

## TECNOLOGÍA DEEP FAKE: EL LADO OSCURO DEL DEEP LEARNING

### Soluciones educativas para un desarrollo sostenible

## DEEP FAKE TECHNOLOGY: THE DARK SIDE OF DEEP LEARNING

### Educational solutions for sustainable development

Mario González Arencibia <sup>1\*</sup>, Emilio Horacio Valencia Corozo <sup>2</sup>, Dagmaris Martínez Cardero <sup>3</sup>

<sup>1</sup> Universidad de las Ciencias Informáticas (UCI). Facultad 3, La Habana, Cuba. Carretera a San Antonio km 2 ½, CP 17830. [mgarencibia@uci.cu](mailto:mgarencibia@uci.cu)

<sup>2</sup> Instituto de Altos Estudios (IAEN), Ecuador, Quito. [emiliohoracio1@hotmail.com](mailto:emiliohoracio1@hotmail.com)

<sup>3</sup> Universidad de las Ciencias Informáticas (UCI). CITEC, La Habana, Cuba. Carretera a San Antonio km 2 ½, CP 17830. [dagmarism@uci.cu](mailto:dagmarism@uci.cu)

\* Mario González Arencibia: [mgarencibia@uci.cu](mailto:mgarencibia@uci.cu)

---

#### Resumen

Las Falsificaciones Profundas (Deep fake) constituyen una evolución tecnológica, que expresa un cambio radical en la manera en que se produce, se divulga y maneja de manera inadecuada la información. Ello podría llevar a que este hecho, mal maniobrado, se convierta en una amenaza cada vez más grave para individuos, organizaciones y la sociedad en su conjunto, al formar parte de un desarrollo sin precedentes en el ecosistema de la desinformación. A partir de esta inquietud, el **objetivo** de este artículo es incursionar en el debate sobre las preocupaciones y soluciones educativas para un desarrollo sostenible frente a los usos maliciosos de la “Tecnología Deep Fake” (TDF), colocando la atención en que los problemas que se puedan generar en estos ámbitos, es desde la transformación y puesta en práctica de la consolidación de valores que enarbolan la ética como principio básico que distingue la naturaleza humana en su expresión más civilizada, en lo cual no se descartan las funciones que podrían desempeñar la tecnología, el componente legal y las políticas gubernamentales y privadas. Se empleó el **método** de análisis documental aplicando la triangulación de autores lo que permitió deslindar las causas y soluciones educativas frente al Deep Fake,

obteniendo como **resultado**, una propuesta de soluciones educativas frente a sus usos maliciosos. La **conclusión** central es que el enfrentamiento a los problemas que causa el inapropiado empleo de la “Tecnología Deep Fake” (TDF), pasa por el trabajo educativo que se realice en los escenarios en que convivimos.

**Palabras clave:** Tecnología, Deep Fake, Ética, Deep Learning, Inteligencia Artificial.

### **Abstract**

Profound falsifications constitute a technological evolution, which expresses a radical change in the way information is produced, disseminated and inadequately used. This could lead to this badly manipulated fact becoming an increasingly serious threat to individuals, organizations and society as a whole, as it is part of an unprecedented development in the disinformation ecosystem. Based on this concern, the objective of this article is to enter into the debate on educational concerns and solutions in the face of the malicious uses of "Deep Fake Technology" (DFT), focusing on the fact that the problems that can be generated in these areas, is from the transformation and implementation of the consolidation of values that raise ethics as a basic principle that distinguishes human nature in its most civilized expression, in which the functions that technology, the legal component and government and private policies could perform are not discarded. The documentary analysis method was used, applying the triangulation of authors, which allowed the causes and educational solutions to be distinguished from the Deep Fake, obtaining as a result, a proposal of educational solutions against its malicious uses. The main conclusion is that the problems caused by the inappropriate use of "Deep Fake Technology" (DTF) are to be addressed through the educational work carried out in the scenarios in which we live.

**Keywords:** Technology, Deep Fake, Ethics, Deep Learning, Artificial Intelligence.

---

## **Introducción**

Cada vez más los desarrollos tecnológicos nacidos en el escenario de la Inteligencia Artificial (IA) son más sofisticados, dando lugar a preocupaciones y soluciones ante los procesos de algoritmización que imitan el comportamiento humano de manera realista. En ello son relevantes las experiencias del “Deep Learning”, conocidas en castellano como “Aprendizaje Profundo”, cuyas expresiones tienen el poder de adquirir autonomía propia a través del aprendizaje de datos de grandes magnitudes, con los cuales, el instrumental tecnológico es capaz de generar procesos de toma de decisiones y de transformaciones positivas, de cualquier contexto que dependa de la ciencia de los datos. Sin embargo, el “Deep Learning” expresión de un alto grado de socialización y de dominio de los algoritmos computacionales de la

IA, ha estado siendo empleado desde su condición de tecnología de doble uso de manera maliciosa, con implicaciones negativas que tienden a quebrantar la dignidad humana, violentando principios éticos (González Arencibia, & Martínez Cardero, 2020). Ello ha hecho notorio, el lado oscuro del “Deep Learning”: La “Tecnología Deep Fake” (TDF), la cual se emplea para falsear la realidad.

El impacto de este uso negativo, ha sido de una magnitud significativa, reflejándose en el cuestionamiento de la individualidad, de organizaciones, poniendo en peligro la sociedad en su conjunto, por las consecuencias implícitas que podrían tener inadecuados manejos de la TDF. Por ejemplo, la ciberdelincuencia está creciendo a un ritmo del 26% anual y se calcula que en 2019 le costará a la economía mundial al menos 2,1 billones de dólares, es decir, el 2% del PIB mundial (Symantec 2018).

En este espacio de inadecuados usos tecnológicos, las falsificaciones que se realizan reproducen violencia doméstica, (Dodge, et, al. 2019) generan procesos pornográficos falseando la identidad de las personas, (Delfino, Rebecca, 2019) lo cual, se ha llevado hasta el escenario audio visual del video, en lo que es posible construir y representar una realidad fraudulenta que ponga en duda a cualquier proceso social, en sus dimensiones legales, éticas, políticas, culturales, económicas, lo cual está generando la posibilidad de desacreditar la realidad misma en todas su expresiones, como resultado de este “cambio tecnológico perturbador”, como lo denominan Chesney, and Citron, (2018 a). Ello por la gran magnitud de su impacto para distorsionar hechos de cualquier índole está en condiciones de convertirse en un problema social. Claro que la falsificación en sí, no es un hecho nuevo, lo novedoso está en el rol de los medios y su alto nivel de “credibilidad” en la distorsión de los hechos. La capacidad de distorsionar acontecimientos sociales ha dado un salto exponencial con la TDF. Esta permite crear audio y vídeo de personas reales que dicen y hacen cosas que nunca han dicho o hecho (Blogger, 2018) y (Vincent, 2019).

Resulta conveniente plantear que el debate sobre este tema, nace en condiciones de un alto desarrollo de las fuerzas productivas ligadas al desarrollo de la IA y emerge en los Estados Unidos de Norteamérica, (Eichensehr, 2018). Sin embargo, la pregunta está: ¿La preocupación por el tema debería ser sólo de los países desarrollados? ¿Tales consecuencias de los usos maliciosos de la tecnología sólo afectan a este grupo de países?

Si partimos de la hipótesis, de que el nivel de conexión y de socialización de la información a través de los instrumentos de trabajo de internet es alto, la temática afecta a toda la comunidad mundial, y más si se calculara el tiempo que demora el traslado de una información de un lado a otro del planeta, que es casi instantáneo. El tratamiento del

asunto planteado es de particular importancia para la sociedad global en su conjunto, lo que exige involucrar a todos en la búsqueda de soluciones éticas frente a los procesos de socialización de tecnologías que mal empleadas se vuelven en contra de los objetivos sociales positivos.

Considerando los aspectos planteados el propósito de este artículo es incursionar en el debate sobre los usos maliciosos de la TDF, colocando la atención en que las soluciones educativas a los problemas que se puedan generar en estos ámbitos, es desde la transformación y puesta en práctica de la consolidación de valores que enarbolan la ética como principio básico que distingue la naturaleza humana en su expresión más civilizada. En este escenario de conflictos la preocupación es ¿Cómo contribuir a la transformación de esta realidad negativa que de continuar construyéndose podría generar conflictos sociales de magnitudes sorprendentes? “Las falsificaciones profundas también tienen ramificaciones políticas” (Beres, and Gilmer, 2018). ¿Cuáles podrían ser estos conflictos? ¿Tiene solución el mal manejo del “Deep Learning”? ¿Esta podría ser desde la tecnología misma? ¿Es la tecnología en sí misma la que genera el daño? Un cuestionamiento de base en este sentido, sería realizar la observación acerca de lo que plantea la literatura sobre el tema, en relación a la conceptualización de la TDF y sus consecuencias políticas, éticas y sociales.

El efecto es que el nuevo escenario que genera el uso de la TDF en su versión oscura “representa un punto de inflexión en la guerra de la información” (Dack, 2019), lo cual se fundamenta en que por un lado se podría aumentar el alcance de las noticias falsas con la velocidad de internet y sus mecanismos de funcionamientos sociales y por otro lado, reducen la comprensión compartida de los hechos, sobre todo cuando se le otorga absoluta credibilidad a los medios de información sin previamente haberlos comprobado sistemáticamente.

## **Materiales y métodos**

Se aplicaron valoraciones desde el paradigma cualitativo, a partir de investigaciones realizadas fundamentalmente en EE.UU. que es donde emerge el debate de la TDF y donde más se ha avanzado en los análisis del tema. Fue empleado el enfoque histórico hermenéutico de tipo documental de análisis de contenido, debido a que la información pertinente estaba contenida en artículos e informes de investigación que requería ser examinada. Para el análisis de contenido, se aplicó el método inductivo y deductivo. Ello permitió identificar las bases conceptuales en las que se apoya la presente investigación, comparando las preocupaciones y soluciones de los autores más relevantes, y a partir

de esto, se logró construir las proposiciones fundamentales que sustentan el artículo, permitiendo una mejor comprensión del fenómeno objeto de estudio.

## Resultados y discusión

### Socialización y contenido de la TDF

La **TDF**, cuando se creó sólo era manejada por académicos e investigadores, actualmente está al alcance de cualquier usuario con conocimientos básicos de informática (Shruti, et al, 2019). Emerge inicialmente como una herramienta para crear pornografía falsa, con las caras de celebridades femeninas mapeadas en los cuerpos de las estrellas porno, representando actos sexuales que de ningún modo se ejecutaron (Chawla, 2019). Su socialización se remonta al mes Diciembre del año 2017, cuando un usuario de Reddit<sup>1</sup> contribuyó a popularizarla (Merrefield, 2019). El mismo se hacía llamar "DeepFake", es decir, debe su nombre a este usuario. La palabra "deepfake" se compone de los términos "Deep Learning" y "Fake". El concepto es producir contenidos multimedia, generalmente videos, que presenten a las personas, en un hecho en el que en realidad nunca han participado (Schwartz, 2018).

La tecnología emplea un método de IA llamado Deep Learning o Aprendizaje Profundo, que consiste en reconocer e intercambiar caras en imágenes y videos, el procedimiento es que se comienza examinando un gran número de fotos o un video de la cara de alguien, adiestrando un algoritmo que manipula la información, y luego se emplea ese algoritmo para mapear la cara de una persona en un video. Dicho algoritmo está entrenado para reconocer patrones en grabaciones reales de audio o visuales de una persona en particular. La TDF, calcula cómo varios puntos de un rostro humano interactúan en la cámara para fabricar de forma convincente un ser humano en movimiento, lo que incluye la capacidad de comunicar verbalmente. Cuanto mayor sea la biblioteca de contenido con la que se alimenta un algoritmo de aprendizaje profundo, más realista puede ser la falsificación. Es una tecnología que es alimentada con una dieta de textos e imágenes faciales<sup>2</sup> para producir un video (Solsman, 2019) y (Floridi, 2018). Permite a los usuarios reproducir la transcripción de texto de un vídeo para añadir, borrar o cambiar los mensajes que salen de la boca de alguien (Vincent, 2019).

---

<sup>1</sup> Es una plataforma social de contenido determinado por la comunidad, en la que los usuarios envían publicaciones que otros usuarios pueden votar -a favor o en contra- según sus preferencias (Boyd, 2018).

<sup>2</sup> La falsificación producida del expresidente Obama supuestamente ofendiendo a Donal Trump, requirió 56 horas de grabaciones de muestra. cuantos más datos de vídeo o audio se introduzcan en la red neuronal, más precisa será la nueva información, falsa, de audio o de vídeo (Howcroft, 2018) y (Shruti Agarwal, et al, 2019).

La creación de falsificaciones profundas se fundamenta en dos avances del aprendizaje automático: las redes neuronales y las Redes Adversariales Generativas (GAN).<sup>1</sup> Las redes neuronales reflejan cómo funciona el cerebro humano; las GAN, trabajan enfrentando dos modelos informáticos de IA entre sí. Una red aprende a identificar los patrones en un clip de medios digitales, como el rostro de un político, y la segunda red sirve como espectador para averiguar si una imagen o un clip de vídeo es real o no. La segunda red da retroalimentación, y la primera red la usa para mejorar la credibilidad del video de Deepfake.

Imagínese a un falsificador tratando de crear una moneda falsa que parezca real, en un escenario en el que también hay policías que intentan detectar la moneda falsa, es decir, tienen el rol de discriminar la falsedad. El objetivo es falsear la realidad lo más objetivamente posible, para producir falsificaciones creíbles. Un modelo de computadora actúa como el falsificador y trata de crear una cara artificial basada en imágenes; mientras el otro modelo actúa como policía y compara las producciones artificiales con las imágenes reales e identifica los lugares donde éstas divergen. El resultado es la creación de modelos que van y vienen muchas veces, hasta que la imagen artificial es prácticamente idéntica a la original (Merrefield, 2019).

El resultado es que cuanto más se expone el cerebro humano a ejemplos de algún acontecimiento, mayor es la posibilidad de su reproducción de manera práctica y precisa. Las redes neuronales utilizan este mismo concepto; cuantos más ejemplos se introduzcan en la red, más exactamente podrá crear un nuevo resultado enriquecido con la imagen anterior (Dack, 2019). Estos videos a veces pueden ser identificados como falsos por el ojo humano; sin embargo, a medida que las redes neuronales se entrenan rigurosamente con más recursos digitales, será más difícil identificar su distorsión (Thaware, 2018).

La lección, es que la materia prima fundamental que se requiere para crear una distorsión de la información que se proyecta, es una compilación de cientos de reproducciones de acciones, o de imágenes de un rostro, para poder intercambiarlos de forma convincente. Esa es una de las razones por las que numerosas figuras públicas son objetivos tan fáciles de reproducir en videos falsos (Solsman, 2019). En Internet existe abundante información de fotos, videos, textos, audios originales para construir las reservas de datos para desarrollar falsificaciones.

---

<sup>1</sup> Las GAN fueron creadas por un estudiante de postgrado, llamado Ian Goodfellow, presentándose a la comunidad en el 2014, como una forma de generar algorítmicamente nuevos tipos de datos, a partir, de los conjuntos de datos existentes. Por ejemplo, una GAN puede mirar miles de fotos y luego producir una nueva foto que se aproxime a esa foto sin ser una copia exacta de ninguna de ellas, como si se tratara de un retrato completamente nuevo que aún no se ha tomado.

El efecto es, que esta tecnología manipulada con intenciones maliciosas tiene la potencialidad de situar a cualquier persona en un escenario que nunca sucedió, o le muestra expresando algo que jamás dijo; por otro lado, puede ser implementada para consumir abusos por parte una pareja íntima, haciendo que ésta aparezca en un video

Las GAN también pueden utilizarse para generar nuevo audio a partir de audio existente, o nuevo texto a partir de texto existente; se trata de una tecnología multiuso (Schwartz, 2018).

pornográfico en el que no estuvo (Chivers, 2019). Con ello indudablemente se genera una atmósfera de contrariedad que podría arruinar la vida de cualquier persona.

### **Preocupaciones sobre el uso indebido de esta tecnología**

Howcroft, (2018), denomina a las creaciones y avances asociados al “Deep Fake” carrera de armamentos tecnológicos, por su nivel de peligrosidad para falsear la realidad. Una preocupación es que universidades y empresas de investigación la están desarrollando con diversos fines: “Investigadores de la Universidad Carnegie Mellon, en EE.UU., crearon un sistema que puede transferir características, tales como expresiones faciales, de un video de una persona a una imagen sintetizada de otra. En China la empresa Baidu, y otras de nueva creación, entre ellas Lyrebird e iSpeech, han estado vendiendo clonación de voz para uso comercial en interfaces hombre-máquina” (Howcroft, 2018).

Es inquietante el hecho de que al igual que otras formas de desinformación, las falsificaciones profundas pueden ser diseñadas para incitar al pánico, sembrar la desconfianza en las instituciones políticas, o producir otros resultados dañinos. Debido a estos daños potenciales a los legisladores de Estados Unidos, les preocupa que se puedan emplear videos falsos para dañar la seguridad nacional. Según la aristocracia norteamericana, a medida que la tecnología para crear falsificaciones avanza y se hace más accesible, podría representar una amenaza para el discurso público y la seguridad nacional de su país, con implicaciones amplias y preocupantes para las campañas de medidas activas ofensivas dirigidas a los Estados Unidos (Westling, 2019).

La intranquilidad en el contexto norteamericano, es que con esta tecnología se pueda destruir indebidamente la reputación de cualquier persona e incluso provocar disturbios. Se maneja como argumento hipotético, “la construcción de un video donde se muestra a un candidato presidencial norteamericano acosando a niños, a un jefe de policía incitando a la violencia contra un grupo minoritario, o a soldados cometiendo crímenes de guerra”. Por ello, se llega a

la conclusión de que las personas de “alto perfil, como los políticos y los líderes empresariales, corren un riesgo potencial, dado el número de grabaciones de las que son de dominio público” (Howcroft, 2018).

La percepción es que el entorno mediático colectivo en el que se inserta la sociedad se ha fracturado y es cada vez más frágil, ya que el contenido de cualquier información falsa puede propagarse como un reguero de pólvora. Un estudio del Instituto Tecnológico de Massachusetts, descubrió que mientras que los robots propagan artículos falsos y artículos verdaderos al mismo ritmo, los artículos falsos alcanzan 1.500 visitas casi seis veces más rápido que los verdaderos (Dack, 2019).

El tema planteado trasciende como preocupación, -también en el contexto norteamericano- cuando se plantean otros efectos que se podrían crear con el inadecuado uso de la TDF, por el hecho de que la difusión de la desinformación es algo que está ocurriendo con creciente regularidad a medida que la gente elige las noticias que se ajustan a su visión del mundo. Por ejemplo, Thaware, (2018) cita situaciones que se pudieran manipular, las cuales pudieran ser válidas para cualquier país sobre todo cuando predomina el caos de la desinformación y fallan los mecanismos de persuasión de la existencia de un problema como la manipulación tecnológica: a) Con el uso malicioso de la tecnología se pueden provocar tensiones no sólo dentro del país, sino que también puede tener efectos dominó en todo el mundo. b) Políticos, expertos, medios de comunicación, agencias de noticias independientes, blogueros y el público en general se han acusado ya unos a otros por las llamadas noticias falsas.

La alerta es que las falsificaciones profundas están progresando rápidamente, y esto puede llevar a problemas de confianza. Incluso el contenido de vídeo auténtico en línea puede suscitar sospechas y preguntas sobre su validez. El resultado está siendo que las falsificaciones que circulan en los medios de comunicación social podrían alimentar campañas de desinformación, lo que podría conducir a una severa violencia y a disturbios sociales.

### **Cañones silenciosos de la desinformación**

Chesney, R. and Citron, D. (2018), en su informe sobre el “Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security”, alertan sobre los aspectos tratados anteriormente, con respecto a cómo las falsificaciones profundas podrían producir soluciones destructivas a la privacidad, la seguridad nacional y la democracia de un país, a ello les denominamos en este artículo, “cañones silenciosos de la desinformación”. A continuación, se podrían estar observando las situaciones hipotéticas que plantean estos autores –partiendo del contexto de EE.UU.- sobre cómo se podrían emplear las falsificaciones de la TDF: (Chesney, and Citron, 2018:10-20) **a) Daños a individuos y organizaciones**

En relación a los daños a individuos y organizaciones plantean que las imágenes que representan el abuso o la violencia, también se pueden usar para amenazar, intimidar e infligir daño psicológico a la víctima representada (o a aquellos que cuidan a esa persona). Las falsificaciones profundas también pueden ser usadas para representar a alguien, falsamente, como avalando un producto, servicio, idea o político. Además de infligir daño psicológico directo a las víctimas, la tecnología falsa puede ser utilizada para sabotear de la reputación de una persona en todos los escenarios de la cotidianidad, lugar de trabajo, romance, deportes, mercado y política. **b) Daño a la sociedad**

La amenaza que representan las falsificaciones profundas tiene dimensiones sistémicas, estas no son sólo una amenaza para individuos o entidades específicas. Tienen la capacidad de dañar a la sociedad de diversas maneras. Considere las siguientes posibilidades:

- El daño puede extenderse, a la distorsión del discurso sobre cuestiones de política; la manipulación de las elecciones; la erosión de la confianza en instituciones públicas y privadas importantes; la mejora y explotación de las divisiones sociales; el daño a operaciones o capacidades militares o de inteligencia específicas; las amenazas a la economía; y el daño a las relaciones internacionales.
- Pueden mostrar a funcionarios públicos que aceptan sobornos, que muestran racismo o que se involucran en adulterio.
- Los políticos y otros funcionarios del gobierno podían aparecer en lugares donde no estaban, diciendo o haciendo cosas horribles que no hacían.
- Se podría mostrar a los soldados asesinando a civiles inocentes en una zona de guerra, precipitando olas de violencia e incluso daños estratégicos a un esfuerzo bélico.
- Una falsificación profunda podría representar falsamente a un oficial de policía blanco disparando a un hombre negro desarmado.
- Un video falso podría retratar a un funcionario haciendo o diciendo algo exagerado como para causar disturbios en los países vecinos, lo que podría perturbar los lazos diplomáticos o desencadenar una ola de violencia.
- El audio falso podría representar de manera convincente a funcionarios que en privado "admiten" un plan para cometer un ultraje en el extranjero.

- Un video falso podría mostrar a los oficiales de emergencia "anunciando" un inminente ataque con misiles o una pandemia emergente, provocando pánico.
- Socavar la seguridad pública: Las falsificaciones profundas aumentan las posibilidades de que alguien pueda inducir un pánico público. En el mejor de los casos, el pánico público real podría simplemente acarrear daños y molestias económicas. En el peor de los casos, podría implicar la destrucción de la propiedad, lesiones personales y/o muerte.

Estas hipótesis podrían parecer exageradas, pero la realidad está indicando la necesidad de asumir un enfoque cada vez más proactivo, en la búsqueda de soluciones frente al pensamiento reactivo que está delineando usos maliciosos, a partir del alto grado de desarrollo alcanzado por la IA.

### **Valoración crítica a las soluciones a que se proponen para enfrentar el uso dañino de TDF**

Este acontecimiento tecnológico ha provocado la propuesta de múltiples soluciones por sus diversas implicaciones en el campo de la información y en las diversas facetas de la sociedad (Simpson, 2018). En este sentido, se evidencia en el debate una postura ética que expresa su preocupación por el tema, lo que hace que el asunto de las falsificaciones profundas sea, como proyectan los autores Chesney y Citron (2018 a), un "desafío inminente en busca de soluciones".

Existen diversas posturas en las propuestas de soluciones potenciales, que están siendo examinadas por compañías, investigadores y políticos, con la finalidad de generar acciones dirigidas a evitar las consecuencias negativas de las falsificaciones profundas. Una primera postura que se evidencia, y que está siendo predominante, es la intencionalidad de construir y aplicar algoritmos creativos dirigidos a la detección de falsedades profundas, lo cual en diversos contextos choca con el asunto de la necesidad de confidencialidad de la información, con la interrogante de ¿Cómo manejar estos datos sin que se filtren?, incluso, a ello se antepone el componente legal, el cual, en naciones como EE.UU. se ha quedado rezagado en relación al desarrollo tecnológico (Dack, 2019).

Una segunda postura, atrincherada en la ideología del "solucionismo tecnológico" (Morozov, 2015) tiene como contenido la creencia de que con la tecnología en si misma se resolverán todos los males, conllevando a la cosificación de las relaciones sociales a través de la tecnología, es decir, se tiende a enfrentar el problema de la falsificación del entorno social, con la generación de algoritmos que detecten tales situaciones, en lo que queda en un lugar secundario la acción del hombre (Thaware, 2018).

Para el caso de las soluciones algorítmicas, el seguimiento de los datos y el uso del concepto de registro de vida planteados como soluciones tecnológicas, el reto que se plantea es cómo garantizar la seguridad y la privacidad de los

datos que van a permitir la emergencia de soluciones. Por ello hasta el momento, según Dack, (2019:1-2), “la tecnología no parece ser la solución a las falsificaciones profundas”. El asunto está en determinar quién creó la falsificación profunda, e incluso si se sabe, puede estar fuera de la jurisdicción del país donde se generó.

Una tercera postura, considera la importancia de que prevalezcan soluciones políticas desde lo legal, y lo gubernamental, valorándose desde una posición violenta, la posibilidad de acciones encubiertas y agresivas de un país a otro, cuando el acontecimiento de falsificación haya sido generado desde el exterior (Eichensehr, 2018). Tal solución es cuestionable por el grado de injerencia política que lleva de un país a otro y por las consecuencias sociales que de ella se derivan.

Desde un enfoque apocalíptico, una cuarta postura hace alusión a que está naciendo una nueva era con este impacto tecnológico con consecuencias negativas, la era de la post-verdad (Palmer, 2018), Chesney, 2019) y Citron, 2019). Vea el tipo de preguntas cuestionadoras de la verdad que se realiza: ¿Qué pasará cuando armonicemos la realidad alternativa? ¿Redefinirá lo que significa ser una superpotencia? ¿Se utilizará para manipular el mercado de valores? ¿Guerra psicológica a escala global? En un mundo donde no podemos diferenciar entre la realidad y la fantasía, ¿qué es la verdad? En la era posterior a la verdad, ¿a quién crearemos?

Ante el enfoque anterior se podría advertir: Tampoco es la totalidad de la sociedad en su conjunto la que está participando en generar problemas sociales asociados al nuevo desarrollo tecnológico, no hay tal caos. Con este enfoque de la era de la post-verdad, se está etiquetando el nuevo desarrollo que involucra un fetichismo tecnológico que elude toda posibilidad de resolver los problemas generados del mal uso de la tecnología con la acción humana.

Chivers, Ch. (2019), plantea: “muchos de nuestros temores sobre la tecnología están exagerados”, por consiguiente, es desmedido el planteamiento acerca de que con los contenidos falsos que se generan desde las falsificaciones tecnológicas, se está llegando a la era de la post-verdad, reconocer este tema como verdadero, es un paso significativo para generar incertidumbre social.

### **Propuesta de soluciones educativas para un desarrollo sostenible**

En este escenario de conflictos y soluciones generados por el mal uso de la tecnología, está quedando en el “vacío” el rol que podría desempeñar el factor psicológico en el enfrentamiento a los problemas que plantea el nuevo reto. Todo parece indicar, que la tecnología es autónoma en sí misma, y que se le está saliendo de las manos a la acción humana. Ante esta situación, es conveniente el planteamiento de varias acotaciones:

Primero, no se elude el rol que podría generar el desarrollo tecnológico en generar y detectar las falsificaciones sociales, pero la solución sería de efecto y no de causa, hay que ir a la raíz del problema, que son los sesgos negativos de los programadores los que hacen actuar a la tecnología., ello exige transformar el componente humano desde los procesos de formación de valores.

Segundo, las soluciones políticas asociadas a agresiones sobre posibles países desde donde se generen falsificaciones empleando la tecnología, serían un fracaso de la ética con comportamientos de barbarie, que sólo harían retroceder las acciones sociales. Al igual que las soluciones basadas en la tecnología, también las soluciones de política y las soluciones jurídicas se valoran como difíciles de conseguir. El argumento es que, para países como EEUU, en el contexto actual de sus leyes, crear y compartir datos generados por la TDF, se clasifica como una acción ilegal en el marco de las leyes de difamación y fraude (Dack, 2019).

Tercero, las soluciones legales y políticas son eficientes en el enfrentamiento coercitivo al problema, estas deberían ir acompañadas de procesos de formación ética en todos los espacios sociales, cuyo contenido, sería desarrollar una cultura de transformación social, que elimine los sesgos negativos de los conductores de los procesos de algoritmización, ello es posible.

Cuarto, en las propuestas de soluciones no ha faltado el enfoque basado en el empleo de las "leyes del mercado", ello se fundamenta en que se anticipa dos reacciones ante las amenazas del Deep Fake. Por un lado, se espera que el sector privado desarrolle y venda servicios destinados a proteger a los clientes de al menos algunas formas de daños basados en la falsificación, ya han aparecido una serie de servicios que responden a las inquietudes de los clientes sobre el robo de identidad. En segundo lugar, se espera que las compañías de medios sociales, por iniciativa propia, tomen medidas para protegerse contra daños falsos en sus plataformas (Chesney, and Citron, 2018).

Quinto, un hecho aislado como el caso de Deep Fake, que incluso se conoce de sus peligros para la sociedad, no debería llevar a etiquetar y a tergiversar la realidad con la afirmación de que se está llegando a la era de la postverdad. Tampoco es la totalidad de la sociedad en su conjunto la que está participando en generar subterfugios sociales empleando la TDF, son individualidades con sesgos negativos y objetivos mal intencionados quienes están participando en la distorsión del escenario social. Por ello, es importante evitar la cultura débil, donde predomine la acción social de grupos aislados, se impone una transformación colectiva donde prevalezca una cultura fuerte guiada, por la ética, los valores, las normas, los símbolos, la identidad cultural que ha hecho avanzar socialmente la humanidad, en lo que deben desempeñar un rol importante, las bases políticas y sociales de cualquier sociedad.

La alerta está, en que TDF está en el entorno social al alcance de todos en Internet, y es de fácil uso para cualquier usuario promedio. Las herramientas de software como FakeApp, están específicamente diseñadas para permitir que personas sin experiencia de programación puedan crear falsificaciones profundas que, empleando la característica de los medios digitales, referida a la facilidad de copiar y almacenar datos en línea hace mucho más difícil eliminar las falsificaciones una vez que se publican y se comparten (Thaware, 2018) Por consiguiente, se impone ante los problemas sociales que podrían afectar la sociedad en su conjunto, conducir transformaciones que emanen e involucren la conciencia social para enfrentar tales usos maliciosos de la tecnología, y más si se tiene en cuenta el planteamiento siguiente: “la guerra contra las falsificaciones profundas” se convertirá en una carrera armamentista evolutiva. Surgirán y se eludirán las herramientas de detección” (Dicho por Kyle Findlay y citado por [Joseph](#), 2019:3).

Un sexto aspecto de esta presentación sugiere, que para detectar si un video es resultado de falsificaciones, sería importante apelar a las herramientas de la inteligencia emocional, particularmente al componente referido a la aplicación de la empatía y de las habilidades sociales, para lo cual se podrían observar varios pasos para detectar una falsificación profunda. Ello podría ser operativo para cualquier usuario promedio, el criterio básico consiste en la importancia de la comunicación y contar con información verificable antes de darle credibilidad a un hecho que resulte inmoral para cualquier persona, institución o de la sociedad en su conjunto.

La explicación lógica está en que si se parte del hecho de que las personas que construyen los argumentos falsos empleando la TDF, cree que el contenido creado inducirá al espectador a confiar en ella - y a compartirla con otras personas en una red social - acelerando así la propagación de la desinformación, entonces la solución a los problemas que se pudieran ocasionar podría estar en procedimientos de asumir comportamientos diferentes asociados al manejo adecuado de la información y la comunicación (Westling, 2019).

Chesney R. y Citron, D. (2019 b). en su obra titulada “DeepFake and the New Disinformation War. The Coming Age of Post-Truth Geopolitics”, partidarios de la era de la “crisis de la verdad” plantean:

“Una imagen puede valer más que mil palabras, pero no hay nada que convenza tanto como una grabación de audio o vídeo de un evento. En un momento en que los partidarios apenas pueden ponerse de acuerdo sobre los hechos, tal persuasión podría parecer como si pudiera aportar una clara bienvenida. Las grabaciones de audio y vídeo permiten a las personas convertirse en testigos de primera mano de un evento, ahorrándoles la necesidad de decidir si confiar o no en el relato de otra persona. Y gracias a los teléfonos inteligentes, que facilitan la captura de contenido de audio y vídeo, y a las plataformas de medios sociales,

que permiten compartir y consumir ese contenido, la gente hoy en día puede confiar en sus propios ojos y oídos en un grado sin precedentes” Chesney R. y Citron, D. (2019 b: 2-3).

Relativamente, es real tal planteamiento hasta tanto no se tenía conocimiento del inconveniente de las falsificaciones con la nueva tecnología, sin embargo, con las nuevas alertas de lo que puede suceder con este acontecimiento, todo se podría comprobar, por ello es importante apelar a los componentes socio-psicológicos para comprobar la realidad de los hechos.

El problema está, en que el contenido que se genera para falsear la realidad depende de factores psicológicos que cualquier medio de engaño podría emplear. Esta afirmación sugiere poner en práctica el componente sociopsicológico y cultural para detectar el nivel de autenticidad de la información, con lo que se pudiera refutar el planteamiento de Chesney R. y Citron, D. (2019 b: 2-3). Para ello, en esta investigación se comparten pasos a considerar para detectar el nivel de credibilidad de un video que pudiera estar o no contaminado, estos son los siguientes: (Joseph, 2019)

1. Confíe en fuentes verificables: No deposite toda su confianza en actores desconocidos que publiquen videos. Por tanto, no comparta información comprometedor, sin antes haber verificado su autenticidad.
2. Efectúe una exploración para observar si alguien más está proporcionando información acerca del contenido del vídeo.
3. Investigue y justifique los hechos para encontrar evidencias de apoyo y fuentes que corroboren lo observado.
4. Sea empático y ejecute el sentido común: En el supuesto video del expresidente Barack Obama Versus Donal Trump cuestionese si realmente Obama presentaría un video en el que acusa a Donald Trump de ser un "imbécil". Y, si sucediera, ¿pasaría esto inadvertido para los medios de comunicación creíbles?
5. Compararla con el video que estás observando.
6. Investigue sobre las afirmaciones más indecorosas para corroborar si realmente fueron ejecutadas por la persona del vídeo.
7. Realice capturas de pantalla del vídeo y, a continuación, efectúe una investigación inversa de imágenes para identificar el vídeo o las imágenes originales.
8. Hágase la pregunta siguiente: ¿Tiene sentido la ubicación del vídeo?
9. Pregúntese. ¿Las artistas famosas de cine se involucrarían de repente en videos pornográficos de bajo presupuesto?

10. Un video sobre un hecho real, tiene grandes posibilidades de que los medios de comunicación creíbles lo hayan divulgado, y probablemente filmado.

11. Pregúntese: ¿Reportaron o publicaron el video? Investigue si existe alguna transcripción del discurso o diálogo.

Se puede concluir parcialmente, que la tecnología por sí misma, y aislada de los procesos sociales no está preparada para evitar y detectar las falsificaciones que se desarrollan desde los procesos tecnológicos, sería prudente apelar al desarrollo de programas de formación pública en todos los espacios sociales. Es decir, este tema no es sólo para los programadores de algoritmos, exige de involucrar a la sociedad en su conjunto.

Hay que reconocer que la revolución tecnológica que se está originando por los efectos de la IA, las máquinas inteligentes, el aprendizaje profundo, la formación del entorno cibernético que se está produciendo, el procesamiento automático de grandes magnitudes de datos, la implementación de recursos analíticos y de nube, por un lado, conduce a un desarrollo acelerado de la sociedad. Sin embargo, la contradicción está, en que estos adelantos tecnológicos, por su uso inadecuado están generando nuevas amenazas, riesgos sociales.

Por consiguiente, ello demanda de nuevas competencias profesionales, inversiones e innovaciones para prevenir los riesgos y proteger internet y los recursos de la red. Una de las maneras de proteger a los ciudadanos es cambiar el enfoque del uso de la tecnología. Las personas, no entienden el lado oscuro de las plataformas hasta que se involucran en acciones que llevan a pérdidas personales y financieras. La idea es que el consumidor de los nuevos usos tecnológicos, tiene más oportunidades de controlar Internet e influir en el estilo de vida en ese entorno (Marinov, 2019).

Clark, J. (2019), Director de Políticas de OpenAI, una compañía de investigación de IA en San Francisco, California, plantea una interesante solución al combate de los usos maliciosos de la TDF, al integrar soluciones tecnológicas, institucionales y políticas; en lo que es notorio el alcance que le otorga a la necesidad que el gobierno forme parte de las soluciones para tomar decisiones sobre la verdad preservando la “salud general de la sociedad”, en lo que asigna un lugar significativo a la “educación integral en IA”, involucrando esquemas educativos que integren a las familias para enfrentar los nuevos retos de la revolución tecnológica.

El efecto es que, se impone el impulso de la alfabetización en red y la alfabetización informativa desempeñando una función significativa en la prevención de la creación de sistemas u orientaciones conceptuales que conduzcan al control masivo de los comportamientos y hábitos de los ciudadanos.

Las lecturas del fenómeno examinado sobre la guerra de la desinformación, es prevenir para estar mejor preparados, no solo en el plano tecnológico y legal, sino, además, en el escenario de las ideas. En este contexto, es sumamente desde el enfoque político, tomar posiciones en relación a generar un sistema de acciones, para enfrentar los retos que pudiera imponer el uso mal intencionado de la “Tecnología Deep Fake. La acción del Estado y los gobiernos locales, es esencial, bajo la idea de pensar globalmente y actuar localmente.

En el plano ideológico, ello requiere incorporar la concepción de la formación y consolidación de valores patrios, frente a la posibilidad que tiene esta tecnología de tergiversar la historia, la cultura y la identidad de entornos locales, regionales y a nivel de país, y el blanco fundamental de esta formación debería ser la juventud actual, la cual representa la generación digital, cuyo aprendizaje es audiovisual. La familia, las instituciones educativas, gubernamentales y no gubernamentales, las empresas estatales y privadas, el gobierno, las organizaciones de masas y políticas deberán tener un rol fundamental en este enfrentamiento.

La práctica política deberá convertirse en el eje de las transformaciones ideológicas que indicarán el curso al enfrentamiento de los inadecuados usos de la tecnología, mediante la instrumentación de un marco legal y códigos éticos que orienten socialmente los caminos a transitar.

Una lectura, con una mirada diferente podría hacer pensar en: ¿Cómo emplear esta tecnología para beneficios sociales? De hecho, existen propuestas que de manera creativa pudieran emplearse con fines políticos, ideológicos y culturales en el escenario de las lecturas anteriores. Por ejemplo:

- Chawla, Ronit (2019) y Vincent, (2019) hacen propuestas de su uso en el cine. Sería de gran ayuda para las industrias cinematográfica y televisiva, ya que les permitiría corregir las líneas mal pronunciadas sin tener que volver a grabar las imágenes, y crear doblajes de actores que hablan diferentes idiomas, ello podría tener un efecto importante en abaratar los costos de la grabación de un programa (Floridi, 2018), sostiene que el empleo de esta tecnología podría permitir el rescate de obras de pintura de múltiples autores.
- Chesney, and Citron, (2018 a), sugieren aplicarlo con un enfoque educativo, para producir videos de figuras históricas que hablen directamente a los estudiantes, dando una nueva oportunidad de vida a una actividad de enseñanza de la historia que de otro modo sería poco atractiva. Los artistas pueden usar la TDF para satirizar, parodiar, criticar y proyectar imágenes diferentes en sus escenarios sociales. Estos beneficios artísticos potenciales se relacionan con los beneficios educativos.

- Howcroft, (2018), toma como positivo el ejemplo de la empresa escocesa CereProc, la cual empleando la TDF crea voces digitales para las personas que pierden las suyas a causa de enfermedades. La idea es que la clonación de voces podría tener un propósito educativo al recrear el sonido de figuras históricas. Un proyecto de la Universidad Estatal de Carolina del Norte sintetizó uno de los discursos no grabados de Martin Luther King Jr. CereProc creó una versión del último discurso escrito por el Presidente John F. Kennedy, quien fue asesinado antes de entregarlo.

Esta innovación tecnológica, que actualmente se emplea en múltiples espacios sociales, como herramienta para la desinformación online y offline, pudiera enriquecer la perspectiva pedagógica de la misma manera que lo es el advenimiento de las tecnologías de producción, y visualización de realidad virtual. Todo ello requiere poner en función la magia del talento y la creatividad, para transformar la práctica educativa de cualquier entorno social.

### **Discusión**

Es un paso de avance importante el nivel alcanzado por la ciencia de los datos, en lo relativo a la generación de algoritmos de aprendizaje que cada vez más se acercan al pensamiento humano lo impropio es su uso negativo, en ello coinciden la totalidad de los autores examinados, lo diferente es, la diversidad de soluciones a esos daños, lo cual indudablemente enriquece la discusión sobre el asunto planteado. El debate sobre el tema es expresión del desarrollo de la conciencia moral ante los nuevos dilemas generados por sesgos negativos que se incorporan a los procesos de algoritmización. Se observa como limitación, un enfoque lineal en los análisis de los autores examinados, no se aprecian análisis críticos de un autor a otro, en este sentido, esta investigación, representa un avance que se pudiera considerar notorio, incorporando cuestionamientos, como el de la crisis del mercado de las ideas y al pensamiento que aborda que se está poniendo en dudas el concepto de la verdad, también aporta un balance crítico y una propuesta de soluciones a los daños que pudiera causar el Deep Fake en su versión oscura.

### **Conclusiones**

Las falsificaciones profundas constituyen una evolución tecnológica, que expresa un cambio radical en la manera en que se produce, se divulga y maneja de manera inadecuada la información.

La socialización por parte de actores nacionales, de la manipulación de los medios de comunicación con fines maliciosos, están dirigidos a ocasionar la guerra de la información para contribuir a ambiciones geopolíticas. Ello podría llevar a que este hecho, mal maniobrado, se convierta en una amenaza cada vez más grave para individuos,

organizaciones y la sociedad en su conjunto, al formar parte de un desarrollo sin precedentes en el ecosistema de la desinformación.

La ciencia en general y en particular la Inteligencia Artificial, debe ponerse en función de proporcionar, conocimientos de los procesos tecnológicos distorsionados que podrían estar afectando el entorno social, y las herramientas para intervenirlos, con conocimiento de causa y efecto. Es importante, a cada desarrollo tecnológico emergente imprimirle un enfoque social, en cuanto a contabilizar los daños y perjuicios que pudieran tener en todas sus dimensiones, es decir, adelantarse a los hechos y que no sea un uso casual como el caso del Deep Fake, el que alerte los problemas que se pudieran generar con el inadecuado uso de la tecnología.

Existen variedad de preocupaciones en torno a la solución de los problemas que ha estado causando el uso malicioso del Deep Fake, en su mayoría, colocan la atención en el componente tecnológico, lo cual no se elude, sin embargo, la gravedad del problema requiere, no sólo soluciones tecnológicas, sino integrales, multifacéticas, lo que debe tener un enfoque antropocéntrico, que regule el comportamiento moral, asegurando coherencia social, en lo que debe prevalecer un sistema de valores, normas, principios e ideales que guíen el proceder de los individuos en la sociedad.

Es importante anotar que el debate que se produciendo, está sesgado por las características propias del país donde nace (EE.UU.), donde existen todo tipo de preocupaciones sobre los daños que podría causar la TDF, para los individuos, las instituciones y la sociedad en su conjunto. No obstante, la inquietud fundamental es sobre el efecto perjudicial que ello pudiera causarle, a políticos y funcionarios del Estado. En este entorno, en el debate de enfrentamiento al Deep Fake, se observa lentitud en el marco legal en relación al desarrollo tecnológico.

La divulgación del problema a nivel local, regional, a nivel de países y continentes debería ser una tarea permanente, ello podría ser una herramienta para la educación moral de la sociedad, en la medida en que se conozcan los hechos distorsionados que genera la desinformación. Preparar a la sociedad en el conocimiento de los riesgos y amenazas de la desinformación, equivale a formar parte de las soluciones.

## **Referencias**

Beres, D. and Gilmer, M. (2018). A guide to 'deepfakes,' the internet's latest moral crisis. Disponible en: <https://mashable.com/2018/02/02/what-are-deepfakes/>.

Blitz, Lies, Marc Jonathan (2018). Line Drawing, and (Deep) Fake News, 71Okl. L. Rev.59 (2018), <https://digitalcommons.law.ou.edu/olr/vol71/iss1/5>.

Blogger, Staff. (2018). Danielle Citron Provides an In-depth Assessment of the Causes and Consequences of “Deep Fake” Technology. Disponible en: <https://www.techpolicy.com/Blog/September-2018/Danielle-Citron-Provides-anIn-depth-Assessment-of.aspx>.

Chawla, Ronit. (2019). Deepfakes : How a pervert shook the world. International Journal of Advance Research and Development. Volume 4, Issue 6). Disponible en: <https://www.ijarnd.com/manuscripts/v4i6/V4i6-1143.pdf>.

Chesney R. y Citron, D. (2019 b). Deepfakes and the New Disinformation War. The Coming Age of Post-Truth Geopolitics. Disponible en: <https://www.foreignaffairs.com/articles/world/2018-12-11/deepfakes-and-newdisinformation-war>.

Chesney, R. and Citron, D. (2018 a). Keats, Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security (July 14, 2018). 107 California Law Review (2019, Forthcoming); U of Texas Law, Public Law Research Paper No. 692; U of Maryland Legal Studies Research Paper No. 2018-21. <https://ssrn.com/abstract=3213954> or <http://dx.doi.org/10.2139/ssrn.3213954>.

Chivers, Ch. (2019). What do we do about deepfake video?. <https://www.theguardian.com/technology/2019/jun/23/what-do-we-do-about-deepfake-video-ai-facebook>.

Clark, J. (2019). Written Testimony of Jack Clark Policy Director OpenAI. HEARING ON. “The National Security Challenges of Artificial Intelligence, Manipulated Media, and ‘Deep Fakes’ . BEFORE THE. House Permanent Select Committee on Intelligence. June 13th, 2019.

Dack, S. (2019). Deep Fakes, Fake News, and What Comes Next. Disponible en: <https://jsis.washington.edu/news/deep-fakes-fake-news-and-what-comes-next/>.

Delfino, Rebecca, (2019). Pornographic Deepfakes — Revenge Porn’s Next Tragic Act – The Case for Federal Criminalization (February 25, 2019). Fordham Law Review, Vol. 88 (Forthcoming, 2019); Loyola Law School, Los Angeles Legal Studies Research Paper No. 2019-08. <http://dx.doi.org/10.2139/ssrn.3341593>.

Dodge, et, al. (2019). Laura’s House & Erica Johnstone, Ridder, Costa. Using FakeVideo Technology To Perpetuate

Intimate. Partner Abuse Domestic Violence Advisory.

[https://www.cpedv.org/sites/main/files/webform/deepfake\\_domestic\\_violence\\_advisory.pdf](https://www.cpedv.org/sites/main/files/webform/deepfake_domestic_violence_advisory.pdf).

Eichensehr, K. (2018). Don't Believe It If You See It: Deep Fakes and Distrust. Technology Law Jotwell. The Journal of Things We Like (Lots). <https://cyber.jotwell.com/dont-believe-it-if-you-see-it-deep-fakes-and-distrust/>.

Floridi, L. (2018). Artificial Intelligence, Deepfakes and a Future of Ectypes. Philosophy & Technology. September 2018, Volume 31, Issue 3. <https://doi.org/10.1007/s13347-018-0325-3>.

González Arencibia, & Martínez Cardero. (Enero-Junio 2020). Impacto económico de la tecnología Deep Learning. Revista Bimestre Cubana. Sociedad Económica Amigos del país. Volumen CXXVII enero-junio, 2020 época III, no. 52 la Habana, Cuba ISSN 1012-9561. P. 6-87. <https://www.seapcuba.cult.cu/wp-content/uploads/2020/10/RBC-52.pdf>.

Howcroft E. (2018). How Faking Videos Became Easy and Why That's So Scary. <https://www.bloomberg.com/businessweek>.

Joseph, R. (2019). Fakebusters strike back: How to spot deep fakes, the manipulated videos that are the newest form of "fake news" to hit the internet. <https://journals.sagepub.com/doi/full/10.1177/0306422019841326>.

Marinov, R. (2019). The new intelligent technologies- the problems and opportunities. IJRDO - Journal of Social Science and Humanities Research (ISSN: 2456-2971), 4(7), 13-23. <https://www.ijrdo.org/index.php/sshr/article/view/3008>.

Merrefield, C. (2019). Deepfake technology is changing fast — use these 5 resources to keep up. <https://journalistsresource.org/studies/society/deepfake-technology-5-resources/#annotations:17478215>.

Morozov, Evgeny. (2015). La locura del solucionismo tecnológico. Madrid; Móstoles, Madrid; Buenos Aires: Clave Intelectual ; Katz.

Palmer, S. (2018). Deepfakes: The Dawn of the Post-Truth Era. <https://www.shellypalmer.com/2018/10/deepfakesdawn-post-truth-era/>.

Schwartz, O. (2018). You thought fake news was bad? Deep fakes are where truth goes to die. The Guardián. <https://www.theguardian.com/technology/2018/nov/12/deep-fakes-fake-news-truth>.

Shruti Agarwal, et al, (2019). Hany Farid, Yuming Gu, Mingming He, Koki Nagano, Hao Li. Protecting World Leaders Against Deep Fakes. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2019, pp. 38-45. These CVPR 2019 workshop papers are the Open Access versions, provided by the Computer Vision Foundation. [http://openaccess.thecvf.com/content\\_CVPRW\\_2019/html/Media\\_Forensics/Agarwal\\_Protecting\\_World\\_Leaders\\_Against\\_Deep\\_Fakes\\_CVPRW\\_2019\\_paper.html](http://openaccess.thecvf.com/content_CVPRW_2019/html/Media_Forensics/Agarwal_Protecting_World_Leaders_Against_Deep_Fakes_CVPRW_2019_paper.html).

Simpson, G. (2018). The societal impact of AI. Experiencing Digital. <https://www.cio.com/article/3273565/thesocietal-impact-of-ai.html>.

Solsman, J. (2019). Deepfakes may ruin the world. And they can come for you, too. <https://www.cnet.com/news/deepfakes-may-try-to-ruin-the-world-but-they-can-come-for-you-too/>

Symantec. 2018. Internet Security Threat Report. Volume 23, March. [www.symantec.com/content/dam/symantec/docs/reports/istr-23-2018-en.pdf](http://www.symantec.com/content/dam/symantec/docs/reports/istr-23-2018-en.pdf).

Thaware, N. (2018). AI Gone Rogue: Exterminating Deep Fakes Before They Cause Menace. Blackhat.com/eu-18/Thu-Dec-6/eu-18 . <https://i.blackhat.com/eu-18/Thu-Dec-6/eu-18-Thaware-Agnihotri-AI-Gone-RogueExterminating-Deep-Fakes-Before-They-Cause-Menace-wp.pdf>.

Vincent, J. (2019). AI deepfakes are now as simple as typing whatever you want your subject to say. Disponible en: <https://www.theverge.com/2019/6/10/18659432/deepfake-ai-fakes-tech-edit-video-by-typing-new-words>.

Westling, J. (2019). Deception & Trust: A Deep Look At Deep Fakes. <https://www.techdirt.com/articles/20190215/10563541601/deception-trust-deep-look-deep-fakes.shtml>.

González Arencibia, M. Martínez Cardero, D. (2019). Deep learning: its social implications. Vol 4 No 7 (2019): IJRDO - Journal of Social Science and Humanities Research | ISSN: 2456-2971. **All Rights Reserved. La India.** <https://www.ijrdo.org/index.php/sshr/issue/view/532>.

Boyd, J. (2018) ¿Qué es Reddit? Brandwatch.: <https://www.brandwatch.com/es/blog/que-es-reddit/>.