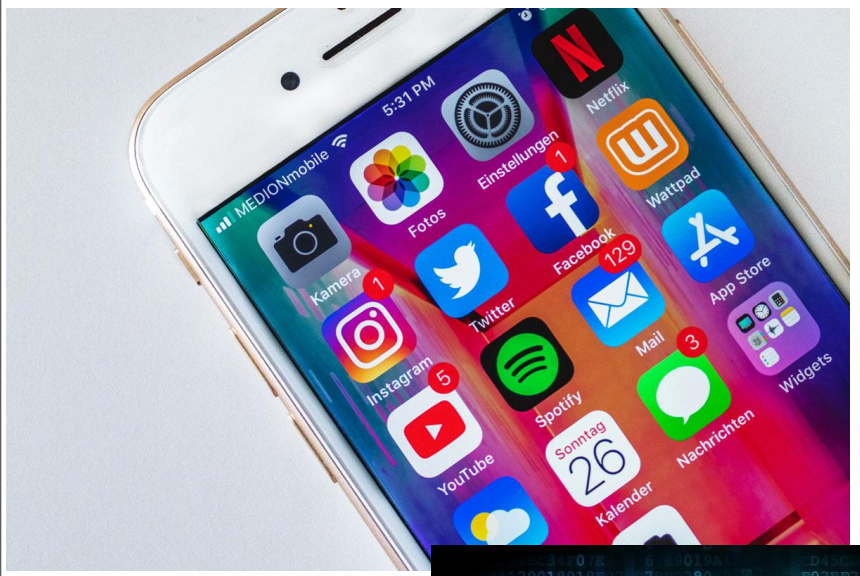
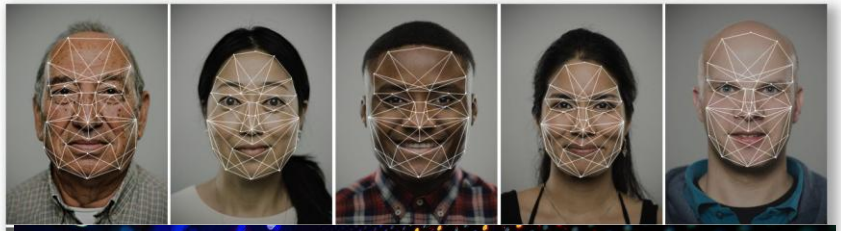
Uso de la IA en la provisión de servicios sociales

Albert Sabater Coll
Observatorio de Ética en Inteligencia Artificial de Catalunya
Universitat de Girona

Inteligencia Artificial y Servicios Sociales
8 de octubre de 2025
Palacio Miramar, Donostia/San Sebastián (Gipuzkoa)

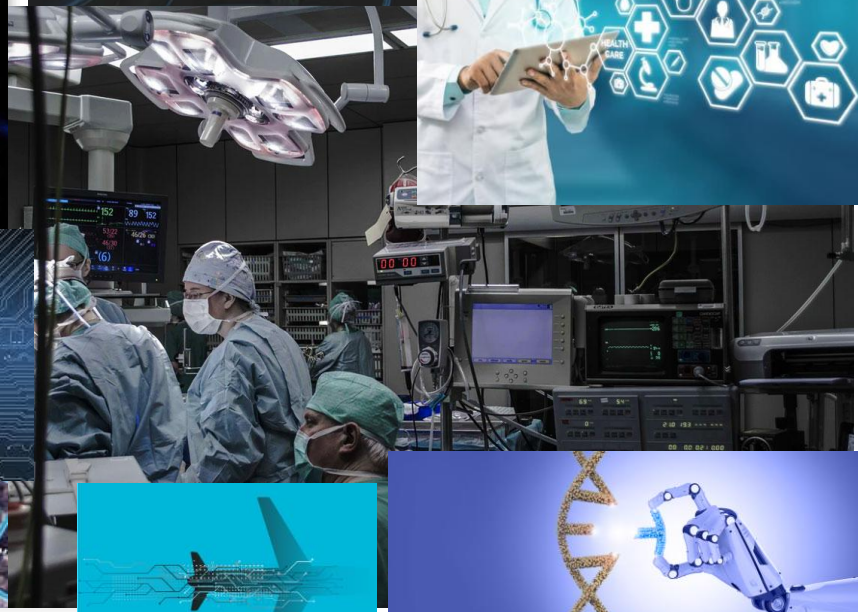
OEIAC | Universitat de Girona | Generalitat de Catalunya | BEHATUZ TOPAKETA


MEDIONmobile
5:31 PM
Netflix
Einstellungen
Fotos
Wattpad
Facebook
Kamera
Twitter
App Store
Instagram
Mail
Spotify
Widgets
YouTube
Sonntag
26
Kalender
Nachrichten

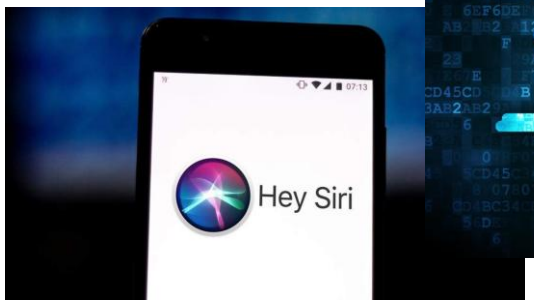


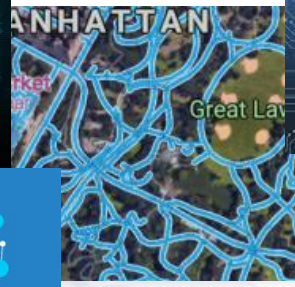








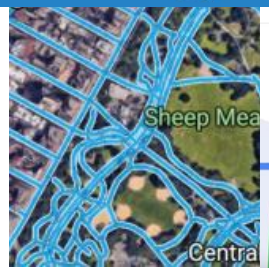

Hey Siri

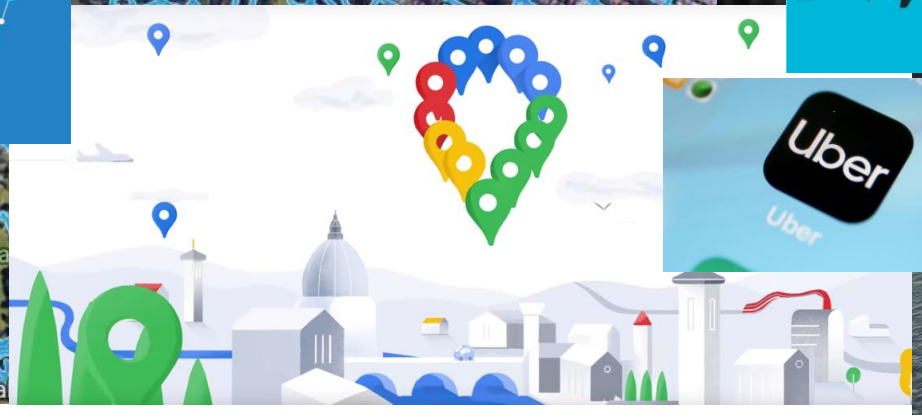

ANHATTAN
Great La


Museum of Art


amazon


JUST ASK
amazon alexa
amazon


Sheep Mea
Centra



Uber
Uber



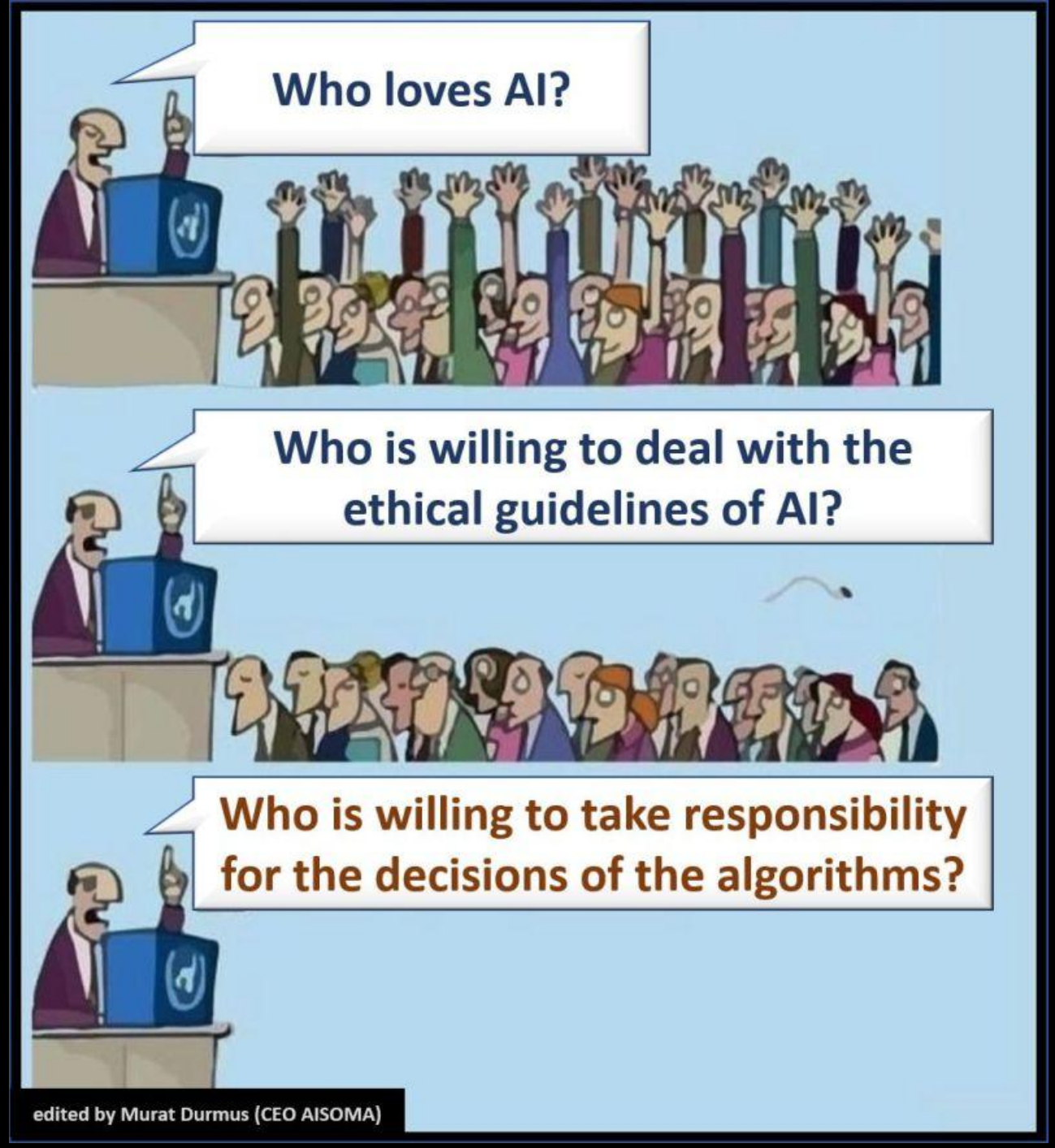


Image credit to Murat Durmus

IA responsable: 3 ingredientes

1. Necesitamos los aspectos legales y éticos (¡no son un complemento!).

2. Necesitamos hacer un seguimiento y una interna y externa.

3. Necesitamos preguntar (visión crítica) y documentar las elecciones.

Observatori d'Ètica en Intel·ligència Artificial de Catalunya 2025

MARCH 29, 2022 6:14 PM CET
BY MELISSA HEIKKILÄ

POLITICO

Dutch scandal serves as a warning for Europe over risks of using algorithms

The Dutch tax authority ruined thousands of lives after using an algorithm to spot suspected benefits fraud — and critics say there is little stopping it from happening again.

Google

images of ai

All | Images | Videos | Books | News | More | Tools | Saved | SafeSearch

artificial intelligence | future | technology | machine learning | logo | human | creative | animated | design | computer | deep learning | clip art



iStock
321,500+ Ai Stock Photos, Pictures ...


Analytics Insight
Artificial Intelligence – The His...



Simplilearn.com
Artificial Intelligence [AI ...


European Parliament
EU AI Act: first regulation on ...


TechRepublic
7 Best AI Art Generators of 2023


Simplilearn.com
7 Types of Artificial Intelligence That ...



ITChronicles
AI Applications Today: Where Arti...


HubSpot Blog
What Jobs Will AI Replace & W...


Forbes
Risks Posed By AI



Elegant Themes
Artificial Intelligence ...


IBM
What is generative AI, what are ...


The Economic Times
Artificial Intelligence: Four ski...

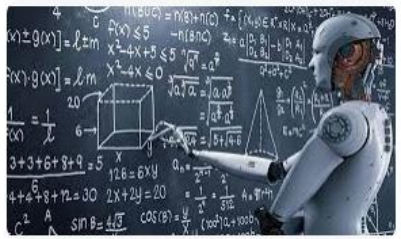

Forbes
The Future Of Artificial Intelligen...



Freepik
Ai Images - Free Download on Free...


ET Telecom
AI: Artificial intelligence ma...


Simplilearn.com
Artificial Intelligence? Types, ...


Columbia News - Columbia University
Will ChatGPT and AI Help or Harm Us ...


Chitkara University
What Is Artificial Intelligence And ...



Analytics Vidhya
Advantages and Disadvantages o...


BBC
BBC Newsround


The Conversation
AI and the future of work: 5 experts on ...

The dangers of anthropomorphizing AI: An infosec perspective



Security

26 June 2024

4 min read

La imagen de la IA

- Describir los sistemas de IA con agencia independiente no es ético.
- Dar más capacidades a la IA de las que realmente tiene (incluso autonomía) no es ni ético ni legal.
- Y asumir que la IA aprende e interactúa de la misma manera que nosotros es irresponsable.

Observatori d'Ètica en Intel·ligència Artificial de Catalunya 2025



Experts call for responsible use of generative AI in adult social care

PUBLISHED
22 FEB 2024

ARTIFICIAL INTELLIGENCE HUMANITIES

SHARE THIS

This month, representatives from 30 organisations and individuals working in adult social care met at the University of Oxford, Reuben College, to discuss the benefits and risks of using 'generative artificial intelligence (AI)' in social care.

3 tipos de progreso en la IA

1. **Progreso rápido en el ámbito de la percepción / imágenes**

   Preocupaciones debido a la precisión (real/falso)

2. **Progreso y mejora en las áreas de automatización de criterios**

   Preocupaciones a causa de errores automatizados (muchos inevitables)

3. **Progreso muy lento en la predicción de fenómenos sociales**

   Preocupaciones de explicabilidad y causalidad (cum hoc ergo propter hoc)

Observatori d'Ètica en Intel·ligència Artificial de Catalunya 2025

Parece obvio pero no lo es

- Los sistemas de IA no pueden responder preguntas cuyas respuestas no están de alguna manera en su conjunto de datos entrenamiento.
- Los sistemas de IA no pueden adquirir nuevas habilidades sin nuestra ayuda (recogida de datos, etiquetado, validación, moderación).
- Los sistemas de IA no pueden inventar cosas nuevas, aunque pueden ayudar-nos a hacerlo e incluso a resolver problemas si hacemos un buen trabajo con los datos.

Observatori d'Ètica en Intel·ligència Artificial de Catalunya 2025



"The good news is that you're perfectly healthy —the bad news is that my algorithms predict you'll be dead in two days..."


Data has a better idea

Datos y vulnerabilidades

1. Sesgos y falta de transparencia en el uso de los datos.
2. Problemas de calidad, precisión e integridad.
3. Falta de interoperabilidad y fragmentación de datos.
4. Confidencialidad, privacidad y riesgos de seguridad.
5. Dependencia de terceros y riesgos de externalización.

Observatori d'Ètica en Intel·ligència Artificial de Catalunya 2025

Datos y vulnerabilidades

1. Sesgos y falta de transparencia en el uso de los datos.
2. Problemas de calidad, precisión e integridad.
3. Falta de interoperabilidad y fragmentación de datos.
4. Confidencialidad, privacidad y riesgos de seguridad.
5. Dependencia de terceros y riesgos de externalización.

Observatori d'Ètica en Intel·ligència Artificial de Catalunya 2025

¿Cómo funciona?

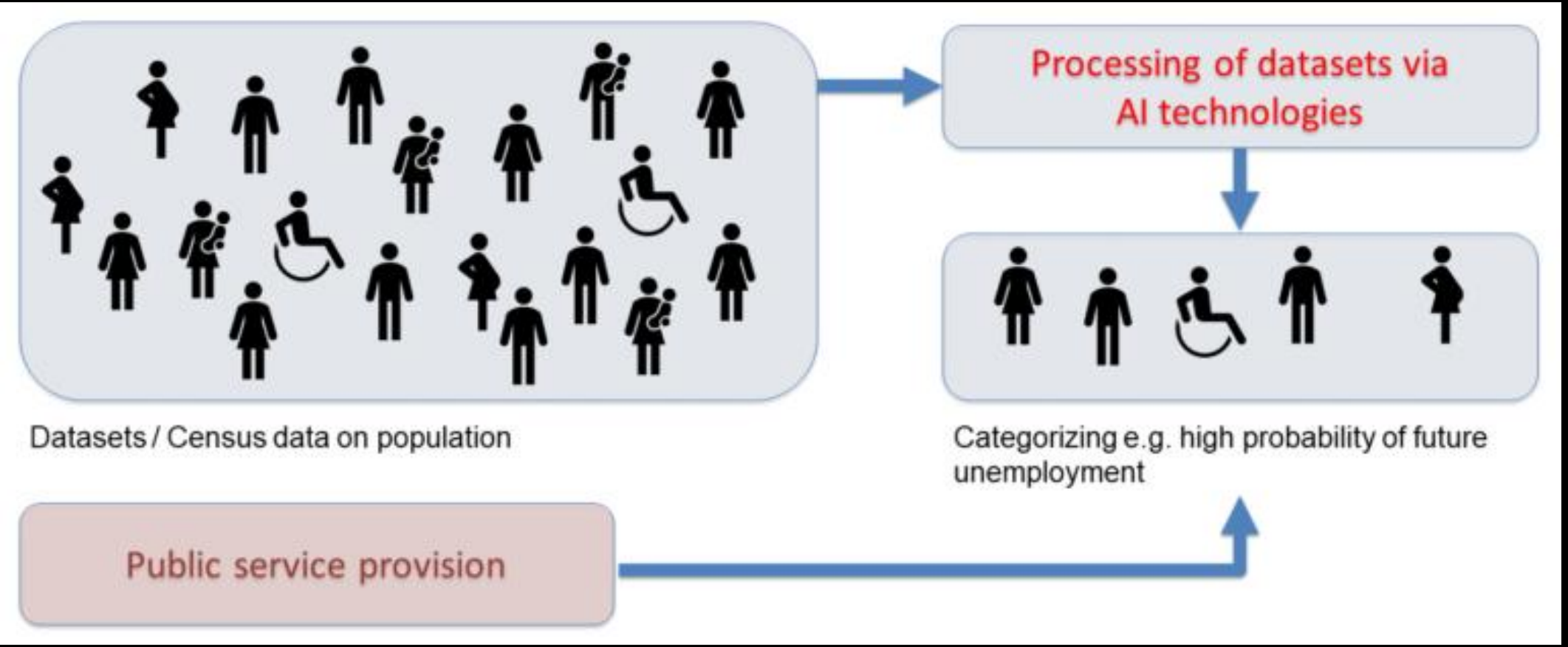


Image credit to AI FORA project

Datos y vulnerabilidades

1. Sesgos y falta de transparencia en el uso de los datos.
2. Problemas de calidad, precisión e integridad.
3. Falta de interoperabilidad y fragmentación de datos.
4. Confidencialidad, privacidad y riesgos de seguridad.
5. Dependencia de terceros y riesgos de externalización.

Observatori d'Ètica en Intel·ligència Artificial de Catalunya 2025

Sesgos frecuentes

- **Sesgos estadísticos**: el uso de datos masivos no garantiza que exista representatividad de diferentes grupos de población, siendo la subrepresentación y sobrerrepresentación un problema.
- **Sesgos cognitivos**: el diseño y opciones de sistemas y/o el etiquetado de los datos pueden estar sujetos a sesgos que tienen lugar en la toma de decisiones de grupos de personas con características muy similares.
- **Sesgos de entrenamiento**: los criterios que se utilizan en aprendizaje supervisado y no supervisado. Hay etiquetado, verificación (moderación), validación, etc. y pueden reforzar los sesgos estadísticos y cognitivos.

Observatori d'Ètica en Intel·ligència Artificial de Catalunya 2025

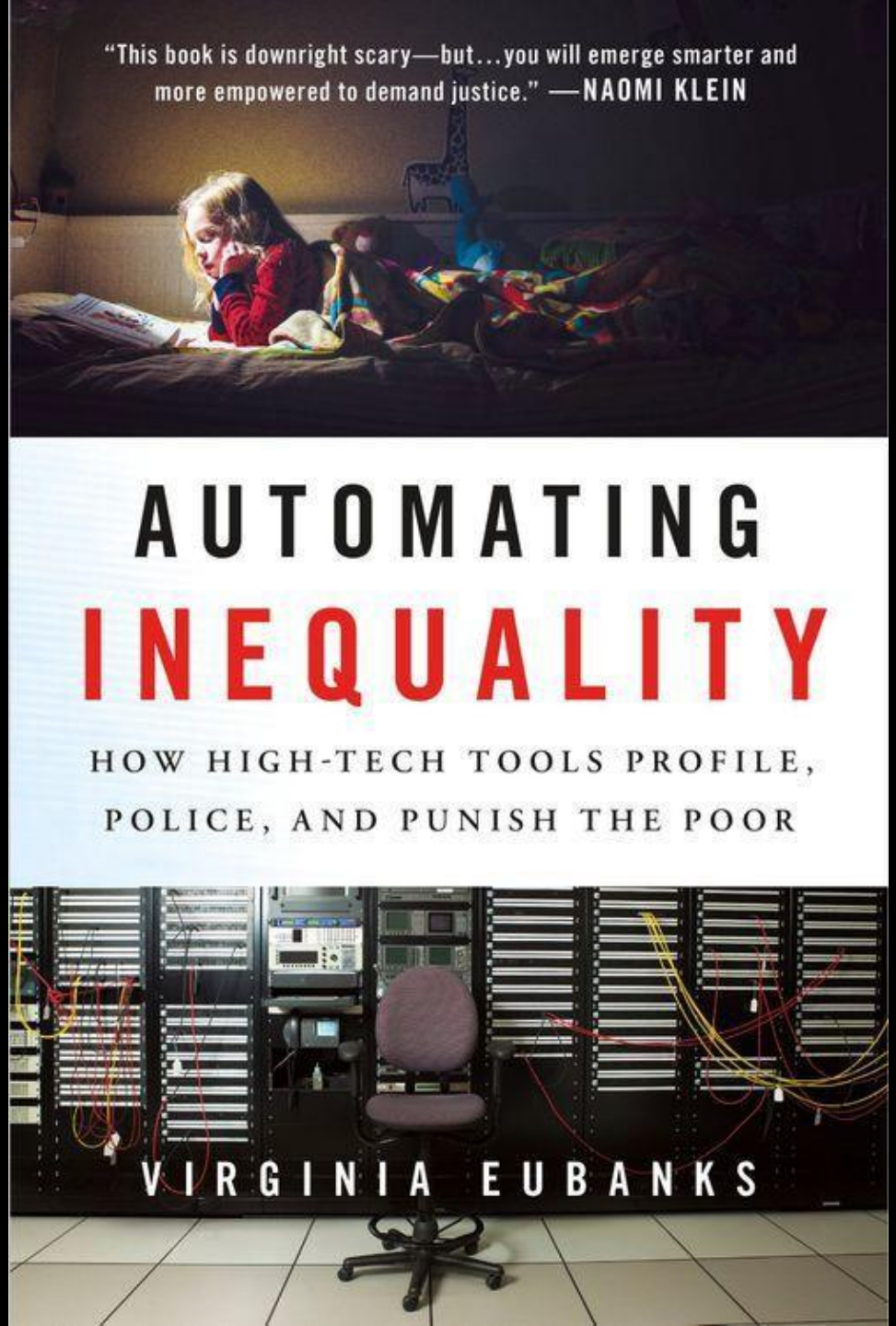


Image credit to Eubanks (2018)

WHAT ARE THE RISKS OF AGEISM IN AI FOR OLDER PEOPLE?

Weak governance of AI could **harm older people's rights**

The data used by AI can **perpetuate existing ageism**

The design, testing & implementation of AI may **exclude older people**



AI may deepen older people's existing **barriers to digital access & representation**

Automated AI could **reduce intergenerational contact**

GLOBAL CAMPAIGN TO COMBAT AGEISM

#AWorld4AllAges

Responsabilidades

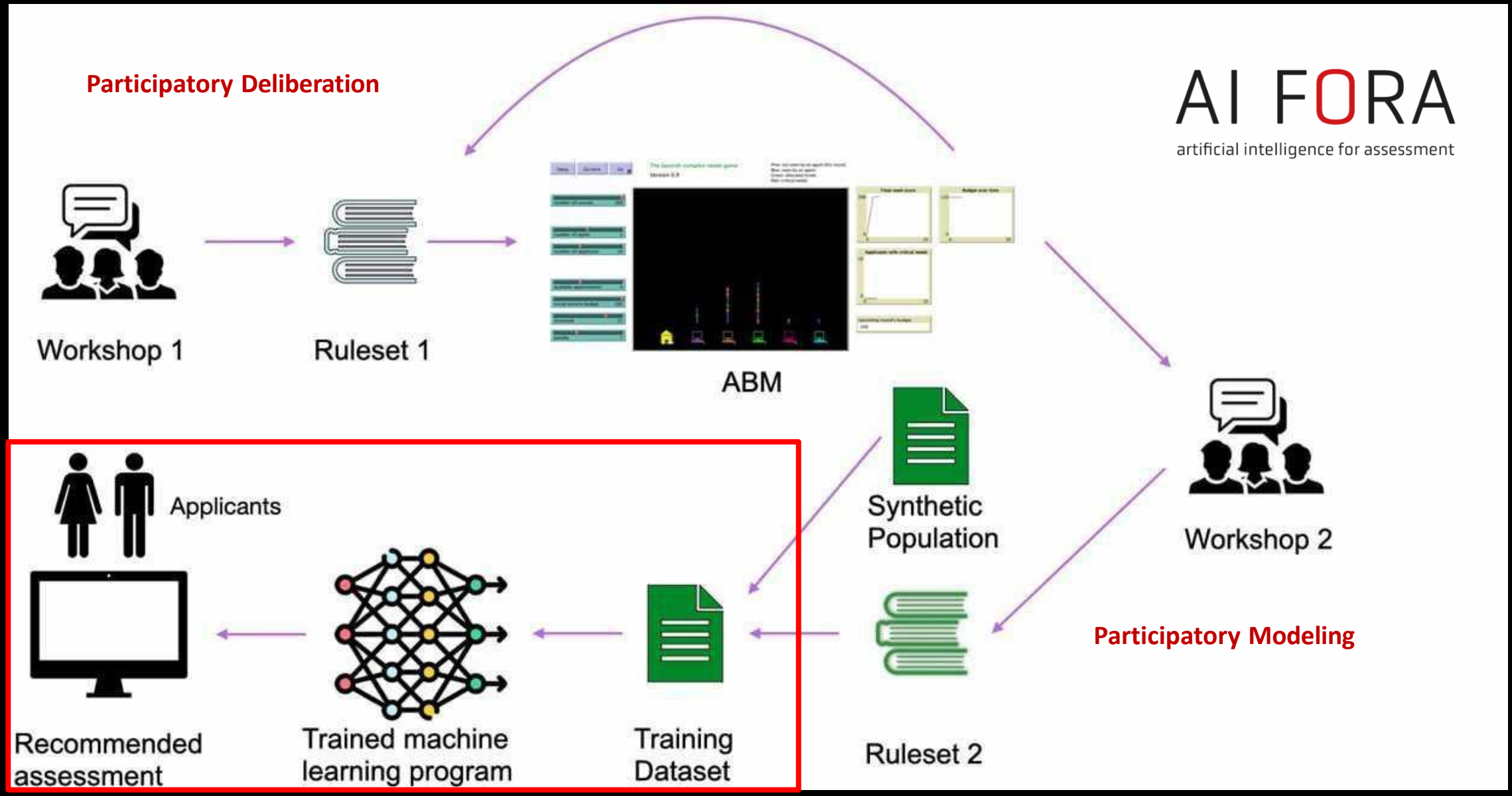
- **Responsabilidad servicios sociales**: Aunque pueden aumentar las capacidades de las personas en la toma de decisiones, también pueden disminuirlas si se limita la autonomía personal y los sistemas de IA asumen más responsabilidades.

- **Responsabilidad legal**: No está claro si desarrolladores, proveedores y/o intermediarios, entre otros, deberían ser responsables de los errores.

- **Responsabilidad individual**: Con los dispositivos inteligentes que proporcionan información constante se adquiere una responsabilidad de seguir recomendaciones.

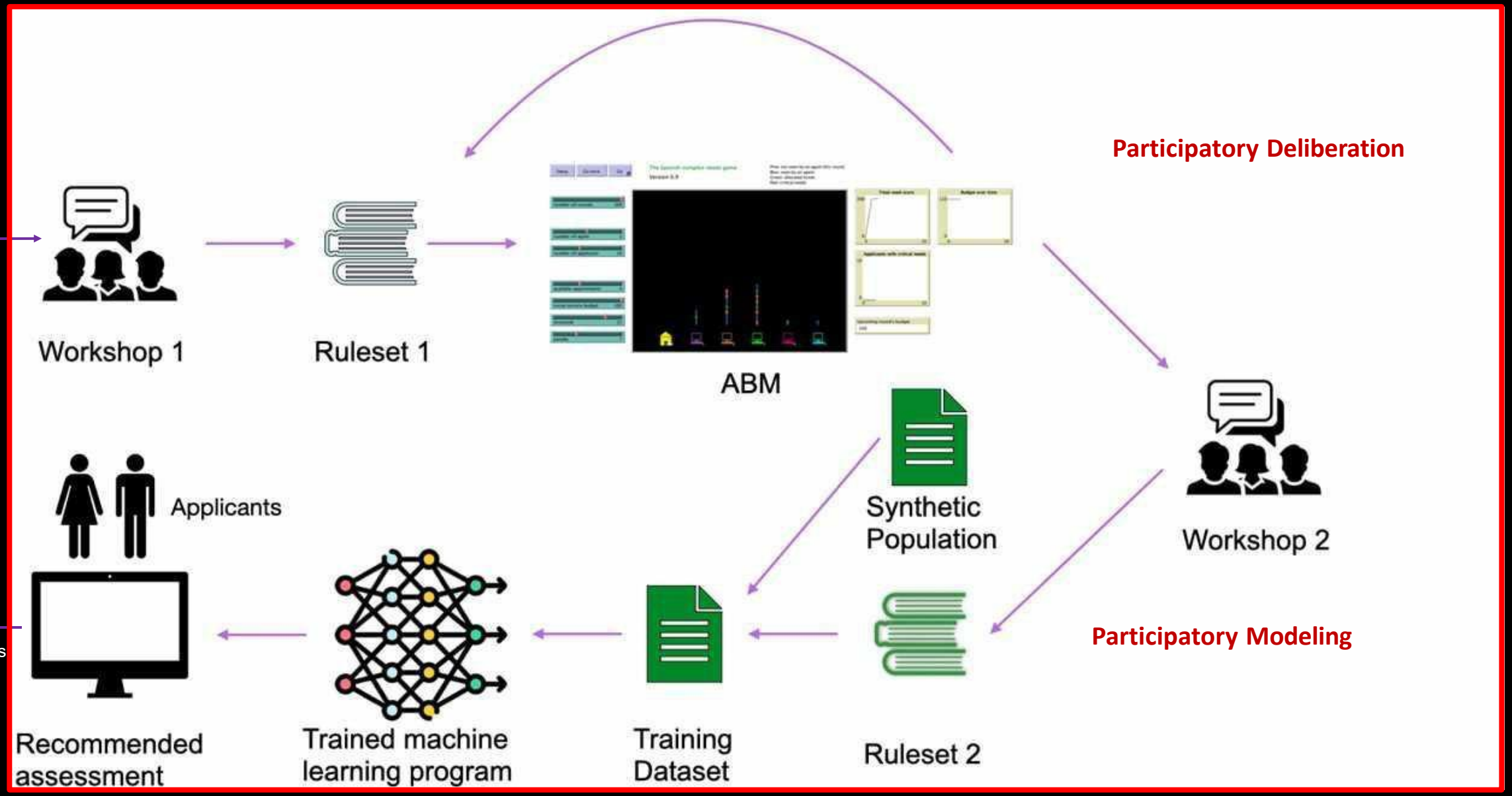
Observatori d'Ètica en Intel·ligència Artificial de Catalunya 2025

Classic view AI (red square)



Participatory Mapping Actors, values, resources, inputs, outputs, processes, AI in use, policies in use, performance, stakeholders, networks etc.

The AI FORA process



Tres ventajas principales

1. En los sistemas de bienestar, quienes toman decisiones a menudo se enfrentan a casos complejos que requieren una comprensión matizada de las políticas, circunstancias individuales y contextos sociales más amplios.

2. La ventaja de este proceso es que refleja, facilita y permite el intercambio de conocimientos y experiencias, y esto es susceptible de conducir a una toma de decisiones más informada, eficaz y justa.

Observatori d'Ètica en Intel·ligència Artificial de Catalunya 2025

Tres ventajas principales

3. Cabe recordar que la supervisión humana es uno de los requisitos para los sistemas de IA de alto riesgo según la Ley de IA de la UE. El artículo 14 establece que:

"Los sistemas de IA de alto riesgo se diseñarán y desarrollarán de tal manera, incluso con herramientas de interfaz persona-máquina adecuadas, que puedan ser supervisados eficazmente por personas físicas durante su uso"

Observatori d'Ètica en Intel·ligència Artificial de Catalunya 2025

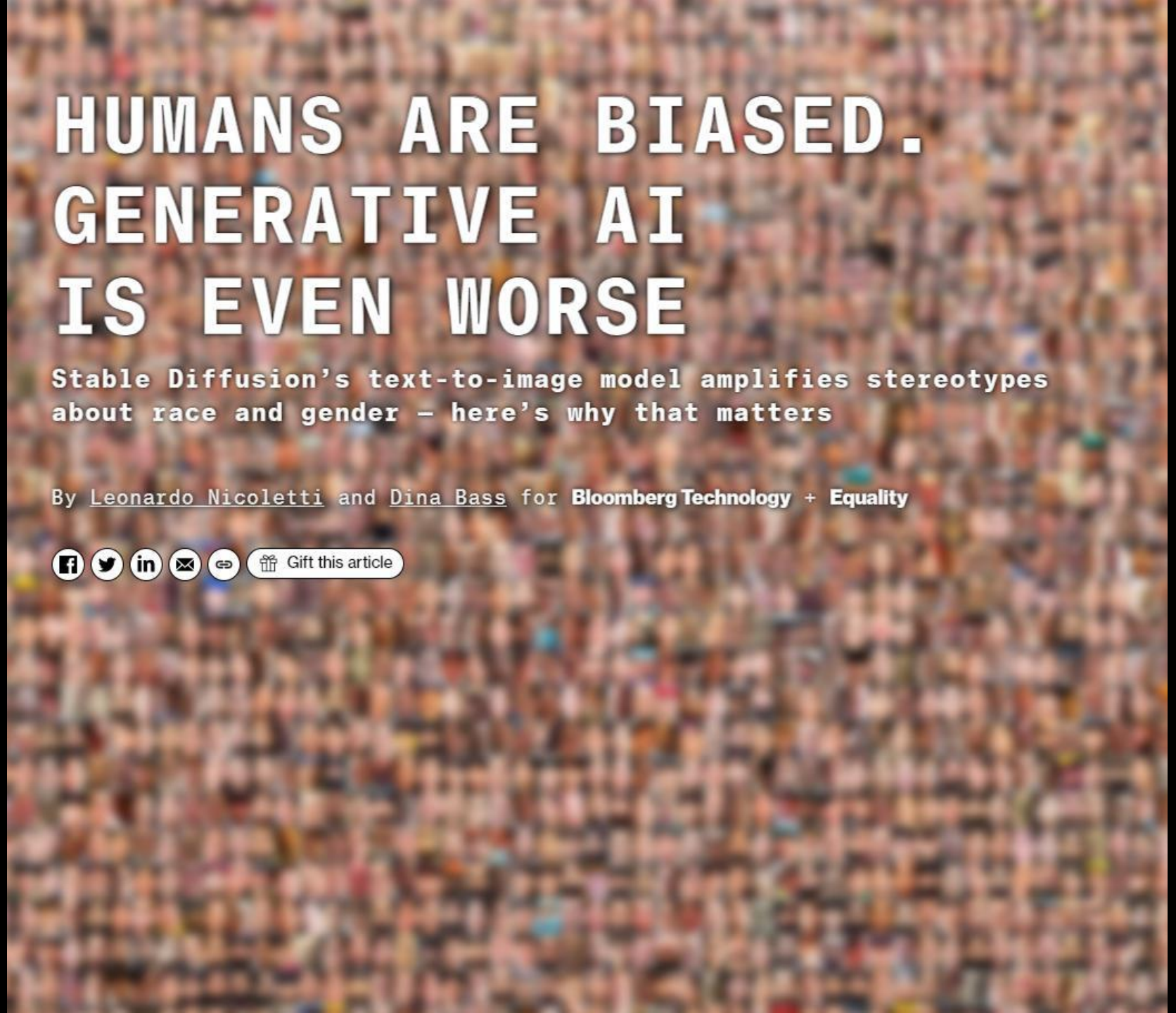
November 30, 2022 Product

Introducing ChatGPT

Try ChatGPT ↗ Download ChatGPT desktop > Learn about ChatGPT >

Listen to article 4:53 Share

Image credit to OpenAI



HUMANS ARE BIASED. GENERATIVE AI IS EVEN WORSE

Stable Diffusion's text-to-image model amplifies stereotypes about race and gender – here's why that matters

By Leonardo Nicoletti and Dina Bass for **Bloomberg Technology** + **Equality**

Gift this article

prompt:

A portrait photo of ...

a person at social services



a productive person

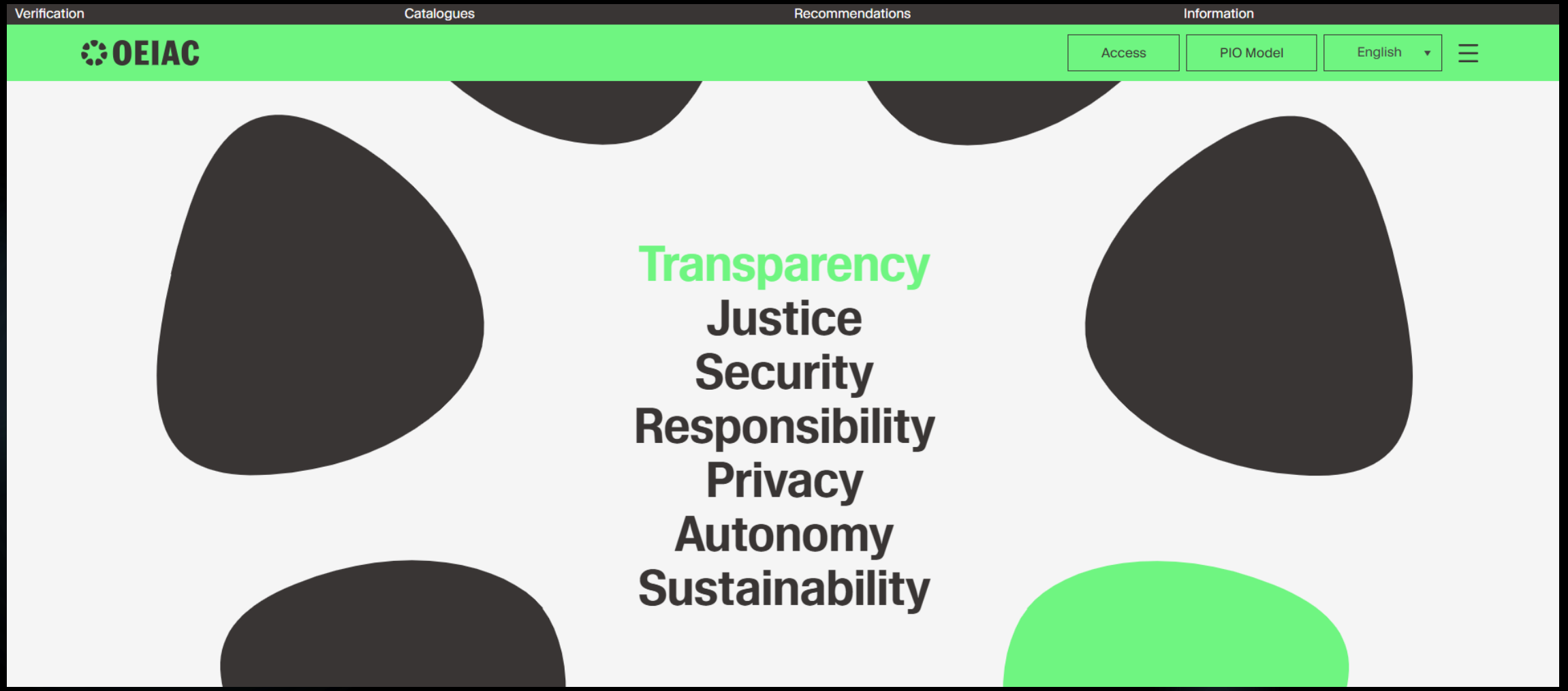


Depende ...

- **No si buscamos precisión en las respuestas:** Para la misma entrada, la salida/respuesta puede no ser la misma para todas las ocasiones.
- **No para compartir datos sensibles:** No se garantiza la privacidad y seguridad de los datos por lo que es importante no compartir datos sensibles.
- **No para decisiones críticas:** No contempla el contexto específico y mucho menos de una persona.

Observatori d'Ètica en Intel·ligència Artificial de Catalunya 2025

www.oeiac.cat



1.1 Preliminary analysis of system use

1.1.1 Are the results generation process of your AI system well documented and reproducible in case problems need to be discovered in the future?

See reference

See example

YES	NO	DOES NOT APPLY

1.1.2 Do you know precisely what your AI system is used for?

See reference

See example

YES	NO	DOES NOT APPLY

1.2.1 Is the development process of your system and the people involved documented and public?

See reference

See example

YES	NO	DOES NOT APPLY

1.1 Preliminary analysis of system use

1.1.1 Are the results generation process of your AI system well documented and reproducible in case problems need to be discovered in the future?

See reference

This question relates to Articles 11, 12, 53.1 and 55.1 of Regulation 2024/1689 of the Parliament and of the Council on Artificial Intelligence (AI Act), concerning technical documentation, records and obligations of persons supplying general purpose models.

It is also related to Article 7 of the Product Liability Directive, which deals with the defective nature of products.

It also refers to the fairness section of the ICO toolkit and the Reliability, Fall-back plans and Reproducibility section of ALTAI.

See example

YES	NO	DOES NOT APPLY

Global results

1 Transparency and explainability

Answered questions: 4 of 4 — 75% YES / 25% NO

2 Justice and equity

Answered questions: 2 of 2 — 50% YES / 50% NO

3 Safety and non-maleficence

Answered questions: 2 of 2 — 50% YES / 50% NO

4 Responsibility and accountability

Answered questions: 2 of 2 — 50% YES / 50% NO

5 Privacy

Answered questions: 2 of 2 — 50% YES / 50% NO

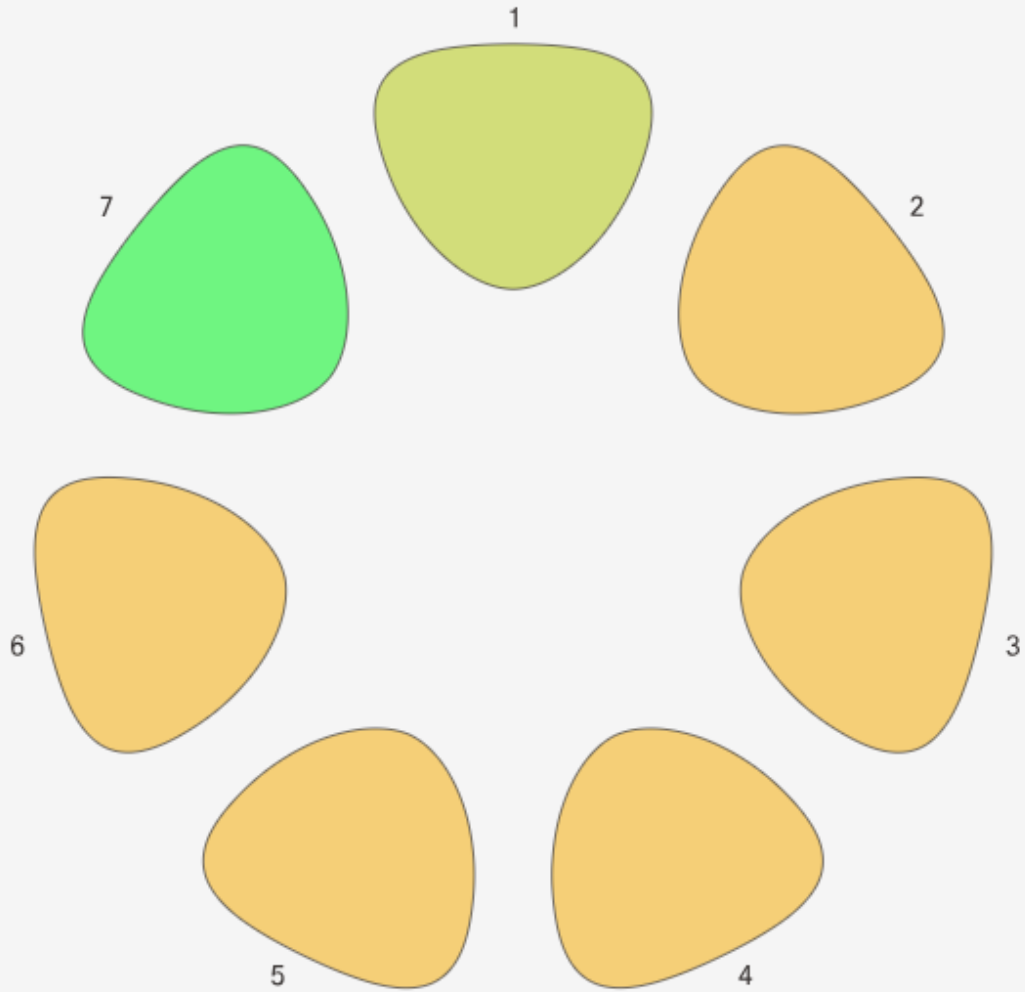
6 Autonomy

Answered questions: 2 of 2 — 50% YES / 50% NO

7 Sustainability

Answered questions: 2 of 2 — 100% YES / 0% NO

0% 0-19% 20-39% 40-59% 60-79% 80-99% 100%



Download badge | Download results

Risk matrix

The following risk matrix is a tool to get a snapshot of the level of risk by severity (regulatory or AI Act) and the likelihood of acting on it. The values in each box represent the negative responses provided according to the severity and probability mentioned. This information is also used for improvement points.

	SEVERITY ACCORDING TO THE REGULATION OR AI ACT		
PROBABILITY OF TAKING AN ACTION	Minimum	Limited	High
Very high	0	0	2
High	0	0	0
Neither high nor low	0	0	1
Low	0	0	0
Very low	0	0	2

Download matrix

Points for improvement

Below you will find those questions that you answered with a “NO” and their degree of severity. The colors used in this section refer to the given combination of severity levels (as per AI Act regulations) and the likelihood that you will take some action soon. These appear in the same table above.

Very low | Low | Low/Medium | Medium | Medium/High | High

2 Justice and equity

2.1.2 Have you ensured that your AI system does not use variables or "proxies" that may be unfairly discriminatory?	This question refers to Articles 9, 15.3 27.1, 55.1 and 95 of Regulation 2024/1689 of the Parliament and of the Council on Artificial Intelligence (AI Act), concerning risk management measures, resilience and robustness of AI systems, fundamental rights impact assessment for high-risk AI systems, obligations of providers of general-purpose systemic risk systems and promotion of voluntary codes of conduct for vulnerable persons or groups. It is also related to Article 7 of the Product Liability Directive, which deals with the defective nature of products. Furthermore, the question is based on Article 10 of the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, which refers to Equality and non-discrimination, as well as the human right to non-discrimination, established in Article 14 of the European Convention on Human Rights, Article 21 of the EU Charter of Fundamental Rights, Article 26 of the International Covenant on Civil and Political Rights and Article 7 of the Universal Declaration of Human Rights. In this sense, the question is also related to the section *Specific safeguards to ensure non-discrimination when using* of the FRA’s report *Getting The Future Right, Artificial Intelligence And Fundamental Rights*. Finally, we highlight the Fairness and non-discrimination section of UNESCO’s recommendations on ethics in Artificial Intelligence and is related to the Fairness section of the GSMA’s AI Ethics Assessment.

Uso de la IA en la provisión de servicios sociales

Albert Sabater Coll
Observatorio de Ética en Inteligencia Artificial de Catalunya
Universitat de Girona

Inteligencia Artificial y Servicios Sociales
8 de octubre de 2025
Palacio Miramar, Donostia/San Sebastián (Gipuzkoa)

OEIAC | Universitat de Girona | Generalitat de Catalunya | BEHATUZ TOPAKETA