



LA INTELIGENCIA ARTIFICIAL GUIANDO AL PUEBLO

ARTIFICIAL INTELLIGENCE LEADING THE PEOPLE

José Manuel Pereira Uzal

Universidad Complutense

Profesional independiente

Lino García-Morales

Universidad Politécnica de Madrid

Universidad Autónoma de Madrid

DOI: [10.33732/ASRI.6584](https://doi.org/10.33732/ASRI.6584)

.....
Recibido: (17 01 2024)

Aceptado: (23 07 2024)

Cómo citar este artículo

Pereira-Uzal, J.M. y García-Morales, L., (2024). La inteligencia artificial guiando al pueblo. *ASRI. Arte y Sociedad. Revista de investigación en Arte y Humanidades Digitales.*, (25), 17-30.

<https://doi.org/10.33732/ASRI.6584>

Resumen

El desarrollo de la inteligencia artificial y la invasión en todos los ámbitos de la vida pública (incluido el arte) ha abierto un debate que anuncia, una vez más, la muerte del arte. Lo artificial se debate contra lo natural por legitimar su existencia a través de la verdad y la mentira.

Abstract

In modern times, the development of artificial intelligence and its invasion in all areas of public life (including art) has opened a debate that announces, once again, the death of art. The artificial fights against the natural to legitimize its existence through truth and untruth. For some

Para algunos la creatividad está en riesgo en pro de la banalización, para otros se trata de meras imitaciones o copias que vulneran los derechos de propiedad intelectual. El objetivo fundamental de este artículo es poner las ideas en contexto mediante las respuestas a las siguientes preguntas: ¿qué es y qué no es la inteligencia artificial?, ¿qué es y qué no es el plagio?, ¿qué es y qué no es útil de todo este conflicto?, ¿estamos, de verdad, ante un nuevo fin del arte?

Palabras clave

inteligencia artificial, arte, derechos de autor, plagio, copia, fakes

people, creativity is at risk in order to trivialize, for others it is just a matter of pure fakes or copies that violate intellectual property rights. The main objective of this article is to put the ideas in context by answering the following questions: what is and what is not artificial intelligence, what is and what is not plagiarism, and what is and what is not useful in this conflict?

Keywords

artificial intelligence, art, copyright, plagiarism, copy, fakes

Introducción

La Inteligencia Artificial Generativa (IAG¹), capaz de generar contenidos originales en forma de imágenes, fue presentada a lo largo de los años 2021 (DALL-E) y 2022 (*Stable Diffusion* y *Midjourney*) provocando un gran revuelo entre la comunidad de creadores y artistas plásticos; no sólo por la sustitución potencial de su labor como productores de imágenes por los algoritmos, sino también por la forma en que estos algoritmos son ‘entrenados’. Primero, trajo a primer plano viejas cuestiones como: ¿puede una máquina ser artista?, ¿puede un algoritmo producir arte?, ¿se trata, en efecto, de arte? Segundo, supuso cierta criminalización dado el uso ‘ilícito’ de imágenes con protección de derechos de propiedad intelectual para el entrenamiento de estas máquinas y también por su relación con las imágenes generadas; acusación que se extiende desde el *plagio* hasta el *robo*.

¿Qué es y qué no es la inteligencia artificial?

Para entender el revuelo que suscitan estas tecnologías es preciso plantear adecuadamente el problema; si se parten de proposiciones falsas, no es posible llegar a conclusiones verdaderas. Noam Chomsky, Ian Roberts, y el filósofo y especialista en inteligencia artificial Jeffrey Watumull, en el ensayo *La falsa promesa del ChatGPT*, publicado originalmente en The New York Times, con afán de establecer límites tanto al entusiasmo desmedido como a las visiones apocalípticas frente a estas tecnologías, que apasionan y desconciertan por igual, escribieron:

La mente humana no es, como el ChatGPT y sus similares, un enorme motor estadístico de comparación de patrones que se alimenta de cientos de terabytes de datos para elaborar la respuesta más factible en una conversación o la respuesta más probable a una pregunta científica. Por el contrario, la mente humana es un sistema sorprendentemente eficiente e

¹ Conjunto de soluciones algorítmicas conocidas como modelos de texto-a-imagen (*Text-to-Image*) en el ámbito del aprendizaje automático (*Machine Learning*) o al aprendizaje profundo (*Deep Learning*).

incluso elegante, que funciona con pequeñas cantidades de información y que no busca inferir correlaciones brutas entre puntos de datos, sino crear explicaciones. (Chomsky, Roberts y Watumull 2023).

En efecto, la inteligencia artificial (IA) no es la inteligencia humana, de la misma manera que una máquina de aprendizaje automático no es, ni funciona igual, que la mente humana; sin embargo, esta 'aclaración' establece una 'asociación' velada entre ambas inteligencias, dado que "la inteligencia es el medio a través del que resolvemos los problemas"; como si se tratara de lo mismo, lo cual constituye, sin duda, un mal planteamiento del problema. Si bien es cierto que ambas inteligencias son medios de resolución de problemas, sus presupuestos de partida y mecanismos operativos son de distinta naturaleza.

La IA es sólo el medio a través del cual las máquinas resuelven problemas mediante mecanismos y procesos diseñados por mentes humanas. El *aprendizaje automático* (*machine learning*) es una estrategia de tipo 'caja negra' para aprender a resolver problemas basados en la experiencia; puede ser supervisada o no, pero, en cualquier caso, 'alimenta' la máquina con datos o contextos para que 'aprenda' a resolver un problema específico. El *aprendizaje profundo* (*deep learning*) es sólo un tipo especial de 'caja negra' monumental. Lo relevante aquí es la introducción en los mecanismos de aprendizaje de determinadas estrategias 'humanas'; como pueden ser refuerzo (premio-castigo), atención, etc. Todas estas incorporaciones han llevado a lo que preocupa a Chomsky y también al mundo del arte. El aprendizaje, en herramientas como *DALL-E*, *Stable Diffusion* y *Midjourney*, se realiza con datos o contextos heterogéneos (por ejemplo, texto a la entrada, imagen a la salida), e incluso híbridos. Tales textos se conocen como prompt y tales soluciones algorítmicas como modelos *Text-to-Image*.

Toda la disciplina de la IA se originó a partir de una idea simple: la simulación del funcionamiento de una neurona natural, real, en una neurona artificial, virtual (McCulloch y Pitts 1943) y la aspiración de emular el comportamiento del cerebro humano. La salida del *perceptrón simple* (primer modelo matemático) es una combinación lineal (excitatoria-inhibitoria) de las entradas que supuestamente llegan por las dendritas y salen por el axón después de una transformación no lineal (técnicamente: *función de activación*). De la misma manera que un ser humano con una sola neurona no es capaz de 'pensar' gran cosa, una máquina de IA con una sola neurona, no es capaz de hacer gran cosa. La complejidad de los sistemas avanzados 'inteligentes' emerge de la interacción de muchas neuronas simples trabajando en conjunto; es decir, las neuronas deben formar redes ordenadas en capas. El *perceptrón multicapa*, una de las estructuras donde operan los algoritmos de aprendizaje automático y profundo, es una estructura más cercana a las interconexiones de neuronas biológicas. En este modelo, la información fluye a través de múltiples capas, desde la entrada hasta la salida. Se podría decir que en las capas intermedias las neuronas se especializan en resolver determinados problemas que se combinan en capas posteriores. El kit diferencial de estas redes es la introducción de la *función de activación*; gracias a la cual la red aprende comportamientos complejos y no lineales (a diferencia de las estructuras lineales frecuentes de 'caja blanca').

Las soluciones *Texto-a-Imagen* son entrenadas con cantidades ingentes de relaciones entre imágenes y textos descriptivos de tales imágenes (*captions*). Proyectos como LAION (<https://laion.ai/blog/laion-5b/>) contienen más de 5 billones de imágenes asociadas a descripciones tomadas de la web y validadas con una estructura de aprendizaje profundo popular de tipo *Transformer* denominada CLIP (*Contrastive Language-Image Pretraining*); es decir, se precisa de una comprobación previa al entrenamiento de la coincidencia entre la imagen y el texto descriptivo.

Herramientas como *Stable Diffusion* utilizan el concepto de *difusión* para generar nuevo contenido; esto es, en la *difusión directa* a una imagen se le añade progresivamente ruido hasta obtener un agregado de píxeles de intensidad aleatoria carentes de forma reconocible, es decir, la imagen se ‘diluye’ o ‘difunde’ en un patrón de ruido. En la *difusión inversa* a una imagen que sólo contiene ruido (rica en información), las intensidades de los píxeles se organizan para producir una forma reconocible acorde a un *prompt*; es decir, una imagen, construida píxel a píxel, siempre nueva e irrepetible.

ChatGPT no funciona como estos algoritmos *Texto-a-Imagen* (no interviene el azar en sus entradas) y estos últimos no son con exactitud máquinas estadísticas, ni generan la respuesta más probable; más bien aproximan una función de probabilidad desconocida a priori de manera no lineal y producen imágenes cuya entropía es menor que la imagen ruido de partida. Si fueran máquinas estadísticas, las imágenes de salida serían siempre las mismas. La mente humana no busca correlacionar información sino inferir conocimiento, exactamente igual que las máquinas, aunque lo hagan de manera diferente, con recursos y dimensiones diferentes. Se podría afirmar que los sistemas de IA son herramientas que funcionan con los mismos principios de la creatividad humana: copia, transformación, combinación y azar; eso sí, carentes de voluntad, conciencia, intencionalidad y finalidad (de momento).

¿Qué es y qué no es el plagio?

Los artistas trabajan con signos que conectan las formas con las ideas de maneras más o menos personales, locales o universales. Todo son imágenes. Entendiendo como ‘imagen’ “una cierta existencia que es más de lo que el idealismo llama una representación, pero menos de lo que el realismo llama cosa, una existencia a medio camino entre la ‘cosa’ y la ‘representación’” (Bergson 2016, p. 25–26).

La imagen fotográfica, desde este punto de vista, es imagen proyectada de una imagen real; una cosa que es representación real cuyo contenido mantiene una relación causal con lo real que captura o congela en el tiempo. Una pintura es una cosa (imagen) que produce una representación en el sujeto (imagen). Una fotografía es una representación de una cosa (imagen) que produce igualmente una representación en el sujeto (imagen). Todo son imágenes. La única diferencia es que esta mediación está inscrita en un soporte que tiene derechos de propiedad intelectual por el solo hecho de corresponder a la ‘voluntad de su creador’.

Según esta distinción, las imágenes que el artista copia, transforma y configura (no sin la intervención del azar) de la realidad (incluidas las imágenes de todas las obras de arte) están libres de derechos de autor; mientras que las imágenes que sirven de entrenamiento a los algoritmos de IA (habitualmente imágenes digitales inscritas en un soporte; ya sea real o virtual) tienen o pueden tener protección de derechos de autor. Esta situación crea una diferencia entre lo que se considera *copia* (referencia de aprendizaje) y *plagio*. Toda copia implica una mimesis en substancia, forma o contenido, teniendo siempre presente, que, en general, se reproduce la forma, no el significado², que este pertenece a la producción de la imagen interior que el espectador hace sobre la imagen exterior. El *plagio* es *copia* en la cual un autor se atribuye substancia de un segundo; mientras que la *falsificación* es *copia* en la cual un autor atribuye substancia a un segundo, con la intencionalidad dolosa de suplantar su autoridad.

² Desde el punto de vista de la apropiación, la recontextualización de la copia la dota de un nuevo significado.

La ley de la propiedad intelectual no protege ideas, sino representaciones (formales) de ideas, según la originalidad o representación de la idea en una substancia de forma singular. Téngase en cuenta que el estilo inscrito en la representación no resulta significativo en la determinación de la singularidad de la imagen; el estilo no es objeto de la ley de la propiedad intelectual; es entendido como producto de una herencia, influencia o contexto social.

Los mecanismos de mimesis o semejanza no son *per se* dolosos ni encaminados a la suplantación o al engaño; son, en primera instancia, mecanismos o referencias de aprendizaje, según Aristóteles (Suñol 2012), mediante los cuales los humanos o las máquinas adquieren habilidades o estructuras útiles en la resolución de problemas. La velocidad con la que una IA procesa e infiere información, y la imparcialidad y neutralidad de sus conclusiones, excede las capacidades humanas. Una IA llega a adquirir, al menos, la misma competencia que un experto humano (García Morales 2020).

La *copia*, en su intertextualidad, es también un mecanismo de vinculación o socialización de la obra (Barthes 1973); un artista inscrito en un determinado estilo (impresionismo, cubismo, expresionismo, etc.) fundamenta su filiación mediante el dominio de los patrones que definen tal estilo o movimiento de forma inequívoca; así, copiar exige abstraer el virtuosismo o comprensión de lo copiado con la pericia suficiente para alcanzar esa semejanza con los referentes. Se *copia* el estilo. En las IAGs, la copia permite tejer una compleja red de relaciones intertextuales entre imágenes, mediante un extenso conjunto de transformaciones e imitaciones que acaba por integrarla en un acervo cultural particular. En definitiva, la historia del arte es una compleja red de enlaces entre artistas e imágenes, basada a menudo en las influencias de otras imágenes, copias parciales e incluso copias completas. La originalidad de una imagen se basa más en la *vacuidad*, en su relación con el resto de las imágenes y en el conocimiento de tales relaciones, que con su *unidad*.

Las imágenes que provienen de las IAGs corresponden a la simulación: cuarta categoría de imagen de Baudrillard (Baudrillard 1978, p. 14); pues “no tienen que ver con ningún tipo de realidad”, es decir, “no corresponden al orden de la apariencia, sino al de la simulación”. No son copias, sino simulacros; son “una suplantación de lo real, por los signos de lo real, es decir, de una operación de disuasión de todo proceso real por su doble operativo” (p. 7). Esto es precisamente lo que las IAGs hacen bien: capturar la operatividad en los procesos.

El aprendizaje en la IA, sea automático o profundo, revela los patrones o estructuras que subyacen en un conjunto de datos, siempre y cuando tales estructuras se presenten de forma recurrente para ser identificadas como patrones significativos. Estos algoritmos aplicados a grandes conjuntos de imágenes contenedores de signos, así como a los *captions* o ideas que transmiten tales signos, permiten determinar la estructura signíca del lenguaje inscrito en la información y, en consecuencia, elaborar modelos que faciliten su codificación y decodificación. La IAG, por tanto, construye imágenes en base al lenguaje signíco deducido del conjunto de datos de entrenamiento. La creación de nuevos signos, o de un nuevo lenguaje signíco para representar ideas en la IAG, es muy limitada o nula; lo primero sólo es posible debido al azar y no por alguna conciencia, voluntad o intencionalidad, lo segundo precisa lo experiencial, social y cultural y eso, por ahora, es exclusivamente humano.

Otro paradigma de la *copia* es la *cita*: *interpretación* o *apropiación* de una imagen; en la cual, aunque el grado de semejanza no es total, hay una referencia clara a su procedencia. La apropiación utiliza elementos, técnicas, estilos o ideas de una imagen como signos o referentes en la producción de una

nueva obra; toma prestada una imagen producida por otro autor y la utiliza como parte de una nueva imagen. La copia en sus formas de cita o apropiación es introducida en la imagen con el fin de ser desvelada a diferencia del *plagio* o la *falsificación* que sólo funcionan cuando estos permanecen ocultos y por lo tanto ligados a un propósito doloso.

La artista Elaine Sturtevant recibió el León de Oro en la 54ª Bienal de Venecia por su brillante carrera apropiacionista. Nunca fue acusada de plagio porque nunca suplantó la autoría de ningún artista. Esta práctica muy común en la posmodernidad, y en el arte en general, no provoca ningún revuelo legal, ético o estético. La motivación de Sturtevant fue: “Si ellos pueden copiar, yo también”; ¿por qué entonces, en un caso extremo, no pueden copiar las máquinas? Si Elaine Sturtevant fue premiada por copiar, ¿por qué se criminaliza a una máquina que no está diseñada para copiar por no copiar?

El artista conecta ideas con signos. La IAG también: conecta ideas, en forma de *prompt*, con signos, en forma de imágenes. La imagen fotográfica conecta con la imagen ‘real’ (aunque no siempre), mientras que la imagen ‘artificial’ que produce una IAG conecta con ‘la idea’, un hecho conceptualmente muy próximo a la creación artística que no puede ser más que alegórico. El artista tiene la voluntad de aproximar lo real o emocional a través de un lenguaje simbólico, mientras que los algoritmos, desprovistos de voluntad, funcionan como meros traductores de ideas hacia un determinado lenguaje articulado por símbolos.

La primera demanda³ contra las IAGs fue presentada por las artistas Andersen, Kelly Mckernan y Karla Ortiz en California y dirigida a las empresas Stability AI LTD, Midjourney Inc. y Deviantart Inc., en enero del 2023. Las alegaciones⁴ interpuestas por las compañías demandadas se basaron en un supuesto mal uso de la idea de collage. El juez, ante la imposibilidad de señalar qué imágenes y derivados fueron objeto de violación e infracción de derechos de autor, solicitó a los demandantes la aportación de mayor ‘claridad’. ¿Cómo las IAGs infringían o podían infringir los derechos de autor de imágenes no registradas con tales derechos? Por último, en octubre de 2023⁵, el juez desestimó la demanda. Este litigio es un ejemplo de cómo la mala comprensión o tergiversación de los fundamentos tecnológicos de una herramienta conduce a una mala interpretación de su alcance o uso y de cómo los derechos de autor sólo pueden ser garantizados en la medida en que se participe de forma íntegra de todos los procesos y normas que estos implican.

Apenas un mes después de este incidente, Getty Images Inc. interpuso una demanda⁶ contra Stability AI por violación del copyright debido al uso de sus imágenes para entrenamiento. En ambas demandas, Stability IA apeló a lo que la jurisprudencia en Estados Unidos describe como ‘uso justo’ (*Fair Use*⁷) o uso legítimo (por ejemplo, cuando la finalidad es la investigación). La propia ley define la ‘sustanciabilidad’

³ Andersen v. Stability AI Ltd. 3:23-cv-00201; District Court, N.D. California <https://www.courtlistener.com/docket/66732129/andersen-v-stability-ai-ltd/>

⁴ Caso 3:23-cv-00201-WHO Documento 52 del 18/04/23; <https://www.courtlistener.com/docket/66732129/52/andersen-v-stability-ai-ltd/> y Caso 3:23-cv-00201-WHO documento 51 del 18/04/23; <https://www.courtlistener.com/docket/66732129/51/andersen-v-stability-ai-ltd/>

⁵ Order on Motion to Dismiss — Documento 117 del 30/10/23; <https://andersen-v-stability-ai-ltd/>

⁶ Getty Images (US), Inc. v. Stability AI, Inc. (1:23-cv-00135), District Court, D. Delaware; <https://www.courtlistener.com/docket/66788385/getty-images-us-inc-v-stability-ai-inc/>

⁷ Apartado 107 de la ley de Copyright de Estados Unidos de diciembre de 2022.

de la parte utilizada como requisito de tal uso. Esta demanda se enfocó en el uso de las imágenes de entrenamiento y apuntó de forma indirecta al proyecto LAION (<https://laion.ai/>), con origen en Alemania, de naturaleza libre (Creative Common CC-BY 4.0) y orientado a la investigación, lo cual añadió a la demanda un importante grado de complejidad.

Contra todo pronóstico, en septiembre de este mismo año Getty Images Inc. introdujo su propio servicio⁸ de IAG entrenado con su propio banco de imágenes bajo la descripción de ‘comercialmente segura’⁹; y una supuesta remuneración a los creadores por la inclusión de sus imágenes en las bases de datos de entrenamiento; cuestión que el propio Craig Peters, CEO de Getty Images Inc., dejó entrever en una entrevista en The Verge¹⁰ como ‘no solucionada’, puesto que la asociación entre el resultado de una IAG y el grado de contribución de ciertas imágenes de entrenamiento no es un problema trivial.

Ante este tipo de polémicas, el ministro japonés Nagaoka respondía¹¹ así a la cuestión del entrenamiento de este tipo de herramientas:

En cuanto a los fines comerciales, el artículo 30-4 de la Ley de Propiedad Intelectual no cuestiona si el análisis de la información tiene o no fines comerciales, por lo que puede utilizarse por motivos que no impliquen el disfrute de las ideas o emociones expresadas en la obra. [...] Además, el uso de una obra con fines de análisis de la información no tiene por objeto disfrutar de los pensamientos y sentimientos expresados en la obra, y aunque se lleven a cabo tales actos, no entran en conflicto con el mercado original para el uso de obras con derechos de autor y no perjudican los intereses de los titulares de los derechos de autor [...]

Caso de estudio

¿Por qué personalidades como Chomsky — indirectamente — e instituciones como las *Big Tech* — directamente — abogan por regular o licenciar la IA, con la peligrosa consecuencia de aplastar la innovación? Las herramientas como extensiones del hombre carecen de ética; es el ser humano quien parece predestinado a eludir las responsabilidades de su propia existencia. Una IA no es más creativa que otra, es la heterogeneidad del conjunto de imágenes de partida, el azar y la cuidadosa escritura de un *prompt* lo que hará que una imagen sea considerada arte o no por un humano.

Los algoritmos de IA funcionan de manera similar a las estrategias de creación de imágenes de los humanos. Ambos parten de imágenes (sean o no libres de derechos) para generar nuevas imágenes (susceptibles o no de protección de derechos). La *copia*, el *plagio*, o la *falsificación*, exige *voluntad e intencionalidad*; atributos de los que la inteligencia artificial carece. Que una imagen sea considerada o no arte seguirá el mismo escrutinio por el que un urinario fue considerado arte. Que una imagen exija mayor o menor rigor técnico (oficio), estético o ético, ya ha sido superado. La historia del arte está

⁸ <https://www.gettyimages.es/ia/generacion/acerca>

⁹ <https://investors.gettyimages.com/news-releases/news-release-details/getty-images-launches-commercially-safe-generative-ai-offering>

¹⁰ <https://www.theverge.com/23903700/getty-images-craig-peters-generative-ai-images-disinformation-payment>

¹¹ 24 de abril del 2023 (<https://kiitaka.net/21312/>)

cargada de un universo de imágenes heterogéneas que dan testimonio del ‘juicio del arte’; que no es otro que el juicio de los humanos.

La Libertad Guiando al Pueblo es un homenaje de Eugène Delacroix a las revueltas producidas en 1830 conocidas como ‘revolución de Julio’ que servirá como laboratorio experimental de las ideas discutidas. Delacroix usó como figura central una representación de Marianne, símbolo de la República Francesa en representación de sus valores: libertad, igualdad y fraternidad. Marianne contiene ciertos símbolos como son el gorro frigio, adoptado por los revolucionarios como símbolo de libertad, y una postura con una cierta alusión a la Victoria de Samotracia.

Figura 1: Eugène Delacroix. 1830. *La Libertad guiando al pueblo*. Óleo sobre lienzo.



Fuente: https://es.wikipedia.org/wiki/La_Libertad_guiando_al_pueblo

Una de las problemáticas en la descripción (*caption*) de esta obra es su asociación constante con la Revolución Francesa de 1789 debido a la simbología utilizada; lo que supone un problema para que los algoritmos asocien con propiedad el contenido de la obra o su posible contexto social. Una búsqueda en LAION (<https://rom1504.github.io/clip-retrieval/>) sobre esta obra ofrecerá un buen número de asociaciones con conceptos muy dispares más relacionados con sus valores, que con su contexto histórico. Es importante matizar que esta imagen no representa un hecho histórico; es una celebración del autor a un hecho en forma de alegoría a ciertos valores a través de símbolos inequívocos.

El experimento propuesto explora las interpretaciones que herramientas como Stable Diffusion (<https://stablediffusionweb.com/>) realizan del *prompt*: “*liberty leading the people*” (título de la obra en inglés). Cuando se alimenta un *prompt* con una idea muy específica la cual se ajusta de forma muy específica

con una serie de datos de entrenamiento, en los cuales además existe una importante redundancia o sobre-representación de un tipo de imágenes, pueden emerger fenómenos denominados de sobreajuste o memorización (*overfitting*) que acaban por arrojar resultados muy próximos a los datos de entrenamiento (Carlini et al., 2023; Somepalli et al., 2022; Yeom et al., 2018) tal y como se muestra en la Figura 3a. Sin embargo, los resultados pueden hacerse dependientes del parámetro conocido como *Classifier-Free Guidance* o *CFG Scale*; una estrategia que usa esta herramienta para ajustarse al *prompt* con más o menos literalidad: valores más bajos producen interpretaciones más libres mientras que valores más altos producen interpretaciones más estrictas.

Figura 2: Imágenes producidas con diferentes *CFG Scales* y *prompt*: *liberty leading the people*. Elaboración Propia.



Lo más literal del *prompt* se presenta como ocurrencias más probables en las bases de datos de entrenamiento, aunque esto no siempre implica mayor literalidad con la idea transmitida escrita en el *prompt*. La escala en este momento oscila entre 0 y 20, con un valor intermedio por defecto de

7.5 (valores cercanos a 0 produce imágenes apenas legibles y valores cercanos a 20 produce imágenes más próximas al contenido del *prompt*).

Se produjeron 20 imágenes con el mismo *prompt* y semilla (*seed*), y diferentes valores de CFG Scale en todo su rango de los cuales la Figura 2 muestra sólo las 4 más significativas para mayor legibilidad.

La variación del parámetro CFG Scale permite una clara transmutación de la figura de Marianne en la *Estatua de la Libertad*; algo que es posible explicar de manera simple desde el omnipresente icono popular de la *Estatua de la Libertad* en la sociedad contemporánea y por lo tanto en las bases de datos usadas en el entrenamiento del algoritmo lo cual introduce un sesgo (Zhou et al., 2024) en el entrenamiento que pondera el resultado hacia dicho icono frente a la obra original de Delacroix. A ello se debe la asociación de las imágenes con símbolos estadounidenses; sin embargo, también es posible apelar a la relación de la *Estatua de la Libertad* con el pueblo francés e incluso con los valores que el autor Frédéric Auguste Bartholdi quiso representar en su obra (inspirados quizá por la de Delacroix). La representación de Marianne converge, al igual que la *Estatua de la libertad*, en la diosa romana Libertas; donde el gorro frigio (símbolo revolucionario) es sustituido por una diadema o con la antorcha con la que finalmente Bartholdi dotó a su imagen, la cual daría el nombre original: *La Libertad Iluminando el Mundo* (*Liberty Enlightening the World*).

Marianne no sólo converge con la *Estatua de la Libertad* (*Lady Liberty*) en la Figura 3b; también con el icono olvidado del personaje Columbia, uno de los primeros íconos patrióticos de Estados Unidos, conocido como *Miss Liberty* o *Lady Liberty*, desarraigado de la cultura popular con la llegada de la *Estatua de la Libertad*. Columbia tomó, muy probablemente, de la propia Marianne, el gorro frigio, la bandera, la espada, y las vestimentas greco-romanas habituales de estas personificaciones con origen en deidades de las culturas clásicas. Es importante también reflexionar sobre el uso del color y las texturas. Para *CFG Scales* bajos se obtienen alusiones explícitas a la paleta y técnica pictórica de Delacroix. Para *CFG Scales* altos, las imágenes adquieren una estética Pop en alusión, muy probablemente, a la cartelería propagandística estadounidense; como el conocido dibujo *Columbia calls-Enlist now for U.S. Army*, diseñado por Frances Adams Halsted (<https://www.loc.gov/resource/ppmsca.50012/>) en la Figura 3c para, curiosamente, 'guiar al pueblo a la guerra'.

Resumiendo, Stable Diffusion es capaz de pasar de la memorización de una obra concreta a una generalización de ideas conectando iconos populares mediante una asociación simbólica e icónica con una cierta pertinencia, tal y como se muestra en la figura 3b y 3c. La asociación entre ideas o *prompts* con los símbolos que produce está condicionada por las asociaciones entre símbolos e ideas (significantes y significados) etiquetadas en las bases de datos de entrenamiento, así como la abundancia de su representación. Estas relaciones son convenciones sociales propias del pulso mediático o modas (*trends*) de cada época o incluso de cada momento que conforman el lenguaje simbólico de cada periodo cultural independientemente de los hechos; es decir, como en el caso de la obra de Delacroix, su asociación reiterada con ciertos valores de la Revolución Francesa establece una asociación al evento más allá del rigor histórico.

Conclusiones

Las imágenes que producen IAGs como Stable Diffusion codifican de manera simbólica el mensaje transmitido en el *prompt*; por ello la apariencia formal de las imágenes se revela ante nuestra

cognición como una posible obra de arte, sin embargo, las imágenes resultantes de la IA, no son una obra de arte *per se*; sino representaciones de ideas (significados) codificadas en términos simbólicos (significantes). La intención creadora o voluntad del sujeto (autor) que utiliza la herramienta (IA) se concentra en la idea, no en su representación; quizá la herencia perfecta de lo postmoderno y quizá también un reto para la gestión de derechos de autor (según sus mecanismos tradicionales de protección).

Conciencia, intencionalidad, deseo, voluntad, intuición, inspiración, etc., son fuerzas determinantes del impulso creador del ser humano y en consecuencia del arte; sin embargo, por ahora las IAs carecen de todas ellas y no está nada claro que las necesiten o se las permitan.

El *prompt* es fundamental para entender la representación propuesta por la IA; el objeto de análisis, desarrollo o prospección se hace a través de la sintaxis del *prompt*; el cual es, en cierta manera, la substancia que da lugar a la representación y a la singularidad de la obra a través de un determinado algoritmo o modelo particular y también en dependencia del conjunto de entrenamiento.

La respuesta a la pregunta: ¿qué es arte y qué no? está ligada a los diferentes periodos históricos, a la manera en la que los artistas lo conciben y trabajan, a múltiples y complejos intereses que superan lo estético, e incluso lo ético, etc. Aun así, el arte entendido como forma de comunicación está siempre contaminado por un componente vivencial ligado a lo cultural que no puede ser sustituido por una máquina.

La fotografía liberó a la pintura de la figuración extrema, permitiéndole evolucionar hacia formas más expresivas y experimentales. La IA es un paso más en la relación del arte con la tecnología, el instrumento con el que el artista doblaba la materia, la energía y la información a su voluntad. En la medida en que la imagen fotográfica se desconecta de la noción de 'lo real' (realidad), su condición de documento se desvanece, el signo destruye al referente (Baudrillard 1978 p.13). Los algoritmos y la instrumentalización

Figura 3: Aproximación icónica de las imágenes generadas a diferentes CFGs con relación a imágenes de entrenamiento.
Elaboración Propia.



(a) A la izquierda: Eugène Delacroix. 1830. *La Libertad guiando al pueblo*. Óleo sobre lienzo.

Fuente: https://es.wikipedia.org/wiki/La_Libertad_guiando_al_pueblo. A la derecha: Imagen generada por Stable Diffusion para CFG Scale: 2.



(b) A la izquierda: Estatua de la libertad. New York.

Fuente: https://es.wikipedia.org/wiki/Estatua_de_la_Libertad. A la derecha. Imagen generada por Stable Diffusion para CFG Scale: 12.



(c) A la izquierda: dibujo de Frances Adams Halsted. C1916. *Columbia calls-Enlist now for U.S. Army*.

Fuente: (<https://www.loc.gov/resource/ppmsca.50012>). A la derecha Imagen generada por Stable Diffusion para CFG Scale: 20.

del documento ‘desadhieren’ el referente de la imagen, el noema barthesiano del ‘ha sido’ se sustituye por el ‘quizás fue’; es decir, la imagen no ofrece una realidad, sino una posibilidad. Las imágenes IAGs no destruyen ni pierden el referente pues, como simulacros que son, nunca han tenido apego a la realidad, no pertenecen a la categoría de las apariencias, sus referentes son exclusivamente modelos computacionales e ideas en forma de *prompts*. Cuando la fotografía deja de servir a la realidad a través de las imágenes para servir a las ideas, la fotografía como instrumento que vincula lo representado con su representación pierde su sentido operativo. La IA, en este sentido, opera de esta manera asociativa desde su génesis.

Por tanto, la confrontación con las IAGs ya no está tanto en la cuestión de la realidad o la verdad, sino en ocultar que la realidad ya no es realidad para salvar el principio de realidad (Baudrillard 1978 p.26); es decir, las IAGs no suponen un cambio de paradigma en la imagen, sino otro paradigma más; tampoco son un cuestionamiento de la realidad, sino que son un cuestionamiento a la instrumentalización de la imagen como medio para justificar lo real.

Las tecnologías sólo son extensiones del hombre, para uso de los hombres. Si su uso es correcto o no depende del juicio ético de los hombres. Nada surge de la nada: ni las imágenes que producen los hombres, ni las imágenes que producen las máquinas. Si la calidad de estas imágenes es buena o mala, o si deben ser inscritas o no en la historia del arte, depende del juicio estético de los hombres. Si la representación de estas imágenes es verdadera o falsa depende del juicio lógico de los hombres. La inteligencia artificial, en la medida que incorpora aspectos del comportamiento humano, permitirá producir imágenes aún más increíbles y guiará al pueblo adonde quiera llegar.

Mucha gente pensará que crear un *prompt* no es ser artista y quizá tenga razón, un *prompt* no es suficiente. Mucha gente pensó que elegir un urinario no era ser artista, un *ready-made* no es suficiente. Pero no se trata de crear un *prompt* o elegir un *ready-made*, se trata de crear un sistema de producción coherente, intencional, reconocible; se trata, en definitiva, no de objetos, sino de procesos.

Referencias bibliográficas

- Barthes, R. (1973). Texte (théorie du). *Encyclopaedia universalis*, 15, 1013-1017.
- Barthes, R. (2002). *El susurro del lenguaje: Más allá de la palabra y de la escritura*. Paidós. Baudrillard, J., Vicens, A., & Rovira, P. (1978). *Cultura y simulacro*. Kairós Barcelona.
- Bergson, H. (2016). *Materia y memoria: Ensayo sobre la relación de cuerpo con el espíritu* (2.ª ed.). CACTUS.
- Carlini, N., Hayes, J., Nasr, M., Jagielski, M., Sehwan, V., Tramèr, F., Balle, B., Ippolito, D., & Wallace, E. (2023). *Extracting Training Data from Diffusion Models* (arXiv:2301.13188). arXiv. <http://arxiv.org/abs/2301.13188>
- Chomsky, N., Roberts, I., & Watumull, J. (2023). *Noam Chomsky: The False Promise of ChatGPT*. The New York Times. <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>
- Genette, G., & Prieto, C. F. (1989). *Palimpsestos*. Taurus Madrid.
- McClulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in neurons activity. *Bulletin of mathematical biophysics*, 5(115-133), 10.
- Morales, L. G. (2020). *Filosofía de la Restauración: Más allá de las cosas* (Vol. 2). BOD GmbH DE.
- Qureshi, M. A., Deriche, M., Beghdadi, A., & Amin, A. (2017). A critical survey of state-of-the-art image inpainting quality assessment metrics. *Journal of Visual Communication and Image Representation*, 49, 177-191.
- Riffaterre, M. (1980). La trace de l'intertexte. *Pensée (La) Paris*, 215, 4-18.
- Riffaterre, M. (1984). Intertextual Representation: On Mimesis as Interpretive Discourse. *Critical Inquiry*, 11(1), 141-162.
- Somepalli, G., Singla, V., Goldblum, M., Geiping, J., & Goldstein, T. (2022). *Diffusion Art or Digital Forgery? Investigating Data Replication in Diffusion Models* (arXiv:2212.03860). arXiv. <http://arxiv.org/abs/2212.03860>

Suñol, V. (2012). *Más allá del arte: Mímesis en Aristóteles*. Universidad de La Plata.

Yeom, S., Giacomelli, I., Fredrikson, M., & Jha, S. (2018). *Privacy Risk in Machine Learning: Analyzing the Connection to Overfitting* (arXiv:1709.01604). arXiv. <http://arxiv.org/abs/1709.01604>

Zhou, M., Abhishek, V., Derdenger, T., Kim, J., & Srinivasan, K. (2024). *Bias in Generative AI*. <https://arxiv.org/abs/2403.02726>

BIO



José Manuel Pereira Uzal es Doctor en Bellas Artes por la Universidad Complutense de Madrid, Máster en Valoración, Gestión y Protección del Patrimonio Cultural de la Universidad de Vigo, Posgrado de Imagen Científica por la Universidad de Alcalá, Máster en Documentación Audiovisual por la UC3M y el Instituto de RTVE y graduado en Conservación y Restauración de Bienes Culturales por la ESCRBC de Galicia. Desarrolla su actividad profesional como profesional independiente, investigador y desarrollador en el ámbito de la digitalización de bienes culturales, así como ejerce de formador para diversas instituciones. info@jpereira.net



Lino García-Morales es Graduado en ingeniería en Control Automático, Máster en Sistemas y Redes de Comunicaciones, en la especialidad de Procesamiento Digital de la Señal y Doctorado en Cancelación de Ecos Multicanal, ambos en la Universidad Politécnica de Madrid (UPM). Su fascinación por el mundo del arte y lo digital lo ha llevado a obtener el Doctorado en Conservación y Restauración de Arte Digital en la Universidad Europea de Madrid así como a mantener una estrecha relación con el ámbito de la conservación del patrimonio y la producción artística, colaborando con proyectos como Grado en Arte Digital para la UEM, el Magíster en Conservación-Restauración de Arte Contemporáneo para la Universidad Complutense de Madrid (UCM) y el Museo Nacional Centro de Arte Reina Sofía (MNCARS), Grado en Diseño Digital para la Universidad Internacional de la Rioja (UNIR) y el Magíster en Acústica Forense de la UPM, el Máster Universitario en Patrimonio Cultural en el Siglo XXI: Gestión e Investigación así como la producción de obras de Hans Haacke o el proyecto “Viaje al interior del Guernica”. Actualmente compagina la docencia en la Universidad Politécnica de Madrid con la investigación y cursa el Doctorado en Filosofía en la Universidad Autónoma de Madrid. linogarcia@upm.es