

AI & serverless systems in war disinformation via LUTs: impact and ethics

Abstract:

The present research consists of the creation and reflection on the role of Generative Artificial Intelligence and its impact on the informational and propagandistic world, with the aim of understanding and mitigating the impact of misinformation in the digital age. The research will be supported by an experiment in which a Generative AI will produce war-themed images from scratch, applying (Look Up Tables) LUTs to these images to increase their realism to a point where they can be perceived as authentic, which could contribute to disinformation campaigns.

The development of technological tools capable of identifying and countering these threats is crucial in a context where the integrity and veracity of information are fundamental elements for security and social stability. In this sense, the importance of this research lies not only in the creation of a generative AI model, but also in its capacity to contribute to the ethical, political and social debate on information manipulation.

Keywords:

Inteligencia Artificial, Propaganda, Fake News, desinformación, redes sociales

Cómo citar este documento:

GARCÍA HUETE, Estíbaliz. et al. *IA generativa y sistemas serverless en desinformación bélica con LUTs: impacto y ética*. Documento de Opinión IEEE 45/2025. [enlace web IEEE](#) y/o [enlace bie³](#) (consultado día/mes/año)

Introducción

Nada de lo que ocurra en la historia se entiende sin una base propagandística y de desinformación. En los contextos bélicos, esta base se engrosa y se mueve de forma mucho más rápida y virulenta cuando entran nuevos medios para calar entre la población y apelar a los sentimientos.

Esto se ha acrecentado con la llegada de la Inteligencia Artificial Generativa, que en menos de dos años ha sido entrenada hasta tal punto que es capaz de hacer desde cero fotos y videos con un grado de realismo lo suficientemente preocupante como para que suponga una amenaza para la seguridad informativa y la opinión pública, provocando así la sensación para los profesionales de la información de «ir a contrarreloj» al intentar analizar la procedencia de dichas creaciones y contrastar su veracidad.

En redes sociales como X (anteriormente llamada twitter) o Instagram se facilita la viralización de información, que puede ser falsa, a través de opciones que te proporcionan dichas redes como retwittear o compartir, respectivamente. En TikTok, una de las redes sociales actuales más jóvenes, se comparten videos en muchas ocasiones sin verificar ni la procedencia, ni el contexto ni la autenticidad.

La Inteligencia artificial (IA) Generativa ha adquirido una relevancia creciente debido a su uso y propagación en diversas plataformas digitales, incluyendo medios de comunicación y redes sociales.

La desconfianza de los medios de comunicación tradicionales, la deriva de los discursos hegemónicos y el auge de la Inteligencia Artificial nos empujan a una vorágine llena de incertidumbre y miedo que conlleva una reflexión sobre su uso en el contexto actual, y cómo y a qué ritmo puede evolucionar su impacto en la sociedad en tiempos venideros.

Debido a su particular relevancia en los escenarios bélicos actuales, donde la percepción pública puede ser manipulada para influir en decisiones políticas o en la moral colectiva de una sociedad, hemos creído necesario ahondar en el papel de la inteligencia artificial en conflictos bélicos y su influencia propagandística a nivel mediático y político

Las Fuerzas Armadas han reconocido en los últimos años el dominio cognitivo como un factor determinante. El motivo es que este dominio de las percepciones, emociones, pensamientos y creencias, trasciende los tradicionales (terrestre, marítimo, aeroespacial y ciberespacial). En el ámbito militar, este concepto exige refinar la precisión operativa, puesto que no depende únicamente de las capacidades tecnológicas o cinéticas, sino también de la influencia ejercida sobre la psique humana. Por este motivo, tanto en sentido metafórico como literal, el dominio cognitivo constituye un nuevo frente de batalla, donde la percepción, las narrativas, la propaganda y la información se convierten en armas tan eficaces como los medios convencionales¹.

Hemos medido su impacto gracias a un experimento que consta de una IA generativa que es capaz de crear mediante patrones y algoritmos imágenes bélicas. La computación de esas imágenes se hace de forma distribuida con arquitecturas *serverless* para mejorar la eficiencia y reducir costes. Se han aplicado LUT para aumentar su realismo percibido. Dichas imágenes han sido evaluadas por el ojo humano para medir su capacidad de persuasión, añadiendo herramientas automatizadas de detección para evaluar la capacidad de los sistemas actuales de identificar la generación (es decir, la no autenticidad) de esas imágenes. Este enfoque ayudará a evaluar de forma efectiva el potencial de la IA Generativa en la aplicación de manipular a la población, y a contribuir al debate sobre las implicaciones éticas de la IA generativa en el campo de la desinformación.

Este trabajo es resultado de la colaboración de grupos de investigación pertenecientes a tres centros de la Universidad Complutense de Madrid: la Facultad de Ciencias de la Información, la Facultad de Óptica y Optometría y la Facultad de Informática. Cada uno de estos grupos aporta un enfoque multidisciplinar que enriquece el desarrollo del trabajo aquí expuesto. El de la Facultad de Ciencias de la Información se centra en los aspectos comunicativos y sociales de la desinformación; el de la Facultad de Óptica y Optometría

¹ PECO YUSTE, Miguel, "Implicaciones del ámbito cognitivo en las Operaciones Militares" Instituto Español de Estudios Estratégicos, pp. 109-137, 2020.

aporta su conocimiento en la percepción visual y los efectos ópticos, que pueden ser simulados y analizados en las imágenes generadas, junto con el color; y el de la Facultad de Informática se encarga de la implementación técnica del modelo de IA, así como de la infraestructura de computación distribuida necesaria para llevar a cabo el experimento a gran escala.

¿Desde dónde partimos?

La mayoría de las redes sociales de la actualidad nacieron al comienzo del nuevo milenio, en la primera década de los años dos mil. En estas se pueden encontrar desde cuentas de medios de comunicación hasta usuarios con cierta reputación, los cuales deben su popularidad a su alto nivel de interacción y a su elevado número de seguidores, y que en algunos casos, pueden ser contactados para la promoción de noticias falsas. Esta promoción de noticias falsas o bulos se suelen utilizar para una defensa férrea y ciega de determinadas ideas o posturas. Debido a los algoritmos que rigen dichas redes sociales, se favorece la aparición en el timeline del usuario de las llamadas «burbujas ideológicas». Este fenómeno, en el que el usuario acaba recibiendo contenidos exclusivamente de otros próximos a su esfera ideológica, está ya ampliamente generalizado.

Por ejemplo, Twitter y Facebook, en gran medida, eran páginas muy parecidas. Las dos ofrecían a los usuarios compartir imágenes, videos y noticias con otras personas. Anteriormente, el universo de *X/Twitter* hacía más sencillo controlar tu filtro y entender qué te aparece y por qué, mientras que en *Facebook* era imposible.² La manera en que los usuarios reciben y procesan la información ha evolucionado dramáticamente y más cuando, en el año 2022, Elon Musk se adueñó de Twitter.

La inteligencia artificial (IA) generativa es definida como: «una disciplina científica y tecnológica que busca crear sistemas capaces de resolver tareas que normalmente requieren de inteligencia humana»³.

² PARISER, Eli, "El filtro burbuja: Cómo la web decide lo que leemos y lo que pensamos." Taurus, 2017.

³ FRANGANILLO, Jorge, "La inteligencia artificial generativa y su impacto en la creación de contenidos mediáticos" *metheados*. revista de ciencias sociales, vol. 11, no 2, p. 10, 2023.

Podría ser una herramienta muy útil en cuanto a fines propagandísticos y promoción de discursos e ideas. Es por ello que se ha propuesto desarrollar y entrenar un modelo generativo de IA que crea imágenes de temática bélica que se perciben como reales.

¿Cómo se ha hecho?

La IA generativa utilizada crea imágenes de temática bélica a partir de una serie de patrones y algoritmos. Para medir y potenciar su impacto, se ha considerado realizar un estudio que aplica LUT a las imágenes generadas para medir su autenticidad percibida y cómo puede medir el riesgo de su utilización en campañas de desinformación. Los LUT replican patrones de color, saturación y un rango dinámico propios de cámaras empleadas por corresponsales de guerra, lo que añade una capa adicional de complejidad a la detección de este tipo de manipulaciones, siendo estas generadas normalmente en las redes sociales actuales.

La correlación entre ciertos patrones de color, estudio de las figuras y ambientes y la probabilidad de que las imágenes parezcan reales, podría ser establecida por el triunfo y reconocimiento de dicho experimento. Esto daría pie a nuevas investigaciones sobre la dinámica de manipulación mediática. Se baraja que la investigación de un marco de Inteligencia Artificial sin servidor en este experimento represente un avance lo suficientemente interesante y significativo tanto en el ámbito de la creación como en el de identificación de noticias falsas. Es por ello por lo que los siguientes objetivos han sido propuestos:

Una contribución sustancial al debate sobre las consecuencias éticas, políticas y económicas y sus posibles contramedidas para evitar la manipulación de la información en la era digital será elaborada. Esto pasará por concienciar a la población mediante imágenes creadas sobre cuál es el contexto actual de la IA generativa, destacar sus principales avances y estudiar sus futuros niveles de perfección.

Técnicas de LUT optimizadas: Integrar y purgar el uso de tablas de búsqueda de color para mejorar la percepción de autenticidad en las imágenes generadas por IA.

Infraestructuras sin servidor (*serverless*) serán implementadas: Desplegar y evaluar arquitecturas *serverless* para la creación y análisis posterior de imágenes y datos falsos a gran escala.

Cómo funciona la Inteligencia Artificial Generativa, cuáles son sus principales características y comprender su funcionamiento será divulgado.

La computación distribuida es un campo de la informática que se centra en el uso de múltiples ordenadores para trabajar juntos como si fueran un solo sistema. *Serverless* es un modelo de computación (actualmente, es la más avanzada que existe) en la nube en la que la infraestructura es gestionada por el proveedor de la nube, lo que permite a los desarrolladores centrarse únicamente en escribir y desplegar código. Consta de grandes cantidades de procesadores que no tienen conexión entre sí y todos resuelven el mismo problema al mismo tiempo. Dicho modelo de computación tiene más tolerancia a los fallos y más problemas pueden ser solucionados en menos cantidad de tiempo.

En el caso del experimento, dos herramientas proporcionadas por *Amazon Web Service* han sido usadas: *AWS Lambda* y *S3*. El *AWS Lambda* es un servicio de computación sin servidor que permite ejecutar código en respuesta a eventos, sin necesidad de administrar servidores. *Lambda* cobra únicamente por el tiempo consumido. El uso de este servicio de computación permite la ejecución paralela y casi inmediata de varias funciones a la vez, mejorando notablemente la eficiencia en este caso de imágenes falsas que se hacen pasar como fotografías reales o como fuentes verídicas.

Se demuestra que cuando un programa se realiza en una computadora local es viable realizar intentos de ejecución una vez por núcleo o incluso en la unidad de procesamiento gráfico para la computación vectorial. Con *AWS Lambda* es posible efectuar este programa hasta mil veces en paralelo y casi de inmediato sin necesidad de comunicación entre los otros programas.

En consecuencia, casi 1000 imágenes pueden ser generadas de forma simultánea. La equivalencia entre ejecutar 1000 funciones a ejecutar una sola función está en la rentabilidad, ya que los cargos de *Lambda* se basan en la capacidad de cómputo utilizada por cada mil ejecuciones.

AWS Lambda es seleccionada entre otras cuestiones por su rentabilidad. *AWS Lambda* permite el despliegue de un modelo de software que incluye entre otras cosas una IA generativa, entrenada para producir imágenes de guerra que pueden ser percibidas como reales. Por otro lado, el S3 es un servicio de almacenamiento en la nube de AWS que ofrece escalabilidad, disponibilidad de datos, seguridad y rendimiento. Sus siglas son *Simple Storage Service*.

Para minimizar costos, una vez que las imágenes del S3 son descargadas, los archivos temporales son suprimidos, excepto el archivo precargado de *Stable Diffusion*. Esta arquitectura sin servidor crea miles de imágenes de guerra falsas en menos de dos minutos. Conviene explicar que *Stable Diffusion* es una inteligencia artificial especializada en la creación de imágenes falsas. Funciona a partir de descripciones textuales. Forma parte de una familia de modelos de difusión como el aprendizaje automático. Esta IA se ha utilizado para fines artísticos.

El S3 tiene similitud con *Google Drive*, donde puedes depositar oraciones. Por ejemplo, la ejecutada en este experimento “genérame las imágenes”. A continuación, bajo la orden de dicha ejecución, las imágenes generadas y depositadas pasan a la función Lambda que hemos definido antes, que le aplica los patrones de color LUT (dichos patrones, como hemos definido anteriormente, son los que hacen que parezcan más realistas, lo que provoca que sea complicado averiguar su autenticidad con herramientas de detección de inteligencia artificial).



Figura 1: Imagen bélica generada, Imagen obtenida del banco de imágenes realizado mediante el algoritmo propuesto

Para este experimento, se empleó una IA entrenada exclusivamente con fotografías reales de ciudades de Chipre y Grecia. A partir de este entrenamiento, se generaron imágenes falsas a las que se aplicaron 20 tipos diferentes de LUT, completando la producción de cada lote en 20 segundos. Gracias a la definición de patrones temáticos específicos, las imágenes generadas representaban escenas bélicas, como columnas de humo, edificios ardiendo o ruinas emblemáticas. Esta descripción temática resultó crucial para asegurar la precisión del modelo de IA en la creación de las escenas. El procedimiento se repite de forma cíclica para producir nuevas imágenes a partir de los patrones establecidos.

A la hora de analizar un modelo fotográfico o generado digitalmente, se deben tener cuenta cierto tipo de conceptos para analizar su posible veracidad o falsedad. Uno de los métodos más comunes es imitar las resoluciones de fotografías auténticas. Este método es común para la generación de imágenes falsas, se toma una fotografía, se elimina resolución y se reescala posteriormente haciendo que pierda o gane calidad, pero que parezca tomada por una cámara y una persona aficionada. Estos

conceptos básicos se explican para su aplicación en la detección de falsificaciones en el modelo final de IA Generativa. La capacidad de un sistema óptico para distinguir entre dos puntos cercanos entre sí se define como la resolución óptica, la cual se ve afectada por varios factores en los que se incluye la longitud de onda y la calidad de las lentes del propio sistema. Con más detalle:

Lentes: Los rayos emitidos por el objeto son recogidos por los elementos esenciales en la óptica y conformados en el espacio imagen. Un ejemplo biológico de este tipo de elementos sería el cristalino del ojo, que se encarga de recoger la luz los elementos más cercanos.

Apertura y difracción: La profundidad de campo y la difracción (desviación de una onda al atravesar ciertos obstáculos) son afectadas por la apertura de la lente y son complementarias. Así, una apertura muy pequeña puede causar difracción, lo que reduce la resolución óptica, mientras que una apertura muy grande puede reducir la profundidad de campo, afectando a la nitidez en ciertas áreas de la imagen.

Longitud de onda de la luz: La resolución óptica también depende de la longitud de onda que recibe. Detalles más finos pueden ser resueltos por la luz de longitud de onda más corta (luz azul) que por la luz de longitud de onda más larga (luz roja).

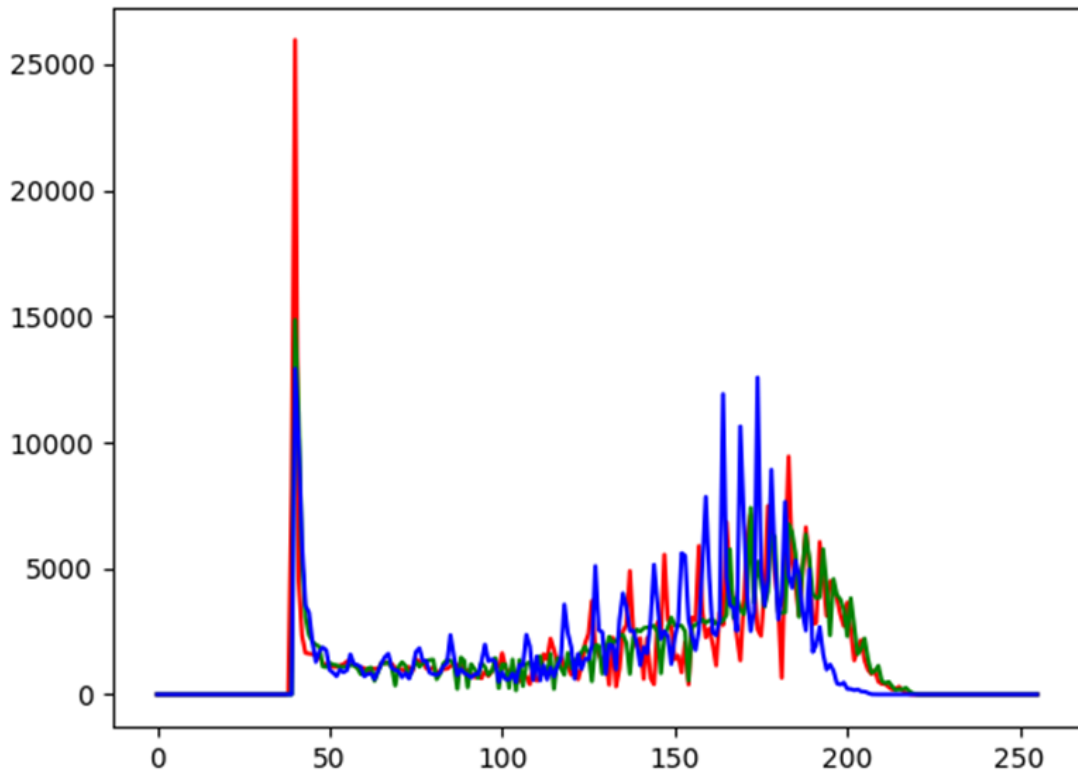


Figura 2: Ejemplo de un histograma de una imagen bélica falsa, gráfica generada con código fuente

Estos factores son esenciales en la realización de fotografías, ya que un fallo en estos elementos puede resultar en una resolución incorrecta de la fotografía, apareciendo elementos extraños (artefactos) en la imagen. Es aquí donde encontramos una gran similitud con las inteligencias artificiales generativas, ya que estos artefactos presentes en la imagen pueden ser imitados tras una descripción correcta de los requisitos o, del mismo modo, burlados. Esto consiste en provocar un «mal uso» conociendo los posibles errores que puede contener una fotografía con el objetivo de que pase por real. La gente suele tener por norma dudar de fotografías que salgan «perfectas». Se suele hacer un paralelismo entre «perfecto» y digitalmente retocado. El ignorar algunas de estas técnicas suele generar resultados difíciles de predecir en las IA Generativas.

El color se percibe de manera distinta en medios ópticos y digitales. Si se trata de forma óptica⁴ tendremos en cuenta elementos como la longitud de onda o la temperatura de color para diferenciarlos, mientras que, por otro lado, en medios digitales tendremos en cuenta valores como la escala, como por ejemplo, la escala RGB (Red, Green, Blue) o la escala HSV (Hue, Saturation, Value) y la luminancia de la pantalla (como referencia el cuerpo negro). Estas escalas son distintas en forma y ejes de referencia. La escala RGB tiene forma de elipse y tiene en cuenta las coordenadas rojo, verde y azul, mientras que por otro lado la escala HSV tiene la misma forma pero tiene en cuenta otras coordenadas, en este caso el tono, la saturación y el color definido como value.

Es preciso que, tras la creación de dichas imágenes, se realice un análisis exhaustivo de los histogramas de color que correspondan a dichas imágenes que tengan más posibilidades de hacerse pasar por fotografías verídicas, pues estos histogramas nos dan pistas muy relevantes en cuanto a la distribución e intensidad de los colores aplicados para extraer más patrones de color y potenciar nuestra IA generativa. Esto, a su vez, fomenta la creación de un sistema de detección de imágenes falsas (sobre todo si se trata de imágenes de temática bélica) y facilita la creación de otras herramientas automatizadas que puedan reconocer una fotografía creada por IA de forma más perfeccionada en un futuro.

La óptica de la cámara tiene una serie de limitaciones y está fundamentada en una serie de parámetros físicos que pueden ser imitables por la inteligencia artificial, entre otras: Deterioro de la cromaticidad debido a la exposición y del contraste en las fotografías y pérdida de resolución de la fotografía.

El uso de la infraestructura sin servidor de *AWS Lambda*, *S3* y *Stable Diffusion* permite la creación de imágenes falsas de guerra y, cuanto más se entrene, más fotografías serán generadas de forma eficiente, proporcionando a su vez una herramienta necesaria para el desarrollo y análisis de técnicas avanzadas de manipulación de imágenes. Aparte de fomentar la creación de imágenes, las bases para futuras investigaciones en la

⁴ HU, Y., YANG, D., MA, D., QI, C., & HUANG, S. Structural Color-Based Smart Liquid Windows Address the Tradeoff Between High Optical Transparency and Brilliant Color. *Advanced Functional Materials*, 2023.

detección de noticias falsas y la comprensión de la manipulación mediática en la era digital serán estimuladas.

Pese a los esfuerzos, la computación distribuida también presenta varios desafíos. La complejidad de diseño y administración es significativamente mayor que en los sistemas centralizados. Coordinar y sincronizar múltiples nodos puede ser complicado, y el riesgo de problemas de coherencia de datos y de comunicación aumenta con la distribución geográfica. La seguridad también es una preocupación importante, ya que los datos y procesos están distribuidos en diferentes ubicaciones, haciendo más difícil proteger la integridad y confidencialidad de la información. Además, el manejo de fallos y la recuperación de errores en sistemas distribuidos requieren estrategias avanzadas que pueden ser difíciles de implementar y mantener.

En el contexto de la generación de imágenes mediante inteligencia artificial, la computación distribuida se utiliza para implementar soluciones de computación sin servidor (serverless computing). Este enfoque permite a los desarrolladores centrarse en la lógica de la generación de imágenes sin preocuparse por la gestión de la infraestructura subyacente que es proporcionada por el proveedor como un servicio que simplemente funciona. Con la computación sin servidor, las funciones de IA pueden ser desplegadas en la nube y ejecutadas automáticamente en respuesta a eventos, como solicitudes de generación de imágenes.

Conclusiones y resultados

Desde que la IA generativa llegó a las redes sociales no ha estado exenta de debate y polémicas. Como definición, el conjunto de tareas realizadas mediante ordenadores que tradicionalmente habrían requerido inteligencia humana es la IA⁵. Se ha utilizado como medio creativo y también como herramienta en el ámbito medicinal (por ejemplo, para la detección de distintos cánceres). También dentro del mundo tecnológico «desde el análisis de datos y la identificación de patrones para comprender fenómenos complejos

⁵ BOULANIN Vincent, et al. "Artificial intelligence, strategic stability and nuclear risk". 2020.

hasta la conducción de vehículos o la generación de contenido visual, textual y sonoro»⁶ y ahora supone un arma arrojadiza en la era de la desinformación.

Gracias a examinar las imágenes creadas y sus propios histogramas de color, hay numerosas ideas sobre el impacto que estas imágenes pudiesen tener si fuesen publicadas en televisión o en una red social como *X/Twitter*. Lo primero a tener en consideración en cuanto a resultados es que varias herramientas que podrían realizar imágenes actualmente se han tenido en cuenta, pero *Amazon Web Service* ha sido elegida. En otras palabras, ha habido constancia y conocimiento de herramientas que podrían ser utilizadas para la fabricación de *fake news* con ayuda de estas imágenes, lo cual conlleva una evaluación sobre los conflictos éticos que esta herramienta supone. Las imágenes creadas en este estudio buscan imitar fotografías reales basándose en imitar conflictos del mundo real actuales, aplicándoles diferentes LUT que contribuyan a obtener conocimientos sobre cómo una imagen puede parecer más auténtica de cara a prensa, redes sociales y público en general. Esta exploración también involucra examinar consecuencias sociales, políticas y económicas que podrían surgir si esta tecnología se vuelve accesible y la ciudadanía es conocedora de esta.

Una computación extremadamente avanzada como la mostrada en el experimento es un peligro en distintos aspectos, entre ellos los que tienen fines propagandísticos y se utilizan como arma arrojadiza contra rivales políticos o culturales. Según Aníbal Monasterio: «esto transformará nuestra forma de comunicarnos, pero también incrementará las vulnerabilidades, convirtiendo la calidad e integridad de la información en algo fundamental para la seguridad y otros valores sociales. Y por calidad e integridad de la información se entiende no solo que la información no sea degradada o corrompida (...) sino también que se proteja y asegure que esa información no se acceda sin consentimiento (privacidad) y que tenga un propietario bien identificado».⁷

La facilidad con la que imágenes falsas pueden ser generadas y distribuidas mediante tecnologías avanzadas plantea una gran responsabilidad para los desarrolladores y

⁶ ÁLVAREZ-VERDUGO, Milagros, "la inteligencia artificial en perspectiva comparada" Revista del Instituto Español de Estudios Estratégicos, n.º 224, pp. 183-215, 2024.

⁷ MONASTERIO ASTOBIZA, Aníbal, "Ética algorítmica: Implicaciones éticas de una sociedad cada vez más gobernada por algoritmos." Dilemata, (24), 185–217, 2017.

usuarios de plataformas de redes sociales. Se necesitan normativas éticas que regulen el uso y la distribución de contenido generativo para prevenir la manipulación informativa.

Esto conlleva un impacto de varios ámbitos: la generación de imágenes falsas puede tener consecuencias significativas en contextos sociales y políticos, influyendo en la percepción pública y potencialmente alterando los resultados de conflictos reales,. Sería interesante proponer, para después desarrollar, métodos efectivos de verificación para asegurar la integridad del contenido compartido en medios de comunicación y plataformas digitales.

Para prevenir esto, se necesita una educación y concienciación pública para la identificación de noticias falsas, especialmente aquellas que provienen de fuentes visuales. La educación puede actuar como un contrapeso al aumento de las imágenes falsas, equipando a los individuos con habilidades críticas para discernir la veracidad del contenido audiovisual.

Por último, cabe destacar la innovación en la detección de *fake news*: el desarrollo de tecnologías para la detección automatizada y eficiente de noticias falsas es vital para contrarrestar la diseminación de desinformación. La integración de sistemas *serverless* para el análisis rápido y el procesamiento de grandes volúmenes de imágenes puede ser una herramienta valiosa en la lucha contra la manipulación mediática.

D^a. Estíbaliz García Huete, Doctoranda, Facultad Ciencias de la Información, UCM

Dr. Sara Ignacio Cerrato, Doctora en óptica, optometría y visión, UCM

Dr. David Pacios Izquierdo, Profesor Ayudante Doctor, Facultad de Informática, UCM

Dr. José Luis Vázquez Poletti, Prof. Titular de Universidad, Fac. de Informática, UCM

Dr. M^a José Pérez Serrano, Prof. Titular de Universidad, Fac. de CC. de la Inf., UCM

Dr. Andrea Donofrio, Prof. Titular de Universidad, Fac. de CC. de la Información, UCM