Desafíos neuro-éticos y brechas digitales

Celia Fernández-Aller

Prof. Univesidad Politécnica de Madrid

Mariacelia.fernandez@upm.es

Universidad Politécnica de Madrid

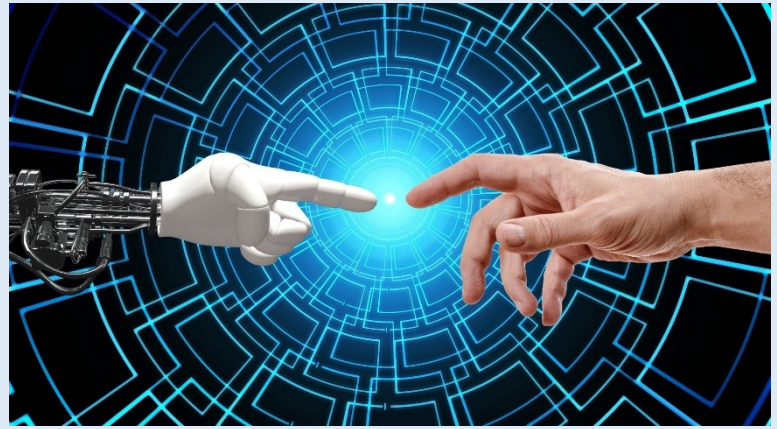
ETSI SISTEMAS INFORMÁTICOS

POLITÉCNICA

UNIVERSIDAD POLITÉCNICA DE MADRID

A pesar de los continuos esfuerzos por atajar los problemas de la humanidad por medio de la investigación científica y la innovación tecnológica, ***las respuestas siguen siendo fundamentalmente de carácter ético…***

Carl Mitcham



https://www.bbvaopenmind.com/articulos/la-tecnologia-y-el-peso-de-la-responsabilidad/

POLITÉCNICA

UNIVERSIDAD POLITÉCNICA DE MADRID

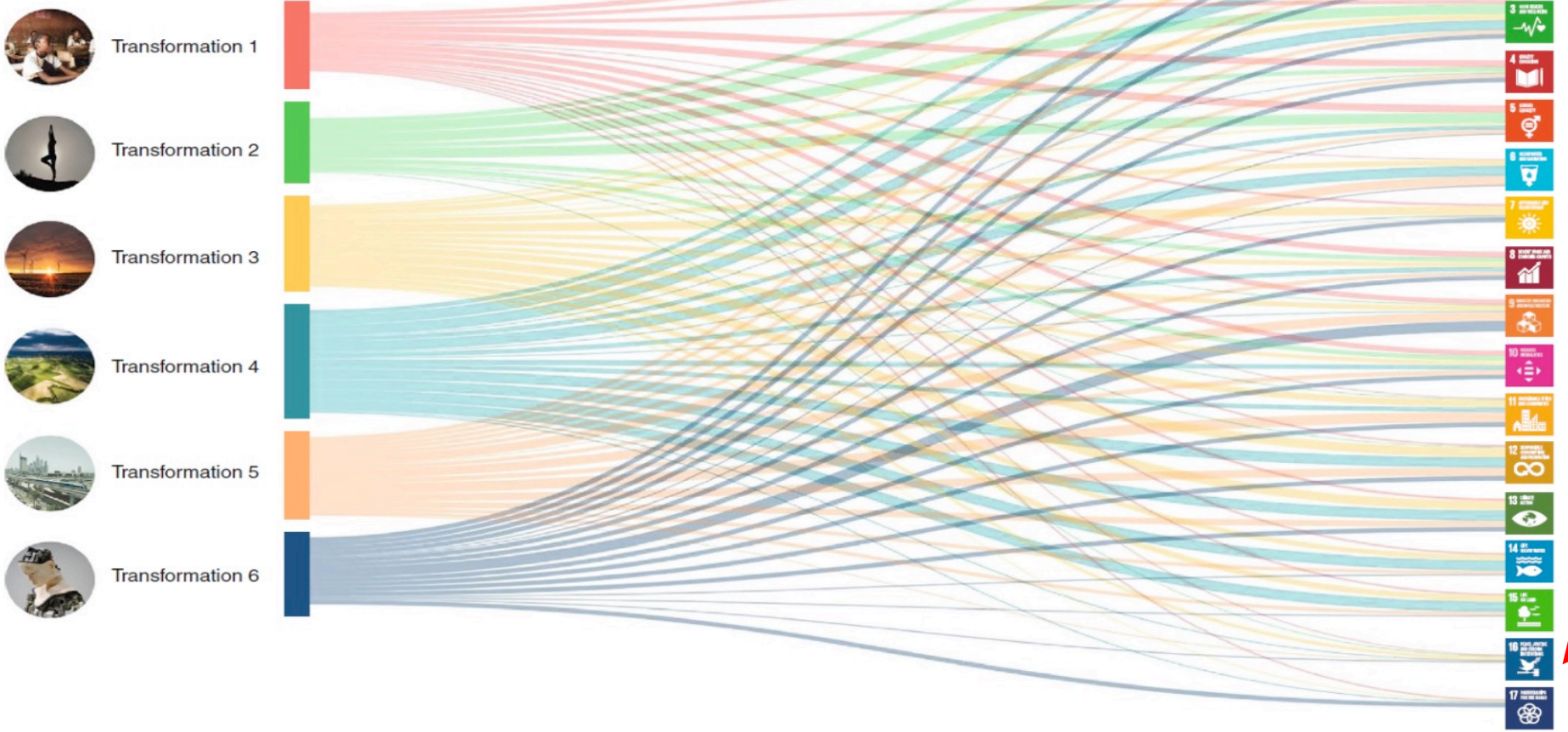
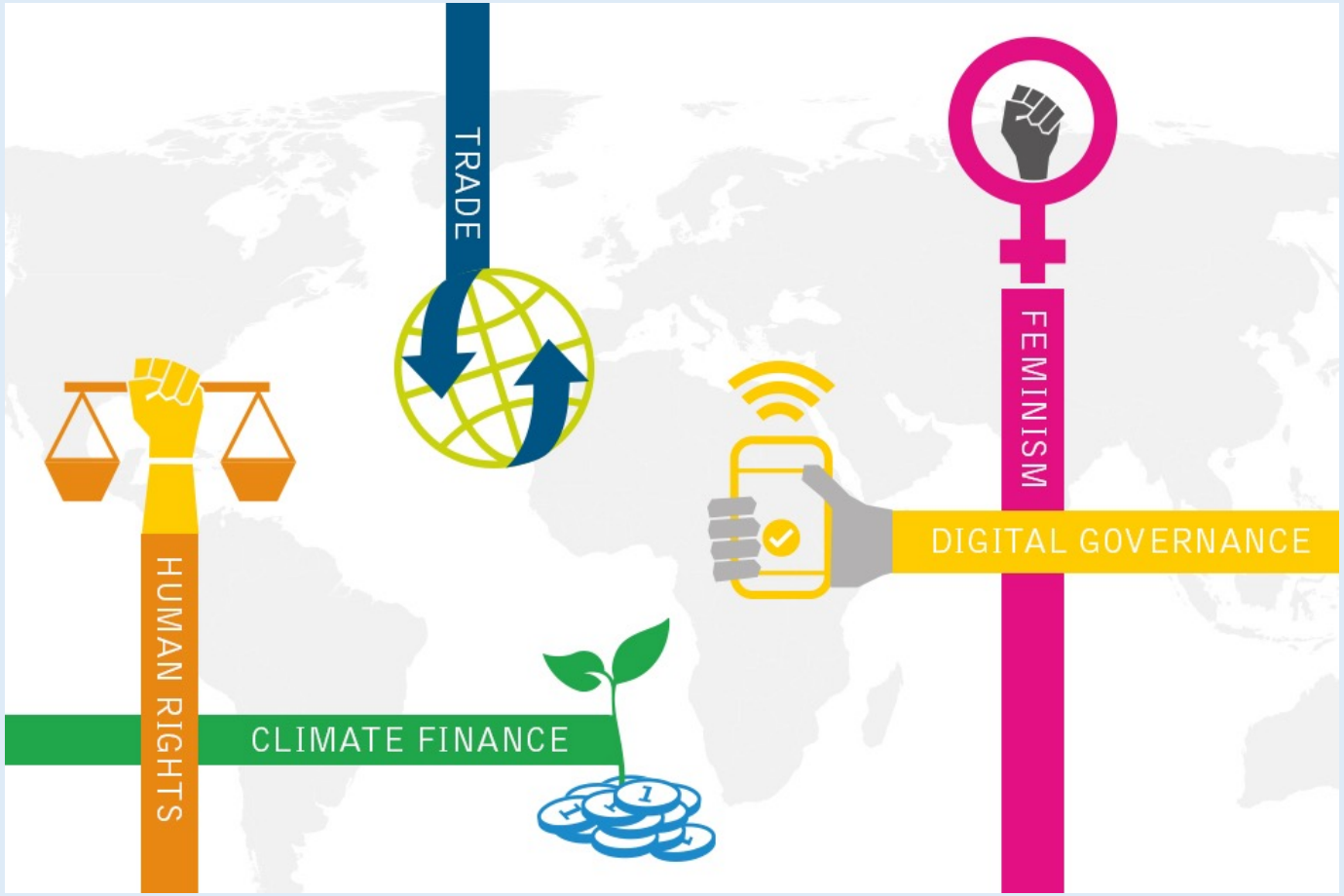


Fig. 2 | Contribution of each SDG Transformation towards the 17 SDGs. Sankey diagram illustrating the relationships between each Transformation and the SDGs (Table 1). The thicker the line, the greater the contribution of that Transformation to meeting the SDGs (see Supplementary Section 5 for methods). SDG icons courtesy of UN/SDG.

Sachs, J.D., Schmidt-Traub, G., Mazzucato, M., Messner, D., Nakicenovic, N., Rockström, J. 2019. **Six Transformations to achieve the Sustainable Development Goals**, Nature Sustainability volume 2, pages 805–814 (2019)

Tres transiciones:
socio-económica, verde, digital

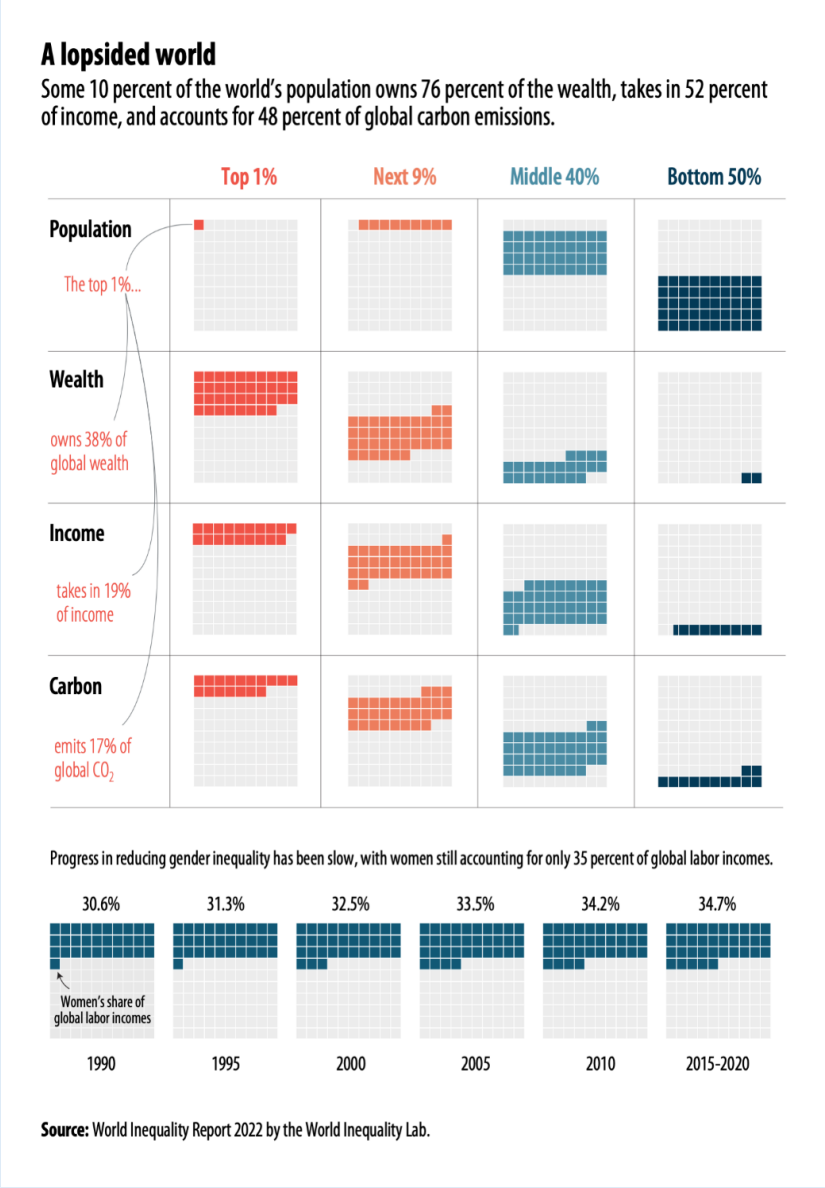


https://eu.boell.org/en/future-of-multilateralism

Shaping the future of multilateralism
Inclusive pathways to a just and crisis-resilient global order

TRANSICIÓN SOCIO-ECONÓMICA

¡DESIGUALDAD!



TRANSICIÓN VERDE

Fracaso de la COP27

Teorías del COLAPSO, decrecimiento, etc. (Illich, Latouch, etc).

Andoni Alonso Puelles. CONTRA EL EVANGELIO DEL DESARROLLO O POR QUÉ VOLVER A LEER A IVAN ILLICH. ARBOR Ciencia, Pensamiento y Cultura 198 (805) julio-eptiembre, 2022, a653 ISSN: 0210-1963, eISSN: 1988-303X

https://doi.org/10.3989/arbor.2022.805001

TRANSICIÓN DIGITAL

James, Debora.
DIGITAL TRADE RULES (2020). A DISASTROUS NEW CONSTITUTION FOR THE GLOBAL ECONOMY, BY AND FOR BIG TECH

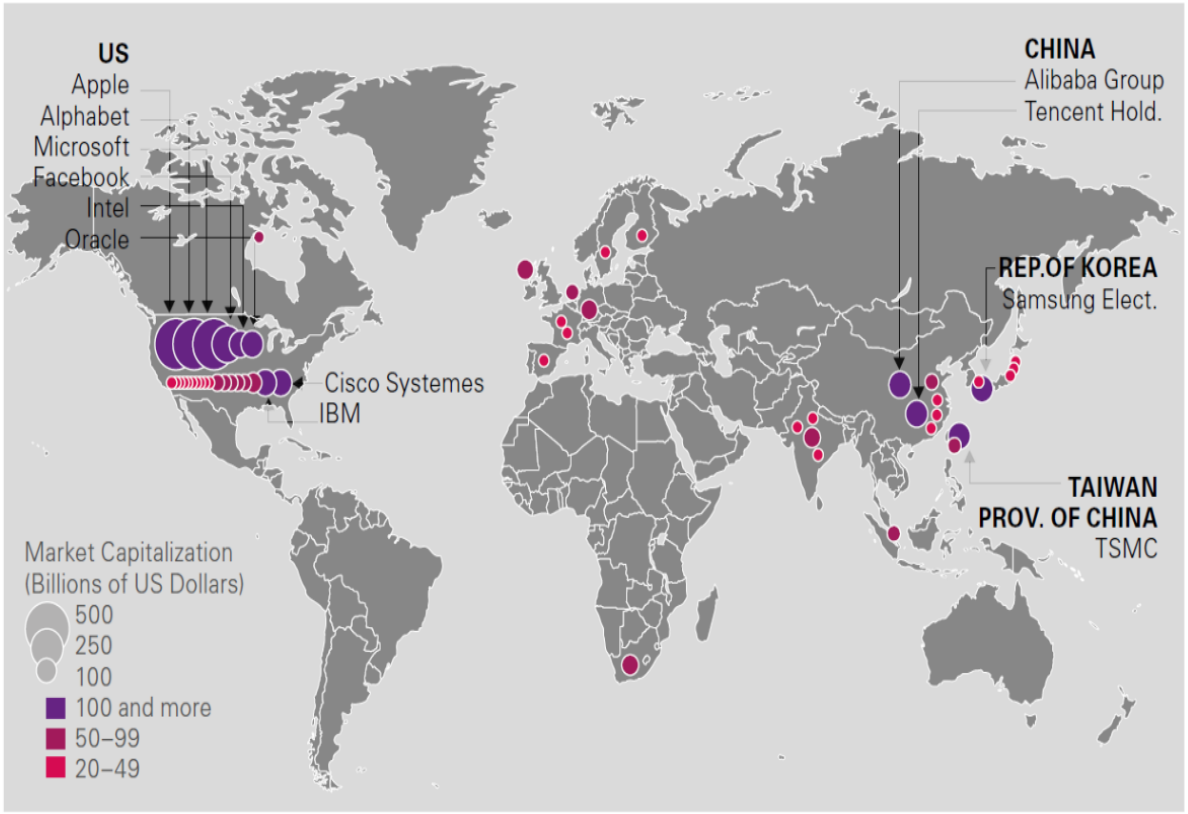


Figure 4 / Source: UNCTAD database of consolidation financial statements, based on Reuters Worldscope

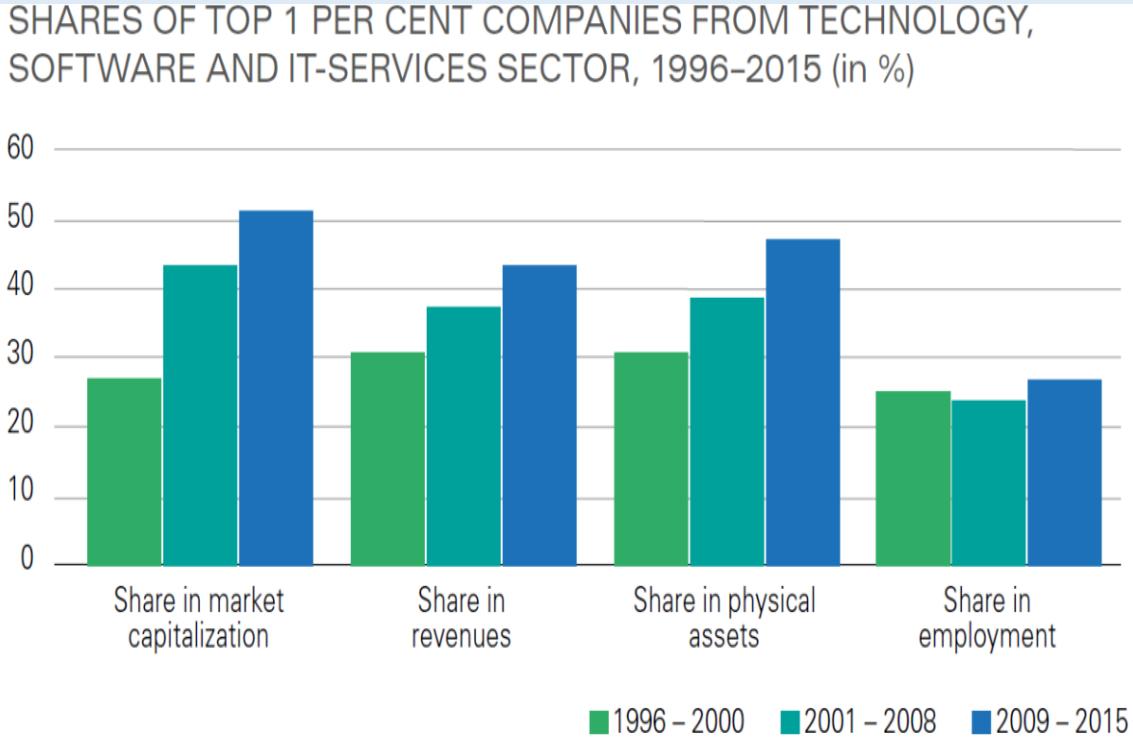


Figure2 / Source: Trade and Development Report, 2018. Note: Top 1 percent companies identified by intangible assets in the sector.

Cada vez más voces subrayan que **el poder está en manos de quien controla los algoritmos**, pues el mundo actual está cambiando radicalmente gracias a ellos, al *big data*, la robótica y la inteligencia artificial (Harari, 2017:429).

NEUROTECNOLOGÍA

Técnicas de neurotecnología:

neurociencia –el estudio del cerebro–

i**ngeniería** –la aplicación de la ciencia y tecnología para resolver problemas–

inteligencia artificial (IA) –la ciencia que estudia y crea sistemas artificiales inteligentes–.

Estas tecnologías reciben el nombre NBIC (nano-bio-info-cogno): nanotecnologías, biotecnologías, tecnologías de la información y ciencias cognitivas (IA, ciencia de datos, robótica, interfaces cerebro-máquina, biología sintética, nanotecnología).

Instituto de Filosofía del CSIC

https://theconversation.com/la-urgencia-de-los-neuroderechos-humanos-176071

Concepto de Neurotecnología

Conjunto de tecnologías que permiten **visualizar, manipular, registrar, medir y obtener información del cerebro y del sistema nervioso con el objetivo de controlar, reparar o mejorar sus funciones.**

La neurotecnología utiliza **distintas técnicas** para registrar o estimular el cerebro. Estas técnicas se pueden dividir principalmente entre aquellas que requieren un contacto directo con el cerebro y el sistema nervioso (invasivas) y aquellas que no necesitan un contacto directo (no-invasivas).

Impactos positivos y negativos

Una tecnología puede ser utilizada para cualquier otro propósito que no había sido previsto originalmente en su diseño.

Puede usarse en el ámbito de la salud (implantes, retraso en el desarrollo de alzheimer), pero también con fines delictivos.

- Salud (apoyo al diagnóstico)
- Acceso a la información (prevención de catástrofes, buscadores) **Global Pulse, UN**
- Proyecto Human Brain (R. Yuste)

https://www.humanbrainproject.eu/en/science-development/scientific-achievements/showcases/

POSTURAS ÉTICAS

TECNO-OPTIMISTA (la tecnología mejorará nuestras vidas, nuestra inteligencia y nos hará más felices)

- Justicia (ciudadanos de primera categoría, brechas, etc)
- Alteración de la identidad de las personas
- Quién decide el rumbo de la innovación científica

NEURO-DETERMINISTA (nuestra estructura ética tiene un componente biológico, físico, químico)

BIO CONSERVADORA (oposición a cualquier modificación de la naturaleza humana)

Fte. Rafael de Asís. "Ethics and Robotics. A first approach" *The age of Human Rights Journal*

Humanismo basado en los derechos humanos

Adela Cortina, Rafael de Asís:

La decisión moral tiene bases psicológicas, sociales y también "cerebrales" (bio físico químico).

- Apertura al conocimiento científico
- Que esté orientado al bienestar de la humanidad respetando los derechos humanos
- Teniendo en cuenta a las futuras generaciones
- Y a las actuales…a todas las personas (diversidad)

REGULACIÓN

No existen muchos ejemplos de países que hayan regulado esta cuestión, salvo el intento de constitucionalización en Chile.

Carta de derechos digitales en España:

– El control de cada persona sobre su identidad

– La autodeterminación individual, soberanía y libertad en la toma de decisiones

– La confidencialidad y seguridad de los datos obtenidos o relativos a sus procesos cerebrales

– El uso de interfaces persona-máquina susceptibles de afectar a la integridad física o psíquica de la persona

– Asegurar que las decisiones basadas en neurotecnologías no sean condicionadas por el suministro de datos.

Una experiencia muy relevante es el caso del proyecto de ley que se está discutiendo en Brasil, en el que se define el **dato neuronal** como "cualquier información obtenida directa o indirectamente de la actividad del sistema nervioso central y cuyo acceso se realiza por medio de interfaces cerebro-ordenador, o cualquier otra tecnología, invasiva o no". PROJETO DE LEI No , DE 2022, que modifica a Lei n° 13.709, de 14 de agosto de 2018 (Lei Geral de Proteção de Dados Pessoais

INTELIGENCIA ARTIFICIAL

Algoritmos

Las líneas de código programado que se necesitaron para poner al ser humano en la Luna en 1969 fueron 145.000 y en el año 2015 las líneas de código programado para gobernar Google fueron de 2 billones. Los sistemas algorítmicos son laberintos incomprensibles, en muchos casos hasta para los ingenieros, matemáticos y físicos que las escribieron (Steiner 2012)

Astobiza, Aníbal (2017) “Ética algorítmica: Implicaciones éticas de una sociedad cada vez más gobernada por algoritmos”. *DILEMATA*, nº24.

Concepto de IA

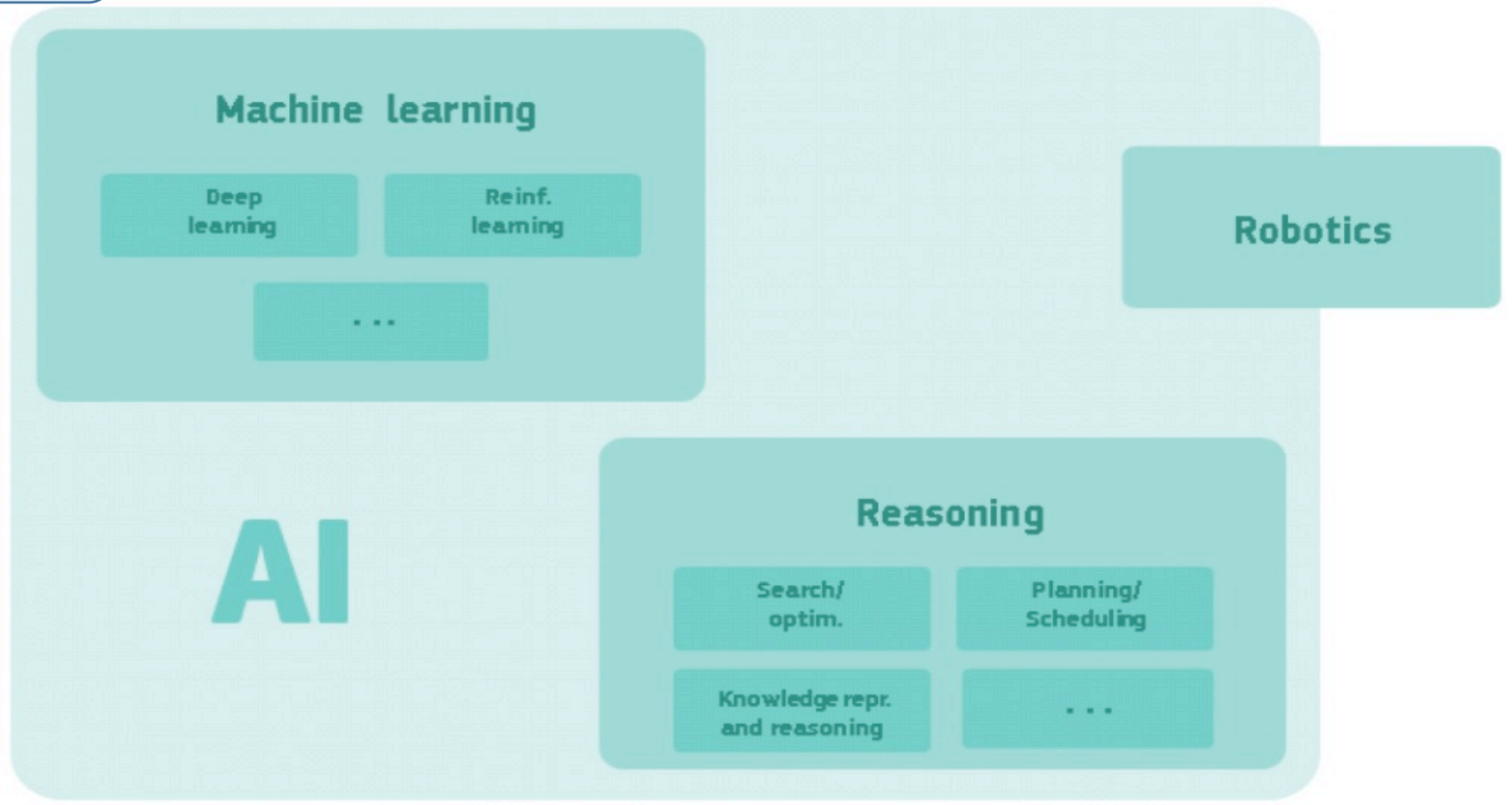


Figure 2: A simplified overview of AI's sub-disciplines and their relationship.
Both machine learning and reasoning include many other techniques, and robotics includes techniques that are outside AI. The whole of AI falls within the computer science discipline.

IBM WATSON

Global Pulse

SPOTLAB

good impact on education

algorithmic injustice

lack of explainability

lack of transparency

bias

Deepl.com

digital divide in ai

unfairness

lack of privacy

positive impact on health

Tucuvi

mediktor

"It is not enough to teach a man a specialty. Through it, he may become a kind of useful machine but not a harmoniously developed personality. It is essential that the student acquire an ***understanding of and a lively feeling for values****. He must acquire a* ***vivid sense of the beautiful and of the morally good****....He must learn to understand the motives of human beings, their illusions and their sufferings, in order to acquire a proper relationship to individual fellow men and to the community".*

Einstein, A. 1952. The New York Times.

“Tenemos que adaptar nuestros principios y valores éticos a las demandas de las tecnologías.

Debemos prestar atención a los sesgos que implícitamente incluimos en los desarrollos tecnológicos”

Colin Allen, Wendell Wallach, Iva Smit, *Why Machine Ethics?*

Vol. 21, No. 4 July/August 2006

La IA se entrena con datos sesgados (datos que incluyen desigualdades), y reproduce dichas desigualdades

https://humanrights.gov.au/about/news/media-releases/infographic-historical-bias-ai-systems

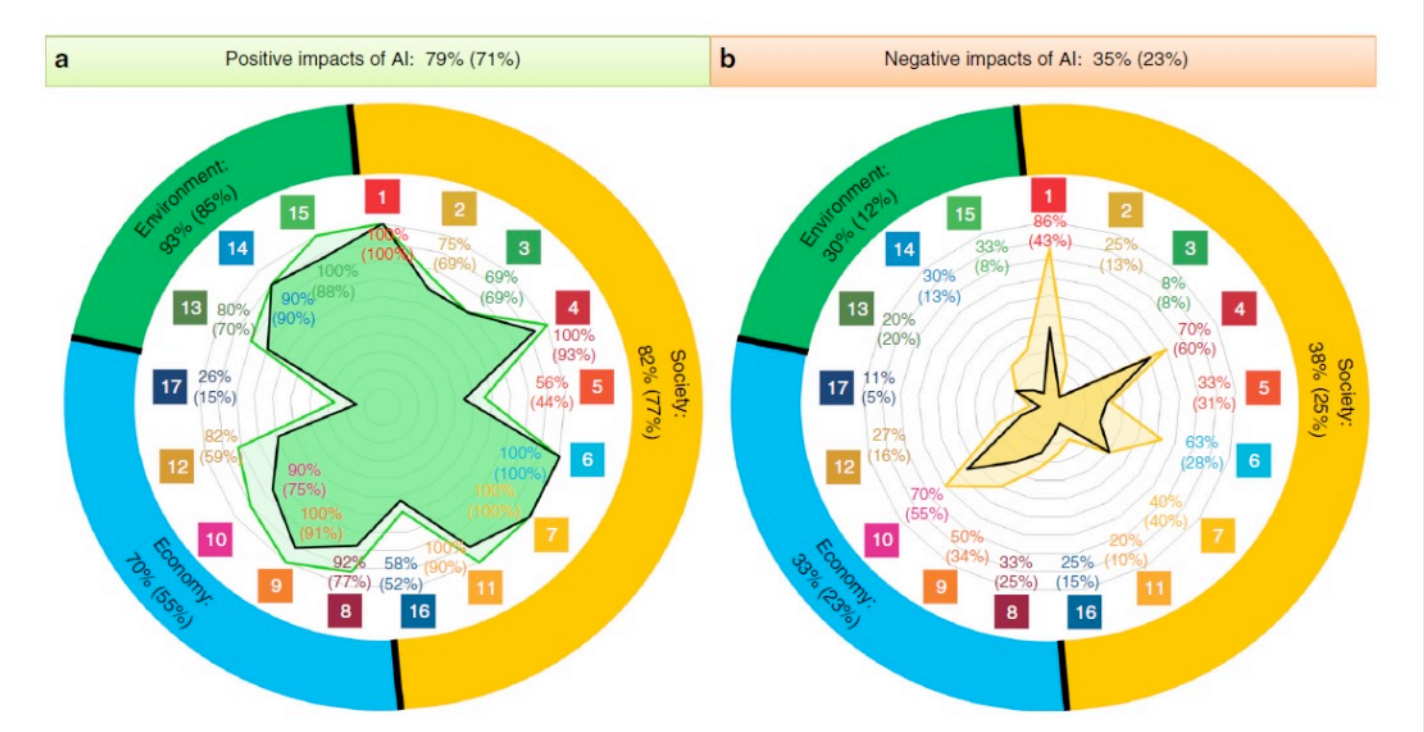
Ethics of AI
https://link.springer.com/book/10.1007/978-3-031-17040-9

Fig. 4.1 Unveiled by whistleblowers

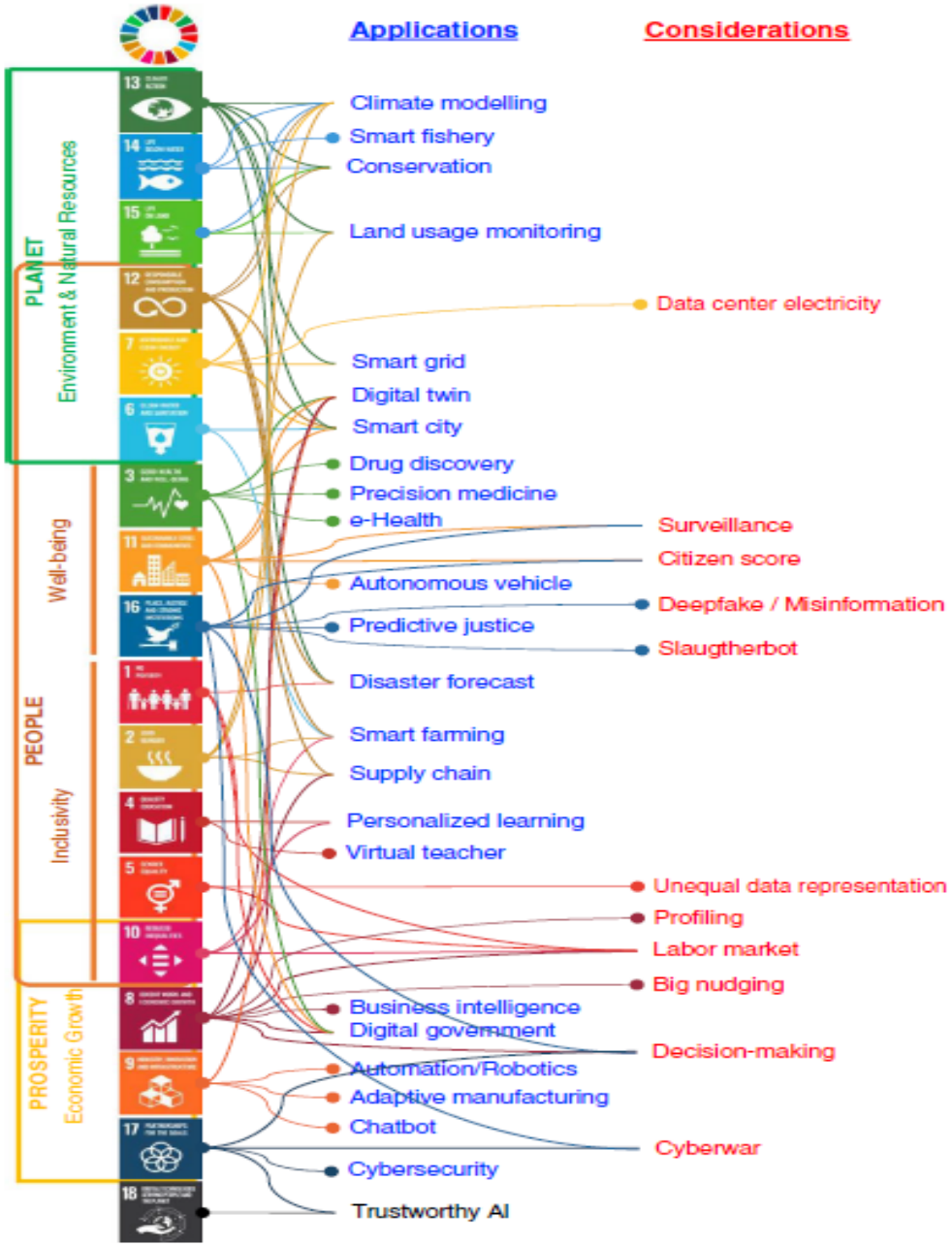
Mitigating Bias in Artificial Intelligence: https://haas.berkeley.edu/equity/industry/playbooks/mitigating-bias-in-ai/.

Timnit Gebru, Cambridge Analytica, IMPACT algorithm
https://www.ted.com/talks/carole_cadwalladr_facebook_s_role_in_brexit_and_the_threat_to_democracy

TECHNO SOLUTIONISM?



Vinuesa, R., Azizpour, H., Leite, I. et al. The role of artificial intelligence in achieving the Sustainable Development Goals. *Nat Commun* **11**, 233 (2020). https://doi.org/10.1038/s41467-019-14108-y



Goh HH, Vinuesa R. **Regulating artificial-intelligence applications to achieve the sustainable development goals.** Discov Sustain. 2021;2(1):52. doi: 10.1007/s43621-021-00064-5. Epub 2021 Nov 29. PMID: 35425914; PMCID: PMC8628838.

Los propios autores reconocen:

El propio interés puede sesgar a la comunidad investigadora y a la industria, puesto que hay presión hacia la publicación de resultados positivos. Encontrarán más fácilmente fondos los proyectos tecnológicos con resultados económicos prometedores.

Descubrir los aspectos negativos de la IA require estudios de largo plazo y faltan metodologías estructuradas para hacerlo.

PRINCIPLED ARTIFICIAL INTELLIGENCE

A Map of Ethical and Rights-Based Approaches to Principles for AI

Authors: Jessica Fjeld, Nele Achten, Hannah Hilligoss, Adam Nagy, Madhulika Srikumar

Designers: Arushi Singh (arushisingh.net) and Melissa Axelrod (melissaaxelrod.com)

HOW TO READ:

Date, Location

Document Title

Actor

COVERAGE OF THEMES:

The size of each dot represents the percentage of principles in that theme contained in the document. Since the number of principles per theme varies, it's informative to compare dot sizes within a theme but not between themes.

The principles within each theme are:

Privacy:
Privacy
Control over Use of Data
Consent
Privacy by Design
Recommendation for Data Protection Laws
Ability to Restrict Processing
Right to Rectification
Right to Erasure

Accountability:
Accountability
Recommendation for New Regulations
Impact Assessment
Evaluation and Auditing Requirement
Verifiability and Replicability
Liability and Legal Responsibility
Ability to Appeal
Environmental Responsibility
Creation of a Monitoring Body
Remedy for Automated Decision

Safety and Security:
Security
Safety and Reliability
Predictability
Security by Design

Transparency and Explainability:
Explainability
Transparency
Open Source Data and Algorithms
Notification when Interacting with an AI
Notification when AI Makes a Decision about an Individual
Regular Reporting Requirement
Right to Information
Open Procurement (for Government)

Fairness and Non-discrimination:
Non-discrimination and the Prevention of Bias
Fairness
Inclusiveness in Design
Inclusiveness in Impact
Representative and High Quality Data
Equality

Human Control of Technology:
Human Control of Technology
Human Review of Automated Decision
Ability to Opt out of Automated Decision

Professional Responsibility:
Multistakeholder Collaboration
Responsible Design
Consideration of Long Term Effects
Accuracy
Scientific Integrity

Promotion of Human Values:
Leveraged to Benefit Society
Human Values and Human Flourishing
Access to Technology

Further information on findings and methodology is available in Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches *(Berkman Klein, 2020) available at* ***cyber.harvard.edu***.

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY
AT HARVARD UNIVERSITY

Necesitamos la perspectiva ética

No single ethical principle common to all documents; an emerging convergence around transparency, justice/fairness, non-maleficence, responsibility, & privacy

Significant underrepresentation of solidarity, (human) dignity, & sustainability

Disagreement on interpretation, domain/stakeholders, implementation etc.

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines, *Nature Machine Intelligence*

- ❑ **Asilomar Principles**

https://futureoflife.org/ai-principles/

- ❑ **Montreal Declaration for Responsible AI**
- ❑ **IEEE Ethically Aligned Design v2**

  https://ethicsinaction.ieee.org/
- ❑ *Ethical Framework for a Good AI Society,* propuesto por el **AI4People** en diciembre de 2018
- ❑ *Ethics Guidelines for Trustworthy AI del High-Level Expert Group on Artificial Intelligence* de la Comisión Europea de abril de 2019
- ❑ **UNESCO Recommendation**
- ❑ **WWW Consortium** has a great proposal on Ethics and Tecnology

Policy challenges for governing AI

- **The operationalization of social values**
  - Public interest
  - Tensions
  - Priorities
- **The scope of stakeholders**
  - Designers and developers
  - Public research vs. industry
  - (Non-)Users
- **The locus of (de-)regulation**
  - Technology & techniques
  - Deployment & use
  - Impact

Necesitamos una metodología ética consensuada

- Explainable AI
- UE (Ethics by Design)
- IEEE (Ehically Aligned and Zinspection)
- Harvard (Embedded Ethics)
- Value-sensitive Design (Cummings)
- Etc.

Ethics, initially, and for a long time, were closely related with religion and politics and more recently focused on communities of practice. The undergoing digital revolution enabled by Artificial Intelligence and Big Data are bringing ethical wicked problems in the social application of these technologies. However, a broader perspective is also necessary. We now face global challenges that cannot be reduced to individual-oriented solutions. Artificial Intelligence and digital technologies are global and make humans more connected and smart but also more homogeneous, predictable and ultimately controllable. The digital revolution is being an ***industry-driven movement****. It is necessary to establish mechanisms to ensure that the society becomes conscious about its own future.*

(David Pastor Escuredo. 1 Computer Science Department, University College London, London WC1E 6EA, UK. [2]LifeD Laboratory, 28010 Madrid, Spain.[3] UNICEF, New York City, NY 10017, USA)

Fernández-Aller, Celia; Fernández, Arturo; Manjarrés, Angeles; Pastor Escuredo, David; Pickin, Salgado Criado, Ausín. *et al*., **"An Inclusive and Sustainable Artificial Intelligence Strategy for Europe Based on Human Rights,"** in *IEEE Technology and Society Magazine*, vol. 40, no. 1, pp. 46-54, March 2021, doi: 10.1109/MTS.2021.3056283.

WORLD VIEW · 07 JULY 2020

Don't ask if artificial intelligence is good or fair, ask how it shifts power

Those who could be exploited by AI should be shaping its projects.

Pratyusha Kalluri

WORLD VIEW | 01 May 2019

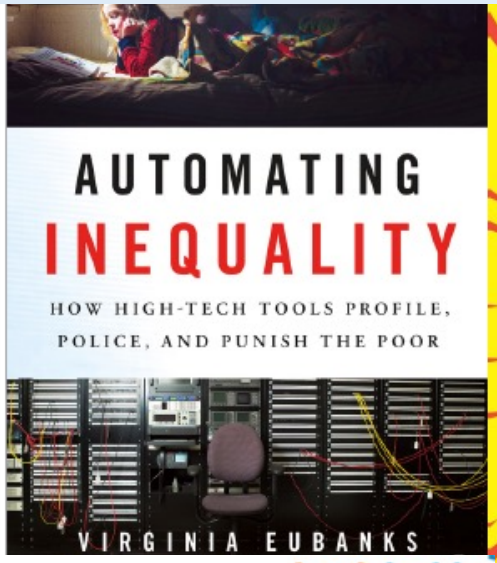
Don't let industry write the rules for AI

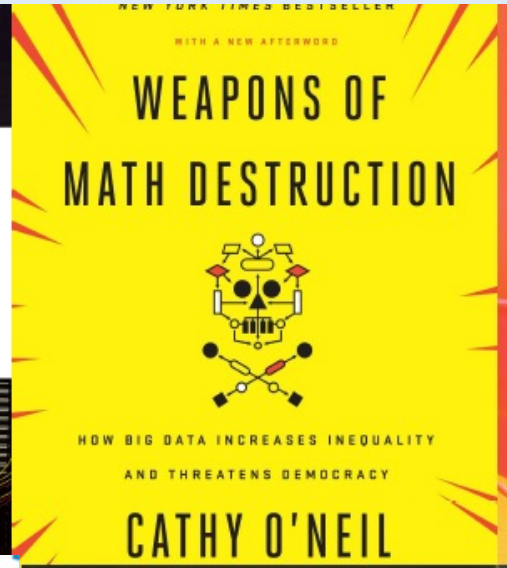
Technology companies are running a campaign to bend research and regulation for their benefit; society must fight back, says Yochai Benkler.

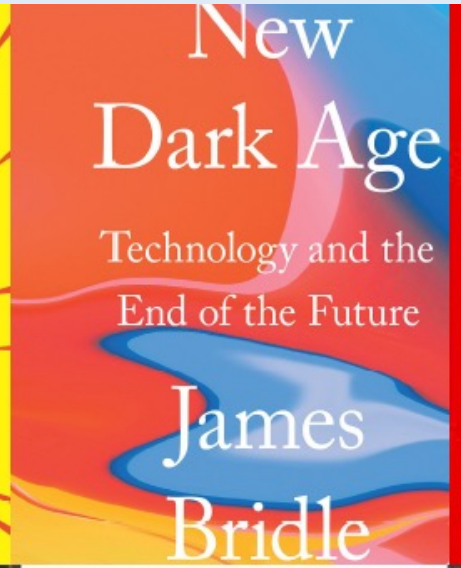
Has your own thinking on the topic evolved?

Our overview of AI ethics guidelines has made clear to me that the devil is in the details. How meaningful is, for example, a pledge to 'human-centred AI' if there are no specifications as to how and by whom this will be defined, implemented, measured and controlled in practice? I have also realized that we should not let discussions about details make us lose sight of the big picture. For instance, it is crucial to pay attention to who gets to define the ethics of AI, and to the processes that decide what counts as ethical AI.

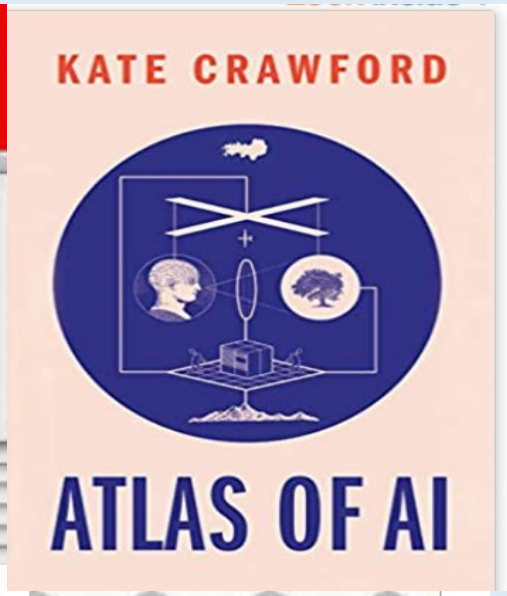
Jobin, A. et al. (2021). AI reflections in 2020, *Nature Machine Intelligence*

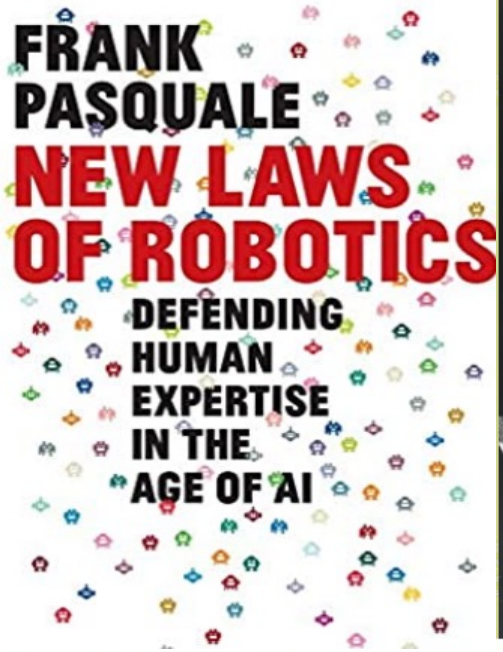

AUTOMATING
INEQUALITY
HOW HIGH-TECH TOOLS PROFILE, POLICE, AND PUNISH THE POOR
VIRGINIA EUBANKS
Look inside

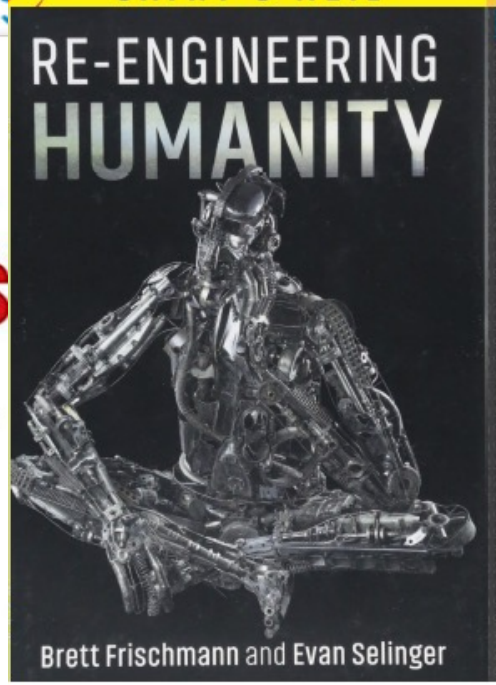

WEAPONS OF
MATH DESTRUCTION
HOW BIG DATA INCREASES INEQUALITY AND THREATENS DEMOCRACY
CATHY O'NEIL

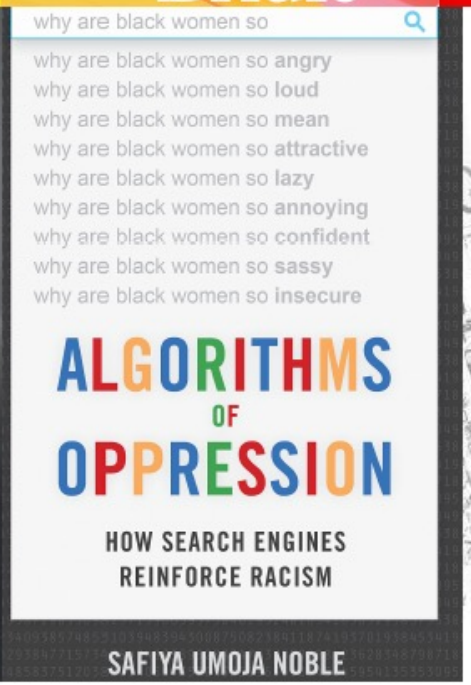

New
Dark Age
Technology and the End of the Future
James Bridle

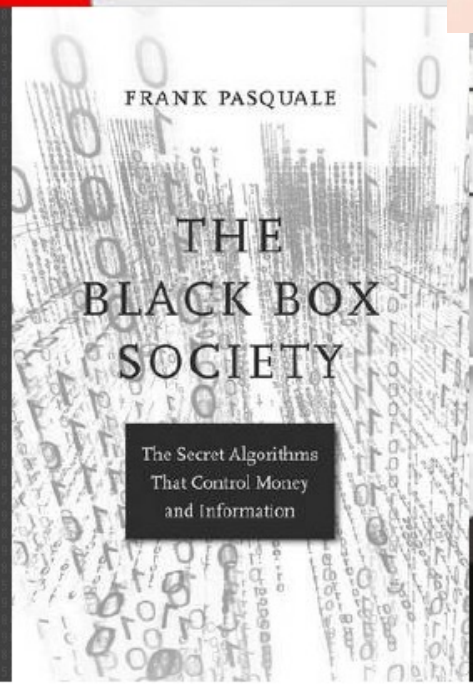

BRUCE SCHNEIER
BEST-SELLING AUTHOR OF DATA AND GOLIATH
CLICK HERE TO KILL EVERYBODY
Security and Survival in a Hyper-connected World

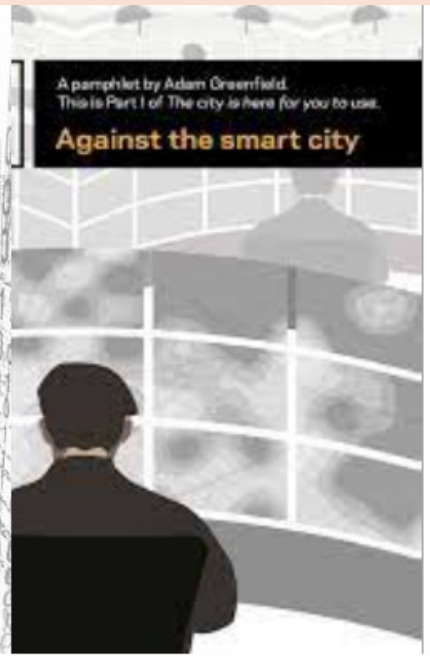

KATE CRAWFORD
ATLAS OF AI


FRANK PASQUALE
NEW LAWS OF ROBOTICS
DEFENDING HUMAN EXPERTISE IN THE AGE OF AI

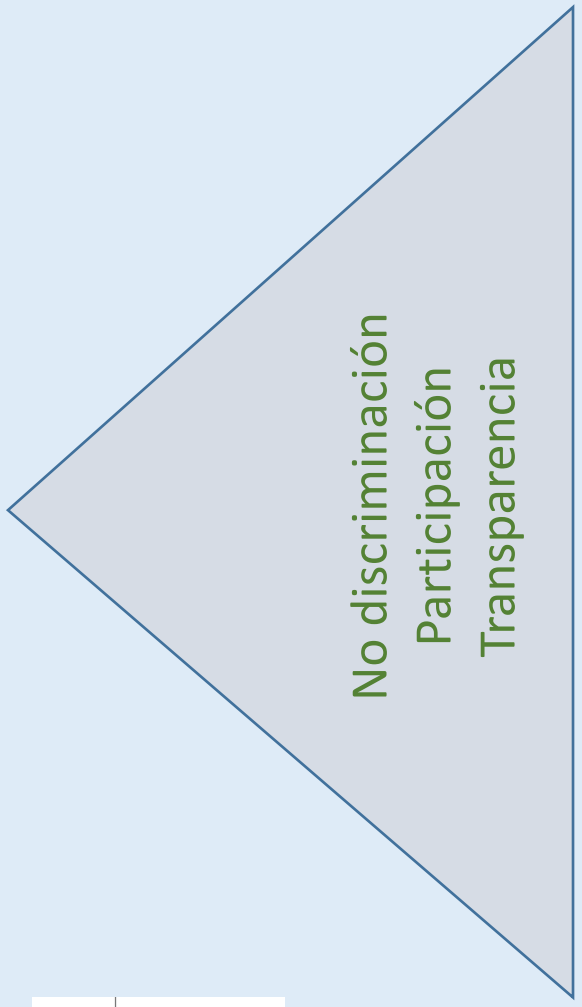

RE-ENGINEERING
HUMANITY
Brett Frischmann and Evan Selinger


why are black women so
why are black women so angry
why are black women so loud
why are black women so mean
why are black women so attractive
why are black women so lazy
why are black women so annoying
why are black women so confident
why are black women so sassy
why are black women so insecure
ALGORITHMS
OF
OPPRESSION
HOW SEARCH ENGINES REINFORCE RACISM
SAFIYA UMOJA NOBLE


FRANK PASQUALE
THE BLACK BOX SOCIETY
The Secret Algorithms That Control Money and Information


A pamphlet by Adam Greenfield.
This is Part 1 of The city is here for you to use.
Against the smart city

Etica y enfoque basado en derechos humanos



ETHICS

HUMAN RIGHTS

No obligatoriedad

Exigible

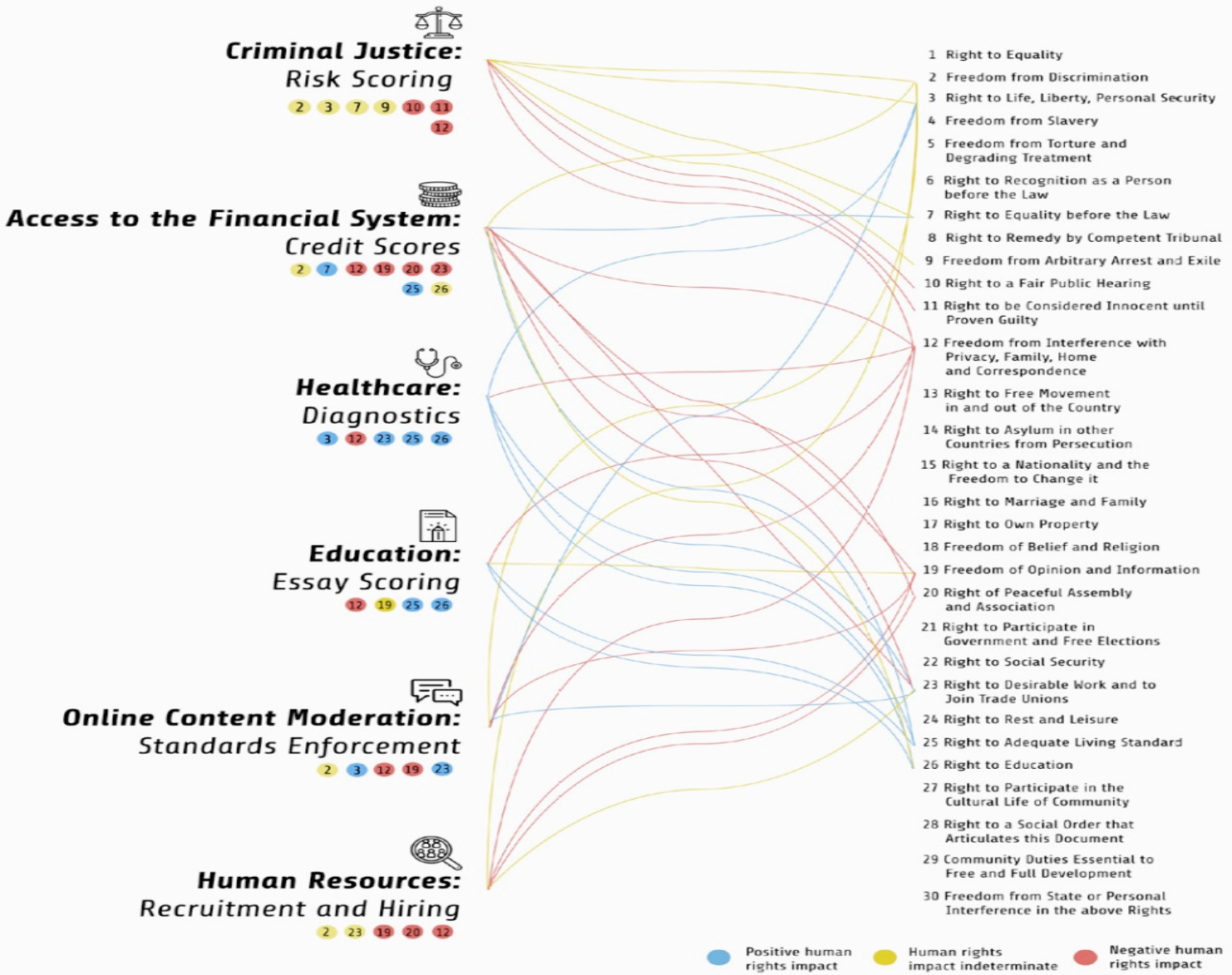
Do not have consensus

Have consensus (HHRR International Law)

Variety of principles

Universal principles

ARTIFICIAL INTELLIGENCE & HUMAN RIGHTS



Sarah case (algorithms, social care)

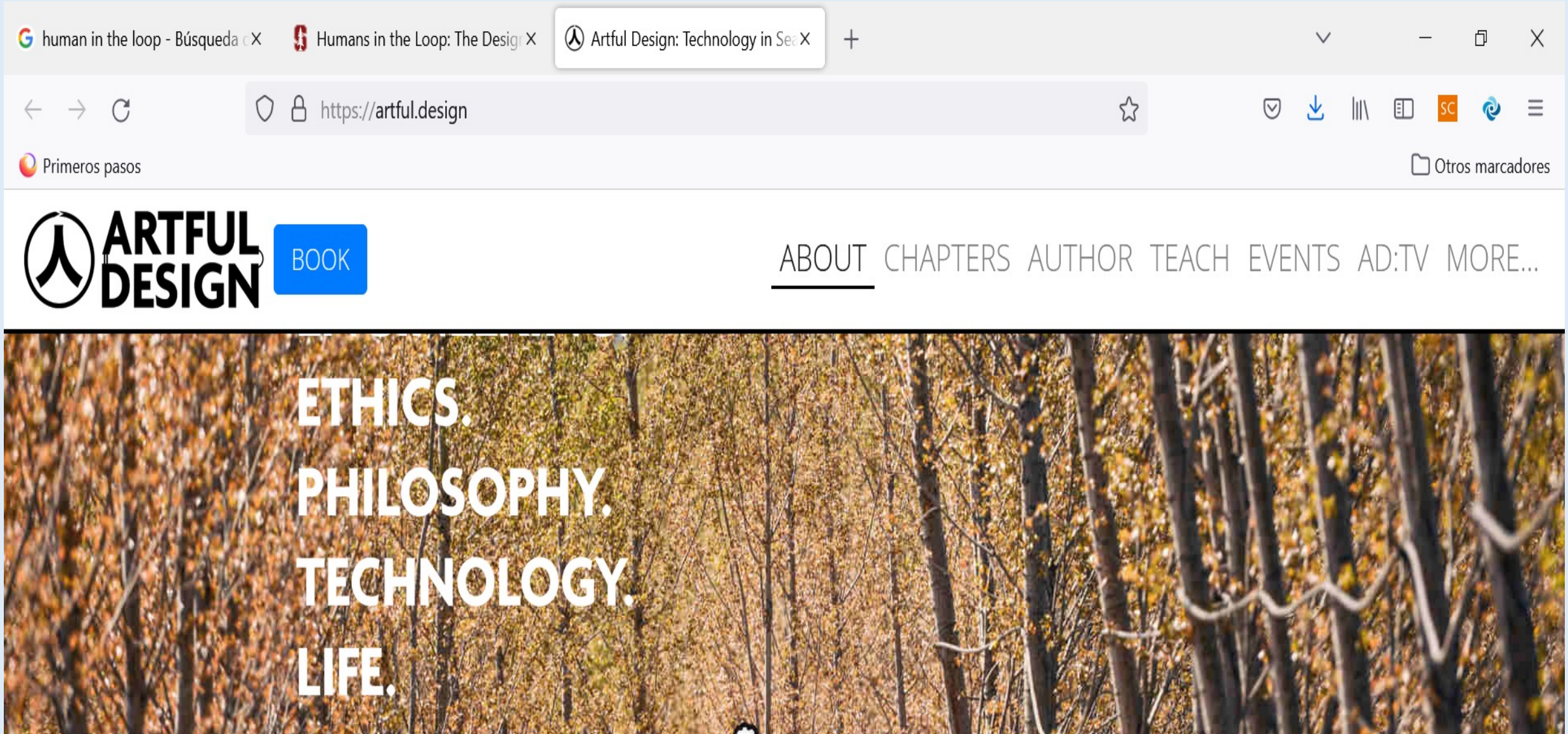
The Dutch tax authority ruined thousands of lives after using an algorithm to spot suspected benefits fraud — and critics say there is little stopping it from happening again.

https://www.politico.eu/article/dutch-scandal-serves-as-a-warning-for-europe-over-risks-of-using-algorithms/

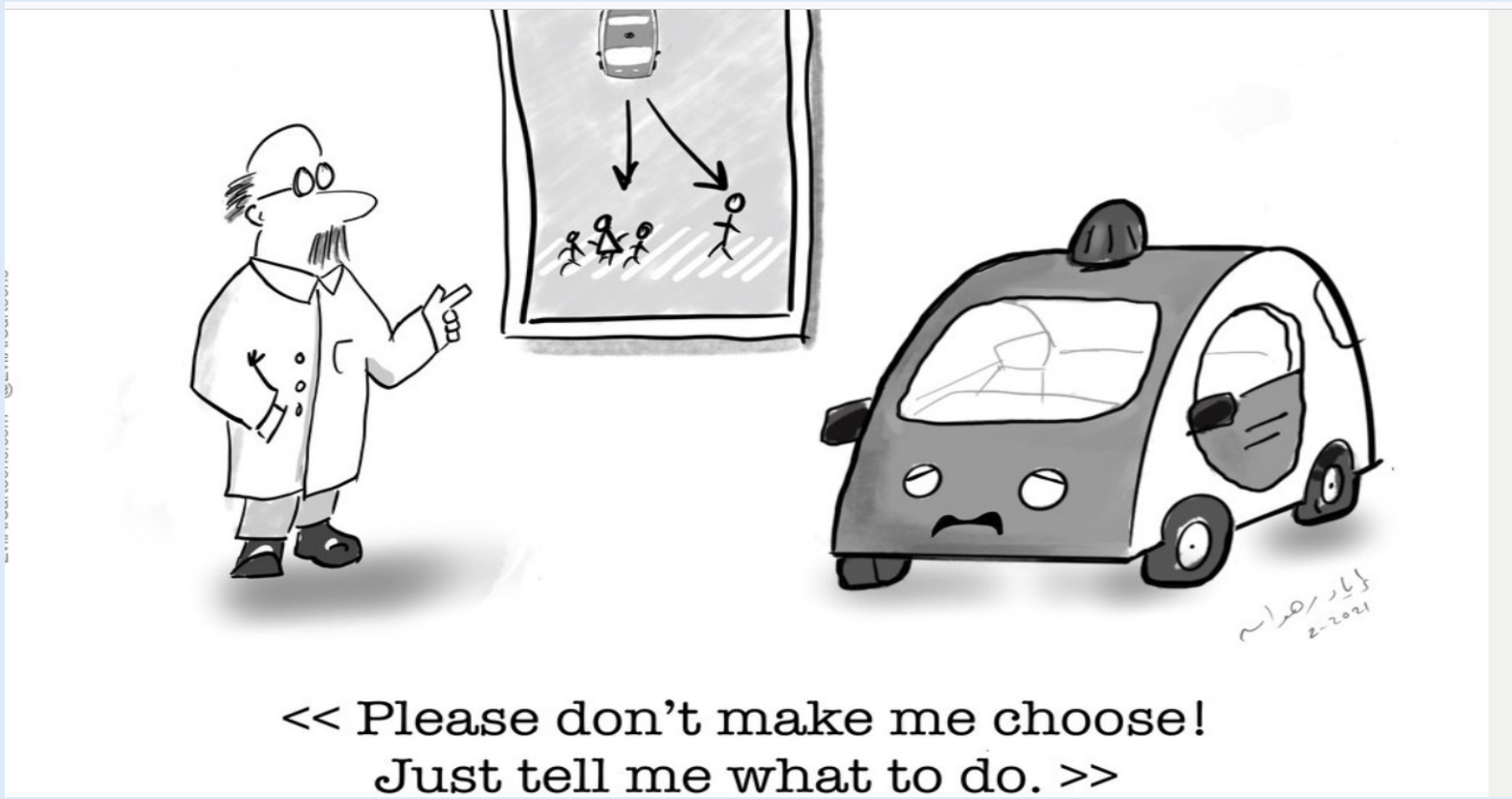
- Sarah was remembering four years ago, when her sister Chantell, who had cerebral palsy and heavily relied on support**, had her home care visits dramatically cut from 56 to 32 hours a week**. A **new algorithm** had reassessed the amount of care her sister would be given.
- When the assessor entered the information about her health status, daily routines and needs for support into the computer, it ran through an algorithm that Brent council had recently approved, determining how many hours of help she would receive.

- And then there was her younger brother Jordan who **had been arrested and charged with burglary** and petty theft for grabbing an unlocked bike and a scooter with his mate. When Jordan was booked into prison, **a computer program** spat out a score predicting the likelihood of him committing a future crime.

Buenas prácticas



Moral machine



- https://www.moralmachine.net/

CONCLUSIONES

- La Política debe escuchar a la **Ciencia y a la Filosofía**
  - Programa europeo S4P (Science for Policy)

COMMISSION STAFF WORKING DOCUMENT *Supporting and connecting policymaking in the Member States with scientific research* (2022)

- Y a la **sociedad civil organizada**

- Y por supuesto…a **otros actores** (pero no sólo y siempre a los mismos)

RETOS

- ✓¿Cómo hacer para que los derechos que hemos consensuado off line también se apliquen adaptados (o nuevos) al mundo on line (sea el entorno que sea: internet, metaversos, etc)
- ✓¿Cómo conseguir que las preocupaciones éticas y de derechos humanos no frenen la innovación tecnológica, sino que fomenten una innovación más segura y participativa?
- ✓¿Cómo conseguir que los robots sean agentes morales?
  - ✓Robot cuidador
  - ✓Coche autónomo
- ✓ ¿Cómo concienciar y educar para una participación responsable en la Cuarta Revolución industrial?