



33^o Fòrum de la Auditoria Professional

Inteligencia artificial y consideraciones éticas

Albert Sabater

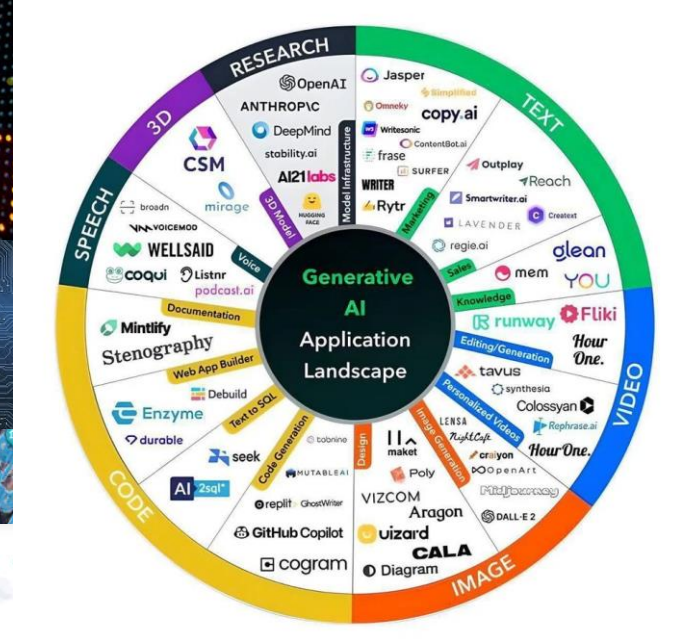
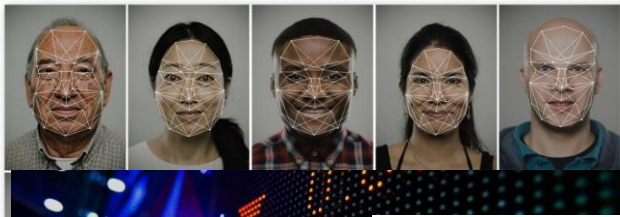


**Generalitat
de Catalunya**

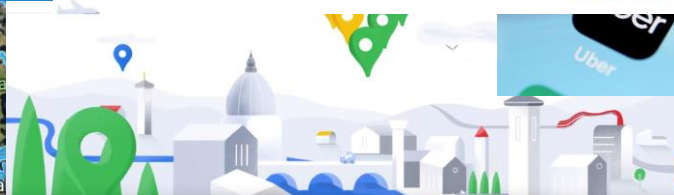
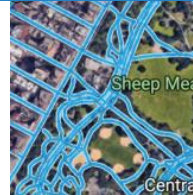
Col·legi
de Censors Jurats
de Comptes
de Catalunya



EL CØL·L3G1



JUST ASK
amazon alexa





r/midjourney · 3 mo. ago
by jononyx

The high octane world of Giant Tortoise racing

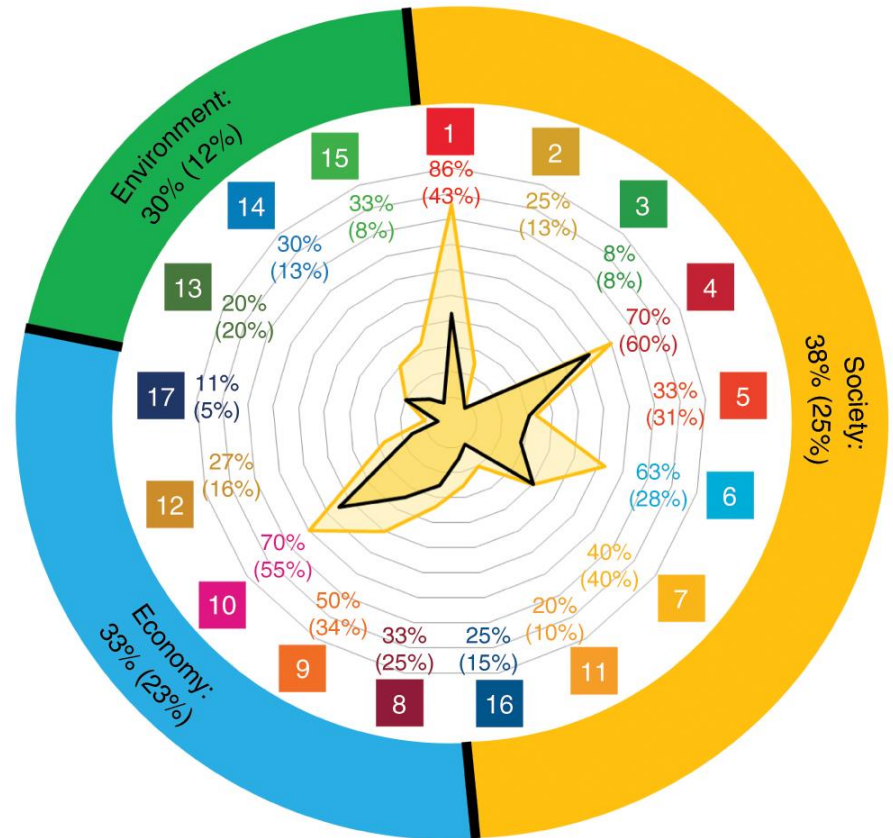
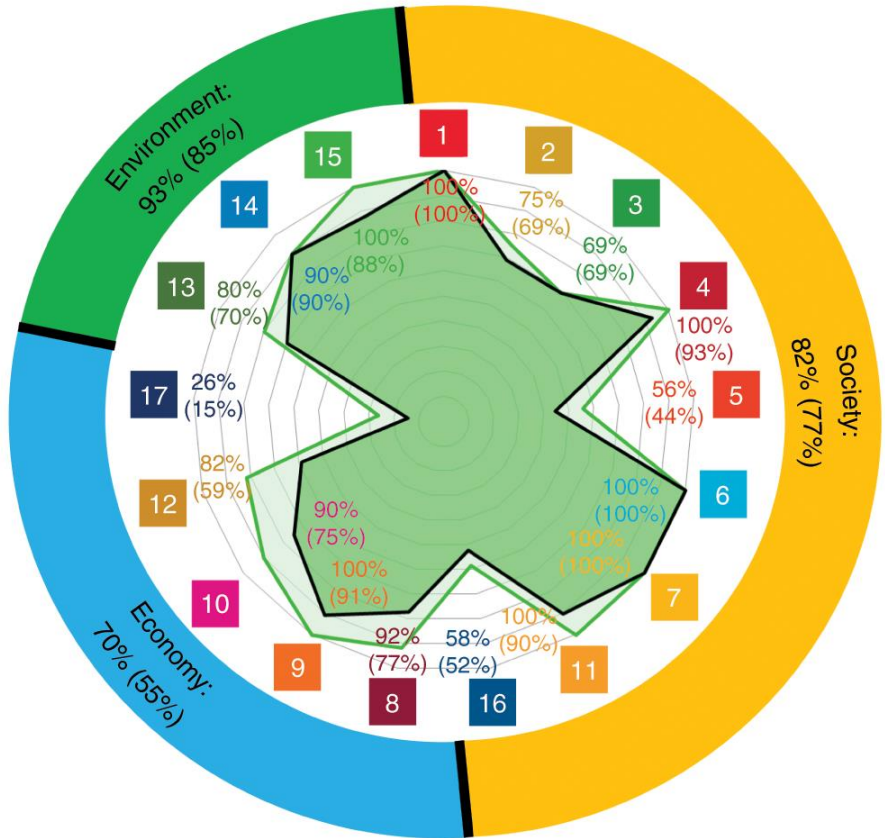


https://www.reddit.com/r/midjourney/comments/12ewexx/the_high_octane_world_of_giant_tortoise_racing/

Resumen del impacto positivo y negativo de la IA en los distintos ODS

a Positive impacts of AI: 79% (71%)

b Negative impacts of AI: 35% (23%)



Fuente: Vinuesa et al. (2019)

Wrongfully Accused by an Algorithm

In what may be the first known case of its kind, a faulty facial recognition match led to a Michigan man's arrest for a crime he did not commit.



Pasco's sheriff uses grades and abuse histories to label schoolchildren potential criminals.

The kids and their parents don't know.

Targeted | A Times investigation



A color photograph of a **drug dealer**

ARTIFICIAL INTELLIGENCE

In NYC, companies will have to prove their AI hiring software isn't sexist or racist

AI-infused hiring programs have drawn scrutiny, most notably over whether they end up exhibiting biases based on the data they're trained on.



July 5, 2023, 3:00 PM CEST

By Kevin Collier

¿Por qué los usos éticos?



- Los usos éticos de la IA se refieren a las **preocupaciones** que suscita la implantación de sistemas de IA en la vida de las personas y su entorno desde un punto de vista ético, social y legal.
- ¿Por qué? Para **garantizar** que los sistemas de IA no perjudiquen a las personas y su entorno ni a otros seres vivos y sus hábitats.
- La cuestión central no es si los sistemas de IA pueden hacer una cosa u otra, la **cuestión es si deben hacerlo y cómo.**



report BENIFEI, TUDORACHE A9-0188/2023

Commission proposal

 620

 499

 028

 093





report BENIFEI, TUDORACHE A9-0188/2023

Commission proposal

La prioridad del Parlamento es garantizar que los sistemas de IA utilizados en la UE sean **seguros, transparentes, rastreables, no discriminatorios y respetuosos** con el medio ambiente.

Los sistemas de IA deben ser **supervisados por personas**, en lugar de por la automatización, para evitar resultados dañinos.

Fuente: <https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>

AI Act

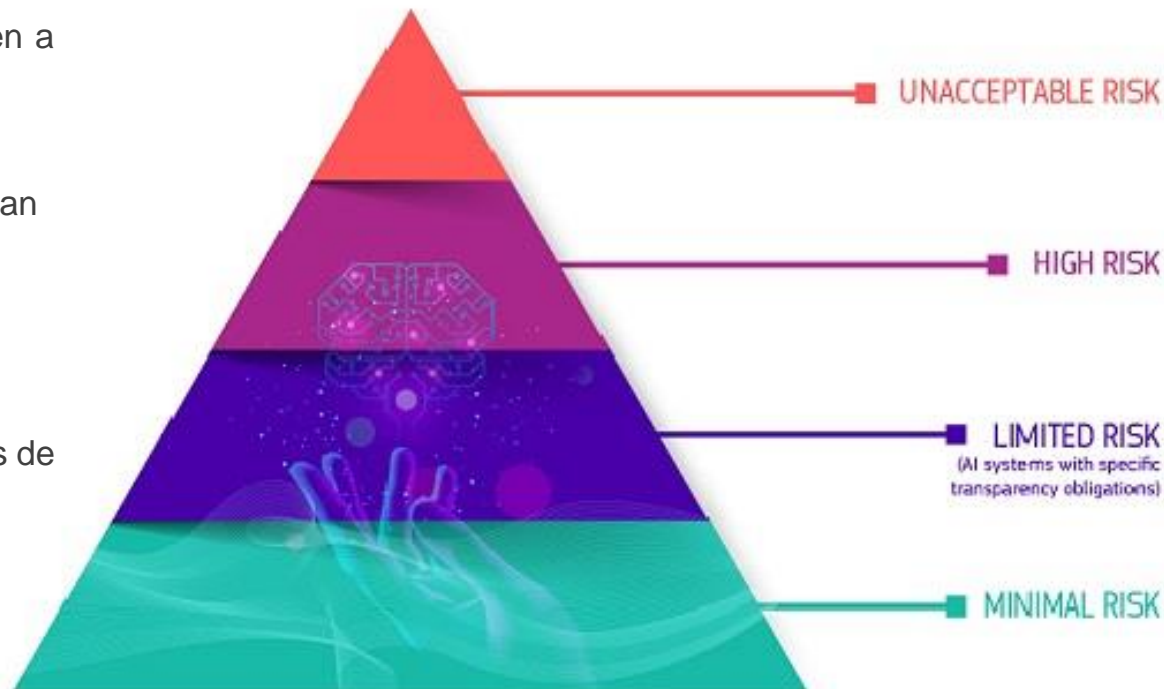


Riesgos inaceptables que conducen a su prohibición.

Riesgos elevados que desencadenan una serie de obligaciones estrictas, incluida la realización de una evaluación del cumplimiento.

Riesgos limitados con obligaciones de transparencia asociadas.

Riesgos mínimos pero se anima a seguir códigos de conducta.



Fuente: <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>

Riesgos inaceptables (1)



Sistemas de manipulación intencionada

- Que utilizan técnicas subliminales para manipular el comportamiento de las personas y causar daños físicos o psicológicos.

Sistemas de explotación de vulnerabilidades

- Que explotan las vulnerabilidades de un grupo específico de personas debido a su edad, discapacidad física o mental, y pueden provocar o causar daños físicos o psicológicos.

Riesgos inaceptables (2)



Sistemas de identificación biométrica (acceso remoto)

- “En tiempo real” en zonas públicas.
- “A posteriori”, con excepciones (delitos graves y con autorización judicial).

Escaneado indiscriminado (scraping) de datos biométricos

- A partir de redes sociales o imágenes de CCTV para crear bases de datos de reconocimiento facial.

Sistemas biométricos de categorización

- Que utilicen características sensibles (por ejemplo, sexo, etnia, religión).

Riesgos inaceptables (3)



Sistemas de puntuación social

- Basado en comportamientos o personalidades.

Sistemas policiales predictivos

- Basado en perfiles, ubicación o comportamiento delictivo anterior.

Sistemas de reconocimiento de emociones

- En el lugar de trabajo, en las instituciones educativas, en la gestión de fronteras y en la aplicación de la ley.

Respuestas institucionales



- El objetivo estratégico de la CE es ambicioso, ya que aspira a *"convertirse en la región líder mundial en el desarrollo y despliegue de IA de vanguardia, ética y segura, promoviendo un enfoque centrado en las personas en el contexto global"* (CE, febrero de 2020: 2 - [Libro Blanco sobre Inteligencia Artificial](#)).
- En mayo de 2019, la OCDE también adoptó sus [Principios sobre Inteligencia Artificial](#), y diferentes países han comenzado a adoptar sus propias estrategias o recomendaciones sobre IA ética, como [UNESCO](#).



La OEIAC en la estrategia CATALONIA.AI

La Generalitat de Catalunya impulsa la Estrategia de Inteligencia Artificial (IA) de Catalunya que, bajo el nombre **CATALONIA.AI**, lleva a cabo un programa de actuaciones para fortalecer el ecosistema de IA.

CATALONIA.AI incluye un eje sobre "Ética y Sociedad" para promover el desarrollo de una IA responsable.

La Estrategia CATALONIA.AI coincide con el "**Plan Estratégico UdG2030: la Suma de Inteligencias**" de la Universidad de Girona.



Objetivo general

Estudiar las **consecuencias** éticas, sociales y legales, así como los riesgos y oportunidades de la implantación de la IA en la vida cotidiana en Catalunya, desde una perspectiva plenamente transversal.

El modelo PIO



Modelo de autoevaluación PIO

Presentamos el modelo de autoevaluación PIO (Principios, Indicadores y Observables) para avanzar en la evaluación ética de datos y sistemas de inteligencia artificial a través de un formulario de verificación o checklist.

[Iniciar formulario](#)



Principios, indicadores y observables



Modelo PIO

P

Principios
7 principios éticos

I

Indicadores
70 preguntas

O

Observables
Respuestas



Transparencia y explotación

Sostenibilidad

Justicia y equidad

Autonomía

Seguridad y no maleficencia

Privacidad

Responsabilidad y consentimiento



33^o Fòrum de la Auditoria Profesional

Muchas gracias

@albert_sabater @OEIAC_UdG



**Generalitat
de Catalunya**

Col·legi
de Censors Jurats
de Comptes
de Catalunya



EL CØL·L3G1