

2024 | Junio

Inteligencia artificial y desinformación

Una guía teórico-práctica para profesores de periodismo y comunicación

Índice

Introducción: objetivos de esta guía	3	Ir a página
Capítulo 1: Nociones básicas sobre inteligencia artificial para periodistas	5	Ir a página
• Nociones básicas sobre inteligencia artificial	5	Ir a página
› Inteligencia artificial tradicional y generativa: diferencias	5	Ir a página
› Datos principales de la historia y el funcionamiento de la IA	6	Ir a página
› Qué son las redes neuronales artificiales y el machine learning en la IA	8	Ir a página
› Qué son y qué determinan los algoritmos	9	Ir a página
› Qué es el ChatGPT y qué son los prompts	10	Ir a página
› Alucinaciones de IA	10	Ir a página
• Inteligencia artificial y periodismo	11	Ir a página
› Usos de la IA para asistir y mejorar el trabajo de los periodistas	11	Ir a página
› Qué son las cajas negras de la inteligencia artificial y por qué requieren interés periodístico	12	Ir a página
• Resumen del capítulo: principales puntos a tener en cuenta	13	Ir a página
• Actividades prácticas para desarrollar en el aula: experimentar y reflexionar sobre IA	14	Ir a página
Capítulo 2: Inteligencia artificial y desinformación	16	Ir a página
• Por qué la inteligencia artificial es un factor clave para la creación y la propagación de desinformación	16	Ir a página
› Imágenes creadas con IA	17	Ir a página
› Deepfakes en audio y video	21	Ir a página
• Herramientas para verificar texto, imágenes y videos	24	Ir a página
› Herramientas para texto	24	Ir a página
› Herramientas verificar imágenes	25	Ir a página
› Herramientas para verificar videos y Deepfakes	26	Ir a página
• Los sesgos informativos de la IA	28	Ir a página
• Resumen del capítulo: principales puntos a tener en cuenta	29	Ir a página
• Actividades prácticas para desarrollar en el aula: verificar desinformación generada con IA	30	Ir a página
Lecturas para ampliar	34	Ir a página

Inteligencia artificial y desinformación

Una guía teórico-práctica para profesores de periodismo y comunicación.

Autores (por orden alfabético): Sebastián Di Doménica, Nira Dinerstein, Ana Laura García Luna, Milena Rosenzvit, Cecilia Vázquez.

Introducción

Desde **Chequeado** impulsamos proyectos de formación continua para profesores, periodistas, y comunicadores de América Latina y el Caribe. En este contexto desarrollamos esta guía dirigida a profesores universitarios de carreras de Periodismo y Comunicación con dos objetivos, por un lado, proporcionarles algunos lineamientos teóricos sobre aspectos fundamentales relacionados con el vínculo entre la inteligencia artificial y la desinformación en el periodismo y por el otro, facilitarles algunas propuestas e ideas de actividades para implementar estas temáticas en el aula.

En el primer capítulo recorreremos nociones básicas sobre inteligencia artificial con un enfoque pensado para profesores de periodismo y compartimos algunas posibilidades de implementación del tema en el aula. En el segundo, describimos las instancias de creación de desinformación con utilización de inteligencia artificial, las diferentes herramientas existentes para verificar la desinformación creada con IA y proponemos también algunos ejercicios para realizar con los estudiantes.

La inteligencia artificial llegó para quedarse, y ya está presente en una innumerable cantidad de actividades productivas y cotidianas de las personas y las instituciones. Por supuesto también ya se utiliza en gran medida en las redacciones de medios de comunicación de todo el mundo, para agilizar y desarrollar numerosas tareas que antes insumían mucho tiempo y que hoy se pueden hacer de manera automática en segundos.

En 2022 el lanzamiento del motor de IA ChatGPT de OpenAI, que funciona como un bot que dialoga con los usuarios, expuso las inmensas posibilidades de la inteligencia artificial generativa para desarrollar todo tipo de acciones basadas en el lenguaje.



Pero así como la inteligencia artificial trajo (y traerá) diversos beneficios y posibilidades para mejorar o automatizar muchos trabajos humanos, también plantea algunos desafíos. Por ejemplo, la amenaza de la multiplicación de contenidos desinformantes, ya que la IA generativa puede sistematizar la producción y distribución de engaños de manera altamente eficiente.

Esta problemática no ha pasado desapercibida a nivel global. En noviembre de 2023, durante el lanzamiento del Alto Grupo Asesor sobre Inteligencia Artificial de la ONU, Antonio Guterres, Secretario General de las Naciones Unidas enumeró los posibles daños de la IA, expresando preocupación por la desinformación y la información falsa, el afianzamiento del sesgo y la discriminación, la vigilancia y la invasión de la privacidad, el fraude y otras violaciones de los derechos humanos. El Foro Económico Mundial la ha destacado como uno de los principales riesgos en su lista del año 2024, advirtiendo sobre el uso malicioso de la IA para propagar desinformación y sus potenciales consecuencias.

Entonces, así como la inteligencia artificial ofrecerá muchos beneficios para asistir a la tarea periodística, también ampliará las posibilidades de multiplicar los contenidos desinformantes. En este contexto, el periodismo puede desempeñar un papel central y el espacio de formación en las aulas universitarias es crucial para que los futuros periodistas adquieran habilidades y herramientas que les permitan verificar y alertar a la comunidad sobre las informaciones falsas generadas por IA, ya sea en forma de texto, imágenes o *deepfakes* en audio y video.

Esta guía ha sido elaborada gracias al intercambio permanente con diferentes actores de la educación y del periodismo. Agradecemos profundamente sus aportes y experiencias, y buscamos fomentar una conversación activa sobre su uso. Valoramos sus comentarios, sugerencias y cualquier tipo de retroalimentación que nos permita mejorar y adaptar nuestros recursos a las necesidades cambiantes de la comunidad educativa. Los y las leemos en: educacion@chequeado.com.

Capítulo 1

Nociones básicas sobre inteligencia artificial para periodistas

Nociones básicas sobre inteligencia artificial

Inteligencia artificial tradicional y generativa: diferencias

La inteligencia artificial (IA) es un desarrollo tecnológico revolucionario: un avance que cambia para siempre las opciones para interactuar con la tecnología; y que abre un mundo de posibilidades inimaginables.

Pero es importante diferenciar entre lo que es referido como IA tradicional y una de las tendencias emergentes más disruptivas en el campo: la IA generativa.

La IA tradicional se utiliza desde hace muchos años en innumerables áreas para automatizar tareas. Por ejemplo en operaciones bancarias, en gestión de servicios o en el funcionamiento de las redes sociales. Es decir, la IA tradicional se enfoca en realizar tareas específicas utilizando algoritmos y reglas predeterminadas, realizando predicciones basadas en patrones detectados a partir del análisis de grandes conjuntos de datos, que en su mayoría se encuentran etiquetados y categorizados, mientras que la IA generativa puede trabajar con volúmenes de datos aún más grandes que no están etiquetados ni categorizados, como el caso de las obras completas de Shakespeare o millones de sitios web.

La inteligencia artificial generativa representa una verdadera evolución en la creación de contenido original y creativo. A diferencia de la IA tradicional, esta nueva forma de IA tiene la capacidad de generar textos, imágenes, audio e incluso código de computadora.

Uno de los ejemplos más conocidos de IA que genera texto es el chatbot ChatGPT de OpenAI, un modelo capaz de producir textos de manera similar a como lo haría un ser humano. También existen otros generadores de texto como Bing de Microsoft y Gemini de Google.

Las implicaciones prácticas de esta tecnología son enormes. La IA generativa está velozmente automatizando en el mundo laboral muchas acciones productivas que hoy realizan las personas. Eso ya ocurre en el periodismo.

Sin embargo, así como la IA ofrece nuevas y valiosas oportunidades para lograr mejores resultados periodísticos también será una herramienta de utilidad para aquellos que quieran sistematizar mecánicas de producción y propagación de engaños y desinformación.

Tanto la IA tradicional como la generativa desempeñarán un papel crucial en la sociedad digital por venir. Reconocer las capacidades únicas de cada forma de IA nos permitirá aprovechar al máximo su potencial e impulsar la innovación en los diferentes sectores. Sin embargo también es decisivo que los estudiantes de periodismo descubran los riesgos que posee la IA generativa y su impresionante poder de producción.

A medida que la IA avanza y se extiende a todas las áreas surgen debates y tensiones en torno a los marcos éticos y legales para un uso responsable.

Al margen de esos riesgos y dificultades, la inteligencia artificial, en todas sus formas representa un avance significativo en nuestra capacidad para automatizar tareas complejas, mejorar la eficiencia y fomentar la innovación en todas las facetas de nuestra vida diaria.

Datos principales de la historia y el funcionamiento de la IA

¿Por qué Alan Turing está considerado el padre de la IA? Alan Turing (1912 – 1954) fue un matemático británico que a inicios de la década del 50 sentó las bases informáticas de lo que se convertiría en la inteligencia artificial (IA). Considerado el padre teórico de la IA, Turing propuso el famoso test que lleva su nombre, una prueba diseñada para evaluar la capacidad de una máquina para imitar respuestas humanas en una conversación.

Su objetivo era determinar si una máquina podía pensar como una persona, y predijo que 50 años más tarde sería posible programar computadoras para superar este test. En la actualidad su visión se ha convertido en una realidad a través de los chatbots con IA, como ChatGPT.

Pero la predicción del científico se hizo realidad en 2014, cuando un equipo de investigadores logró un avance significativo al presentar la primera IA que pasó el test de Turing. Este logro marcó un hito en la historia de la IA y demostró que las máquinas son capaces de imitar respuestas humanas de manera convincente.

Desde los primeros planteos teóricos de Turing hasta la actualidad fueron muchos los avances y desarrollos que marcaron el camino hacia la IA del presente.

A continuación una breve cronología de ese recorrido:

- 1950** ● Alan Turing propone el test de Turing, un hito en la definición de la inteligencia artificial (IA) y su capacidad para imitar respuestas humanas en una conversación.
- 1956** ● En la conferencia Dartmouth Workshop (Hanover, New Hampshire, EEUU), se acuña el término "Inteligencia Artificial" y se establecen los fundamentos de la disciplina.
- 1966** ● ELIZA (Instituto Tecnológico de Massachusetts), un programa de computadora se destaca como una de las primeras IA en simular una conversación humana, origen de los chatbots modernos.
- 1972** ● PROLOG se introduce como un lenguaje de programación especialmente diseñado para la IA, facilitando el desarrollo de sistemas basados en reglas lógicas.
- 1997** ● Victoria de Deep Blue, la supercomputadora de IBM, sobre el campeón mundial de ajedrez Garry Kasparov

- 1997** ● Victoria de Deep Blue, la supercomputadora de IBM, sobre el campeón mundial de ajedrez Garry Kasparov
- 2011** ● IBM Watson compite y gana contra campeones humanos en el programa "Jeopardy", y demuestra su capacidad para procesar grandes cantidades de información y proporcionar respuestas precisas.
- 2014** ● Test de Turing superado; ya que por primera vez un chat informático logró convencer a más del 30% de los jueces de que era "genuinamente humano"
- 2019** ● GPT-2 de OpenAI se destaca al generar texto prácticamente indistinguible del escrito por humanos, abriendo nuevas posibilidades en la generación de contenido.
- 2022** ● GPT-3 de OpenAI revoluciona la generación de texto, código y poesía con calidad similar a la humana, impulsando aplicaciones en diversos campos.
- 2023** ● OpenAI Codex asiste a los programadores al generar código automáticamente, agilizando el proceso de desarrollo de software.
- 2024** ● El desarrollo de GPT-4 de OpenAI, que permite la entrada de imágenes además de texto, y Google Bard de Google, que ofrece una alternativa en el mundo de los chatbots, continúan impulsando la IA hacia adelante.

Qué son las redes neuronales artificiales y el machine learning en la IA

Un modelo de inteligencia artificial (IA) se basa en algoritmos y técnicas matemáticas que permiten a las máquinas procesar información, extraer patrones, tomar decisiones y realizar una variedad de tareas. Como resultado del entrenamiento en la lectura de millones de datos, los motores de IA predicen a partir de métodos estadísticos y probabilísticos la sucesión posible de palabras según el contexto de una frase en un texto. Los motores de IA no serían viables sin el acceso a la inmensa cantidad de datos e información disponible en la web o en otros medios con los que fueron entrenados.

Algunos modelos pueden aprender de los datos generados durante su uso, es decir, retroalimentarse y mejorar su desempeño con la experiencia, lo que los hace capaces de adaptarse a cam-

bios y realizar mejoras constantes. Las redes neuronales artificiales son un tipo de modelo computacional inspirado en la estructura del cerebro humano. Están compuestas por capas de nodos interconectados, llamados neuronas artificiales, que procesan y transmiten información a lo largo de la red. Cada conexión entre las neuronas se ajusta durante el entrenamiento: las redes neuronales son capaces de incorporar como dato sus propias respuestas, y “aprender” a realizar tareas específicas sin ser explícitamente programadas para ello. Esto es lo que conocemos como aprendizaje automático o machine learning.

A su vez, el aprendizaje profundo o *deep learnin* es, una subdisciplina del machine learning que involucra redes neuronales profundas, con múltiples capas y tiene mayor capacidad de aprendizaje, lo que otorga gran poder para tareas especialmente complejas como el procesamiento de lenguaje natural, reconocimiento de imágenes y otras que estamos viendo en los últimos avances de estos meses.

Qué son y qué determinan los algoritmos

Un algoritmo es un conjunto de instrucciones o reglas que se siguen para resolver un problema o realizar una tarea específica. En la inteligencia artificial juegan un papel fundamental porque son los programas informáticos que utilizan datos para aprender, tomar decisiones y realizar tareas de manera autónoma.

Los algoritmos que se utilizan en el campo de la Inteligencia Artificial son de los más complejos, y refieren a conjuntos de instrucciones que realizan operaciones matemáticas sobre los datos para obtener así la capacidad de identificar patrones y tomar decisiones. Estos algoritmos suelen tener una alta complejidad.

También se los considera el núcleo de la inteligencia artificial, ya que son responsables de procesar la información, aprender de ella y tomar decisiones basadas en esos datos. La elección del algoritmo adecuado es crucial para el desarrollo y el desempeño de los sistemas de inteligencia artificial.

- | | |
|--|---|
| <ul style="list-style-type: none">• El ChatGPT es un motor de inteligencia artificial que ha ganado popularidad mundial desde su versión gratuita 3.5 en 2023.• Funciona a través de un bot que permite a los usuarios interactuar con él mediante un chat conversacional, respondiendo preguntas y realizando tareas como la escritura de textos.• Aunque es un ejemplo revolucionario de inteligencia artificial generativa basada en lenguaje, su creciente uso plantea desafíos, especialmente en el periodismo y otras actividades que requieren información precisa y confiable.• Si bien puede automatizar tareas y acelerar procesos, carece de la capacidad para mencionar fuentes de información y analizar el contexto, lo que podría afectar la profundidad y precisión de los artículos generados. | <ul style="list-style-type: none">• Prompt es otra de las palabras a tener en cuenta en la comprensión de los motores de IA generativa porque para poner a nuestro servicio a los modelos de Inteligencia artificial tendremos que saber hacer buenas preguntas y dar órdenes e instrucciones precisas. Todas las que sean necesarias para responder a nuestra búsqueda y objetivos.• El Prompt es una construcción o un conjunto de palabras que tiene como fin lograr las respuestas o la generación de contenidos de parte de un software de inteligencia artificial.• Por ejemplo: redactar una noticia simple con estructura de pirámide invertida que incluya un accidente de autos y tres heridos. |
|--|---|

Qué es el ChatGPT y qué son los prompts

Alucinaciones de IA

Las alucinaciones son otro de los puntos vulnerables de los motores de IA generativa, y a tener en cuenta cuando está al servicio del trabajo periodístico.

Se producen cuando ante una pregunta o una orden de un usuario el modelo de IA genera desinformaciones, información falsa, incorrecta, incoherente o inventada. Estas alucinaciones pueden

manifestarse en forma de respuestas erróneas (hechos, personas o historias que no existen), imágenes distorsionadas (personas con un solo ojo o con manos de ocho dedos), o datos generados de forma imprecisa.

Algunas de las posibles causas de las alucinaciones ocurren cuando la IA se entrena con datos limitados o no representativos, y de esa manera aprende patrones incorrectos y también a falta de contexto que se puede plantear cuando el modelo no logra descifrar de manera completa el marco o la naturaleza del problema o temática que aborda.

Por eso, en el periodismo y también en toda actividad en la que se requiera la veracidad de los datos, al utilizar los motores de IA será necesario implementar técnicas de validación y comprensión profunda del problema que se aborda.

Inteligencia artificial y periodismo

Usos de la IA para asistir y facilitar el trabajo de los periodistas

La integración de la Inteligencia Artificial al periodismo ya está en pleno avance. Y plantea una evolución muy significativa en la forma de gestionar el trabajo periodístico en los medios de comunicación.

La IA al servicio del periodismo se utiliza o se puede utilizar enfocada en varias áreas: producción y gestión del trabajo periodístico, definición y diseño de producto según tendencias de consumo o distribución para ampliar audiencia.

Entre otras muchas cosas la IA ya se usa para la traducción, desgrabación de entrevistas, armado de resúmenes de documentos, escritura de borradores de texto, selección o creación de imágenes, búsqueda de fuentes, o automatización de tareas.

También se utiliza para redactar propuesta de titulares efectivos, edición según intereses de audiencias, estrategias SEO (optimización para motores de búsqueda, en español), relevamiento de reque-

rimientos de suscriptores, predicción de nuevos temas de interés o para hacer textual contenidos en audio. Por último, con IA se desarrollan estrategias en redes y en otros canales de distribución.

Más allá de estas inmensas posibilidades, es crucial no perder de vista el verdadero propósito del periodismo: informar con precisión para que las sociedades puedan tomar decisiones fundamentadas. En un contexto donde la desinformación puede propagarse como nunca antes en la historia, la responsabilidad de los medios de comunicación será aún mayor. La IA puede ser una aliada poderosa en esta lucha, pero nunca debe reemplazar el juicio humano y la ética periodística.

En el caso de Chequeado, desde 2016, desarrollamos tecnología en base a inteligencia artificial que nos ayuda a chequear más rápido y mejor. Nuestro Chequeabot, centro de aplicaciones basadas en inteligencia artificial, evoluciona día a día y ya es usado en 7 países. Actualmente utilizamos IA para realizar monitoreos, sistemas de alertas, identificación de frases chequeables, detección de datos, entre otros usos. Toda esta información se integra en una base de datos estructurada, que no solo registra los chequeos realizados, sino también el proceso, las fuentes y los contenidos.

Qué son las cajas negras de la inteligencia artificial y por qué requieren interés periodístico

Para finalizar, hay que remarcar que las redes neuronales artificiales son consideradas las “cajas negras” de la inteligencia artificial, ya que su complejidad y cantidad de parámetros hacen que sea difícil interpretar cómo llegan a sus decisiones. A menudo, los procesos de una red neuronal no pueden explicarse fácilmente en términos humanos, lo que lleva a que se perciban como cajas negras en las que se introduce información y se obtiene una salida sin comprender por completo el proceso interno de toma de decisiones que llevó a las respuestas a las solicitudes de los usuarios de IA.

Por eso será una tarea central del periodismo indagar en el funcionamiento de la inteligencia artificial, y descubrir los secretos de esas cajas negras. El objetivo será transparentar el funcionamiento de una tecnología que será cada vez más relevante en las diferentes áreas productivas y en la vida de todas las personas.



Resumen del capítulo: principales puntos a tener en cuenta

- La inteligencia artificial es un avance tecnológico revolucionario.
- La inteligencia artificial tradicional analiza datos y realiza predicciones basadas en patrones existentes (a partir de algoritmos).
- La inteligencia artificial generativa va más allá y crea nuevos datos y productos que surgen del entrenamiento (con sistemas predictivo, machine learning y redes neuronales artificiales).
- ChatGPT es un bot que puede responder preguntas de todo tipo y resolver solicitudes de los usuarios (a partir de prompts).
- ChatGPT es uno de los modelos de lenguaje más grandes y avanzados en la actualidad entrenado a partir de millones de conversaciones.
- La IA generativa ofrece posibilidades inmensas para automatizar o acelerar todo tipo de tareas incluidas algunas tareas periodísticas, y también genera riesgos como la posibilidad de multiplicar contenidos desinformantes.
- ChatGPT no expone fuentes, carece de análisis de contexto y también puede ofrecer información errónea (alucinaciones).



Actividades prácticas para desarrollar en el aula: experimentar y reflexionar sobre IA

A partir de las siguientes actividades, proponemos hacer una puesta en común con los estudiantes respecto de los posibles usos de **ChatGPT**: generación de ideas, búsqueda de información, corrección ortográfica y gramatical, traducción, resúmenes, resolución de problemas, entre otros, y evaluar la utilidad y riesgos de estos usos.

✓ Utilizar al ChatGPT como buscador de información: pedirle que presente los principales datos de la vida de un personaje reconocido a nivel mundial. Luego contrastar la información ofrecida por ChatGPT con otras fuentes obtenidas en Internet.

✓ Utilizar al ChatGPT como buscador de fuentes: pedirle que te recomiende fuentes primarias sobre alguna temática y luego comparar y realizar la búsqueda de los documentos originales.

✓ Utilizar a ChatGPT como herramienta de redacción: pedirle que con los principales datos de la vida de un personaje, escriba una nota periodística de seis párrafos y que tenga un título que capte el interés del lector. Analizar los aciertos y errores en el resultado.

✓ Utilizar a ChatGPT como herramienta de redacción y jerarquización de criterios de noticiabilidad: suministrarle una serie de datos desordenados que describan un hecho policial, (por ejemplo, un choque entre dos autos, en el que fallece un conductor y tres testigos participan de la escena llamando a la policía, etc.) y solicitarle que redacte con esa información una noticia con título, bajada y cuerpo respetando los criterios de noticiabilidad. Esa nota puede compararse con las que escribieron previamente los estudiantes y debatir sobre la importancia de la jerarquización.

✓ Pedirle a ChatGPT que asuma un rol en el momento de redactar: solicitarle que escriba como si fuera un periodista una serie de preguntas para realizarle una entrevista periodística al actual presidente de un país a elección.

✓ Obtener información y enfoques temáticos con ChatGPT: formularle una pregunta que la respuesta implique una valoración subjetiva. Por ejemplo: ¿cuál es la mejor ciudad para vivir en Argentina? Repetir con otras preguntas del mismo estilo. Evaluar los criterios utilizados para llegar a las respuestas. Buscar en Google información relacionada para reconocer posibles fuentes utilizadas por el motor de IA. Por último compartir en clase los resultados obtenidos y definir en conjunto si tienen relevancia o rigurosidad para ser incluidas en una nota periodística.

✓ En el caso de querer trabajar con imágenes se puede utilizar Copilot designer, la IA de Bing. Solicitarle la creación de una imagen que contenga un grupo de alumnos en el aula junto al maestro. Verificar la corrección de la imagen con relación a las formas y características de las personas que aparecen (manos, ojos, definición caras, detalles, contexto). A su vez observar si expone algún tipo de sesgo con relación a la representación de las personas incluidas. Por último compartir en clase los resultados obtenidos y definir si son o no publicables en un medio.

Capítulo 2

Inteligencia artificial y desinformación

Por qué la inteligencia artificial es un factor clave para la creación y la propagación de desinformación

La Inteligencia Artificial facilita muchas tareas productivas y diarias de la ciudadanía, y en particular en el periodismo, es un asistente efectivo para mejorar, acelerar y multiplicar tareas.

Sin embargo, la IA también puede facilitar los procesos de creación de contenidos desinformantes, falsos, descontextualizados y los mecanismos existentes para su propagación.

A continuación detallamos algunos ejemplos:

1. **Generación de contenido falso:** los modelos de IA generativa pueden ser utilizados para crear contenidos falsos y construir relatos que parecen auténticos, pero que son inventados en parte o en su totalidad. Personas o grupos pueden aprovechar la capacidad de estas herramientas de generar contenido verosímil en distintos formatos para desinformar con algún fin en particular.

Por otro lado, aún sin la intención de una persona de generar un contenido falso, las mismas herramientas pueden responder erróneamente a consultas de los usuarios, generando desinformación. En el caso de ChatGPT, por ejemplo, no tiene acceso a Internet: por lo tanto no puede proporcionar información en tiempo real, lo cual puede llevar a generar información descontextualizada. Además, al preguntar por fuentes de información, los resultados que ofrece pueden no existir.
2. **Automatización de la producción de falsedades:** los sistemas de IA pueden ser utilizados para crear noticias falsas a gran escala, y al aumentar la cantidad se hace más difícil la diferenciación entre información verdadera y engañosa. Las llamadas “granjas de contenido” son sitios web similares a los periodísticos pero que solo presentan información falsa o de muy baja calidad, y al servicio de los intereses de quienes las arman. No necesariamente se realizan con IA, pero la generación automática de contenido a gran escala

sin dudas puede ayudar a su propósito.

3. Creación de perfiles falsos y bots para distribuir: la IA también puede ser utilizada para crear perfiles falsos en redes sociales y bots que difunden desinformación y propaganda con el objetivo de influir en la opinión pública. Es decir, es una manera de propagar el contenido falso que se genera en las granjas de desinformación. En esta etapa también puede sumarse la utilización de los algoritmos de recomendación y de los criterios de SEO para amplificar el alcance de la desinformación. Por último, los actores desinformantes pueden generar con IA correos electrónicos persuasivos de manera automatizada con la intención de multiplicar el engaño.

Imágenes creadas con IA

A los textos se suman las imágenes creadas con motores de inteligencia artificial generativa, que abrieron un nuevo capítulo de desinformación.

Los motores de IA diseñados para crear imágenes, como **DALL·E** y **MidJourney**, tienen un gran potencial para generar desinformación a través de la manipulación de imágenes preexistentes o la creación de imágenes completamente nuevas, que pueden ser prácticamente indistinguibles de fotos. Estas herramientas utilizan redes neuronales para generar imágenes realistas a partir de descripciones textuales, lo que plantea varias complicaciones en términos de desinformación. Algunas formas en las que estos motores de IA pueden contribuir a la generación de desinformación incluyen:

1. Creación de imágenes falsas convincentes: los motores pueden generar imágenes que representan objetos, escenas con personas o situaciones inexistentes pero visualmente realistas, lo que dificulta discernir entre imágenes auténticas y falsificadas.
2. Manipulación de la realidad visual: la capacidad de alterar y manipular imágenes de manera convincente puede utilizarse para crear escenas engañosas que promuevan narrativas falsas o distorsionen la percepción de eventos reales.

Existen numerosos ejemplos de imágenes falsas de políticos (y otras figuras públicas) y eventos ficticios que han circulado en línea. Estas falsificaciones pueden ser muy convincentes, y requieren de una rápida y robusta verificación para informar a la comunidad sobre el engaño. A continuación algunos de los casos que más impacto causaron en 2023:

1. El falso arresto de Trump

Circularon por Twitter (ahora X) supuestas fotos que mostraban a Donald Trump arrestado por la policía. Sin embargo, las imágenes eran falsas y creadas con inteligencia artificial. El tuit original con las supuestas fotografías fue publicado inicialmente por Elliot Higgins, fundador del sitio de periodismo de investigación Bellingcat.



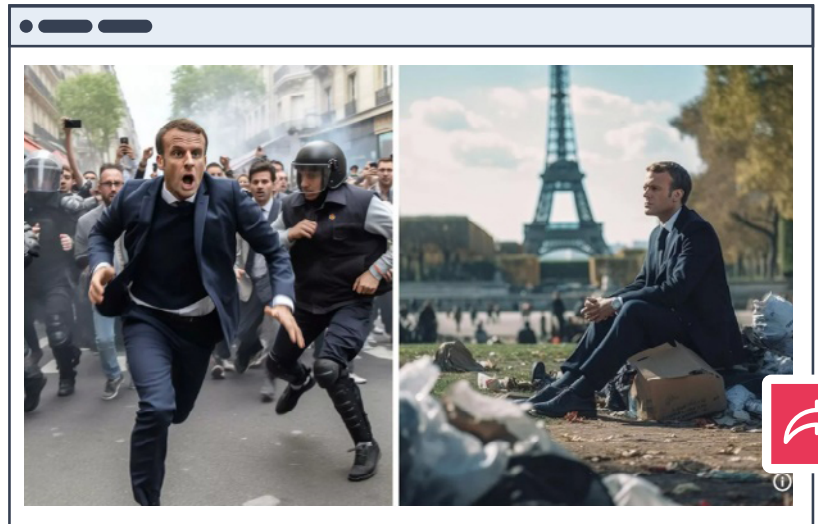
2. Papa Francisco con una llamativa campera blanca

Circuló en redes sociales una falsa imagen del Papa Francisco con una campera blanca. Varios usuarios de redes preguntaron si se trataba de una imagen real. Según la revista Forbes, la imagen que fue tendencia en Twitter se creó utilizando Midjourney, se hizo viral y confundió a innumerable cantidad de personas en todo el mundo.



3. Macron en protestas, arrestado y en un basurero

Se hicieron virales en redes sociales unas imágenes que mostraban al mandatario francés Emanuel Macron como un manifestante en las calles parisinas en el contexto de las protestas contra el gobierno de aquel país por la reforma a la ley de pensiones. Estas imágenes también fueron generadas con la herramienta Midjourney.



Recomendaciones para detectar si una imagen fue creada con inteligencia artificial

La primera recomendación frente a una imagen sospechosa (creada o no con IA) es dudar, evitar compartir e intentar verificar. Y cuando se cree que son imágenes creadas por inteligencia artificial, es clave observar los detalles. Porque aunque las imágenes desarrolladas a través de Dall-E o Midjourney (y otros motores) pueden ser muy precisas y similares a las fotos verdaderas, aún así pueden tener errores o detalles que permitan sospechar que son una creación de IA. De todos modos, luego de la sospecha inicial siempre es necesario hacer un chequeo o verificación. Los pasos del método de Chequeado incluye: buscar el contenido original utilizando herramientas digitales, consultar a distintas fuentes, contextualizar y calificar.

Algunos puntos a tener en cuenta:

- **Marcas de agua:** algunos motores de IA ubican en el margen de la foto una marca de agua, sin embargo, con un programa de edición puede ser borrada o tapada.
- **Dedos, manos o dentaduras:** manos, dedos y dientes son los elementos a observar en detalle al analizar las imágenes que generan dudas. En especial las manos porque a los motores de IA les cuesta imitarlas y es muy común que les sume dedos a la mano o cambie las formas naturales. En las imágenes creadas por inteligencia artificial los dientes no suelen

estar bien alineados o se ven más piezas dentales que las que posee un ser humano.

- **Expresiones exageradas o rostros completamente tersos:** el motor de IA Midjourney suele representar las expresiones solicitadas de forma exagerada, con pliegues de la piel muy pronunciados, ya sea para la sonrisa o para el enojo. A su vez las caras que generan los sistemas de IA pueden mostrar la piel muy tersa, las orejas a diferentes alturas o texturas raras. También, el brillo de los ojos puede ser excesivo en las imágenes creadas con IA.
- **Sombras:** en las imágenes generadas con IA, las sombras pueden tener ángulos errados o una extensión equivocada.
- **Entorno de la imagen:** los defectos del entorno que muestra la imagen, más allá del principal foco de atención, también se deben observar porque pueden tener defectos y errores. Eso ocurre porque la inteligencia artificial pone foco en el objeto central de la imagen y no en lo que lo rodea. Eso hace que en el contexto se puedan observar los detalles y fallas que permiten sospechar que la imagen fue creada de manera artificial.

1 — 8

Qué tener en cuenta para identificar una imagen creada con inteligencia artificial

2 — 8

Circulan imágenes de personas conocidas que parecen creíbles, pero que son generadas con inteligencia artificial.

Te mostramos los casos de Donald Trump y del Papa Francisco y cómo darse cuenta de que las fotos son falsas.

3 — 8

Manos

Las manos parecen un dibujo y además, una de ellas es completamente blanca y con los dedos más cortos.

4 — 8

Manos

A este policía pareciera que le faltara el dedo índice.

5 — 8

Ojos

Este policía pareciera no tener el iris del ojo.

6 — 8

Sombras y texturas

Sombras en lugares inesperados y textura de la piel que no coincide con la edad.

7 — 8

Manos y orejas

Orejas extrañas y detalles inusuales.

8 — 8

Recordá revisar y centrarte en esto para no caer en desinformaciones:

- Manos
- Ojos
- Oreas
- Vello facial
- Sombras
- Texturas

• Buscá si hay detalles inexactos u objetos extraños.

Deepfakes en audio y video

Los *deepfakes* son contenidos creados con tecnología basada en inteligencia artificial que recrean la apariencia y la voz de personas (existentes o inexistentes) de una manera convincente. Los resultados pueden ser extremadamente realistas y difíciles de diferenciar de los que son verdaderos, y plantean desafíos muy significativos en la lucha contra la desinformación y la manipulación.

Los que se crean en video pueden copiar de manera precisa los movimientos faciales y las expresiones de una persona, y colocarla en situaciones o contextos falsos. En tanto, los de audio generan grabaciones de voces sintéticas que imitan (de manera muy precisa) la voz y el habla de una persona real. En estas recreaciones la víctima de la imitación puede decir palabras o frases que nunca dijo.

Las *deepfakes* generan gran preocupación en el ámbito político, porque las falsas imágenes, videos y audios generados por inteligencia artificial pueden distorsionar la realidad y tienen el potencial de afectar a candidatos y procesos electorales, y golpear la confianza en la información. A su vez pueden ser utilizados como herramientas de manipulación política a gran escala y tener un impacto significativo en la opinión pública y en los resultados electorales.

Algunas de las recomendaciones a tener en cuenta frente a posibles *deepfakes* son: verificar la fuente de la información, consultar al involucrado, buscar evidencia adicional que respalde la veracidad de una imagen o video, prestar atención a los detalles de la imagen, buscar inconsistencias en la apariencia física o en los movimientos.

En enero de 2024, se viralizó en redes sociales un [video](#) que muestra a Javier Milei hablando en inglés en el Foro Económico Mundial de Davos, aunque el presidente dio su discurso en español.



No es la primera vez que se viralizan diferentes fragmentos de personalidades que fueron doblados a través de la inteligencia artificial. Desde el futbolista [Lionel Messi](#) hablando en inglés hasta el empresario [Elon Musk, en japonés o alemán.](#)

Los audios falsos son los más difíciles de detectar. Estas falsificaciones son más fáciles y baratas de crear que los videos *deepfakes*, lo que aumenta su potencial para ser utilizados como herramientas de

desinformación. Además tienen menos pistas contextuales detectables al ser escuchados.

Otro aspecto complejo de los audios falsos creados con IA es que pueden ser utilizados en llamadas automáticas para difundir desinformación de manera masiva. Además los audios pueden ser distribuidos de forma simple a través de plataformas de mensajería como WhatsApp, y así pueden llegar a un gran número de personas y generar un impacto significativo.

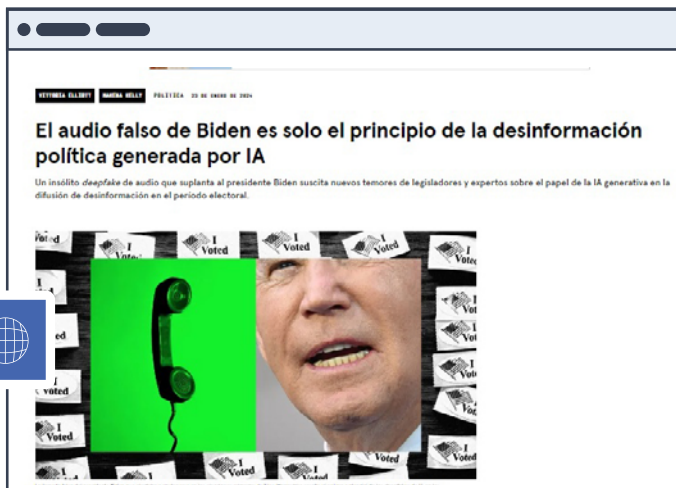
Ejemplo de [deepfake](#) en Argentina

En abril de 2024 circuló en redes sociales un supuesto audio del presidente de los Estados Unidos, Joe Biden, en una entrevista con el comediante Conan O'Brien, en el que hablarían sobre Javier Milei. Pero el audio es falso. No hay registros de que Biden haya hecho esas declaraciones. Ciertos indicios demuestran que la pieza podría estar hecha con inteligencia artificial.

Ejemplo de [deepfake](#) en audio en las elecciones 2024 en EEUU



Durante enero de 2024 en el marco de las primarias de EEUU para la elección presidencial, los votantes de New Hampshire y New England recibieron una llamada automática de una voz que se parecía mucho a la de Biden, y en la que el supuesto presidente y candidato demócrata a la reelección hablaba de la “farsa” de las elecciones y los desanimaba a votar en las primarias del martes. “Los representantes deben apresurarse a establecer medidas de protección, o nos enfrentaremos a un caos electoral. El deepfake de New Hampshire es un recordatorio de las muchas formas en que estos modelos siembran la confusión y perpetúan el fraude”, afirmó a la revista Wired Robert Weissman, presidente del grupo Public Citizen con relación al hecho, cuya falsedad fue informada en primer lugar por la NBC News.



Mirar los detalles

Si bien hay herramientas tecnológicas que pueden ayudar a descubrir un deepfake (en el próximo apartado), al igual que con las imágenes creadas por IA, existen detalles a observar para dudar sobre la veracidad de un video.

Son varias las señales que se pueden detectar en este tipo de contenido manipulado. En un video es importante prestar atención a anomalías visuales, como distorsiones en la apariencia facial o movimientos poco naturales, así como problemas de sincronización labial entre el audio y los movimientos labiales. Además, la identificación de inconsistencias en la calidad de imagen, falta

de expresiones faciales sutiles, el parpadeo (en los videos deepfake la persona parpadea menos de lo normal) y el contexto inapropiado, dentro del video, en comparación con otras fuentes.

La combinación de todas estas estrategias puede ayudar a fortalecer la capacidad de los estudiantes para identificar contenido falso y creado con IA, en un contexto cada vez más confuso y desafiante.

Herramientas para verificar texto, imágenes y videos

Herramientas para texto

Google Fact Check Tools

- **Fact Check Explorer** de Google para temas de artículos que plantean dudas:
Ante un texto que genera sospechas de contener desinformación, el Fact Check Explorer de Google puede servir para poner foco en el tema central de la nota y reconocer los chequeos ya existentes al respecto. Es una herramienta que busca verificaciones. Solo es necesario ingresar palabras clave relacionadas con el tema, el nombre de una persona o incluso una frase literal. Y si hay alguna verificación relacionada, se mostrará junto con el motivo, el veredicto y un enlace al artículo correspondiente.

GPTZero

- **GPTZero** para verificar si un texto fue escrito por un motor de inteligencia artificial:
Existen varias herramientas para analizar si un texto puede haber sido escrito o no con motores de IA como ChatGPT. No plantean resultados cerrados, pero pueden ofrecer señales para fortalecer o reducir las sospechas. Una de esas herramientas es GPTZero. Funciona de la siguiente manera: se recorta y pega el texto sospechoso, y la herramienta plantea una probabilidad sobre si está escrito por ChaGPT o por un humano. Fue creado por Edward Tian, un estudiante de la Universidad de Princeton, y utiliza la medida de “perplejidad” o aleatoriedad del texto como criterio. El mayor nivel en este indicador mostrará

un aumento en el caos de la escritura y por lo tanto una probabilidad mayor sobre su creación con una herramienta como ChatGPT.

Herramientas para verificar imágenes

Si nos encontramos con una imagen cuya autenticidad nos genera dudas porque puede haber sido generada por inteligencia artificial, un primer paso es ver si ya fue publicada, en qué fecha, en qué contexto y por quién.



- [Google Imágenes](#)

Con la extensión de Google se puede subir la imagen, o copiar la URL en el buscador, y explorar los resultados.



- [Tineye](#)

Es una herramienta similar. Se puede subir la imagen, o copiar la URL y explorar los resultados. En algunos casos, podemos encontrar que la imagen corresponde efectivamente a una foto, que fue publicada por la fuente que estaba en duda; o podemos descubrir que la foto fue publicada de manera previa en otro contexto, o bien que la foto que recibimos está editada, manipulada, o que nunca fue publicada porque fue creada con IA. También podemos ver si hay otras imágenes del mismo suceso y compararlas. Esta información nos ayudará a pensar si la imagen es o no una foto real.

HIVE MODERATION

- [Hive Moderation](#)

Es una extensión para Google Chrome que permite detectar imágenes y textos realizados con Inteligencia Artificial (IA). Es una herramienta abierta y gratuita, aunque tiene una versión paga más completa. Detecta imágenes producidas con plataformas que generan imágenes con IA, como Midjourney, DALL-e, o Stable Diffusion.



Hugging Face

- [AI Image Detector](#)

Con esta herramienta se puede obtener una probabilidad de que la imagen de una persona haya sido creada o no con inteligencia artificial. El detector permite cargar o arrastrar y soltar imágenes cuestionables; y luego del análisis expone dos números porcentuales. Uno refleja la probabilidad (en porcentaje) obtenida de que la imagen sea artificial; y otro, la probabilidad de que la imagen refleje a un ser humano real. No ofrece resultados concluyentes pero puede servir para realizar una primera evaluación.

Herramientas para verificar videos y deepfakes

Existen numerosas herramientas que pueden ser de ayuda para intentar verificar un deepfake en audio o video. Sin embargo, al igual que las mencionadas para el caso de las imágenes, nunca son concluyentes, porque plantean análisis de probabilidad. Más allá de eso y frente a la multiplicación de este tipo de desinformación pueden ser de enorme utilidad y son herramientas de uso simple que ofrecen alternativas de chequeo, para los periodistas y también para la comunidad en general.

deepware

- [Deepware scanner](#) (videos)

Es un motor que permite escanear y verificar videos y busca detectar los *deepfakes*. Los videos se deben subir a través de la página y luego el motor realiza el análisis. El resultado se presenta a través de los colores de un semáforo. Si la aguja marca verde señala que no se trata de un deepfake “No *deepfakes* detected”. Mientras que el amarillo plantea dudas y el rojo señala una alta probabilidad de ser un deepfake creada o manipulada a través de inteligencia artificial. Al igual que otros detectores de *deepfakes*, los modelos de deepware buscan signos de manipulación en el rostro humano.

Eleven Labs

- [AI Speech Classifier de EvenLabs](#) (audios)

Es una herramienta gratuita de la empresa ElevenLabs (creadora de un software para emular voces humanas) que busca colaborar con los usuarios en la identificación de audios creados con IA. En realidad ha sido creada en especial para detectar los audios generados con la tecnología de esa empresa. Sin embargo al subir un audio a la página, si el resultado es negativo, y fue creado con voz humana real, señala que es poco probable que haya sido creado con la tecnología de EvenLabs o con otra de creación de voz artificial.

AI OR NOT

- [AI or not](#) (audios)

El motor permite subir imágenes y audios. Cuando se suben audios, realiza un análisis y luego señala si la grabación parece exponer una voz sintética creada por inteligencia artificial o si puede tratarse de voz humana. Es un servicio pago, pero al suscribirse, se pueden realizar 10 verificaciones gratuitas por mes.

InVID

- [Invid](#) para indagar sobre videos sospechosos

La herramienta Invid es gratuita y está disponible como complemento de Chrome y Firefox. Está pensada en primer lugar para periodistas y medios de comunicación y permite obtener información contextual sobre videos de Facebook y YouTube, además de realizar búsqueda de imágenes inversas (activa la búsqueda a partir de cuatro miniaturas extraídas de un video de Youtube). Además sobre los videos de plataformas hace posible fragmentar fotogramas clave, para mejorar y explorar las imágenes a través de una lupa.

Los sesgos informativos de la IA

Los sesgos informativos en los algoritmos es otro de los aspectos que de manera indirecta puede sumar desinformación en los textos o imágenes creados con motores de IA.

Los sesgos informativos en la inteligencia artificial surgen por las distorsiones que ya estaban en los datos, algoritmos o resultados generados y pueden influir en la toma de decisiones y en la presentación de información. Estos sesgos tienen el potencial de perpetuar y amplificar injusticias, discriminación y desigualdades.

Uno de los factores que contribuyen a que existan estos sesgos es el uso de conjuntos de datos no representativos de la población mundial. Al entrenar los modelos de IA generativa con datos limitados y sesgados, se corre el riesgo de reflejar únicamente los puntos de vista de minorías. Esto puede llevar a la exclusión y la discriminación de grupos con menor representación. Además, los programadores también pueden estar expuestos a los sesgos de los algoritmos. Sus propias creencias y prejuicios pueden filtrarse en el diseño y desarrollo de la IA generativa, lo que podría resultar en la reproducción de estereotipos y discriminación en el contenido generado.

Para abordar los sesgos informativos en IA, se deben implementar prácticas y medidas correctivas como por ejemplo auditorías de sesgos, que consiste en realizar evaluaciones sistemáticas para identificar y mitigar sesgos en los conjuntos de datos, algoritmos y resultados generados.



Resumen del capítulo: principales puntos a tener en cuenta

- La IA contribuye a la labor periodística sistematizando algunos procesos, sin embargo, también puede ser utilizada para generar desinformación, engaños y falsedades
- IA al servicio de la desinformación: puede generar contenido falso, automatizar su producción y distribución masiva en redes con bots y perfiles falsos
- Resulta clave promover procesos de verificación de fuentes y reconocimiento de herramientas tecnológicas para facilitar la detección de contenido falso creado con IA
- Frente a imágenes creadas con IA es central mirar detalles en facciones, dedos, manos, ojos, dentaduras o sombras, que pueden tener errores
- Los *deepfakes* son contenidos en video o audio creados con IA que recrean la apariencia y la voz de personas de una manera convincente con el objetivo de engañar
- Aunque no ofrecen resultados concluyentes, existen varias herramientas tecnológicas y de IA para detectar *deepfakes* de audio y video



Actividades prácticas para desarrollar en el aula: verificar desinformación generada con IA



Verificaciones de desinformaciones creada con IA

1. Desde el celular utilizar Google Fact Check Explorer para detectar notas verificadas por organizaciones de fact-checking en las que se utilizó IA para desinformar. Identificar las temáticas más recurrentes en la desinformación generada por IA. Analizar los formatos utilizados en la desinformación generada por IA (por ejemplo, imagen, videos, etc.).
2. Seleccionar algunos casos específicos de desinformación generada por IA para un análisis más profundo. Elaborar un informe que resuma los hallazgos del análisis e identificar las principales estrategias utilizadas para generar desinformación con IA.



Creación y detección de imágenes creadas con IA

1. Utilizar una aplicación de creación de imágenes con IA, como [Copilot de Bing](#) o similar.
2. Crear una imagen original a partir de un tema o concepto de su elección, luego guardar la imagen en un dispositivo móvil.
3. Instalar la aplicación [AI Image Detector](#) en su dispositivo móvil y subir la imagen creada previamente.
4. Analizar los resultados de la detección de IA proporcionados por la aplicación.
5. Elaborar un artículo periodístico que describa la experiencia de creación y verificación de imágenes con IA. También puede incluir casos reales de desinformación o manipulación utilizando imágenes creadas con IA.

✓ Deep fakes con videos creados con IA

1. Investigar sobre los *deepfakes* y su impacto en el periodismo.
2. Explorar las características y funcionalidades de la herramienta Deepware Scanner.
3. Formar grupos de trabajo y seleccionar un video de su propio celular que pueda ser analizado por la herramienta
4. Analizar el video seleccionado con la herramienta y luego interpretar los resultados proporcionados.

✓ Regulación de la inteligencia artificial

1. Realizar un relevamiento sobre los principales puntos del Proyecto de Ley de Inteligencia Artificial Europea.
2. Analizar los potenciales riesgos de la IA para la sociedad que detalla el proyecto.
3. Evaluar las obligaciones y prohibiciones que imponen los diferentes artículos para las organizaciones tecnológicas que trabajan con IA.
4. Reflexionar sobre la utilidad del proyecto para reducir el impacto de la desinformación creada con IA.
5. Debatir sobre la necesidad de la regulación de la inteligencia artificial.

✓ Formato de nota de verificación

1. Leer en detalle la **nota** de Chequeado sobre las imágenes creadas con IA sobre el supuesto arresto de Donald Trump.
2. Analizar el recorrido de la verificación, las fuentes utilizadas y el contexto de publicación de la imagen creada con IA.
3. Identificar los diferentes datos que componen los formatos de verificación.



Reconocimiento de imágenes creadas con IA



1. A partir de la visualización de la imagen del papa Francisco, realizar una observación guiada con las siguientes categorías:

- **Patrones repetitivos**

Las imágenes generadas por IA pueden contener patrones repetitivos o irregularidades que no se encuentran típicamente en fotografías reales. Esto puede deberse a las características de entrenamiento del modelo.

- **Errores de textura**

A veces, las imágenes generadas por IA

pueden tener errores en la textura o detalles, especialmente en áreas complejas como cabello, ojos, orejas o detalles finos.

- **Incoherencia global**

Las imágenes generadas pueden tener objetos o escenarios que no tienen sentido con el contexto del hecho que relata.

- **Iluminación**

Las imágenes generadas pueden tener problemas para mantener una iluminación coherente en todos los objetos.

- **Problemas de perspectiva**

Las imágenes generadas pueden mostrar problemas con la perspectiva, especialmente cuando se trata de objetos complejos en el espacio tridimensional.



- **Excesiva perfección**

Las imágenes generadas a veces pueden parecer excesivamente perfectas, sin los pequeños errores o imperfecciones que a menudo están presentes en imágenes reales.

2. Luego de esta primera observación, realizar una búsqueda inversa.
3. A partir de los resultados de la búsqueda reflexionar sobre:
 - ¿Cuáles son los primeros resultados que arroja la búsqueda?
 - ¿Cuáles son los resultados que se obtienen con coincidencias visuales con la imagen propuesta?
 - ¿Qué se puede decir del contexto de dicha imagen?
4. Luego, ingresar al sitio web de la herramienta [Hive Moderation](#) y realizar la prueba para detectar si el contenido fue creado por IA

Esta herramienta proporciona resultados que incluyen porcentajes de probabilidad de que la imagen que está evaluando haya sido generada con IA y también indica la plataforma utilizada. Es importante tener en cuenta que ninguna de estas herramientas es completamente infalible.

Lecturas para ampliar

- Aldea de Periodistas. (2023, 27 de septiembre). Ventajas y limitaciones de la inteligencia artificial ([link](#))
- Artopoulos, A. (2023). Imaginarios de IA generativa en educación. Revista Hipertextos, 11 (19), e070 ([link](#))
- Ballarino, F. y Gadel, L. (2023, 21 enero). Qué es el ChatGPT y qué oportunidades y desafíos genera. Chequeado ([link](#))
- Ballarino, F y Gardel, L. (2023, 17 de junio). Cómo cuidarse de las desinformaciones con inteligencia artificial en las elecciones de 2023. Chequeado ([link](#))
- Corti, D. (2023, 27 de marzo). La foto del Papa Francisco con una campera blanca no es real, fue creada con inteligencia artificial. Chequeado ([link](#))
- Di Domenica, S. (2023, 14 de febrero). Tres reflexiones sobre ChatGPT, inteligencia artificial y periodismo. Cronista con anteojos ([link](#))
- Factchequeado. (2023, 25 de febrero). Qué son los *deepfakes* y qué se sabe sobre los presentadores de noticias creados con inteligencia artificial ([link](#))
- Fernández, P. (2023, 10 de marzo). ¿Desinteligencia artificial: otra puerta a la desinformación? Chequeado ([link](#))
- Fresneda, C. (2016, 10 de junio). Un ordenador logra superar por primera vez el test de Turing. El Mundo ([link](#))
- Higuera, A. (2023, 23 de junio). Quién fue Alan Turing y por qué ha sido una figura clave para la inteligencia artificial. 20 minutos ([link](#))
- Huwyler, L. Navarro, M. (2023, 28 de marzo). El anorak del Papa, la detención de Trump y otras imágenes falsas generadas por Inteligencia Artificial: claves para detectarlas. RTVE ([link](#))
- Instituto Internacional de la UNESCO para la Educación Superior en América Latina y el Caribe. (2023). ChatGPT e Inteligencia artificial en la educación superior: guía de inicio rápido ([link](#))

- Kassova, L. (2024, 26 febrero). Por qué es fundamental mitigar los sesgos de la IA generativa. Red Internacional de Periodistas ([link](#))
- Maldita. (2023, 17 de febrero). Cómo una imagen creada con inteligencia artificial se viralizó como si fuera una foto real de una manifestación en Francia ([link](#))
- Mansour, J. (2023, 30 de marzo), Cómo reconocer imágenes creadas por Inteligencia Artificial: algunas pistas y mucho sentido común. AFP Factual ([link](#))
- Martínez, L. (2023, 21 de abril). Cómo saber si una imagen fue creada con inteligencia artificial. Chequeado ([link](#))
- National Geographic España. (2020, 2 de diciembre). Breve historia visual de la inteligencia artificial ([link](#))
- Infobae. (2024, 10 de enero). La desinformación impulsada por la IA, el nuevo gran riesgo global con la crisis climática ([link](#))
- Parra-Sánchez, J. S. (2022). Potencialidades de la Inteligencia Artificial en Educación Superior: Un Enfoque desde la Personalización. Revista Docentes 2.0, 14(1), 19–27. Recuperado de ([link](#))
- Secretaria de Cultura de la Nación. (2020, 22 de junio) Alan Turing, el padre de la inteligencia artificial ([link](#))
- Swenson, A. y Weissert, W. (2024, 23 de enero). New Hampshire investigating fake Biden robo-call meant to discourage voters ahead of primary. APNews ([link](#))
- Therani, A. (2023, 25 de septiembre). Lecciones clave de nuestra herramienta de IA para periodistas. Red Internacional de Periodistas ([link](#))
- Vera, F. (2023). Integración de la Inteligencia Artificial en la Educación Superior: Desafíos y oportunidades. Revista Transformar ([link](#))
- Recursos: Caja de Herramienta de Chequeado ([link](#))

2024 | Junio

Desinformación e inteligencia artificial

Una guía teórico-práctica para profesores
de periodismo y comunicación

Información de contacto

Chequeado | Programa de Educación

educacion@chequeado.com

chequeado.com

[f](#) [@](#) [in](#) [🎧](#) [🎵](#) [📞](#) [✖](#) [📺](#) /chequeado