



ÉTICA PARA ALGORITMOS

RADOSLAV DEPOLO

OCTUBRE 2019

Kennedys



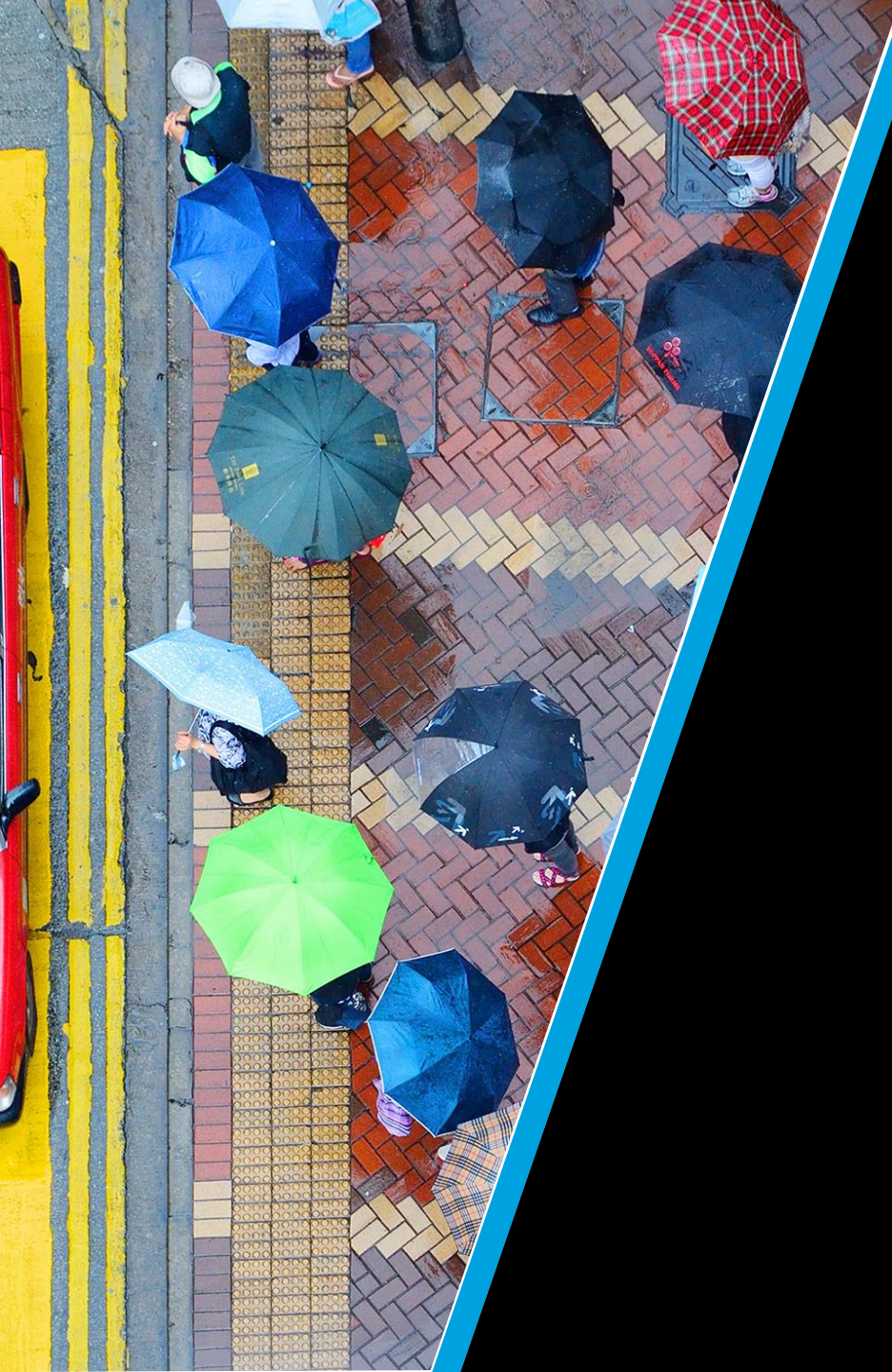
IA: CONCEPTO Y CONTEXTO

CONCEPTO

- En general, es la capacidad que tiene una **máquina** de **pensar** mediante la utilización de **algoritmos** que le proporcionan la capacidad de interpretación, decisión y **resolución de problemas** de forma **autónoma** ante las señales o **datos** que recibe o capta.
- Un **algoritmo** es un conjunto de reglas y fórmulas matemáticas o de programación que, aplicado sistemáticamente a unos datos de entrada apropiados, resuelve un problema en un número finito de pasos elementales.
- Los procesos de AI incluyen el **aprendizaje** (adquisición de información y reglas para usarla), el **razonamiento** (uso de reglas para llegar a conclusiones aproximadas o definitivas) y la **autocorrección**.
- Se habla de **IA débil** (sistema de IA diseñado y programado para una tarea en particular - Siri) y de **IA fuerte** (sistema de IA con habilidades cognitivas humanas generalizadas).

CONTEXTO

| | | | | |
|---|--|--|--|--|
|  <p>#1 Artificial Intelligence AI /Machine Learning / Deep Learning</p> |  <p>#2 Internet of Things IOT, IIOT, Sensors & Wearables</p> |  <p>#3 Mobile/Social Internet Advancements - Search/Social/Messaging/Livestreams</p> |  <p>#4 Blockchain Distributed Ledger Systems, Cryptocurrencies & DApps</p> |  <p>#5 Big Data Apps, Infrastructure, Technologies + Predictive Analytics</p> |
|  <p>#6 Automation Information, Task, Process, Machine, Decision & Action</p> |  <p>#7 Robots Cons./Comm./Indus., Robots, Drones & Autonomous Vehicles</p> |  <p>#8 Immersive Media - #VR/ #AR/ #MR/ 360°/ Video?Gaming</p> |  <p>#9 Mobile Technologies Infrastructure, networks, standards, services & devices</p> |  <p>#10 Cloud Computing SaaS, IaaS, PaaS & MESH Apps</p> |
|  <p>#11 3D Printing Additive Manufacturing & Rapid Prototyping</p> |  <p>#12 CX Customer Journey, Experience Commerce & Personalization</p> |  <p>#13 EnergyTech Efficiency, Energy Storage & Decentralized Grid</p> |  <p>#14 Cybersecurity Security, Intelligence Detection, Remediation & Adaptation</p> |  <p>#15 Voice Assistants Interfaces, Chatbots & Natural Language Processing</p> |
|  <p>#11 Nanotechnology Computing, Medicine, Machines + Smart Dust</p> |  <p>#17 Collaborative Tech. Crowd, Sharing, Workplace & Open Source Platforms & Tools</p> |  <p>#18 Health Tech. Advanced Genomics, Bionics & Health Care Tech.</p> |  <p>#19 Human-Computer Interaction Facial/Gesture Recognition, Biometrics, Gaze Tracking</p> |  <p>#20 Geo-spatial Tech. GIS, GPS, Mapping & Remote Sensing, Scanning, Navigation</p> |
|  <p>#21 Advanced Materials Composites, Alloys, Polymers, Biomimicry, Nanomanufacturing</p> |  <p>#22 New Touch Interfaces Touch Screens, Haptics, 3D Touch, Paper, Feedback & Exoskeletons</p> |  <p>#23 Wireless Power</p> |  <p>#24 Clean Tech. Bio-/Enviro-Materials + Solutions, Sustainability, Treatment & Efficiency</p> |  <p>#25 Quantum Computing + Exascale Computing</p> |
|  <p>#26 Smart Cities + Infrastructure & Transport</p> |  <p>#27 Edge/Computing + Fog Computing</p> |  <p>#28 Faster, Better Internet Broadband incl. Fiber, 5G, Li-Fi, LPN and LoRa</p> |  <p>#29 Proximity Tech Beacons, RFID, Wi-Fi, Near-Field Communications & Geofencing</p> |  <p>#30 New Screens TVs, Digital Signage, OOH, MicroLEDs & Projections</p> |



CASOS

Kennedys

| | # | Summary Nr | Event Date | Report ID | Fat | SIC | Event Description |
|--------------------------|----|---------------------------|------------|-----------|-----|------|--|
| <input type="checkbox"/> | 1 | 112475.01 | 01/10/2019 | 0454712 | X | | Employee Is Crushed And Killed By Wind-Up Machine |
| <input type="checkbox"/> | 2 | 101894.01 | 12/12/2017 | 0418600 | | | Worker Sustains Fractured Hip When Struck By Robot |
| <input type="checkbox"/> | 3 | 105315.01 | 11/27/2017 | 0214500 | | | Employee Is Injured When Struck By 100 Pound Bag Of Flour |
| <input type="checkbox"/> | 4 | 96664.015 | 01/17/2017 | 0950644 | | | Employee Is Struck By Robot Arm And Sustains Fractured Stern |
| <input type="checkbox"/> | 5 | 200627032 | 06/16/2013 | 0454712 | X | 3711 | Employee Is Struck By Axis Arm< Later Dies |
| <input type="checkbox"/> | 6 | 202570313 | 03/07/2013 | 0452110 | X | 3714 | Maintenance Worker Is Struck And Killed By Robot |
| <input type="checkbox"/> | 7 | 201392776 | 12/15/2012 | 0552652 | X | 3465 | Robot Crushes And Kills Worker Inside Robot Work Cell |
| <input type="checkbox"/> | 8 | 202509683 | 11/29/2012 | 0950622 | | 5099 | Employee Suffers Head Injures In Fall On Energized Track |
| <input type="checkbox"/> | 9 | 200785285 | 08/02/2011 | 0626300 | X | 2051 | Employee Is Killed When Caught In Equipment |
| <input type="checkbox"/> | 10 | 202475737 | 07/21/2009 | 0950636 | X | 5141 | Employee Is Killed By Robotic Palletizer |
| <input type="checkbox"/> | 11 | 200052975 | 05/13/2007 | 0728900 | X | 2821 | Employee Dies After Being Struck By Robotic Arm |
| <input type="checkbox"/> | 12 | 201391489 | 10/31/2006 | 0552652 | X | 3542 | Worker Is Killed, Lockout Procedures Not Followed |
| <input type="checkbox"/> | 13 | 200631406 | 07/24/2006 | 0316400 | X | 3489 | Employee Is Killed When Crushed By Robot |
| <input type="checkbox"/> | 14 | 200623742 | 03/22/2006 | 0454712 | X | 3465 | Employee Dies When Struck By Robotic Equipment |
| <input type="checkbox"/> | 15 | 201634839 | 11/16/2004 | 1054112 | | 3325 | Employee Fractures Chest When Crushed By Robot |
| <input type="checkbox"/> | 16 | 202075727 | 03/30/2004 | 0453730 | X | 3714 | Employee Was Killed By Industrial Robots |
| <input type="checkbox"/> | 17 | 200991081 | 12/13/2003 | 0551800 | X | 3714 | Employee Is Killed When Caught In Robotic Arm |
| <input type="checkbox"/> | 18 | 171063340 | 07/28/2003 | 0854910 | X | 3949 | Employee Is Killed When Crushed By Equipment |
| <input type="checkbox"/> | 19 | 202313102 | 01/15/2002 | 0950612 | | 3674 | Employee Amputates Fingers While Examining Scanner |
| <input type="checkbox"/> | 20 | 200101236 | 12/29/2001 | 0522300 | X | 3714 | Employee Killed When Robot Pinned His Neck |



FACEBOOK'S ARTIFICIAL INTELLIGENCE ROBOTS SHUT DOWN AFTER THEY START TALKING TO EACH OTHER IN THEIR OWN LANGUAGE

Facebook abandoned an experiment after two artificially intelligent programs appeared to be chatting to each other in a strange language only they understood.

The two chatbots came to create their own changes to English that made it easier for them to work – but which remained mysterious to the humans that supposedly look after them.

ARTIFICIAL INTELLIGENCE, ALGORITHMIC PRICING AND COLLUSION[†]

EMILIO CALVANO^{*‡}, GIACOMO CALZOLARI^{&‡§},
VINCENZO DENICOLÒ^{*§}, AND SERGIO PASTORELLO^{*}

APRIL 2019

Amazon scraps secret AI recruiting tool that showed bias against women

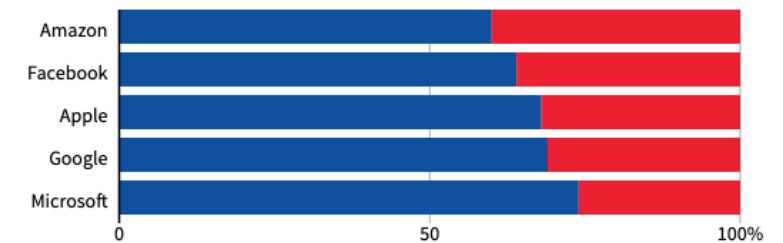
The team had been building computer programs since 2014 to review job applicants' resumes with the aim of mechanizing the search for top talent, five people familiar with the effort told Reuters.

But by 2015, the company realized its new system was not rating candidates for software developer jobs and other technical posts in a gender-neutral way.

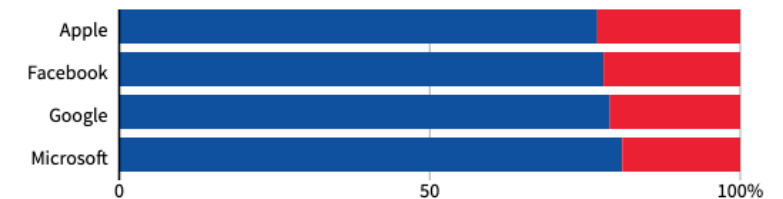
That is because Amazon's computer models were trained to vet applicants by observing patterns in resumes submitted to the company over a 10-year period. Most came from men, a reflection of male dominance across the tech industry.

GLOBAL HEADCOUNT

■ Male ■ Female



EMPLOYEES IN TECHNICAL ROLES



Experimental evidence of massive-scale emotional contagion through social networks

Adam D. I. Kramer^{a,1}, Jamie E. Guillory^{b,2}, and Jeffrey T. Hancock^{b,c}

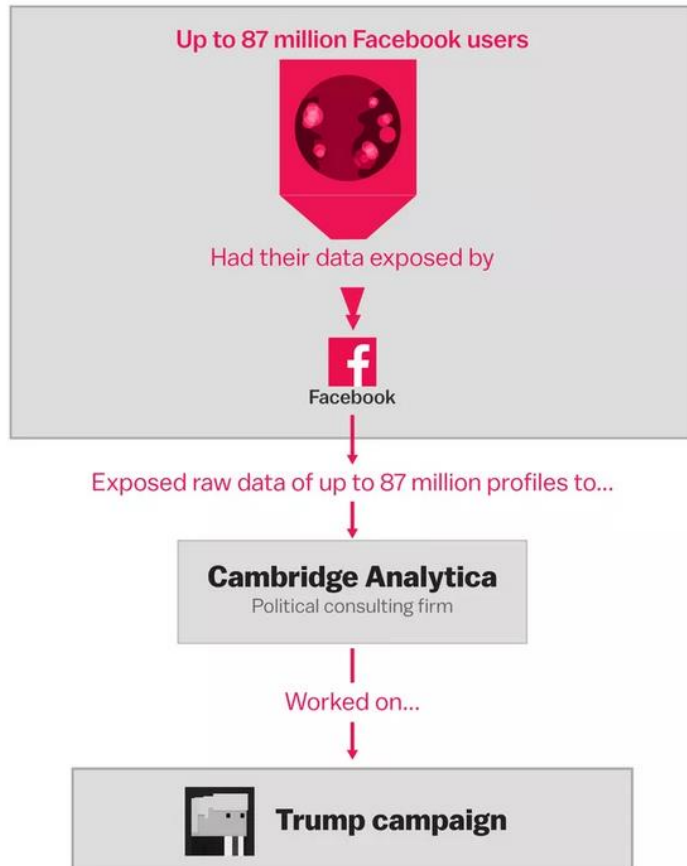
^aCore Data Science Team, Facebook, Inc., Menlo Park, CA 94025; and Departments of ^bCommunication and ^cInformation Science, Cornell University, Ithaca, NY 14853

Emotional states can be transferred to others via emotional contagion, leading people to experience the same emotions without their awareness. Emotional contagion is well established in laboratory experiments, with people transferring positive and negative emotions to others. Data from a large real-world social network, collected over a 20-y period suggests that longer-lasting moods (e.g., depression, happiness) can be transferred through networks [Fowler JH, Christakis NA (2008) *BMJ* 337:a2338], although the results are controversial. In an experiment with people who use Facebook, we test whether emotional contagion occurs outside of in-person interaction between individuals by reducing the amount of emotional content in the News Feed. When positive expressions were reduced, people produced fewer positive posts and more negative posts; when negative expressions were reduced, the opposite pattern occurred. These results indicate that emotions expressed by others on Facebook influence our own emotions, constituting experimental evidence for massive-scale contagion via social networks. This work also suggests that, in contrast to prevailing assumptions, in-person interaction and non-verbal cues are not strictly necessary for emotional contagion, and that the observation of others' positive experiences constitutes a positive experience for people.

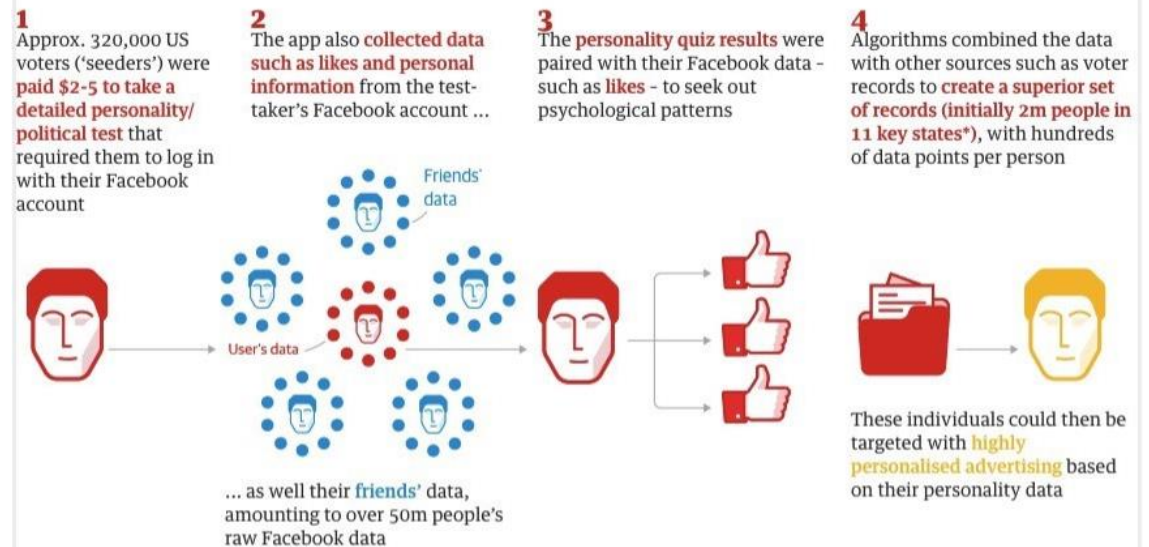
The Facebook and Cambridge Analytica scandal, explained with a simple diagram

A visual of how it all fits together. They're now shutting down.

By Alvin Chang | @alv9n | alvin@vox.com | Updated May 2, 2018, 3:25pm EDT



Cambridge Analytica: how 50m Facebook records were hijacked



Guardian graphic. *Arkansas, Colorado, Florida, Iowa, Louisiana, Nevada, New Hampshire, North Carolina, Oregon, South Carolina, West Virginia



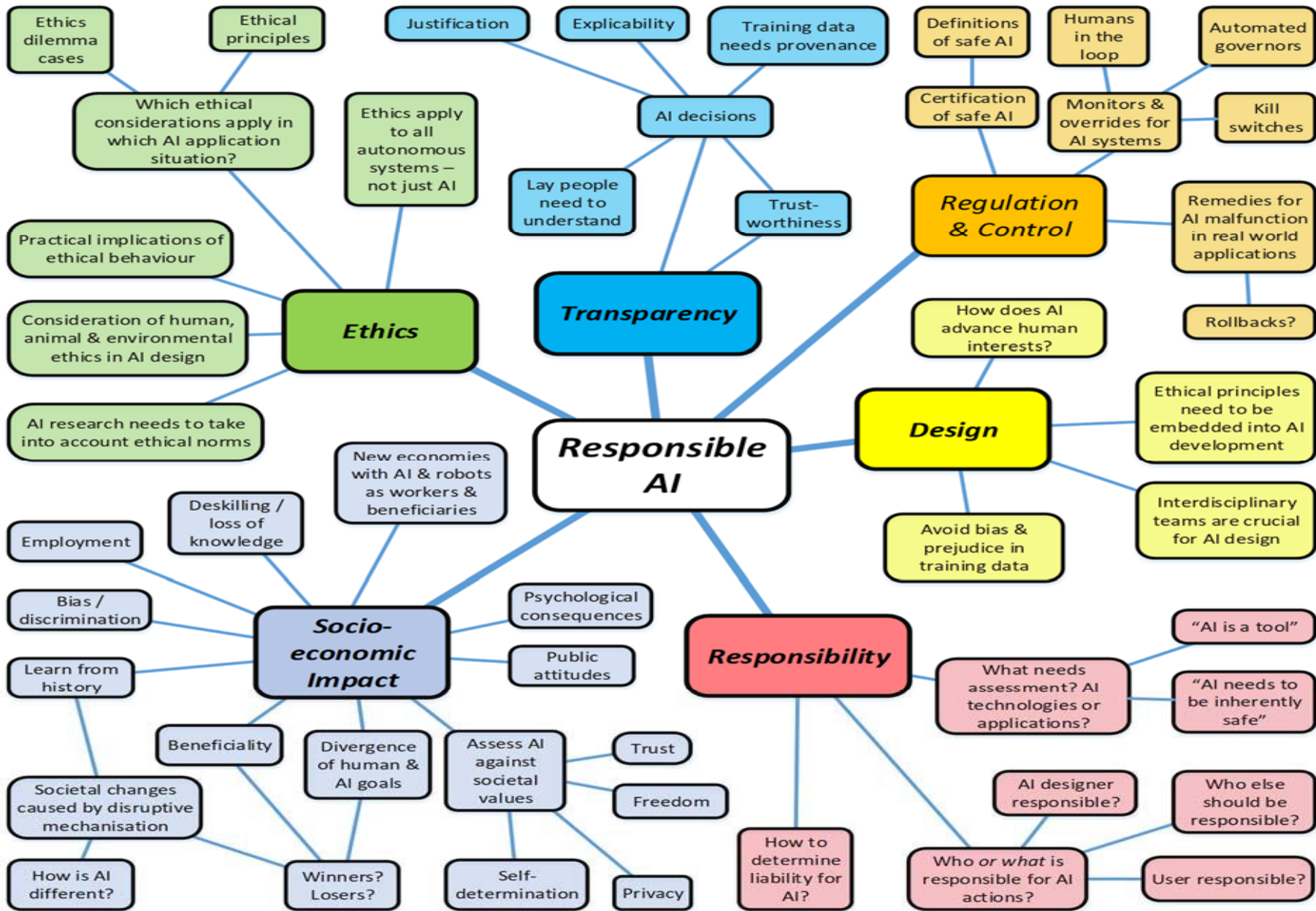
HACIA UNA IA RESPONSABLE

¿QUIÉN RESPONDE?

¿SE DEBE REGULAR?

¿SIRVE LA REGULACIÓN
ACTUAL?

¿ÉTICA O/Y DERECHO?





IA Y **DERECHO**: LA LIEBRE Y LA TORTUGA

DIAGNÓSTICO

- **Inercias largas** (proceso político/legislativo/administrativo)
- Desarrollo **ultra-veloz** de la IA
- **Aplicación ineludible** de normas **vigentes** en caso de conflicto
- Normas vigentes parecen ser insuficientes: ¿**Brecha normativa?**

¿ES NECESARIO REGULAR LA IA?

- Normas de **libre competencia**
- Normas de **protección al consumidor**
- Normas de **privacidad y datos personales**
- **Sanciones penales** (estafa/discriminación/amenazas/delitos especiales)
- Régimen de esponsabilidad **extracontractual** (Art. 2329 CC: “Por regla general todo daño que pueda imputarse a malicia o negligencia de otra persona, debe ser reparado por ésta” - ej. El que dispara imprudentemente un arma de fuego)
- Régimen de responsabilidad **contractual** (autonomía de la voluntad)

¿ES NECESARIO REGULAR LA IA?

- Normas de **libre competencia**
- Normas de **protección al consumidor**
- Normas de **privacidad y datos personales**
- **Sanciones penales** (estafa/discriminación/amenazas/delitos especiales)
- Régimen de esponsabilidad **extracontractual** (Art. 2329 CC: “Por regla general todo daño que pueda imputarse a malicia o negligencia de otra persona, debe ser reparado por ésta” - ej. El que dispara imprudentemente un arma de fuego)
- Régimen de responsabilidad **contractual** (autonomía de la voluntad)

RESPUESTA: SÍ

- Al involucra un conjunto de **problemas nuevos no encasillables** en la normativa vigente
- Al requiere **normas** generales o **estándares normativos básicos** de aplicación **universal**, pero que permitan liberar su potencial (propiedad de datos, mitigación de sesgo humano)
- Al requiere de **capacidades de supervisión y fiscalización** que **exceden las disponibles** actualmente (know how, capacidad de procesamiento, updates)
- Al implica un **desafío** para el concepto de **soberanía**, **territorialidad** y **jurisdicción** (cross border enforcement, tributación, GDPR, procesos electorales)

¿CÓMO?

ESTADO

- Constitución
- Leyes
- Reglamentos / Decretos / Guidelines

ONGs

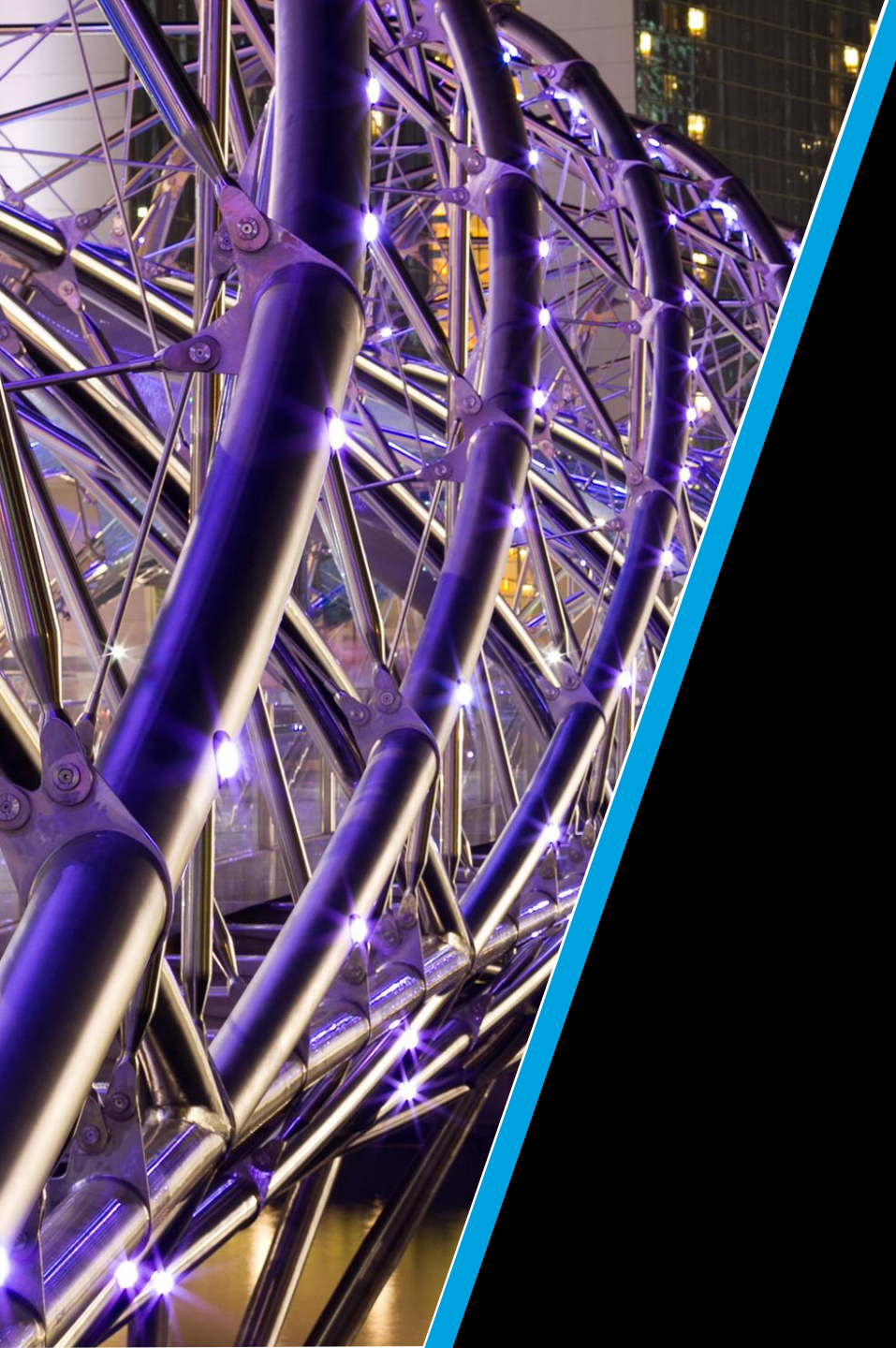
- Organismos o asociaciones internacionales
- Asociaciones gremiales y empresariales: políticas y procedimientos
- Normas técnicas y estándares (ISO, INN)

B2B/B2C

- Contratos de adhesión
- Contratos bilaterales
- Smart contracts

EMPRESAS

- Códigos de conducta
- Códigos de ética
- Sistemas de supervisión y control de riesgos de AI

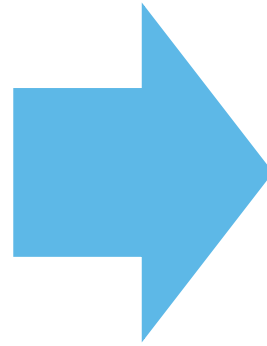


ÉTICA Y AI

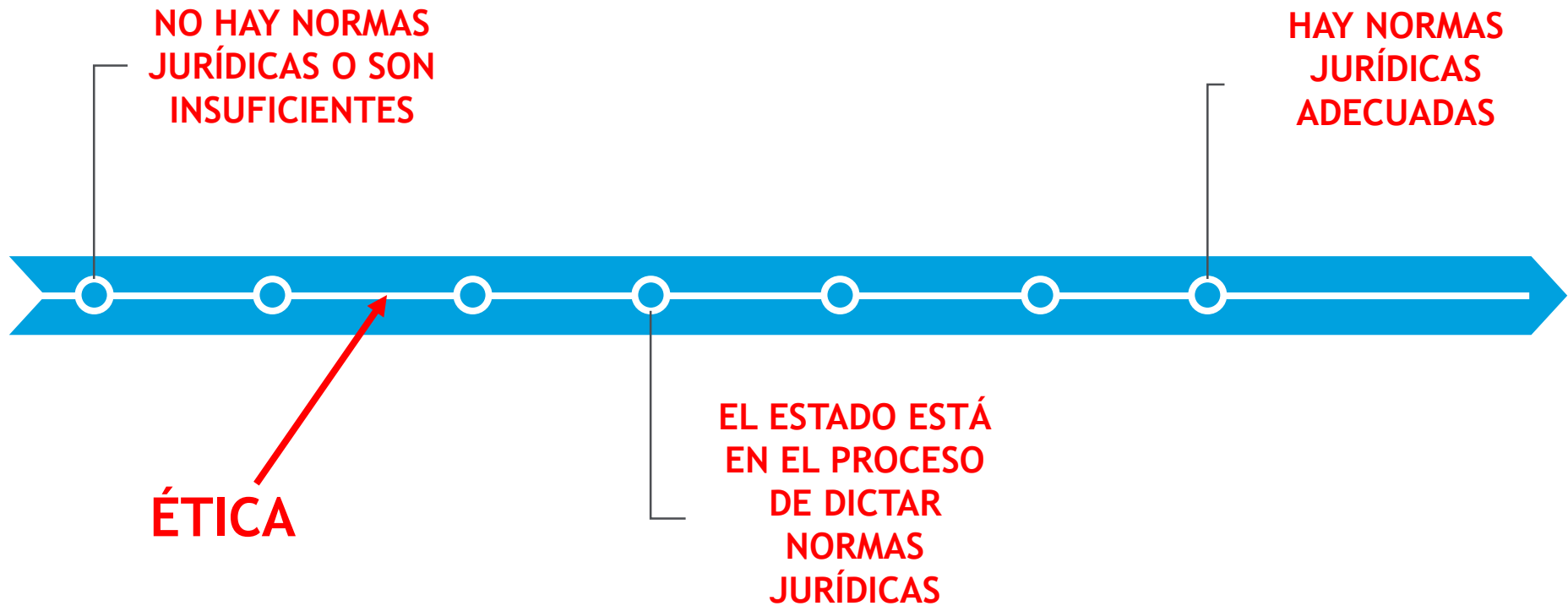
ÉTICA Y PROFUNDIDAD DE CAMPO



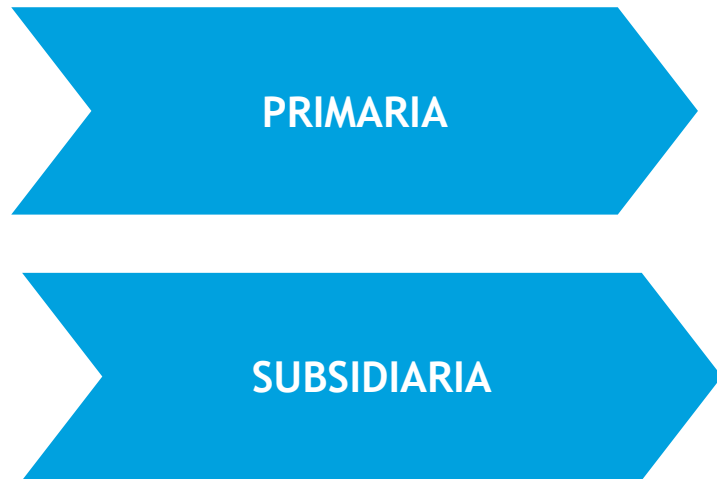
ÉTICA Y PROFUNDIDAD DE CAMPO



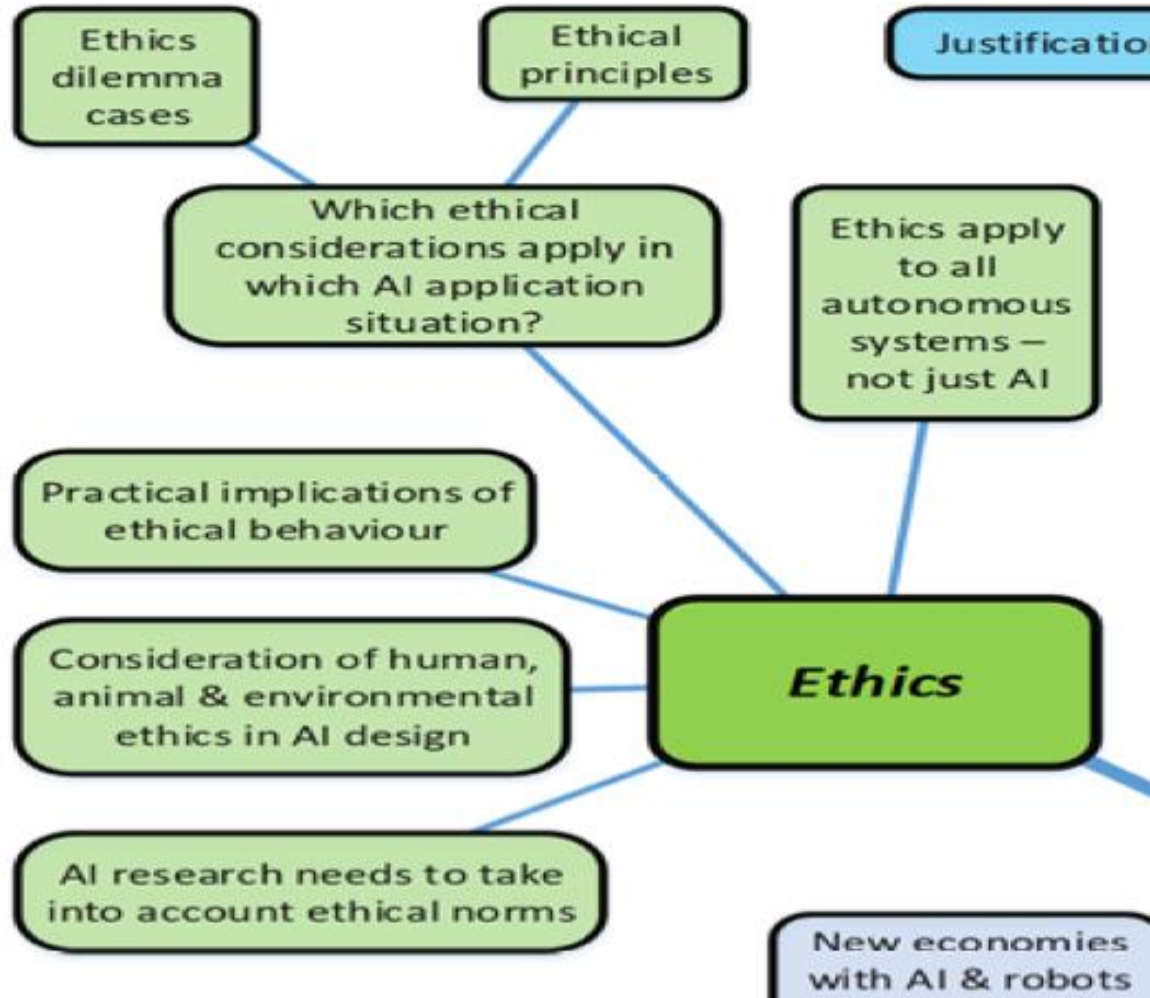
ÉTICA SIN / CON DERECHO



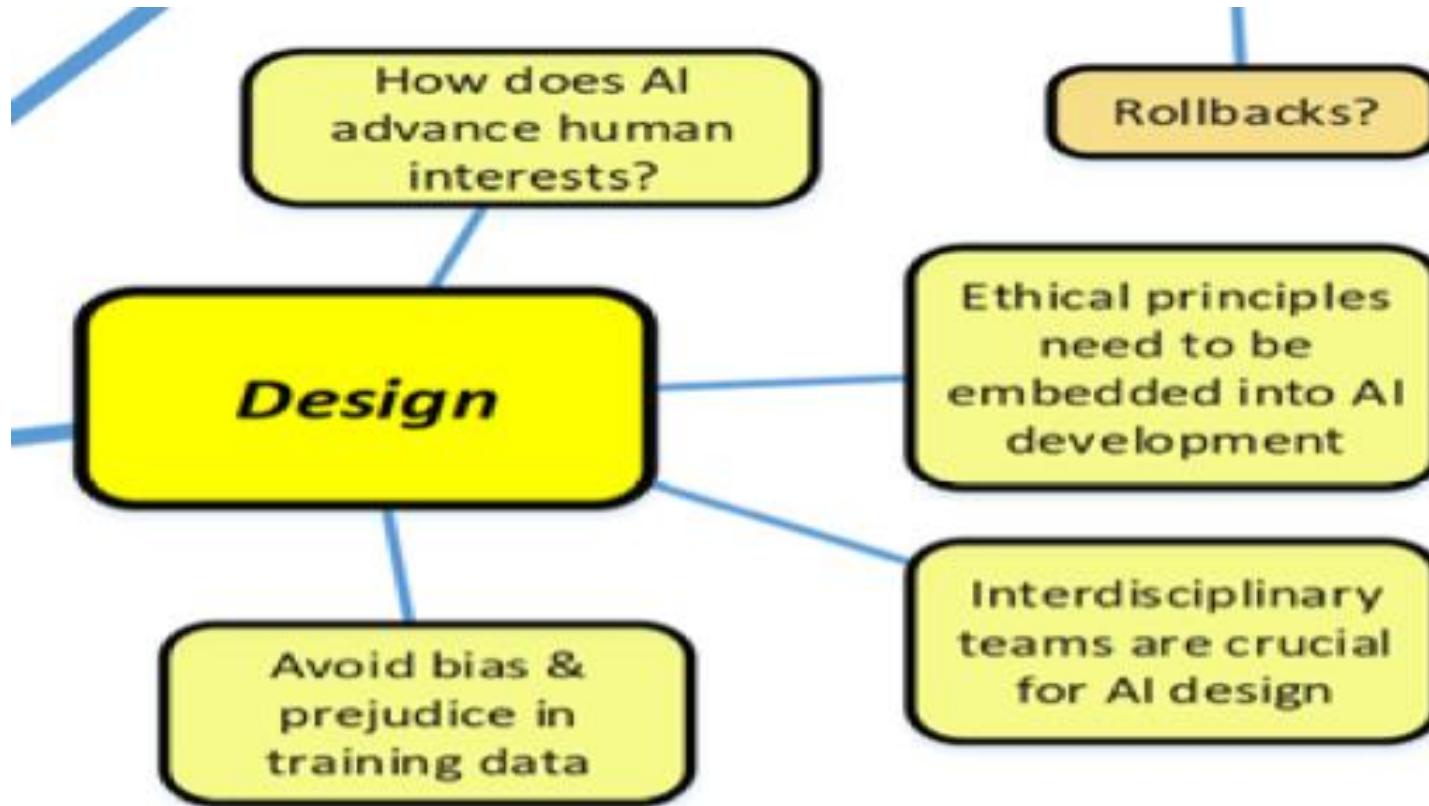
FUNCIONES



PRINCIPIOS Y TAREAS



PRINCIPIOS Y TAREAS



ALGUNAS FORMULACIONES

Asimov's three laws of Robotics (1950)

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.

Murphy and Wood's three laws of Responsible Robotics (2009)

1. A human may not deploy a robot without the human-robot work system meeting the highest legal and professional standards of safety and ethics.
2. A robot must respond to humans as appropriate for their roles.
3. A robot must be endowed with sufficient situated autonomy to protect its own existence as long as such protection provides smooth transfer of control which does not conflict with the First and Second Laws.

ALGUNAS FORMULACIONES

The European Commission's High Level Expert Group on AI Ethics Guidelines for Trustworthy AI (April 2019)

1. **Human agency and oversight** AI systems should support human autonomy and decision-making, as prescribed by the principle of respect for human autonomy.
2. **Technical robustness and safety** A crucial component of achieving Trustworthy AI is technical robustness, which is closely linked to the principle of prevention of harm.
3. **Privacy and Data governance** Closely linked to the principle of prevention of harm is privacy, a fundamental right particularly affected by AI systems.
4. **Transparency** This requirement is closely linked with the principle of explicability and encompasses transparency of elements relevant to an AI system: the data, the system and the business models.
5. **Diversity, non-discrimination and fairness** In order to achieve Trustworthy AI, we must enable inclusion and diversity throughout the entire AI system's life cycle.
6. **Societal and environmental well-being** In line with the principles of fairness and prevention of harm, the broader society, other sentient beings and the environment should be also considered as stakeholders throughout the AI system's life cycle.
7. **Accountability** The requirement of accountability complements the above requirements, and is closely linked to the principle of fairness

ALGUNAS FORMULACIONES

Google AI Principles (June 2018)

1. **Be socially** beneficial.
2. **Avoid creating or reinforcing unfair bias.**
3. **Be built and tested for safety.**
4. **Be accountable to people.**
5. **Incorporate privacy design principles.**
6. **Uphold high standards of scientific excellence.**
7. **Be made available for uses that accord with these principles.**

¿EL FUTURO DE LA IA?



¿EL FUTURO DE LA IA?

El caso de Sandra es el que ha llegado más lejos hasta el momento. Nacida en cautividad en el zoológico de Rostock en 1986, fue trasladada a Buenos Aires a los nueve años. La asociación derecho animal presentó una demanda para que le fuese reconocido su estatuto como "persona no humana". Una primera sentencia en diciembre de 2014 se lo reconoció y fue confirmada el 21 de octubre por otro juzgado.



¿EL FUTURO DE LA IA?

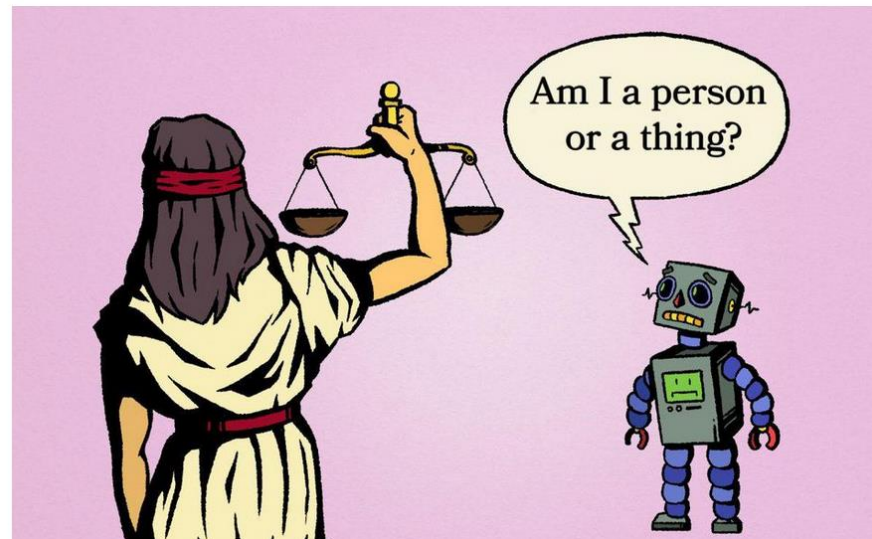
We Need To Talk About Sentient Robots



Andrea Morris Contributor 
Science
Science, Robots & The Arts

TWEET THIS

-  If robots do become sentient, how would we know? Are we disincentivized to recognize sentient machines?
-  we have no criteria for recognizing sentience in beings without biological brains and nervous systems.



Gracias!

Radoslav.Depolo@kennedyslaw.com

 [@KennedysLaw](https://twitter.com/KennedysLaw)

 [linkedin.com/company/Kennedys](https://www.linkedin.com/company/Kennedys)

 [facebook.com/KennedysTrainees](https://www.facebook.com/KennedysTrainees)

[kennedyslaw.com](https://www.kennedyslaw.com)

Kennedys