesús, R. (2022). Noticias falsas y pseudocientíficas. *Revista del Hospital de Juárez de México*. Vol. 89, (N°3) pp. 145-152. <https://doi.org/10.24875/RHJM.22000030>

175

2024

DERECHOS

2017