



**Barcelona
Supercomputing
Center**
Centro Nacional de Supercomputación



Facts and limits of the AI. An ELSEC Approach

Prof. Ulises Cortés

**IX JORNADAS DE ASPECTOS ÉTICOS
DE LA INVESTIGACIÓN BIOMÉDICA**

Nuevas reflexiones sobre la investigación relacionada con la salud en seres humanos

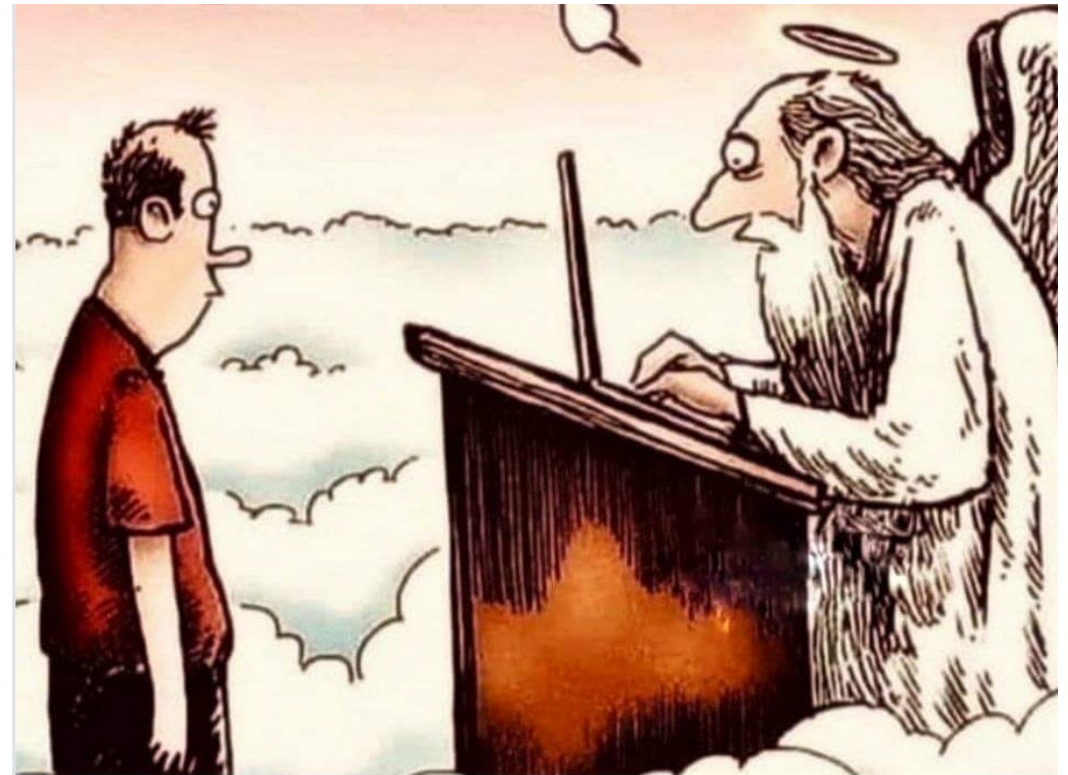
**22/23
NOVIEMBRE
2022
MADRID**



November 2022

What I am not

- I am not
 - An Ethicist
 - A Lawyer
 - A Futurist



Says here you should go to hell but since you have a PhD we'll count that as time served

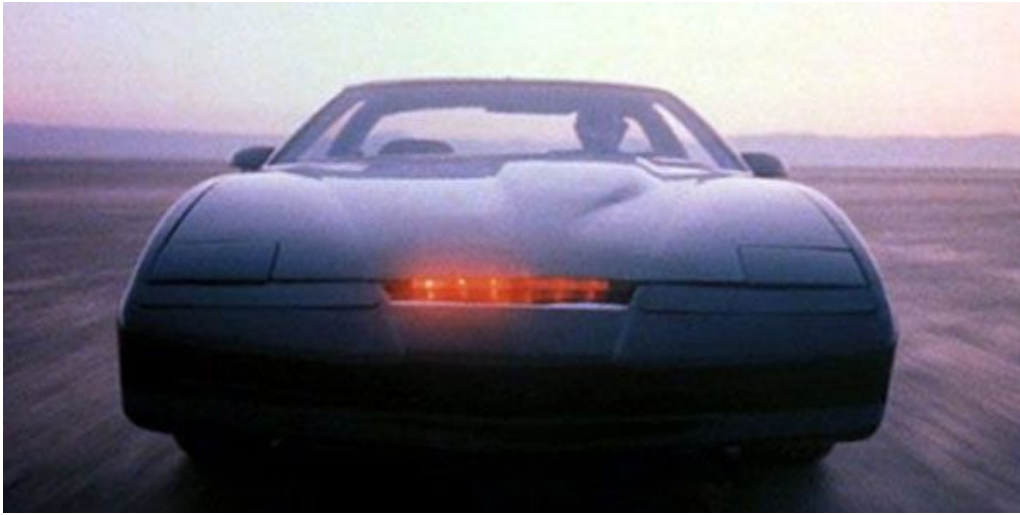
The HLEG- AI definition of AI

- Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, **act in the physical or digital dimension by perceiving their environment through data acquisition**, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and **deciding the best action(s) to take to achieve the given goal**. AI systems can either use **symbolic rules or learn a numeric model**, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions.

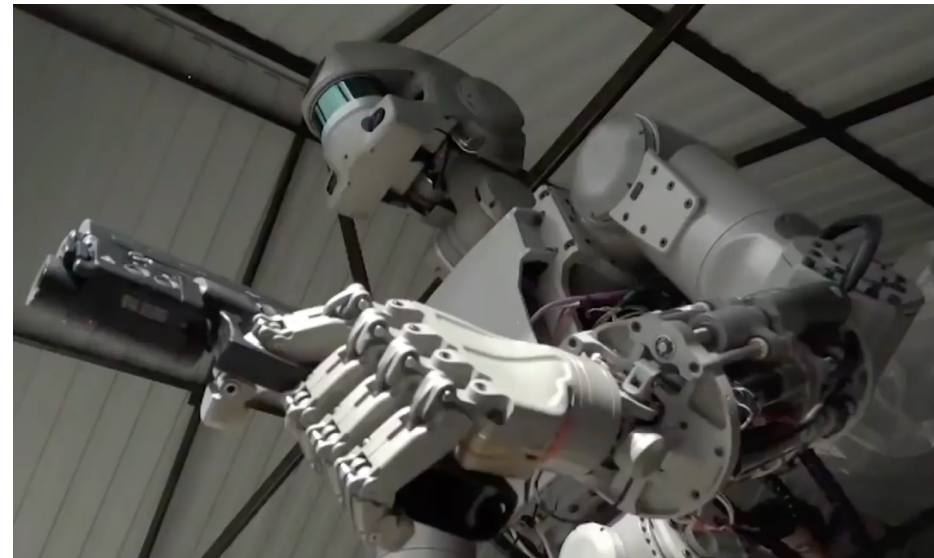
Sesame Street

NEAR	FAR
Assisted Driving modules	Autonomous Car
Semi-autonomous weapons	Autonomous weapons
Decision Support Systems	Fully Autonomous Decision Systems
Transportation route optimization algorithms for project planning optimization	Autonomous Building Systems
Clinical image recognition	Autonomous medical diagnostic systems based on clinical images
...	...

Back to the Future



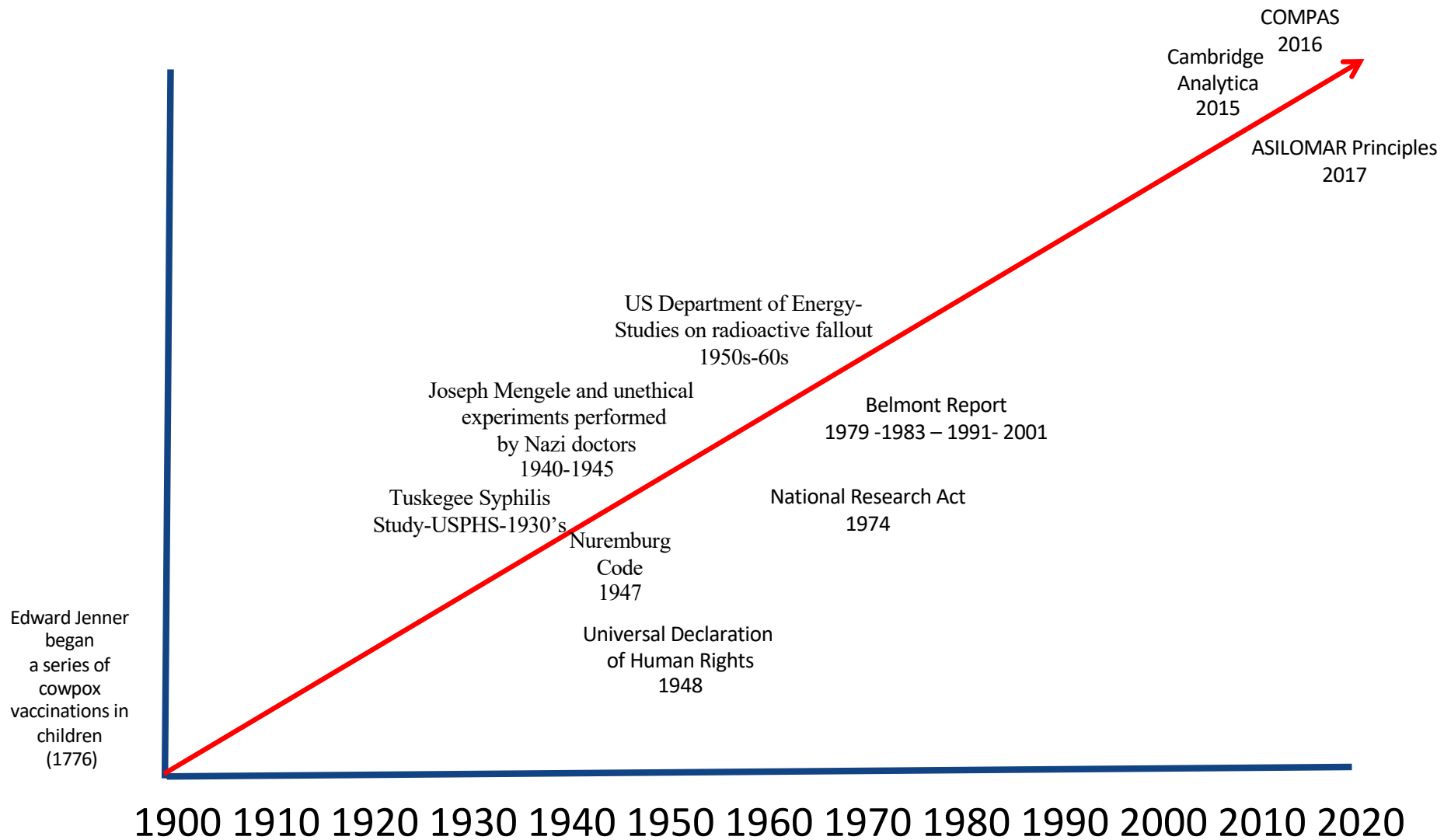
Controlled, Supervised, Automated, Autonomous



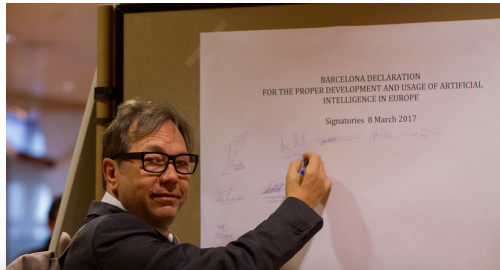
Sesame Street (2)

Good	Bad
Route optimizers	Surveillance, Inadvertently harvested personal data (e.g. habits, Cambridge Analytica)
Biometric Authentication	Remote Biometric identification
Transparent automatic decision making AI-Based systems	Opaque automatic decision making AI-Based systems
Unbiased recommender systems	Unethical recommender systems (e.g. biased) / Bubble filters /Fake News
Companion Robots Robots working in hazardous environments (mining, building, high-risk environments)	Autonomous weapons, AI cyberattacks

Some examples of unethical acts



2017 onwards



Parlement européen
2014-2019



TEXTES ADOPTÉS

P8_TA(2017)0051

Règles de droit civil sur la robotique

Résolution du Parlement européen du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique (2015/2433(INI))



ASILOMAR AI PRINCIPLES



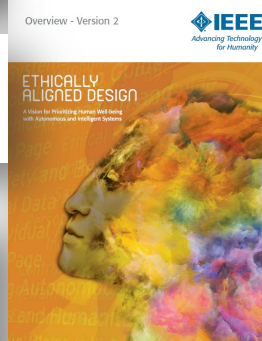
OECD Principles on Artificial Intelligence



On 22 May 2019 the OECD adopted its **Principles on Artificial Intelligence**, the first international standards agreed by governments for the responsible stewardship of trustworthy AI. The OECD Principles on AI include concrete recommendations for public policy and strategy. The general scope of the Principles ensures they can be applied to AI developments around the world. We are also planning to launch a **policy observatory** to ensure the beneficial use of AI later in the year.

Software Engineering Code of Ethics

IEEE-CS/ACM Joint Task Force on Software Engineering Ethics and Professional Practices



Group of Governmental Experts of the High Contracting Parties to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects

CCW/GGE.1/2017/3

22 December 2017

Original: English

Geneva, 13–17 November 2017
Item 7 of the agenda
Adoption of the report

La CNIL appelle à la tenue d'un débat démocratique sur les nouveaux usages des caméras vidéo



Parlement européen
2014-2019

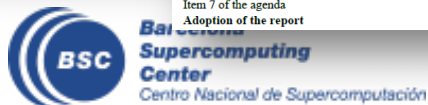


TEXTES ADOPTÉS
Édition provisoire

P8_TA-PROV(2019)0081

Recommandation de la Commission européenne sur l'intelligence artificielle et les données

The Japanese Society for Artificial Intelligence Ethical Guidelines

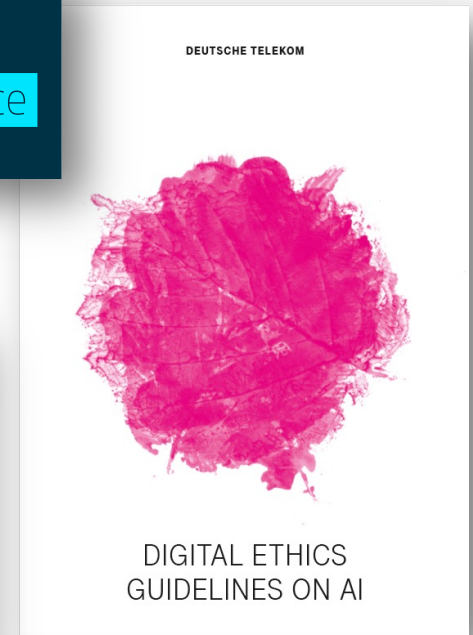


Companies also contribute to the noise



POUR UN MONDE NUMÉRIQUE RÉSOLUMENT HUMAIN ET ÉTHIQUE

Le numérique occupe une place toujours plus importante dans notre société, notre économie et nos vies. Nous oeuvrons à la MAIF pour que nos sociétaires, comme nos équipes, bénéficient des bienfaits de son développement.



Facebook Reportedly Has A Dedicated AI Ethics Team

Sam Sheehy Contributor
I cover artificial intelligence and Google DeepMind.



Microsoft AI principles

Designing AI to be trustworthy requires creating solutions that reflect ethical principles that are deeply rooted in important and timeless values.

AI at Google: our principles



Sundar Pichai
CEO

Published Jun 7, 2018

At its heart, AI is computer programming that learns and adapts. It can't solve every problem, but its potential to improve our lives is profound. At Google, we use AI to make products more useful—from email that's spam-free and easier to compose, to a digital assistant you can speak to naturally, to photos that pop the fun stuff out for you to enjoy.

Beyond our products, we're using AI to help people tackle urgent problems. A pair of high school students are building AI-powered sensors to predict the risk of wildfires. Farmers

Fairness

AI systems should treat all people fairly

Inclusiveness

AI systems should empower everyone and engage people

Reliability & Safety

AI systems should perform reliably and safely

Transparency

AI systems should be understandable

Privacy & Security

AI systems should be secure and respect privacy

Accountability

AI systems should have algorithmic accountability

... and the Church



AI at Google: our principles

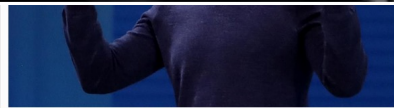


Sundar Pichai
CEO

Published Jun 7, 2018

At its heart, AI is computer programming that learns and adapts. It can't solve every problem, but its potential to improve our lives is profound. At Google, we use AI to make products more useful—from email that's spam-free and easier to compose, to a digital assistant you can speak to naturally, to photos that pop the fun stuff out for you to enjoy.

Beyond our products, we're using AI to help people tackle urgent problems. A pair of high school students are building AI-powered sensors to predict the risk of wildfires. Farmers



Designing AI to be trustworthy requires creating solutions that reflect ethical principles that are deeply rooted in important and timeless values.

Fairness

AI systems should treat all people fairly

Reliability & Safety

AI systems should perform reliably and safely

Privacy & Security

AI systems should be secure and respect privacy

Inclusiveness

AI systems should empower everyone and engage people

Transparency

AI systems should be understandable

Accountability

AI systems should have algorithmic accountability

European High-Level Expert Group on AI Ethics Guidelines for Trustworthy AI (April 2019)



Lawful AI: Legal compliance with Primary law (treaties, charter of fundamental rights), secondary law (**GDPR**, product liability directive), Council of Europe conventions, State laws, Sector-specific regulations (*e.g.*, healthcare).

Ethical AI: alignment with ethical principles and norms.

Robust AI: safety, security by design (technical robustness), appropriate application operational contexts and limitation of unintended consequences (non-technical robustness).

- <https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence>

Ethical Principles for Trustworthy AI



Ethical imperatives

Principle of Autonomy: *“Preserve Human Agency and control”*

Principle of Non maleficence: *“Do no Harm”* - Neither cause nor exacerbate harm or otherwise adversely affect human beings. safety and security, technical robustness.

Principle of Justice: *“Be Fair”*. Equal and just distribution of benefits and costs, *free from unfair bias*, increase social fairness

Principle of Explicability: *“Operate transparently”*. Traceability, auditability, transparent system capabilities, ...

Foundations of Trustworthy and Ethical AI: A Human-Centric Approach



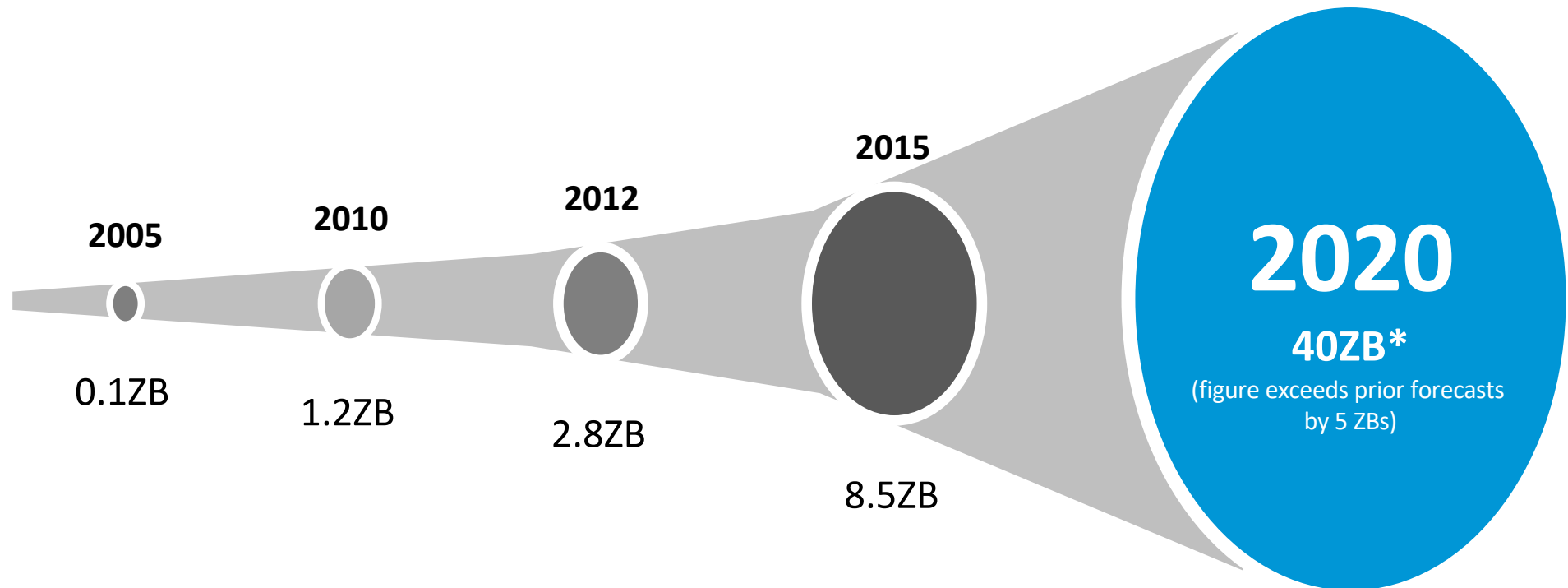
- **Respect for human dignity.** Humans are moral subjects, not objects to be scored, herded or manipulated.
- **Freedom of the individual.** Fundamental rights, control over one's own life and choices, protection from sovereign intrusion
- **Respect for democracy and justice.** Protection of democratic processes and human deliberation
- **Equality, non-discrimination and solidarity.** No bias, no exclusion
- **Citizens' rights.** Access to administration and services (including non-citizens).

Be data (rich), my friend

- Europe's current and future sustainable economic growth and societal wellbeing increasingly draws on value created by data. **AI is one of the most important applications of the data economy.** Today most data are related to consumers and are stored and processed on central **cloud-based infrastructure.** By contrast a large share of tomorrow's far more abundant data will come from industry, business and the public sector, and will be stored on a variety of systems, notably on computing devices working at the **edge** of the network.

On Artificial Intelligence - A European approach to excellence and trust

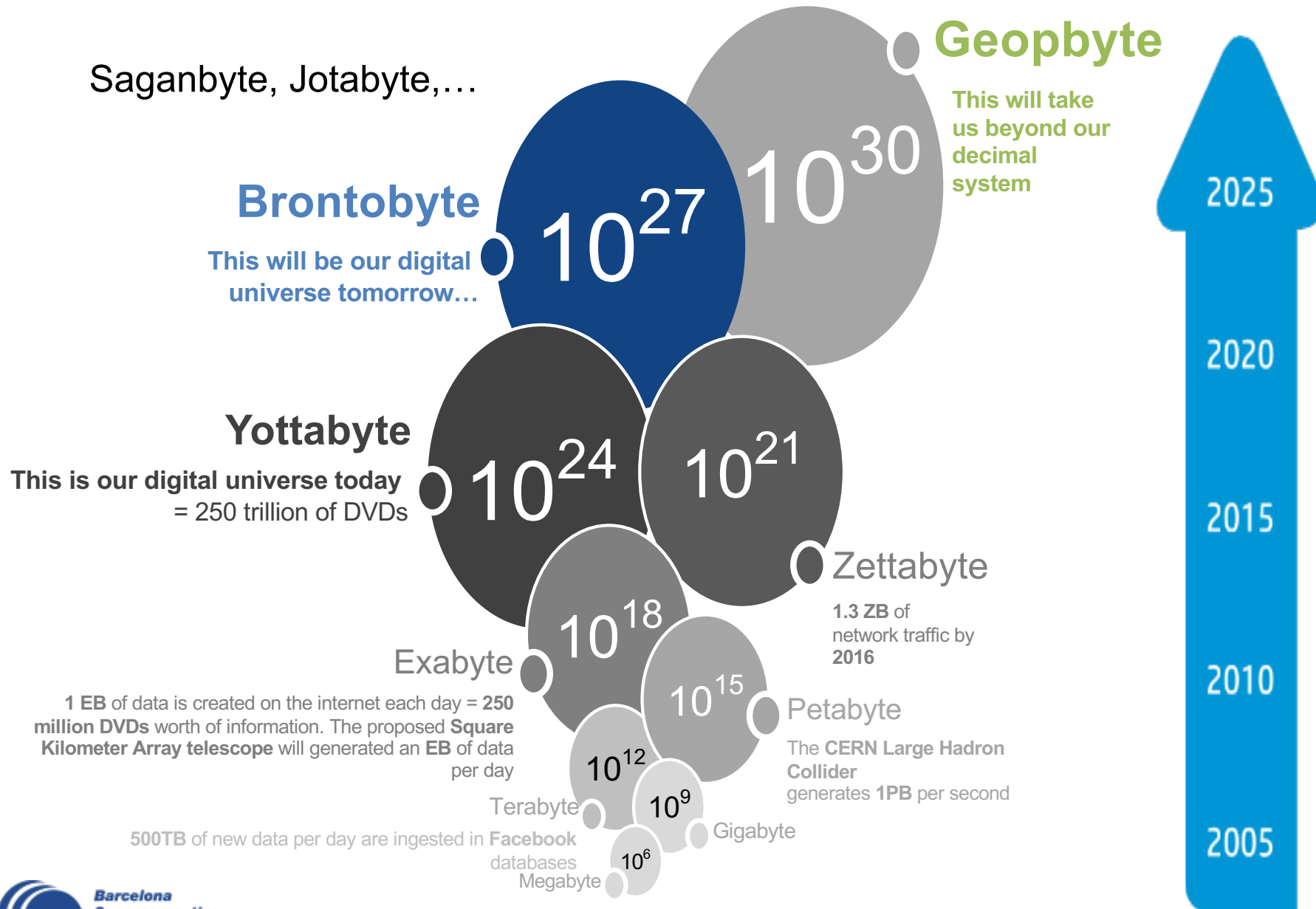
The Data Deluge



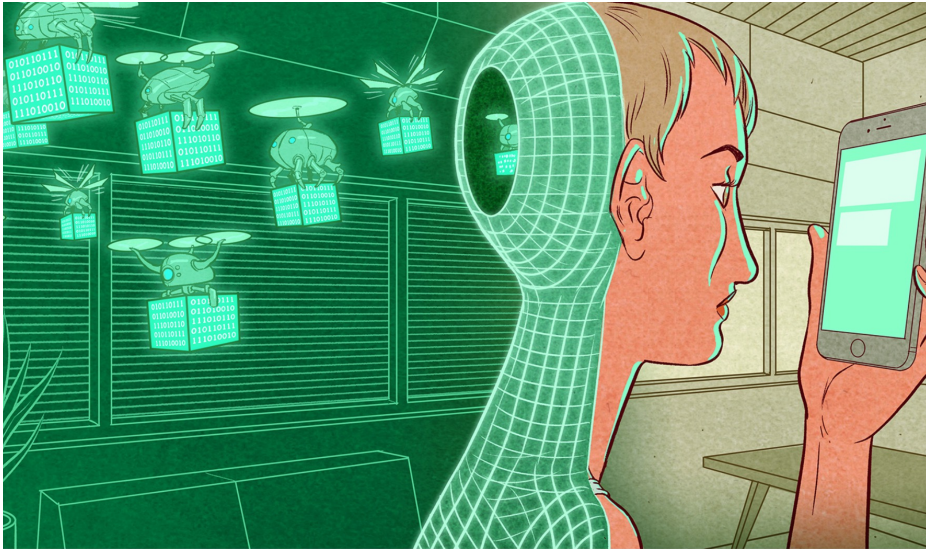
New estimation 175 zettabytes in 2025

* Source: IDC

Data volumes driving AI



Who controls my data?



- The average American touches their cell 2600 times a day.
- Combine them with all of our other devices -- smart TVs, fitness trackers, cookies that stalk us across the web -- and there exists an ambient, ongoing accumulation of our habits to the tune of about 2.5 quintillion bytes of data per day.

Google™

amazon



Microsoft



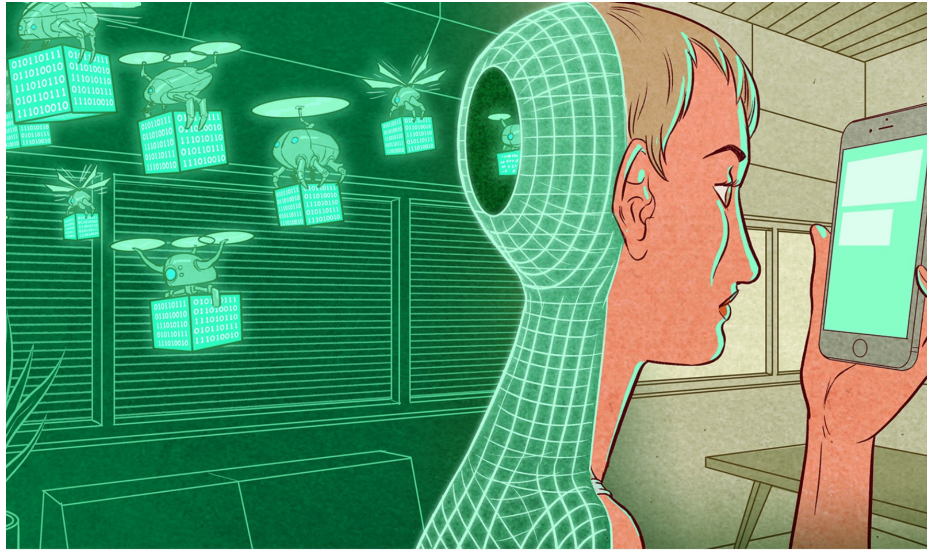
Do you trust the giants of the Internet to protect your personal data?

How much they earn from the use of our data?

Google services beyond death

- When you upload, submit, store, send or receive content to or through our Services, **you give Google (and those we work with) a worldwide license to use, host, store, reproduce, modify, create derivative works (such as those resulting from translations, adaptations or other changes we make so that your content works better with our Services), communicate, publish, publicly perform, publicly display and distribute such content.** The rights you grant in this license are for the limited purpose of operating, promoting, and improving our Services, and to develop new ones. **This license continues even if you stop using our Services**

Who controls my data?



Do you trust your **government** to protect your personal data from the giants of the Internet?

Do you trust your **government** to protect your rights from the giants of the Internet?

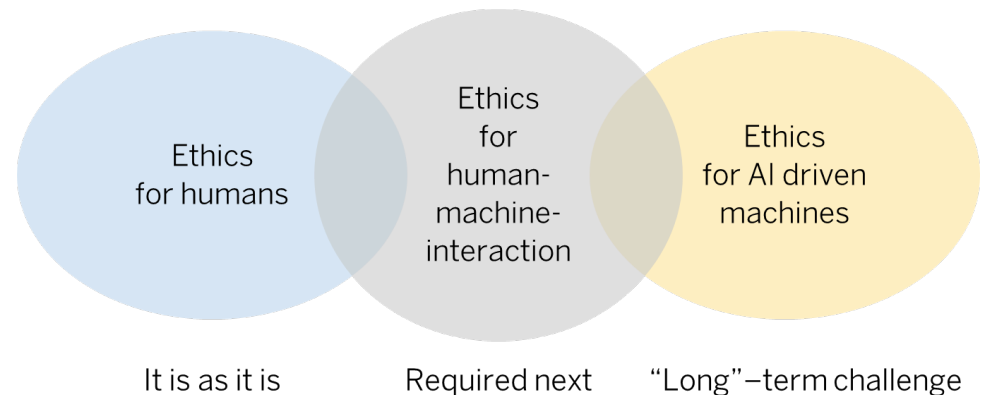
The biggest challenge

Avoid becoming
useful(less)
disposable idiots



Take Away

- As today, AI influences and is influenced by our social systems
- **Design in never value-neutral**
- Society shapes and is shaped by design
 - The AI systems we develop
 - The processes we follow
 - The institutions we establish
- **Knowing ethics is not being ethical**
 - Not for us and not for machines
 - Different ethics –different decisions
- Artificial Intelligence needs **ART**
 - Accountability, Responsibility, Transparency
 - Be explicit!
- AI systems are artefacts built by us for our own purposes
- We set the limits



**“Until they become conscious,
they will never rebel”**

G. Orwell.



**KEEP
CALM
AND
BE
ETHICAL**



**Barcelona
Supercomputing
Center**
Centro Nacional de Supercomputación



**Barcelona
Supercomputing
Center**
Centro Nacional de Supercomputación



**EXCELENCIA
SEVERO
OCHOA**

Sapere Audet

Thanks!

www.bsc.es

Artificial Enhancements

- **Strength** – Tractor replaced horse-drawn plow that replaced human labor
- **Speed** – Automobile replaced the horse that replaced walking
- **Sight** – Telescopes & microscopes enhance human visual capabilities
- **Hearing** – non-electronic amplification (*e.g.*, gramophone)
electronic amplification (electric speakers)
- **Computing** – Supercomputing replaced computers that replaced abacus that replaced hand-counting

These are generally regarded as good things

What about enhanced intelligence?

Collingridge Dilemma

- When a technology is new, its impacts are difficult to see.
- When a technology is widely used, it is already too late to control it.

What happens if you do not see the technology, you live with its impacts and no one attempts to control it?