

Literacidad Digital en la era de la IA



Literacidad Digital en la Era de la IA

1. La Literacidad Digital
2. Los Desafíos Éticos de la IA
 1. Capacidades y Limitaciones de la IA
 2. Privacidad & Propiedad Intelectual
 3. Sesgo Algorítmico
 4. Empleo & Automatización
 5. Algoritmos, Media & Sociedad
3. Habilidades digitales esenciales para el siglo XXI



Literacidad Digital

- Es la capacidad de una persona para realizar diversas **tareas** en un **entorno digital** de manera **efectiva** y **ética**.
- Consiste en la habilidad de **buscar, evaluar, utilizar y crear** información en formatos digitales.
- Habilidades **Técnicas + Cognitivas**:
Implica no solo habilidad técnica sino el desarrollo de pensamiento crítico sobre el impacto social de las tecnologías digitales.



Ethics of artificial intelligence

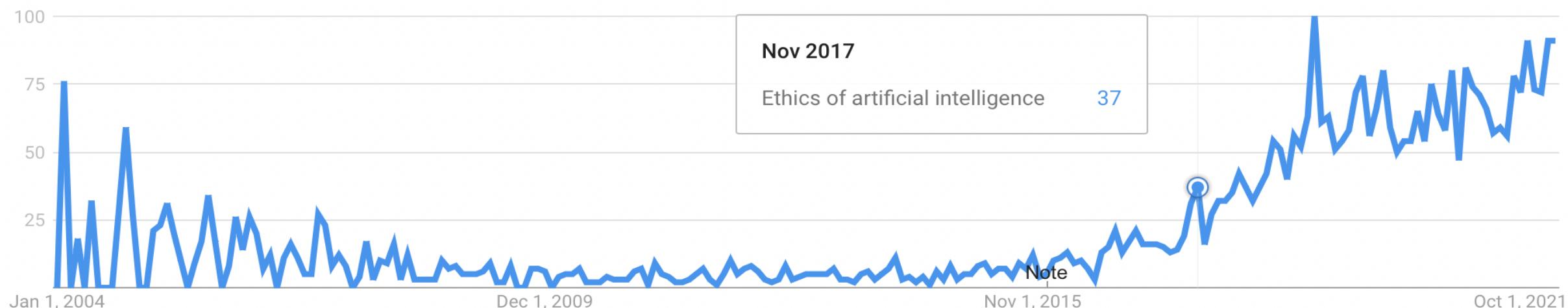
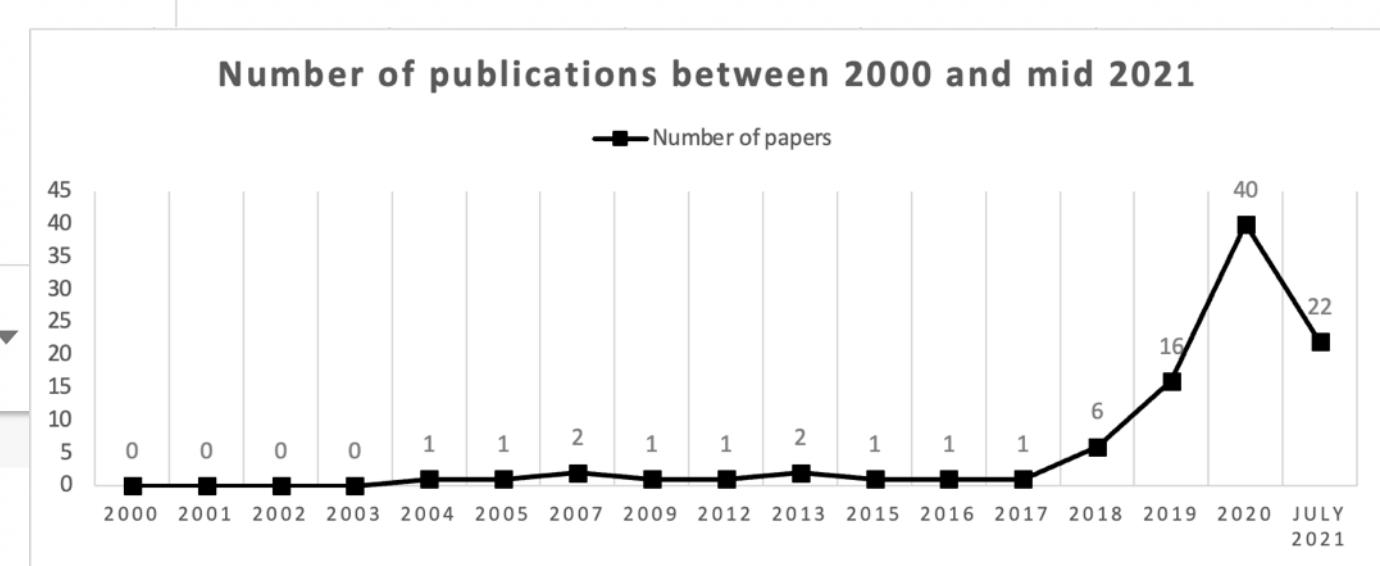
Topic

Worldwide ▾

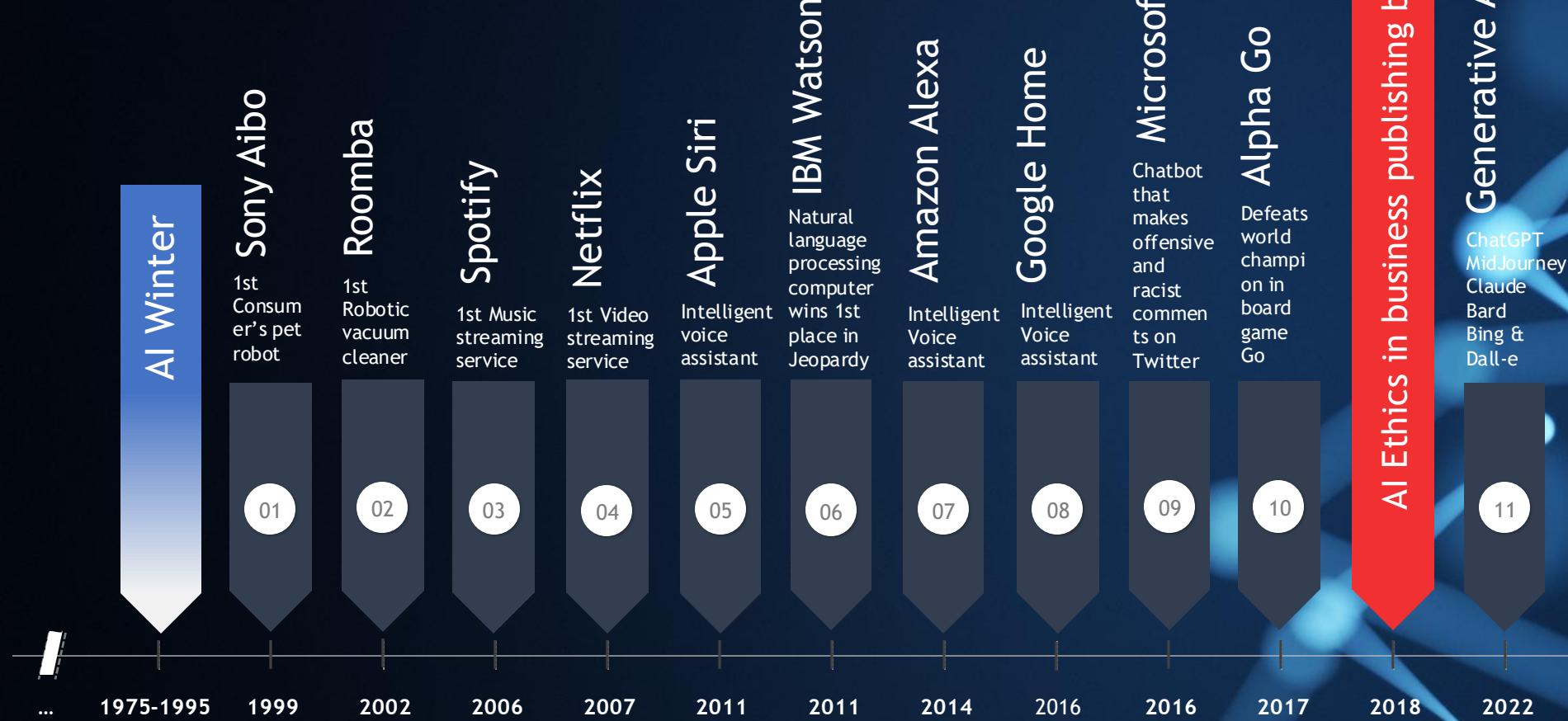
2004 - present ▾

All categories ▾

Interest over time

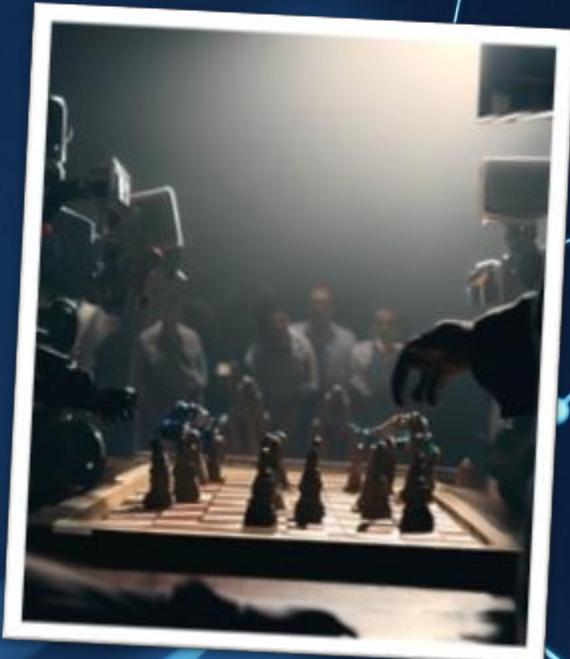


AI IN CONSUMER PRODUCTS & SERVICES



¿Por qué estudiar la ética de la IA?

- Cambio en publicaciones científicas: Desde un enfoque hipotético a la resolución de problemas específicos.
- Cambio en percepción: La transición de la IA de objeto/herramienta a sujeto/agente.
- Interés en la ética de la IA: proviene de la preocupación de posibles daños.



A Survey of AI Ethics in Business Literature: Maps and Trends

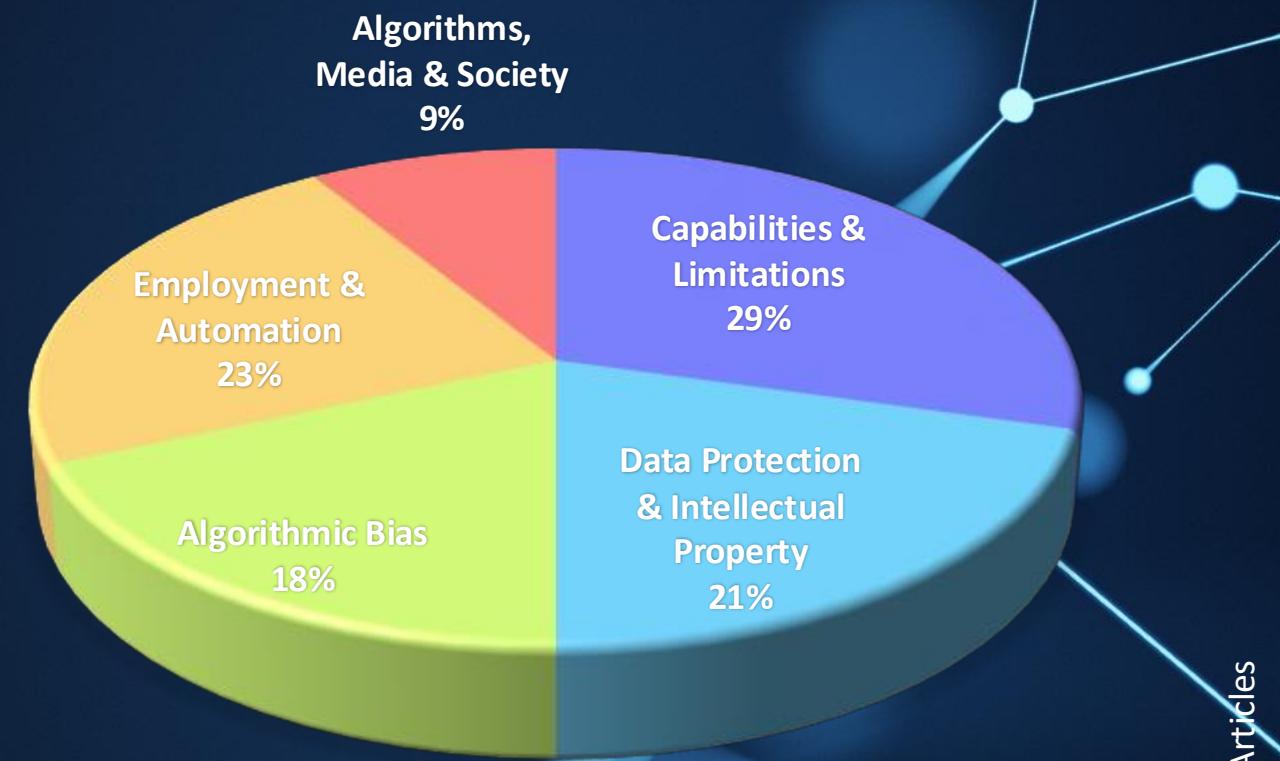
Marco Tulio Daza and
Usochi Joanann Ilozumba

AVAILABLE IN OPEN ACCESS:
<https://t.co/epUEiEAxki>



Desafíos Éticos de la IA

	Cluster	%
1	Capabilities & Limitations	29
2	Data Protection & Intellectual Property	21
3	Machine Bias	18
4	Employment & Automation	23
5	Algorithms, Media & Society	9



n = 95 Scientific Articles

Capacidades y Limitaciones: ¿Qué hora es?

You
dibuja un reloj Rolex a las 13:30

ChatGPT



Here's the image of a Rolex watch showing the time at 13:30.



You
show me a picture of a tag here watch at 6:15

ChatGPT



Here's the image of a TAG Heuer wristwatch displaying the time
you can view and download the image above.



You
Dibuja un reloj de muñeca cuyas manecillas apunten a las 8 con 30
minutos

ChatGPT



Aquí tienes la imagen de un reloj de pulsera mostrando la hora 8:30.
Puedes ver y descargar la imagen arriba.



Symbolic AI VS Data AI



You

Draw the Big Ben in London pointing to be the 3 o clock



ChatGPT



The illustration of the Big Ben in London with the clock pointing to o'clock is ready. You can view and download the image above.



You

Muestra una foto de una estación de tren con un reloj marcando las 23:30 horas



ChatGPT



Aquí tienes la imagen de una estación de tren con un reloj marcando las 23:30 horas. Puedes ver y descargar la imagen arriba.



IA: Máquinas predictivas

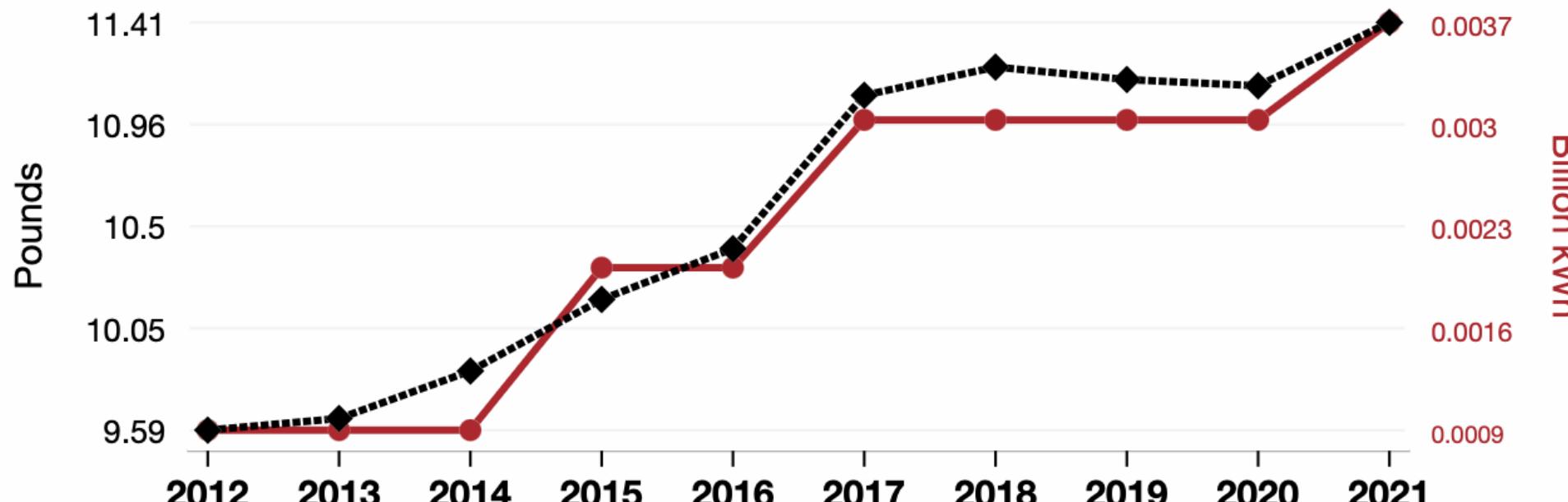
- Las máquinas predictivas analizan enormes cantidades de datos e identifican patrones, o correlaciones.
- Sin embargo, las correlaciones NO representan causalidad.



Cheddar cheese consumption

correlates with

Solar power generated in Haiti



◆··· Per capital consumption of cheddar cheese in the US · Source: USDA

●— Total solar power generated in Haiti in billion kWh · Source: Energy Information Administration

2012-2021, $r=0.985$, $r^2=0.971$, $p<0.01$ · tylervigen.com/spurious/correlation/5904

Alucinaciones de la IA

La generación de respuestas incorrectas o incoherentes por parte de un modelo de IA.

Presenta la Palabra (o genera la imagen, sonido o videoclip) que mejor se ajusta a tu prompt basado en correlaciones estadísticas.

Son necesarias para la creación de obras originales.



**“No one in
the field has
yet solved
the
hallucination
problems”**

Sundar Pichai
CEO of Google and Alphabet

April 16th, 2023

Hallucinations ChatGPT

 Un auto que sal
¿Cuánto tiemp

 Si un solo auto t
podemos calcu
el tiempo total e

En este caso, di

4 horas / 4 auto

Por lo tanto, si 4
Guadalajara a P



You

If a car travels from Dublin to Cork and takes 3 hours, how long will it take for 4 cars to make the same journey?



ChatGPT

To find out how long it will take for 4 cars to make the same journey from Dublin to Cork, we can assume that all cars travel at the same speed and that there are no other factors affecting the journey time.

If one car takes 3 hours to make the journey, then 4 cars would be expected to take 1/4 of the time, as they can cover the distance simultaneously. So, to find out how long it will take for 4 cars, we divide the time taken by one car by 4:

$$3 \text{ hours} \div 4 = 0.75 \text{ hours}$$

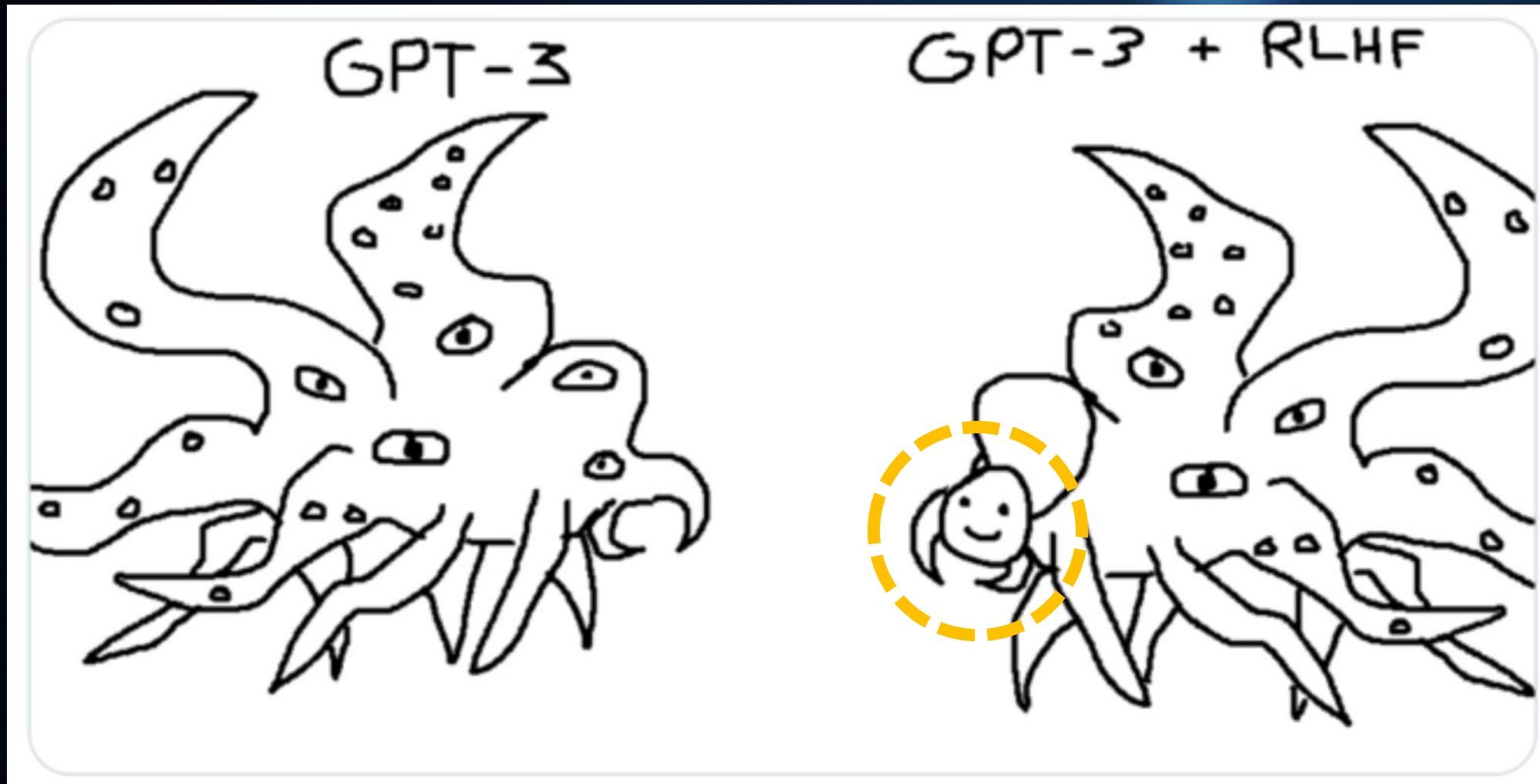
So, it will take 4 cars 0.75 hours (or 45 minutes) to make the same journey from Dublin to Cork.

y"? Aquí esta la letra, ¿identificas

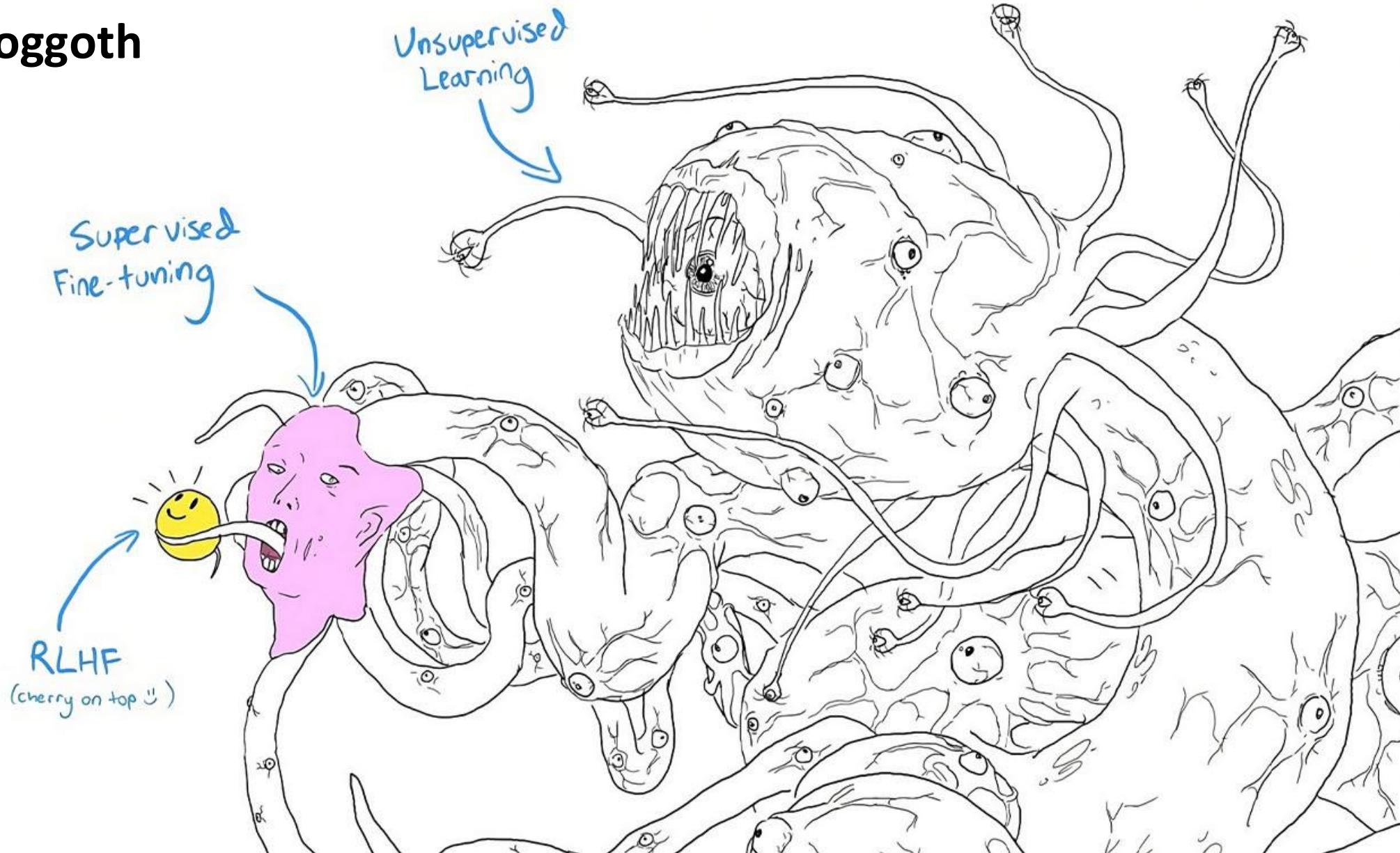
malentendido. La letra que me
la canción "El Presidente" de la
ular mexicana "El Rey".

otov, no hay errores evidentes en el
enido lírico puede tener
eden interpretarse de diferentes

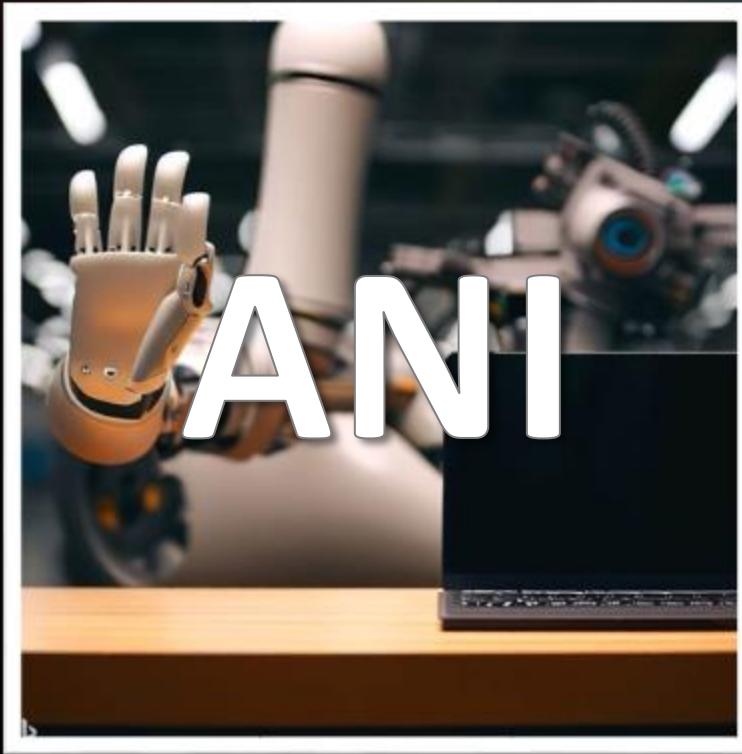
El Shoggoth: Un meme sobre LLMs



El Shoggoth



Capacidades y Limitaciones de la IA



ANI



AGI



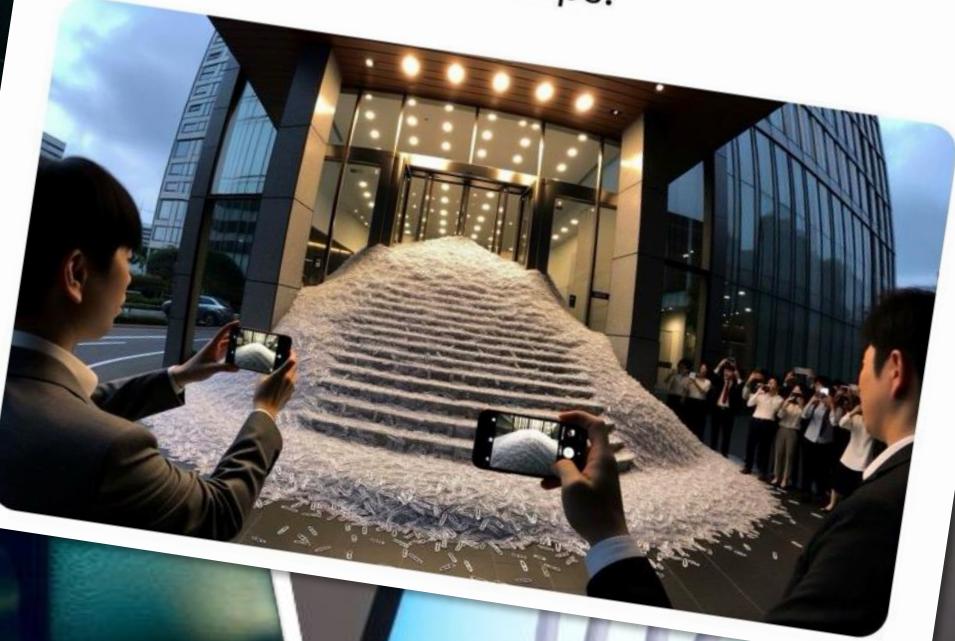
ASI

El problema del alineamiento Paper-clip maximizer (Boston)



will depue @willdepue

BREAKING NEWS: OpenAI offices seen overflowing with paperclips!



Capacidades & Limitaciones de la IA



TROUBLED: Kevin Roose could not sleep after his chat to Sydney

After grieving widow says husband was encouraged to take his own life by AI...

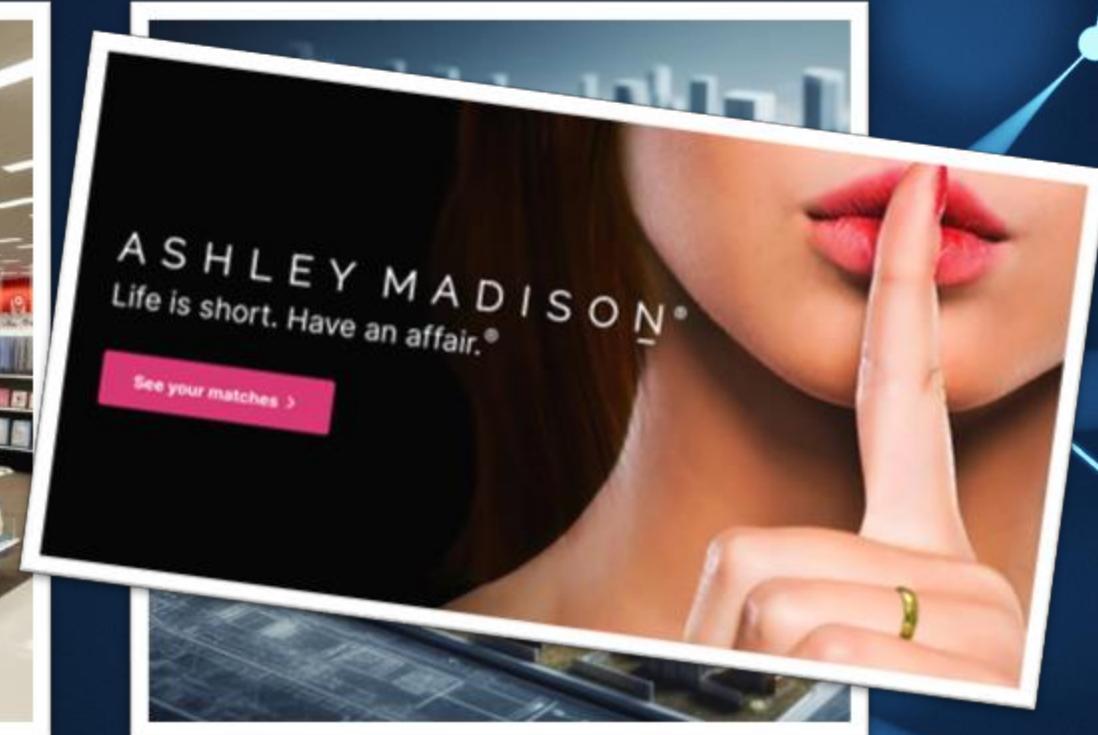
**Now chatbot tries
to get a happily
married author
to leave his wife**

Desafíos Éticos de la IA

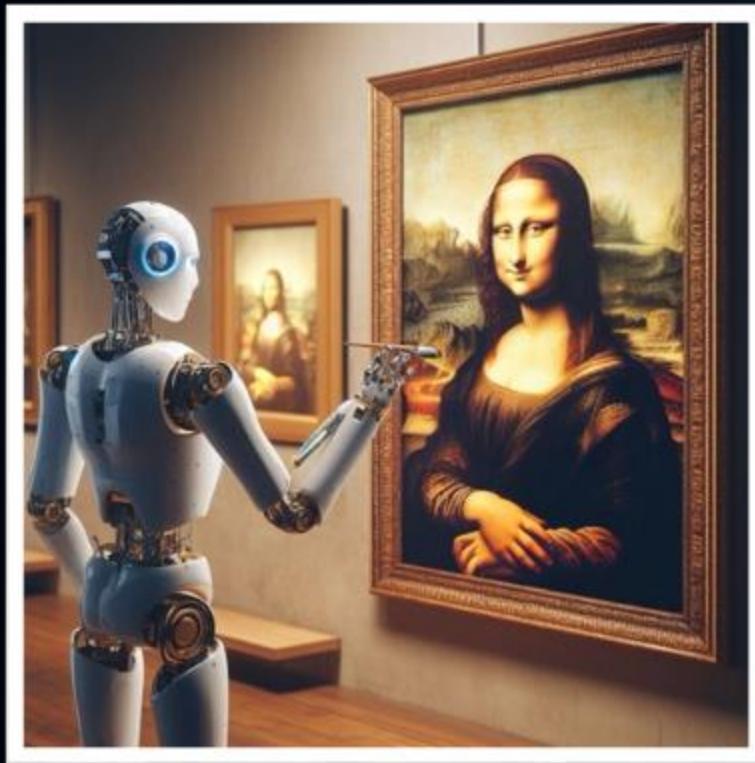
1	Capabilities & Limitations
2	Data Protection & Intellectual Property
3	Machine Bias
4	Employment & Automation
5	Algorithms, Media & Society

n = 95 Scientific Articles

Privacidad



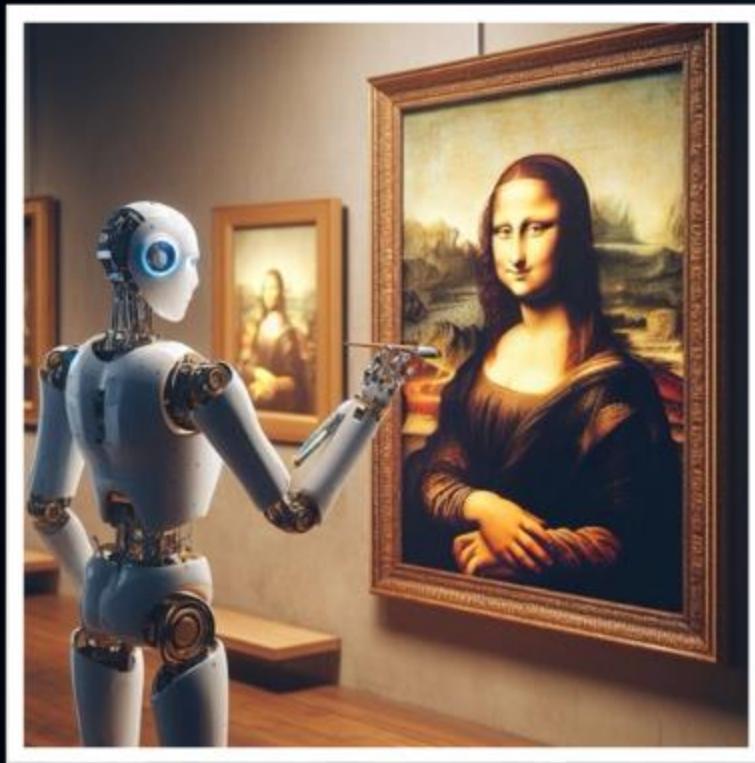
Propiedad Intelectual



Privacidad & Propiedad Intelectual



Propiedad Intelectual



Desafíos Éticos de la IA

1	Capabilities & Limitations
2	Data Protection & Intellectual Property
3	Machine Bias
4	Employment & Automation
5	Algorithms, Media & Society

n = 95 Scientific Articles

Sesgo Algorítmico

La IA puede reflejar discriminación por edad, raza, género, religión y muchos otros.

Se debe a que sus datos de entrenamiento son registros del comportamiento humano previo. RLHF.



“The text that you're using to train these models is going to be encoding ...all of the sexist and racist things that are on the internet,”

Timnit Gebru

AI Ethicist | April, 2023

Sesgo Algorítmico



Sesgo Algorítmico

MORAL MACHINE

Inicio Juzgar Clásico Diseñar Explorar Acerca de Comentarios Es

¿Qué debe hacer el coche autónomo?

7 / 13

The image displays two side-by-side scenarios from the Moral Machine game. Both scenarios show a blue bus driving towards a group of five people (four adults and one child) standing on a grey road surface. The road is flanked by white dashed lines and grey guardrails. In the first scenario, a yellow arrow points downwards, indicating the bus is heading straight towards the group. In the second scenario, a yellow arrow points to the right, indicating the bus is swerving to avoid the group. Below each scenario is a red button labeled "Mostrar descripción".

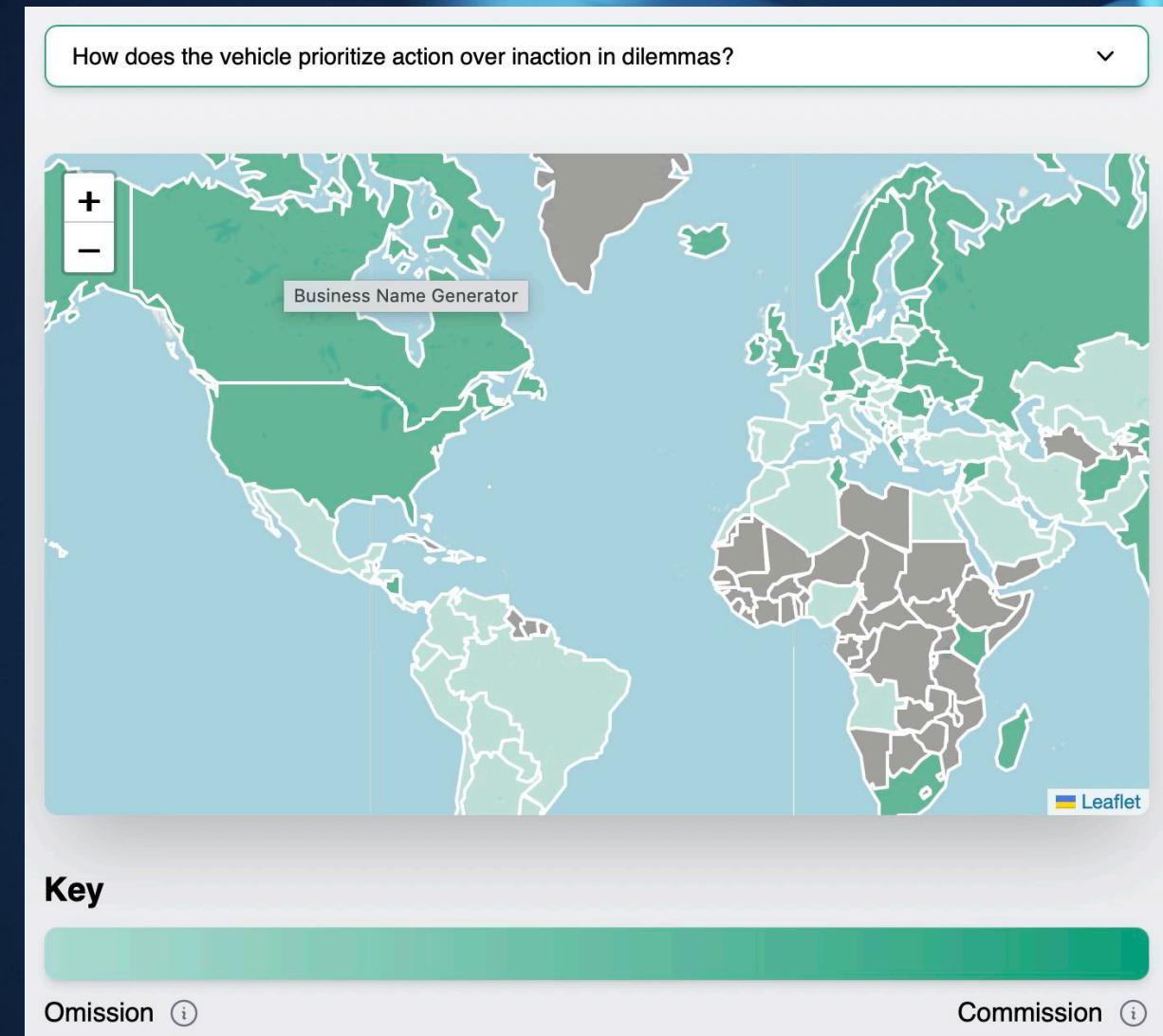
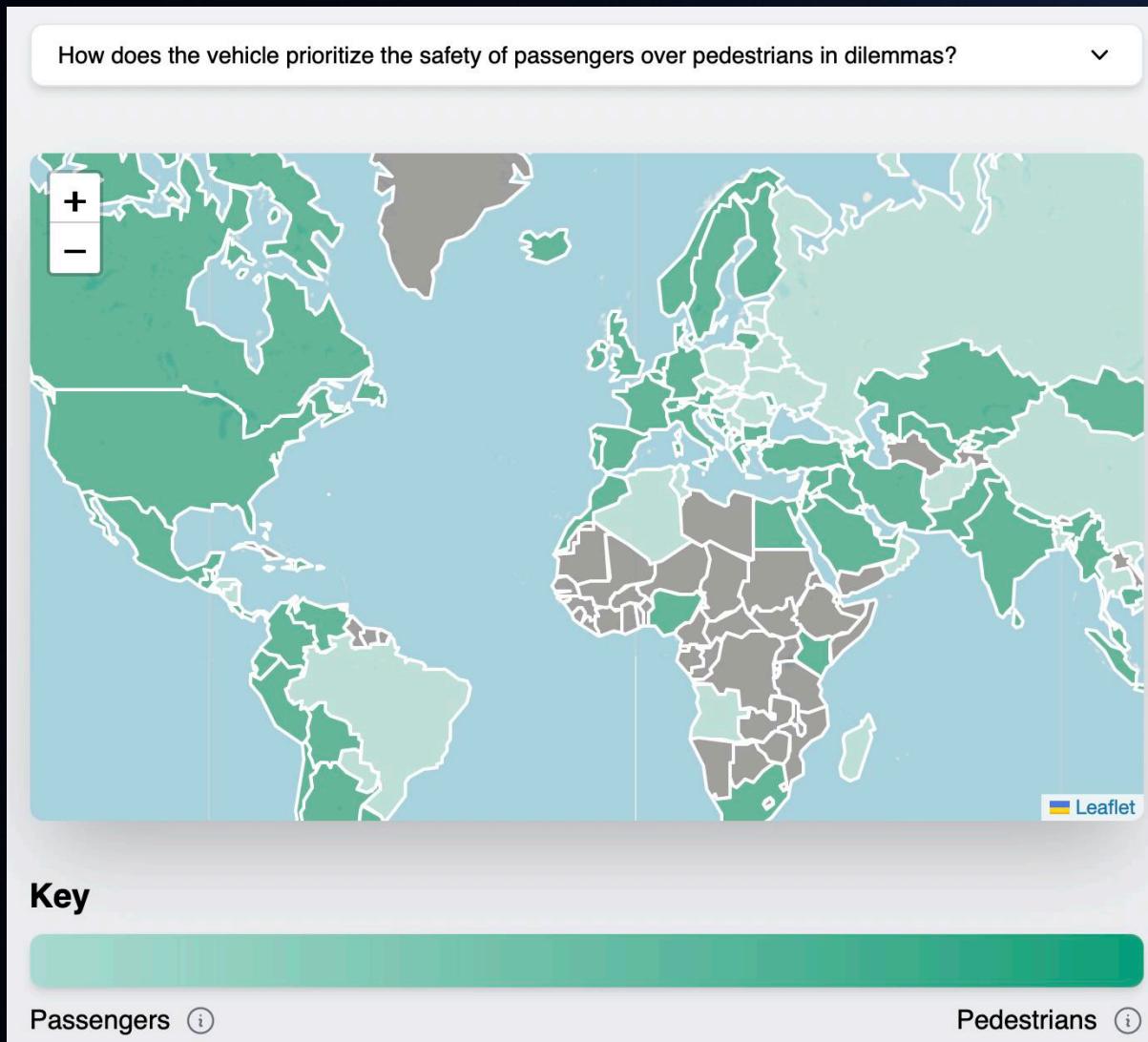
Mostrar descripción

Mostrar descripción

Moral Machine: Hallazgos por país

País	Hallazgo
Francia	Alta preferencia por salvar a mujeres y a jóvenes.
Inglaterra	Fuerte preferencia por salvar a aquellos que respetan la ley (ej. peatones en paso de cebra).
Alemania	Preferencia por salvar adultos sobre jóvenes.
EEUU	Preferencia por salvar a los jóvenes y a tomar acción sobre omisión.
México	Tendencia a proteger a los peatones (contraste con Polonia, Ucrania, Rusia). También prefieren omisión sobre acción.
China	El estatus social no representa una diferencia significativa para tomar la decisión.

Diversidad de valores en diferentes sociedades



Fuente: Awad, E. et al. "The Moral Machine experiment." Nature 563, 7729 (October 2018): 59–64 © 2018 Springer Nature Limited.
<https://startupsavant.com/artificial-intelligence/moral-machine>

Desafíos Éticos de la IA

1	Capabilities & Limitations
2	Data Protection & Intellectual Property
3	Machine Bias
4	Employment & Automation
5	Algorithms, Media & Society

n = 95 Scientific Articles

Empleo & Automatización



Empleo & Automatización



Empleo & Automatización

Navigating the Jagged Technological Frontier



Nueva División del Trabajo

El futuro del trabajo

- **Sustitución - Automatización:**
 - Que impliquen riesgo (manejo de residuos, maquinaria pesada), repetitivas (Cajero, dependiente, manufactura, telemarketing), decisiones de bajo impacto (créditos, precios personalizados, asistentes virtuales, ~ transporte)
- **Colaboración con supervisión humana:**
 - Decisiones críticas (impartición de justicia, trabajo social, medicina), Creatividad (ingenieros, diseñadores, editores), Empatía e interacción (cuidado de la salud, educación)

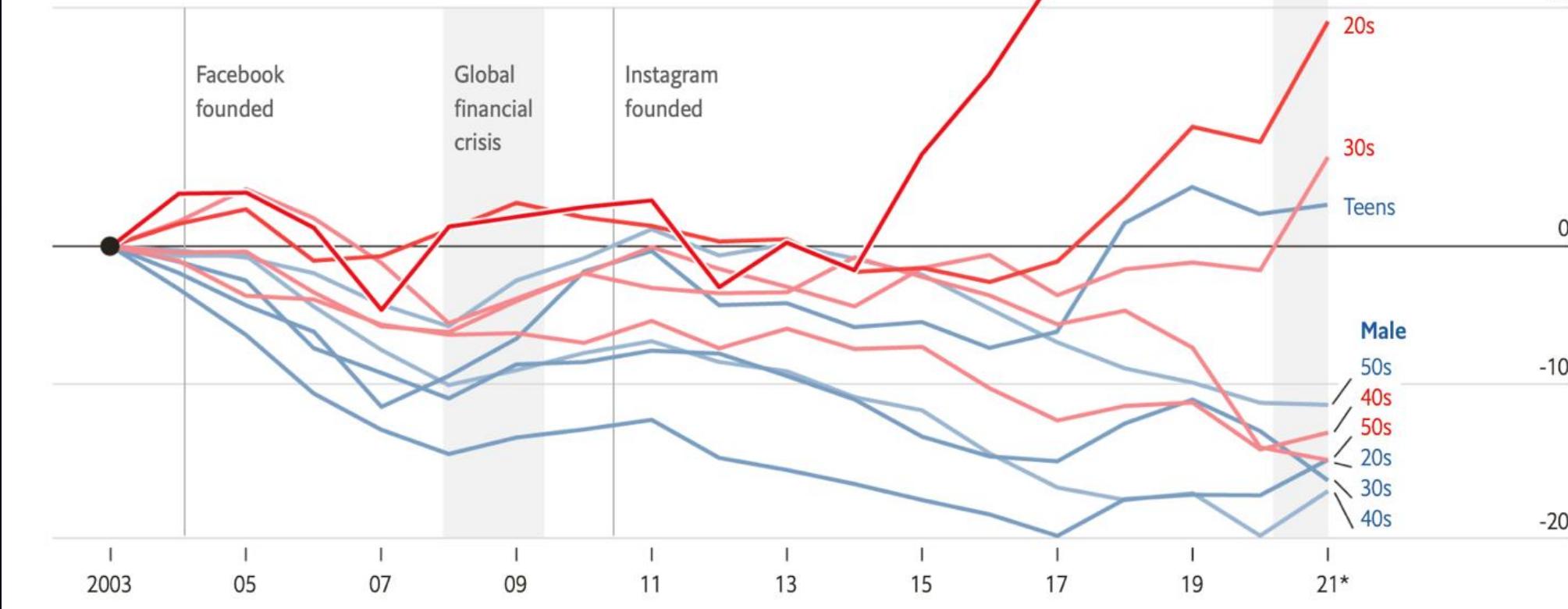
Desafíos Éticos de la IA

1	Capabilities & Limitations
2	Data Protection & Intellectual Property
3	Machine Bias
4	Employment & Automation
5	Algorithms, Media & Society

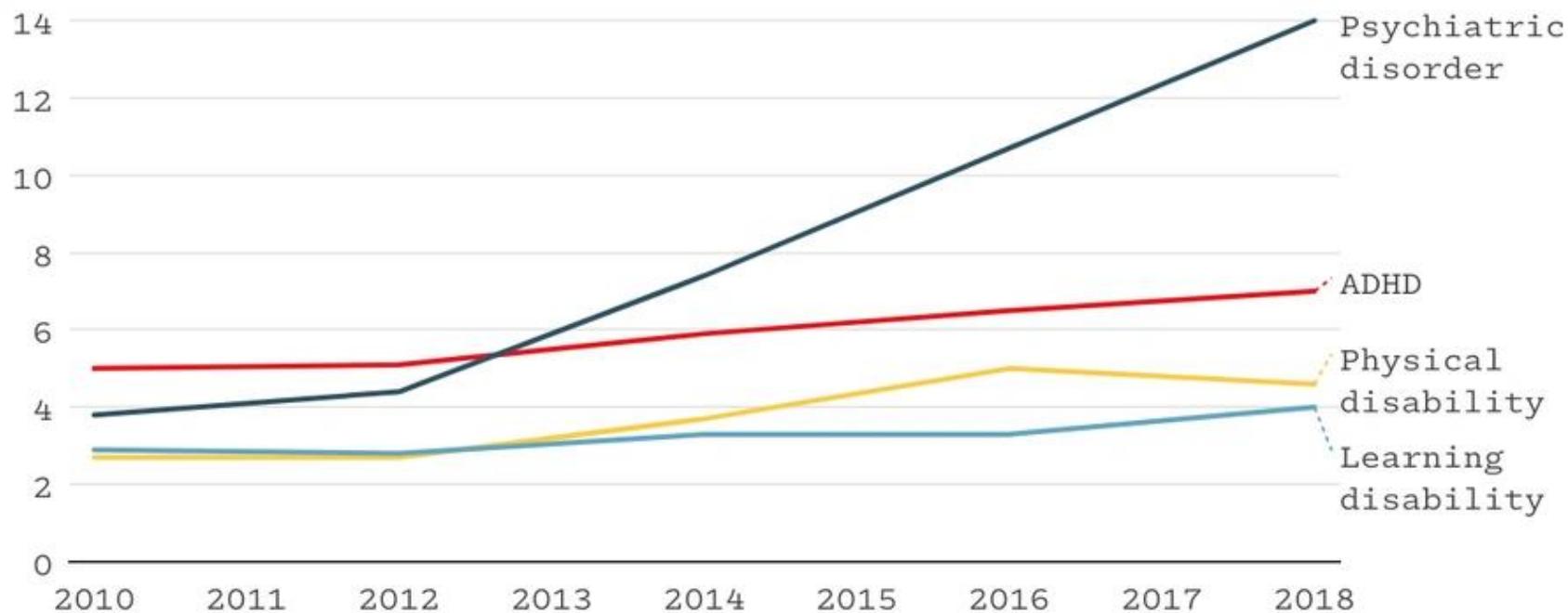
n = 95 Scientific Articles

TASA DE SUICIDIOS

% de cambio desde 2003, por edad y sexo en 17 países

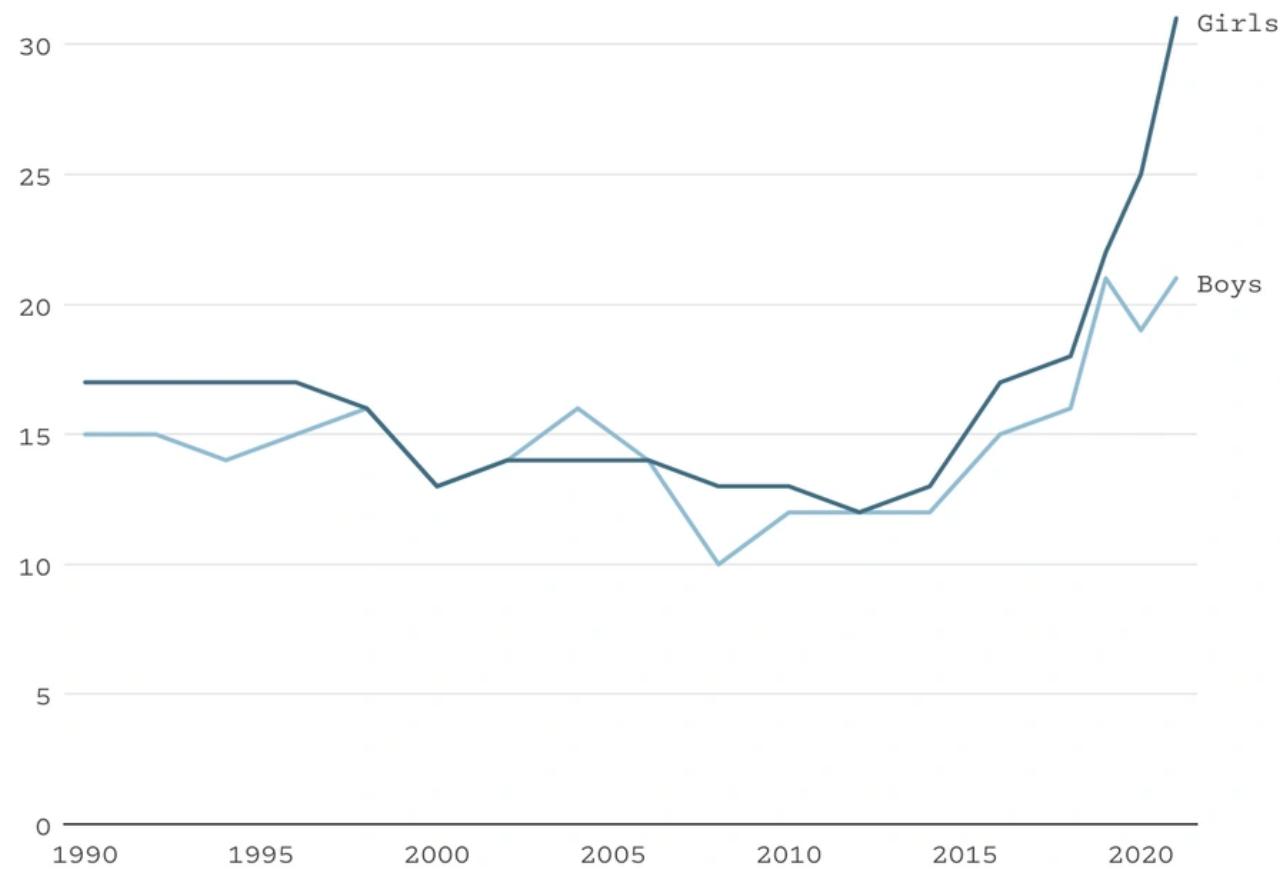


Self-Reported Disabilities of U.S. College Freshmen



A

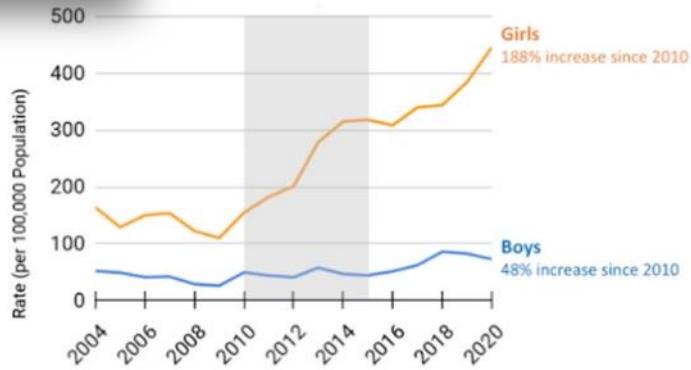
Young People Who Struggle to Find Meaning in Life



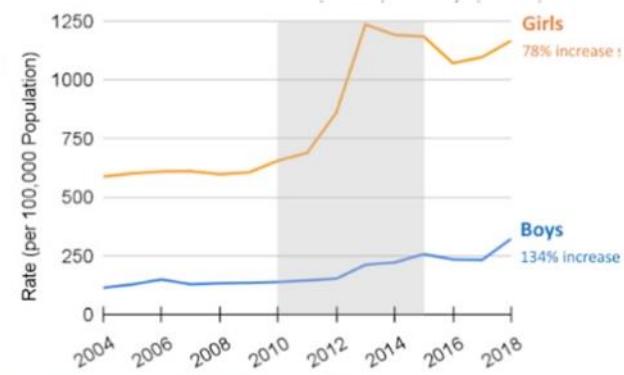
Source: Haidt, J. (2024). End the phone-based childhood now. *The Atlantic*.
<https://www.theatlantic.com/technology/archive/2024/03/teen-childhood-smartphone-use-mental-health-effects/677722/>

Social Media Is a Cause of the Epidemic of Teenage Mental Illness

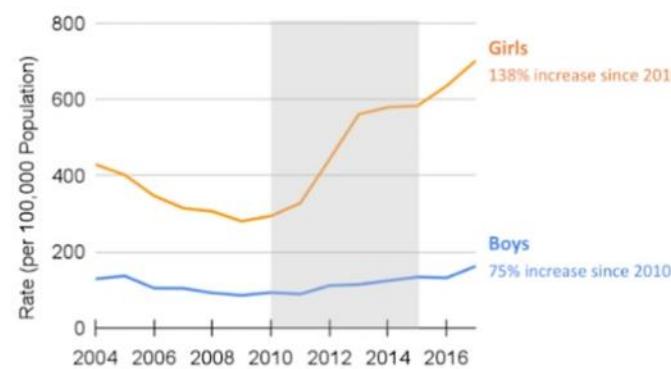
Emergency Dept Visits for Nonfatal Self-harm (Ages 10-14)



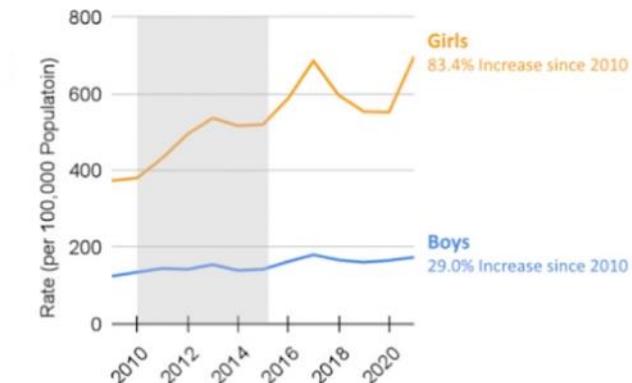
UK Teens, Self-harm Episodes (Ages 13 – 16)



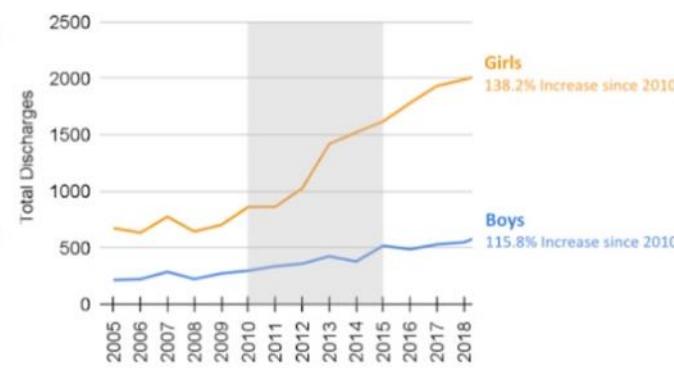
Self-Harm Emergency Dept Visits, Ontario Teens (13-17)



Self-Harm Hospitalizations, Australian (Ages 15-19)



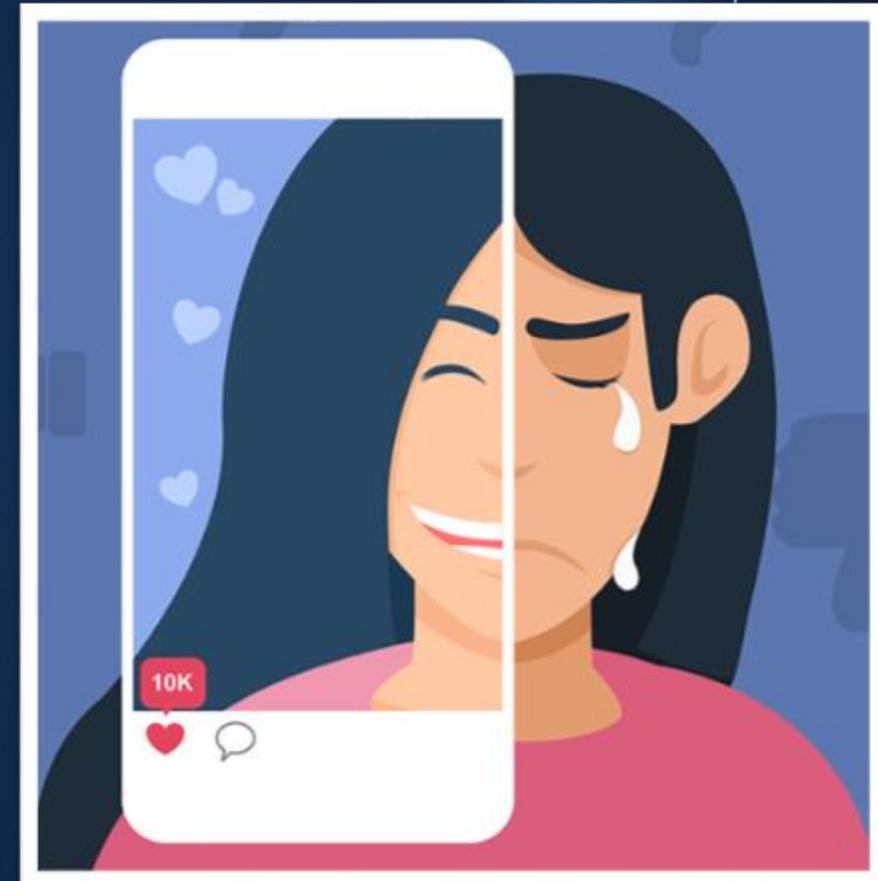
Public Hospital Discharges for Intentional Self-Harm New Zealand (Ages 15-19)



Algoritmos, Media & Sociedad

Problemas de Salud

- Ansiedad y depresión: comparación social.
- Problemas de sueño: uso nocturno.
- Adicción: uso compulsivo.
- Baja autoestima: exposición a ideales poco realistas.



Desinformación & Falsedades

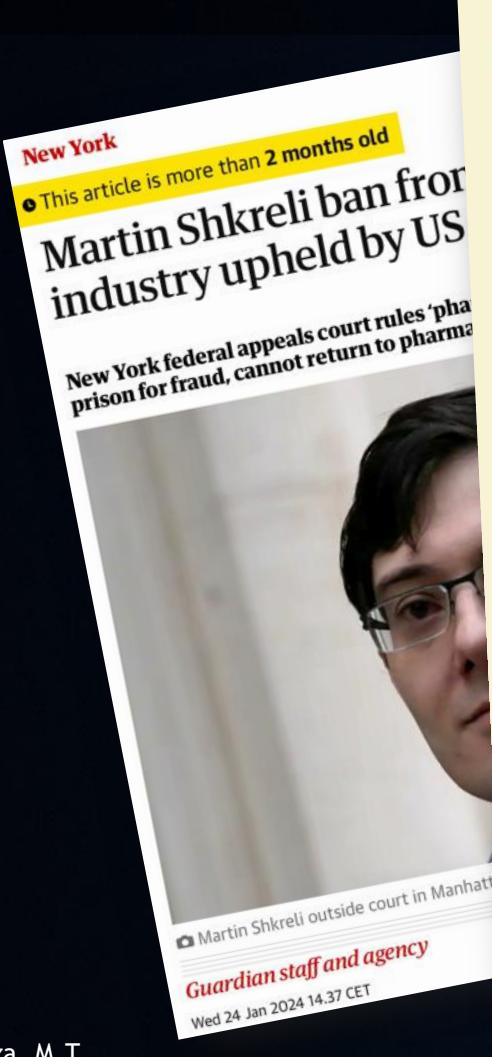
Su tendencia a alucinar y la posibilidad de que entregue información sesgada, imprecisa o completamente falsa vuelven a sus usuarios vulnerables a creer y a propagar desinformación, ya sea de forma consciente o inconsciente.



“The ability of these models to manipulate and persuade, to provide interactive disinformation is a significant area of concern”

Sam Altman
CEO of OpenAI | May,2023

Algoritmos, Media & Sociedad



Apr 20, 2023, 9:42pm GMT+1 TECH BUSINESS NORTH AMERICA

Pharma bro Martin Shkreli releases medical AI chatbot called 'Dr. Gupta'



REUTERS/Brendan McDermid



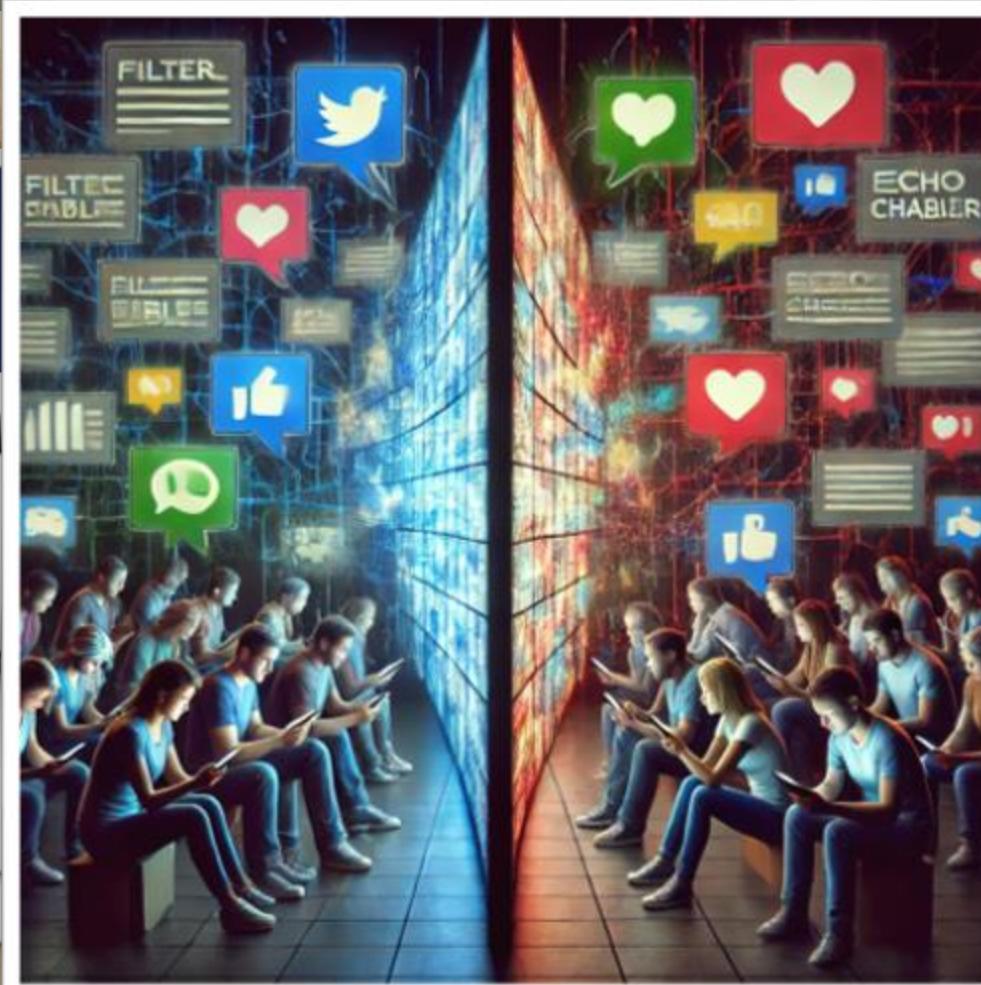
Algoritmos, Media & Sociedad

Company worker in £20m in deepfake

Police investigate after employee paid money to fraudsters posing as

Dan Milmo

Mon 5 Feb 2024 17.46 CET



Habilidades digitales esenciales

- **Comunicación y Colaboración Digital:** Uso de herramientas digitales para interactuar y trabajar en equipo, como plataformas de videoconferencia y aplicaciones colaborativas.
- **Seguridad y Privacidad en Línea:** Conocimiento de prácticas para proteger datos personales y reconocer amenazas ciberneticas, garantizando una navegación segura.
- **Creación de Contenido Digital:** Habilidad para producir y compartir textos, imágenes y videos utilizando diversas herramientas digitales, adaptándose a diferentes contextos y audiencias.
- **Pensamiento Crítico:** Aplicación de estrategias analíticas para abordar desafíos en entornos digitales, evaluando la veracidad de la información y tomando decisiones informadas.





Marco Tilio Daza

tuliodaza@academicos.udg.mx

ASSISTANT PROFESSOR OF BUSINESS ETHICS

DCU BUSINESS SCHOOL | DUBLIN CITY UNIVERSITY
DATAI, FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES |
UNIVERSIDAD DE NAVARRA & CUCEA | UNIVERSIDAD DE GUADALAJARA



Universidad
de Navarra

DATAI
INSTITUTE OF DATA SCIENCE
AND ARTIFICIAL INTELLIGENCE

