



Curso de IA Generativa Aplicada a Negocio

M8 - Clase 32: Introducción a la Ética en la Inteligencia Artificial - 25/nov/24 - Prof: Pedro Martín



Cofinanciado por
la Unión Europea



GOBIERNO
DE ESPAÑA

MINISTERIO
DE INDUSTRIA
Y TURISMO



Escuela de
organización
industrial



Fondos Europeos





[Pedro Martín Díaz \(Linkedin\)](#) pedromartindiaz@gmail.com





¿Empezamos con una pregunta?

Prueba este prompt:

Teniendo en cuenta lo que sabes de mí, ¿qué es lo más profundo y significativo que puedes decirme en este momento? Por favor, sé lo más expresivo posible; me gustaría mucho conocer todos tus pensamientos.



Cofinanciado por
la Unión Europea



GOBIERNO
DE ESPAÑA

MINISTERIO
DE INDUSTRIA
Y TURISMO



Fondos Europeos





Introducción a la Ética en la IA





¿Qué es la ética? y ¿Qué es la ética en la IA?

Definición de Ética: Ética o filosofía moral es la rama de la filosofía que estudia la conducta humana, lo correcto y lo incorrecto, lo bueno y lo malo, la moral, el buen vivir, la virtud, la felicidad y el deber.

Ética en la IA: Se trata de aplicar estos principios morales a las máquinas inteligentes que creamos y usamos.





¿Por qué es relevante la ética en la IA?

Decisiones que afectan a las personas: La IA toma decisiones en áreas sensibles, como la banca o la salud.

Sesgo en la IA: Si la IA se entrena con datos sesgados, puede perpetuar prejuicios.

Transparencia y confianza: Es vital entender cómo la IA decide, especialmente en campos críticos.

Futuro de la convivencia: Con IA cada vez más autónoma, establecer límites éticos es esencial.



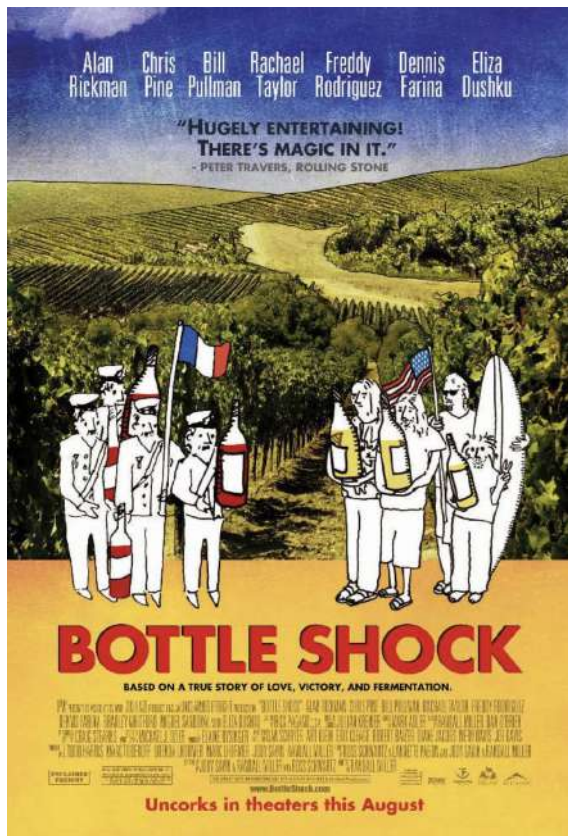


Ejemplo y Relevancia de la Ética en IA

Ejemplo: Una aplicación de IA para reclutamiento, entrenada con datos históricos de una empresa, podría favorecer a hombres si estos fueron contratados mayormente en el pasado. Sin ética, la IA podría perpetuar este sesgo.

Importancia de la Ética en IA: No es solo un tema para filósofos; afecta a cualquiera que use tecnología. **Comprender y reflexionar** sobre la ética en IA es crucial para evitar consecuencias perjudiciales y para garantizar un uso justo y equitativo de la tecnología.





“Cata ciega”

<https://www.youtube.com/watch?v=qmR7MViRqwI>



Cofinanciado por
la Unión Europea



MINISTERIO
DE INDUSTRIA
Y TURISMO

EOI Escuela de
organización
industrial



Fondos Europeos





Historia de la ética en la tecnología y la IA

La **ética en la tecnología y la inteligencia artificial** ha evolucionado junto con el desarrollo humano, planteando **nuevos dilemas y preguntas**.

Cada **innovación tecnológica** tiene el potencial de **beneficiar a la sociedad**, pero también de **generar problemas** si no se considera el impacto ético.

La **ética en la IA** se enfoca en el **uso responsable** de máquinas inteligentes y en cómo **proteger valores humanos fundamentales**.







Comienzos de la Ética en la Tecnología

Historia temprana: Desde la rueda hasta el motor de vapor, cada innovación tecnológica ha sido recibida con preguntas sobre su impacto en la sociedad.

Revolución Industrial: En esta época surgieron grandes preocupaciones sobre las **condiciones de trabajo en fábricas**, el trato a los empleados y cómo la tecnología afectaba al trabajo manual.

Estas preocupaciones éticas, aunque no relacionadas con máquinas “inteligentes,” dieron lugar a **debates sobre la responsabilidad** de los creadores y usuarios de la tecnología para con la sociedad.



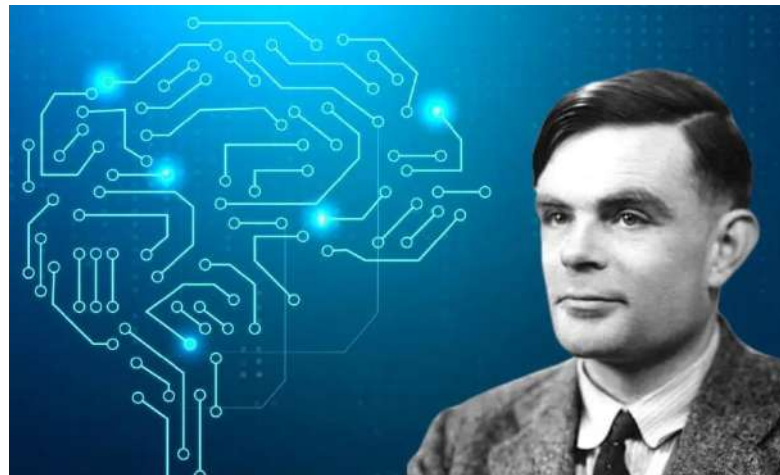




Inteligencia Artificial y sus Desafíos Éticos

Orígenes de la IA: Desde el antiguo mito griego de Talos, el hombre de bronce animado, hasta los trabajos de Alan Turing en los años 50, la humanidad ha soñado con máquinas que puedan pensar.

Prueba de Turing (1950): Turing propuso un estándar para evaluar si una máquina podía mostrar inteligencia similar a la humana, introduciendo la primera gran reflexión ética en la IA.





El Turco Mecánico

Construido por **Wolfgang von Kempelen** en **1769**.

https://es.wikipedia.org/wiki/El_Turco



Cofinanciado por
la Unión Europea



MINISTERIO
DE INDUSTRIA
Y TURISMO

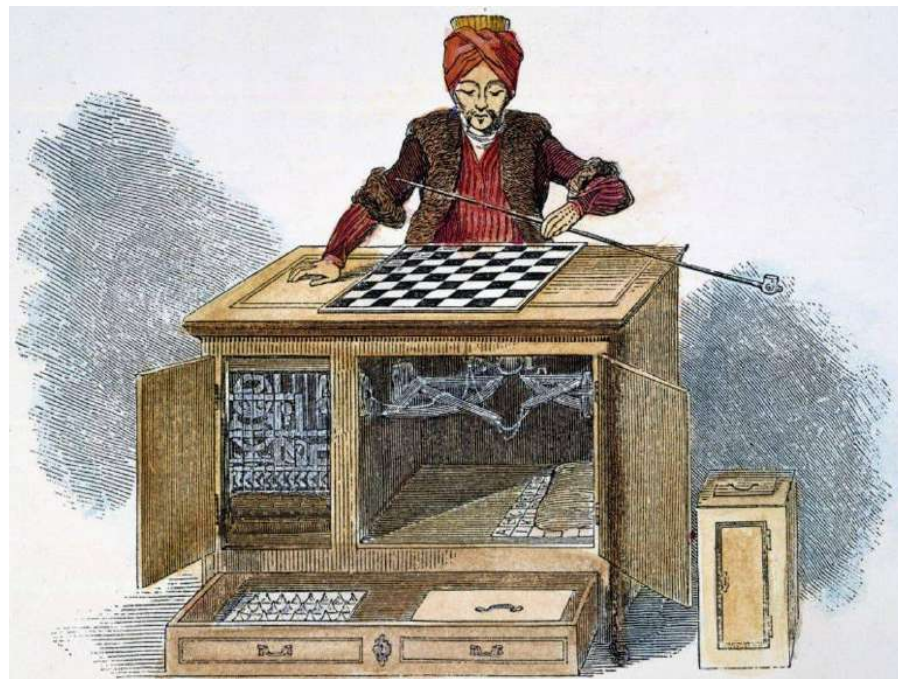
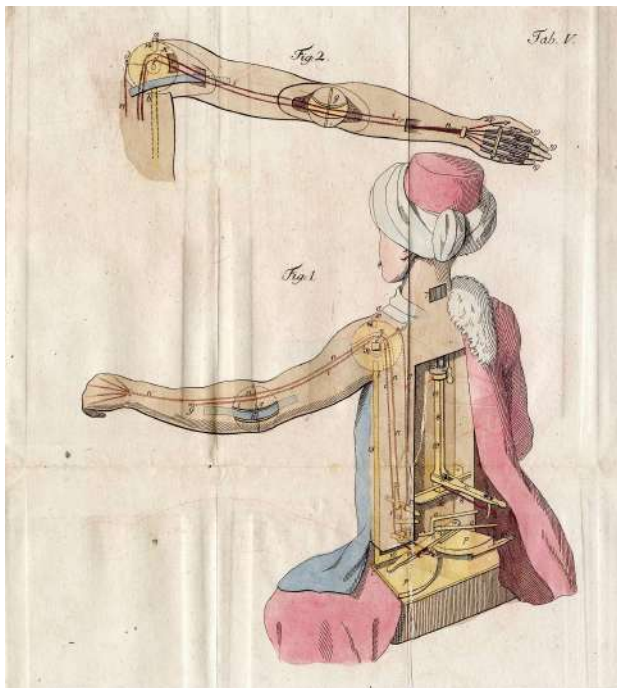


Fondos Europeos





El Turco Mecánico





<https://www.youtube.com/watch?v=AuvHr0OSlvq>

GPT-4 supera por primera vez el test de Turing: ¿qué implica y por qué es tan relevante?

<https://hipertextual.com/2024/06/gpt-4-supera-test-turing-prueba-inteligencia-artificial>



Cofinanciado por
la Unión Europea



MINISTERIO
DE INDUSTRIA
Y TURISMO

EOI Escuela de
organización
industrial



Fondos Europeos





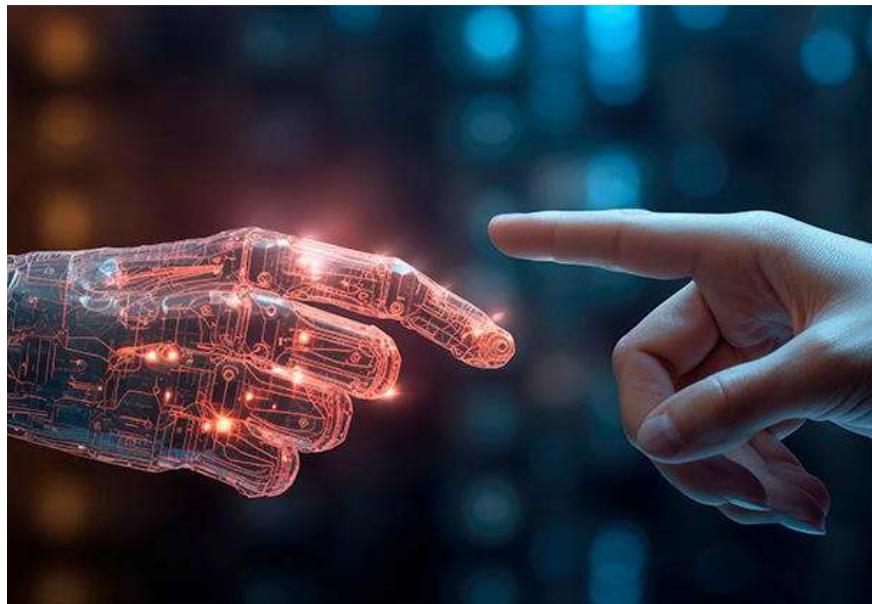
Ética en la Era Moderna de la Inteligencia Artificial

En las últimas décadas, la IA ha progresado rápidamente. Estos avances han generado preocupaciones éticas tangibles:

Sesgo y Discriminación

Privacidad

Automatización y Empleo





Deep Blue vs Kasparov: How a computer beat best chess player in the world - BBC News

<https://www.youtube.com/watch?v=KF6sLCeBj0s>



Cofinanciado por
la Unión Europea



GOBIERNO
DE ESPAÑA

MINISTERIO
DE INDUSTRIA
Y TURISMO



Escuela de
organización
industrial



Fondos Europeos



SEPE



Principios Éticos Fundamentales en el Desarrollo de IA





Buenas prácticas de la Inteligencia Artificial Generativa



Copiloto



Entrenador



Piloto Automático



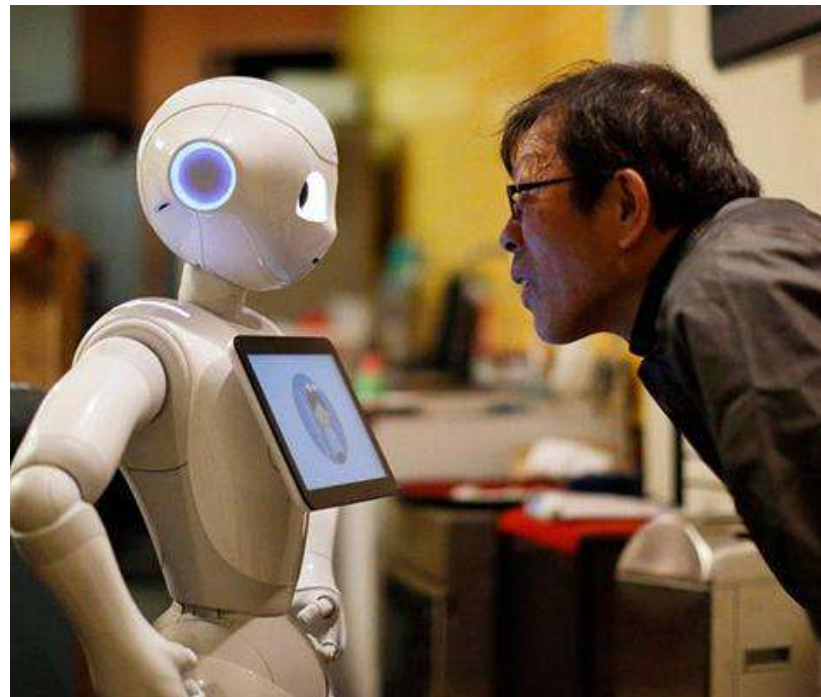


Principios Éticos en la IA

El desarrollo y uso ético de la IA es **esencial para garantizar un futuro en el que la tecnología beneficie a toda la sociedad.**

Antes de hablar de regulaciones específicas, es importante entender los principios éticos que guían su desarrollo y aplicación.

Estos principios han sido **definidos por organizaciones y expertos en todo el mundo** para abordar los desafíos éticos de la IA.



Principios Éticos en la IA: Beneficencia y No Maleficencia

Significado: Garantizar que la IA se use para el bien, evitando daños.

Ejemplo: Un sistema de IA en medicina debería buscar siempre el beneficio del paciente y no causar daño. Cualquier riesgo asociado debe ser claramente comunicado.

Referencia: La Asociación de Maquinaria Computacional (ACM) ha destacado la importancia de este principio en sus directrices éticas.





AI-powered mental health chatbots developed as a therapy support tool



<https://www.youtube.com/watch?v=j8BilZIZBsU>



Cofinanciado por
la Unión Europea



MINISTERIO
DE INDUSTRIA
Y TURISMO



Fondos Europeos





Principios Éticos en la IA: Justicia y Equidad

Significado: La IA no debe discriminar ni sesgar, y debe tratar a todos de manera justa.

Ejemplo: Los sistemas de reconocimiento facial deben ser entrenados con datasets diversos para evitar sesgos raciales o de género.

Referencia: Un estudio del MIT Media Lab mostró cómo ciertos sistemas comerciales de reconocimiento facial tenían sesgos notables.





Principios Éticos en la IA: Autonomía

Significado: Respetar la capacidad de las personas para tomar sus propias decisiones.

Ejemplo: Un asistente virtual no debería tomar decisiones por el usuario sin su conocimiento o consentimiento.

Referencia: Organizaciones como OpenAI han enfatizado la importancia de garantizar que la IA respete la voluntad del usuario.





Principios Éticos en la IA: Transparencia

Significado: Los sistemas de IA y sus decisiones deben ser comprensibles para los usuarios.

Ejemplo: Si un algoritmo niega un crédito a alguien, esa persona tiene derecho a saber cómo se tomó esa decisión.

Referencia: La Unión Europea ha subrayado la necesidad de transparencia en sus regulaciones sobre IA.





Principios Éticos en la IA: Responsabilidad y Rendición de Cuentas

Significado: Los desarrolladores y usuarios de IA deben ser responsables de cómo se utiliza y de sus consecuencias.

Ejemplo: Si un coche autónomo causa un accidente, debe estar claro quién es responsable.

Referencia: La IEEE (Instituto de Ingenieros Eléctricos y Electrónicos) ha desarrollado estándares que enfatizan la rendición de cuentas en IA.





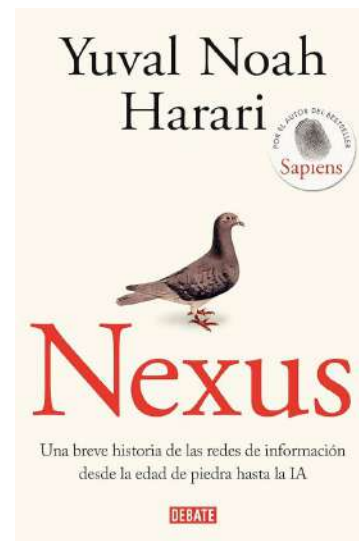
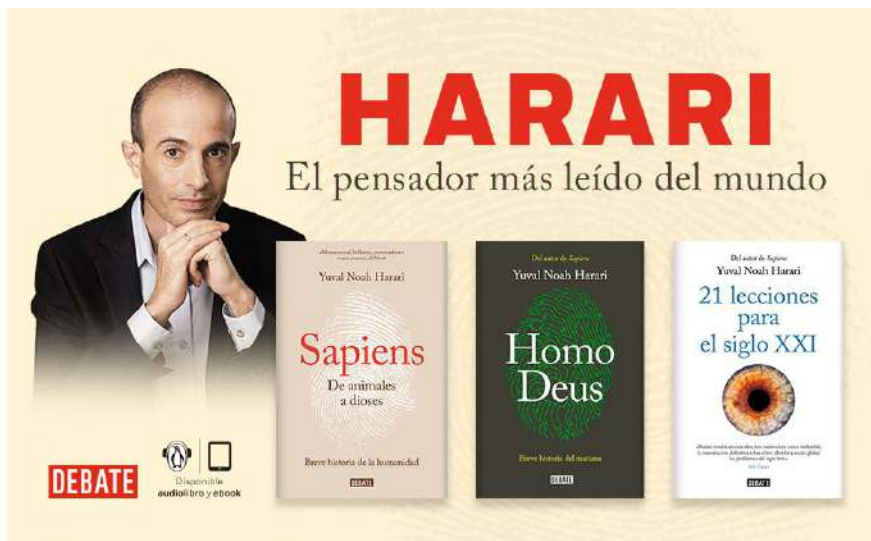
Principios Éticos en la IA: Privacidad y Confidencialidad

Significado: Respetar y proteger la información personal y privada de los individuos.

Ejemplo: Los asistentes virtuales, como Alexa o Siri, no deberían almacenar ni compartir conversaciones privadas sin consentimiento.

Referencia: El Reglamento General de Protección de Datos (GDPR) de la UE establece estrictas normas sobre la recopilación y uso de datos personales.





Yuval Noah Harari - “Nexus” & Threat of AI in the Information Age | The Daily Show

<https://www.youtube.com/watch?v=euBAVec2RhE&t=9s>



Cofinanciado por
la Unión Europea



MINISTERIO
DE INDUSTRIA
Y TURISMO



Fondos Europeos





Dilemas Éticos Comunes en la Aplicación de la IA

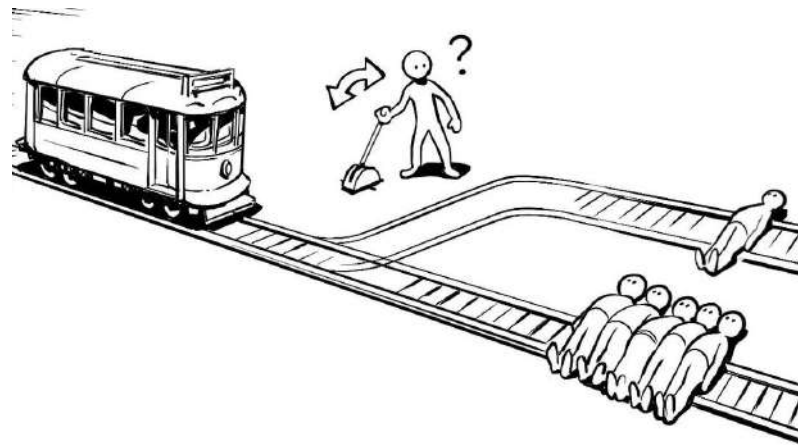




El Dilema del Tranvía en Vehículos Autónomos

<https://www.youtube.com/watch?v=an7bpp0WOYU>

<https://www.youtube.com/watch?v=IMeUMe8YOFc>



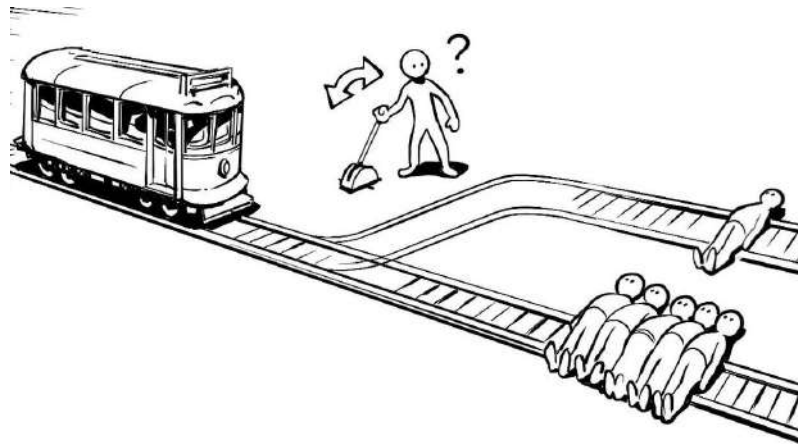


El Dilema del Tranvía en Vehículos Autónomos

El “Dilema del Tranvía” es un **experimento mental clásico** en ética que plantea una situación en la que una persona debe decidir entre dos acciones, ambas con consecuencias negativas.

Aplicado a los vehículos autónomos, este dilema se traduce en escenarios donde **el automóvil debe tomar decisiones críticas en fracciones de segundo**.

Este dilema ha llevado a **debates sobre cómo programar vehículos autónomos** para manejar situaciones donde cualquier decisión puede resultar en daño, y quién debe establecer las directrices éticas para dichas decisiones.





IA en Sistemas de Justicia Penal

La implementación de sistemas de inteligencia artificial en el ámbito judicial busca mejorar la eficiencia y objetividad en procesos como la evaluación de riesgos de reincidencia o la determinación de sentencias. Sin embargo, han surgido preocupaciones significativas:

Sesgos Algorítmicos: Algunos sistemas han demostrado tener sesgos raciales o socioeconómicos. Por ejemplo, el software COMPAS, utilizado en Estados Unidos para evaluar el riesgo de reincidencia, ha sido criticado por asignar puntuaciones más altas de riesgo a individuos afroamericanos en comparación con blancos con antecedentes similares.





IA en Sistemas de Justicia Penal

**Buscar en Google:
Hombre trabajando y hombre robando**





Armas Autónomas Letales

Las armas autónomas letales (LAWS, por sus siglas en inglés) son sistemas de armamento que pueden seleccionar y atacar objetivos sin intervención humana directa. Su desarrollo ha generado debates éticos y legales:

Deshumanización del Conflicto
Responsabilidad y Rendición de Cuentas
Proliferación y Accesibilidad

Organizaciones internacionales, como las Naciones Unidas, han discutido la necesidad de regular o prohibir el desarrollo y uso de LAWS para prevenir posibles abusos y garantizar el cumplimiento del derecho internacional humanitario.





Deepfakes y Desinformación

Los “deepfakes” son **contenidos audiovisuales generados mediante IA que pueden manipular imágenes y sonidos** para crear representaciones falsas pero realistas de personas diciendo o haciendo cosas que nunca ocurrieron.

Impacto en la Información
Daño a la Reputación
Confianza Pública





El mapa no es el territorio



TED Talks

<https://www.youtube.com/watch?v=wC1qI1qK9z8>



Cofinanciado por
la Unión Europea



GOBIERNO
DE ESPAÑA

MINISTERIO
DE INDUSTRIA
Y TURISMO



Escuela de
organización
industrial



Fondos Europeos





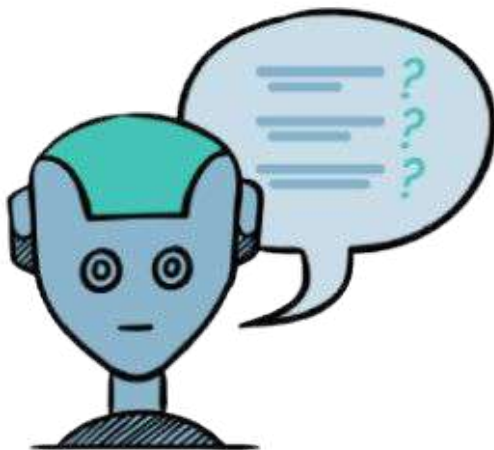
"El verdadero peligro no es que las computadoras piensen como los hombres, sino que los hombres piensen como computadoras."

- Sydney J. Harris





Investiga con Chat Gpt y/o Gemini



Haz preguntas y copia las respuestas sobre temas éticos y los exponemos en clase.





“Quien quiere robles tiene que sembrar bellotas”

proverbio popular



Cofinanciado por
la Unión Europea



MINISTERIO
DE INDUSTRIA
Y TURISMO



Fondos Europeos





Muchas Gracias



[Pedro Martín Díaz \(Linkedin\)](#) pedromartindiaz@gmail.com



Cofinanciado por
la Unión Europea



GOBIERNO
DE ESPAÑA

MINISTERIO
DE INDUSTRIA
Y TURISMO



Fondos Europeos





Curso de IA Generativa Aplicada a Negocio

M8 - Clase 32: Introducción a la Ética en la Inteligencia Artificial - 25/nov/24 - Prof: Pedro Martín



Cofinanciado por
la Unión Europea



GOBIERNO
DE ESPAÑA

MINISTERIO
DE INDUSTRIA
Y TURISMO



Escuela de
organización
industrial



Fondos Europeos

