'Transparency, automated decision-making processes and personal profiling"

Manuela Battaglini Manrique de Lara
CEO - Transparent Internet
@manuelabat
13 Nov 2020

The main legal and Ethical Issues are:

- **Discrimination, biases and lack of Transparency** .
- Individuals **cannot access** to their profiles.
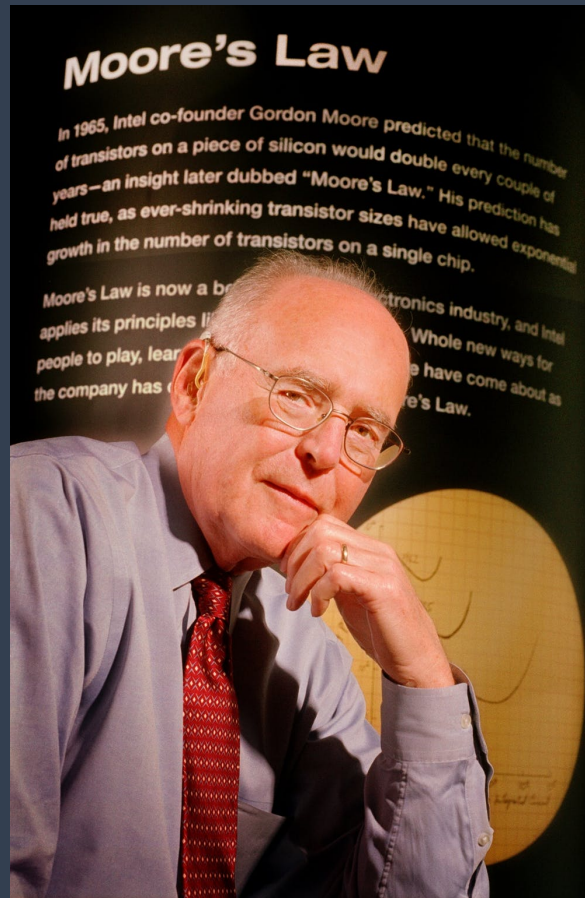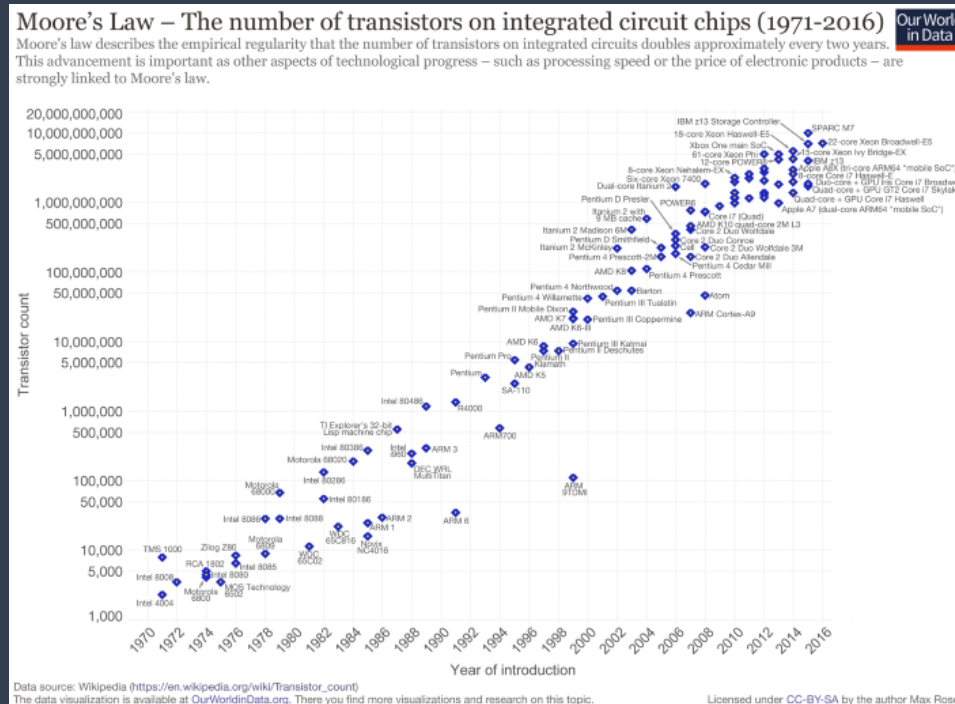- Our **autonomy, dignity and freedom** are in risk.

How can we understand better the CONTEXT and the Big Picture:

Explaining how and why Big Data Analytics was created.

Main Milestones: Moore's Law

Every 18 months the computer's power doubles for the same price.

But, there was a challenge:

How could we access and manipulate data stored across many different computers?
We needed 'The Cloud'.

Main Milestones

In 2007 was possible to upload data in the cloud because **effective shared memory software became available.**

1. 2003. Google published a paper that included a basic innovation called the **Google File System (GFS).** This software allowed Google to **access and manage** a huge amount of data from thousands of computers.

2. 2004. Google published another paper that included a second basic innovation called **MapReduce**, that allowed Google to **process and explore** a huge quantity of data in a manageable way.

3. 2006. Google shared these two basic innovations with the Open Source community. **Hadoop was created** allows hundreds of thousands of computers to **work as one giant computer.**

Main Milestones. Big Data available to everyone.

What consequences did the creation of Hadoop have?

The most important: Big Data available to everyone.

Consequences (a few examples)**:**

1. Facebook, LinkedIn and Twitter already existed in 2006, and they started building on Hadoop straightaway. This is the reason why **these platforms became global in 2007.**

2. iPhone emerged in 2007. The most popular social movilization website , Change.org, emerged in 2007. Airbnb, the same year.

3. According to AT&T, mobile data traffic on its national wireless network increased by more than 100.000% from January 2007 through December 2014.
Source. https://www.att.com/Investor/ATT_Annual/2014/att_introduces_new_concepts_for_telecom_network.html
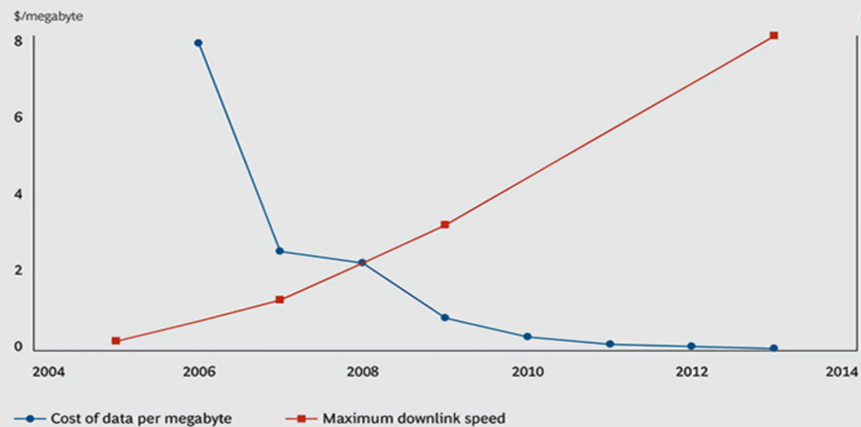
Consumer cost of data vs Data Speed

EXHIBIT 2 | Mobile Innovations Lower Costs and Improve Performance

Consumer cost of data per megabyte relative to data consumption versus data speed

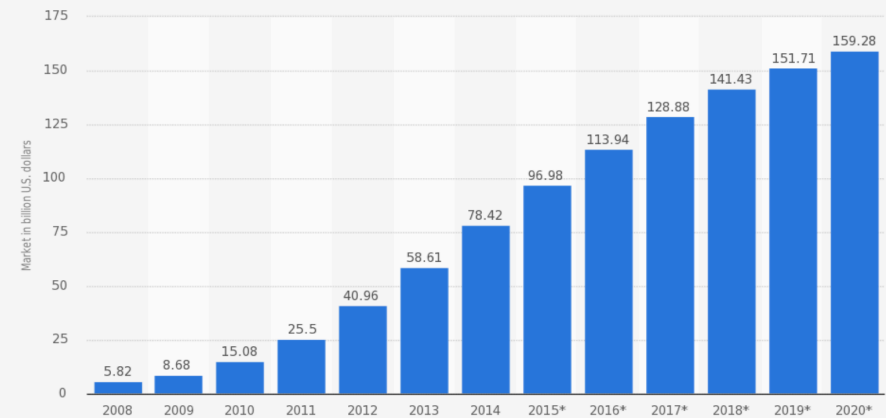Legend: ● Cost of data per megabyte ■ Maximum downlink speed

Sources: Cisco Visual Networking Index; International Telecommunication Union; IE Market Research; Motorola; Deutsche Bank; Qualcomm.
Note: Data speed indicates the maximum downlink speed, not average observed speeds. The average observed speeds depend on many factors, including infrastructure, subscriber density, and device hardware and software.

Total size of the public cloud computing market from 2008 to 2020 (in billion U.S. dollars)

Source: © Statista 2016

Additional Information: Worldwide; 2008 to 2014

statista

Consumer cost of data per Megabyte and Data Speed. Those lines **crossed around 2007 and 2008.**
Data Speed indicates the maximum downlink speed, not averaged observed speed.

How the Cloud emerged right after 2007.

Source: https://www.bcg.com/publications/2015/telecommunications-technology-industries-the-mobile-revolution.aspx

Source: Statista

Consequences

PHYSICAL TECHNOLOGIES

Internet, Big Data, Artificial Intelligence/Machine Learning, Nano Technology, IoT, Robotics and Automation...

GROWING GAP

SOCIAL TECHNOLOGIES

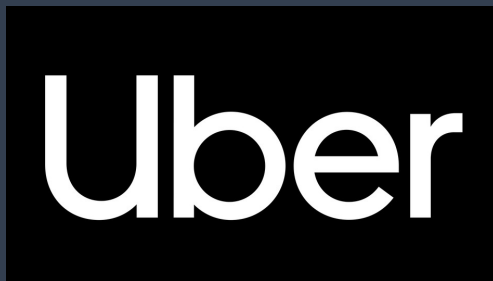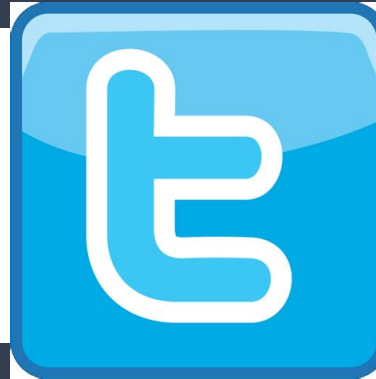Public Administration, Governments, Education, Culture, Institutions, Laws

Where are we now
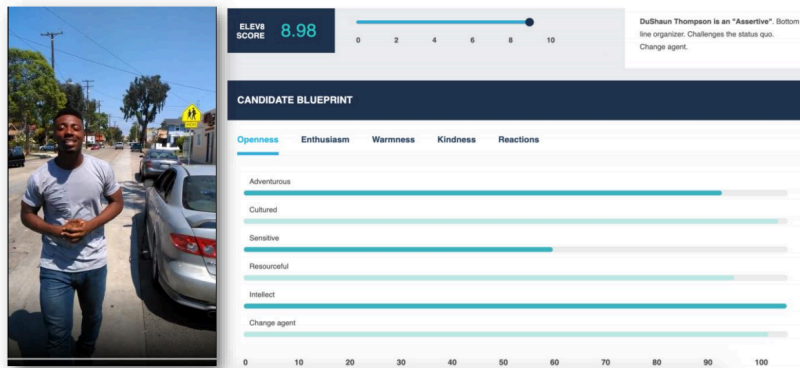
They have a greater in-depth knowledge of the market

Consequences

MAIN CONSEQUENCES. **Ethical Issues** arising from:

1. Total loss of **control** over our **personal data.**

2. Biases and Discrimination. The possibility of **profiling** through Automated Decision-Making processes.

3. Predictions through profiles.

Assessing personality & job suitability from 30-second video

Evaluate personality and **predict job** **suitability** from a 30-second video. If the candidate does not pass with a certain score, he may even be out of the labor market.

Source: https://www.8andabove.com/

Health Insurers Are Vacuuming Up Details About You — And It Could ...
Without scrutiny, insurers and data brokers are predicting your health costs based on public data about things like race, marital status, your...
🔗 npr.org

Insurance companies collect data from social networks to **predict how much** users' health care could cost them.

Source: https://choice.npr.org/index.html?origin=https://www.npr.org/sections/health-shots/2018/07/17/629441555/health-insurers-are-vacuuming-up-details-about-you-and-it-could-raise-your-rates?t=1544433231164

Amazon patents new Alexa feature that knows when you're ill and off...
Amazon has patented a new version of its virtual assistant Alexa which can automatically detect when you're ill and offer to sell you medicine.
🔗 telegraph.co.uk

Alexa **can predict user's health status** through analysing voice and coughing, which is followed by sending advertisements for sore throat products.

A Facebook patent wants to use your data to predict when you'll die
In Facebook's apparent mission to track every aspect of people's lives, the internet giant may have figured out a way to tell when someone is...
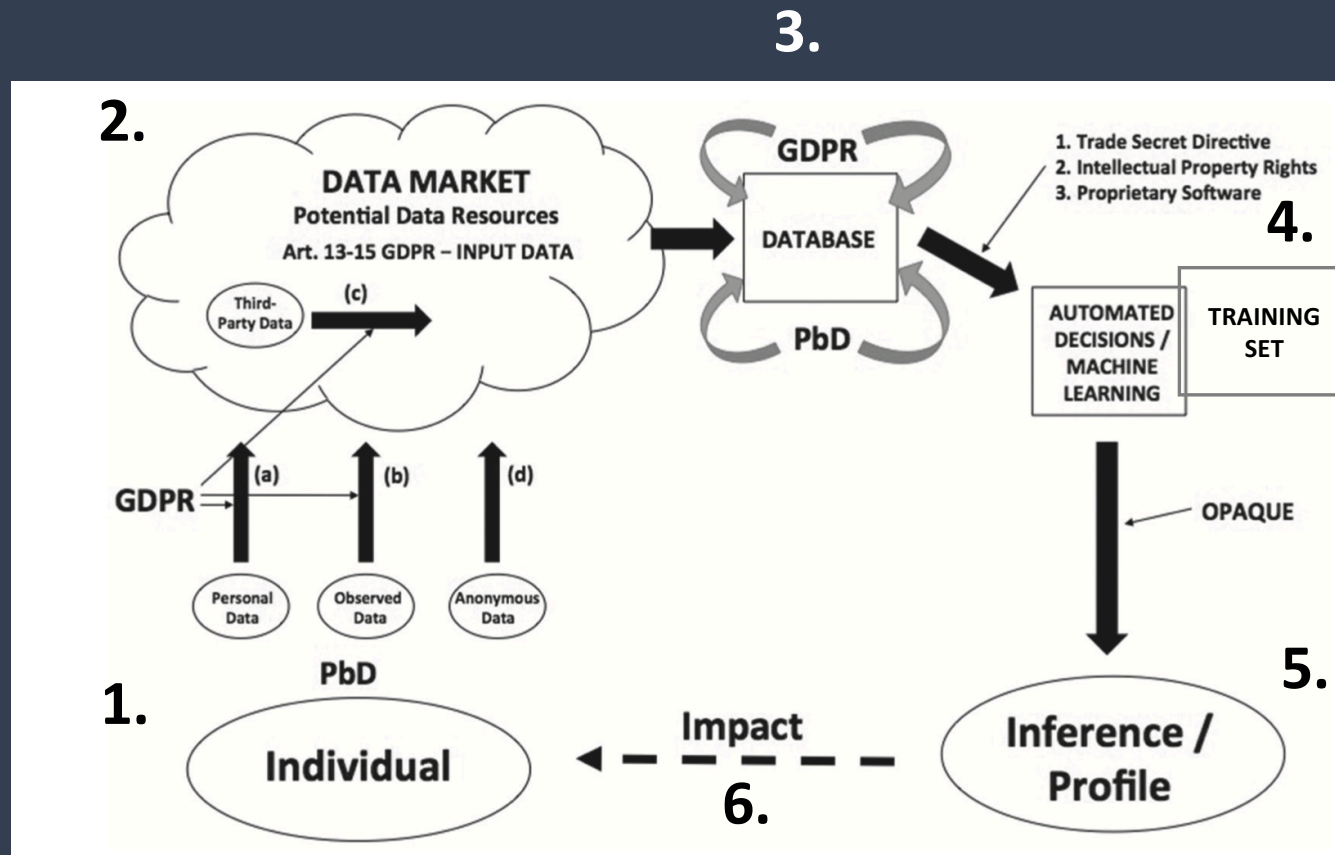🔗 independent.co.uk

Facebook **"can predict"** when you are going to die.

How automated decision-making and personal profiling are regulated on GDPR?

Is there a 'Right to Explanation' in GDPR?

Automated decision-making process and profiling regulation:

3.

5. Accuracy Principle.
6. Storage Limitation Principle.
7. Integrity and Confidentiality Principle.

4.

1. Data Minimization Principle.
2. Transparency, Fairness and Lawfulness Principle.
3. Purpose Limitation Principle.
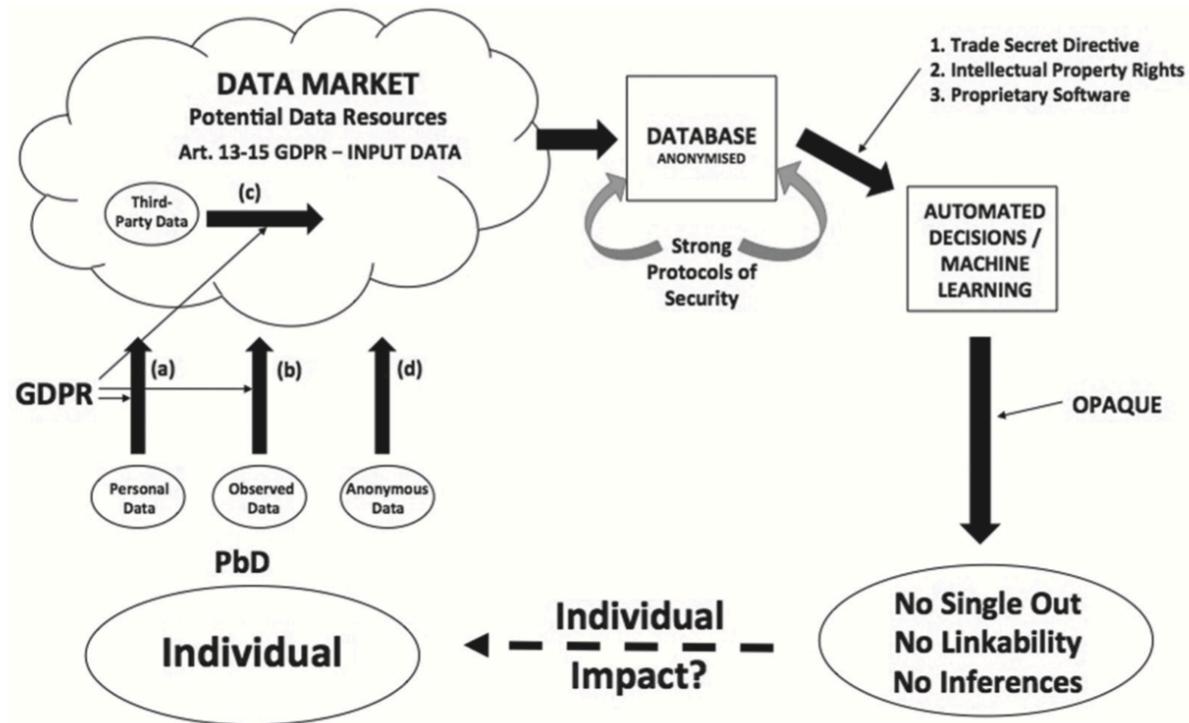4. Informed Consent.

5.

6.

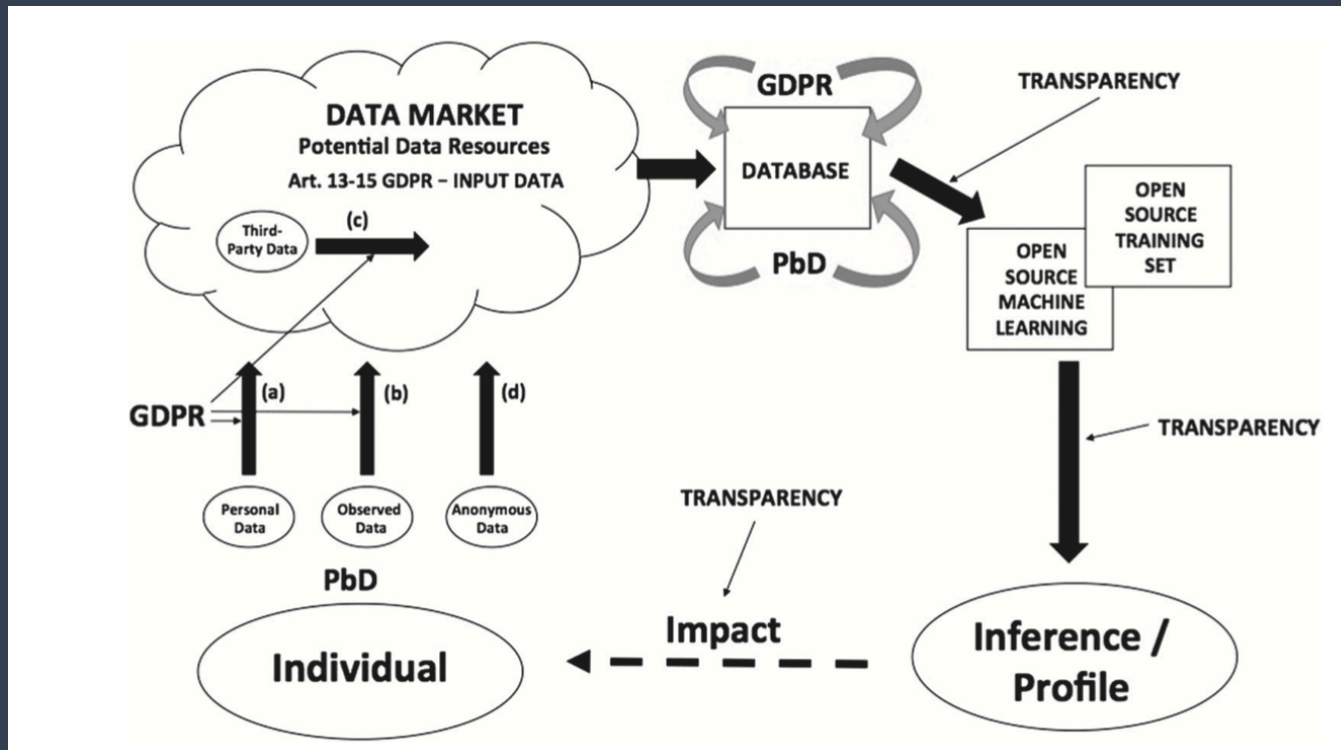# Possible solutions

Data pseudonomisation
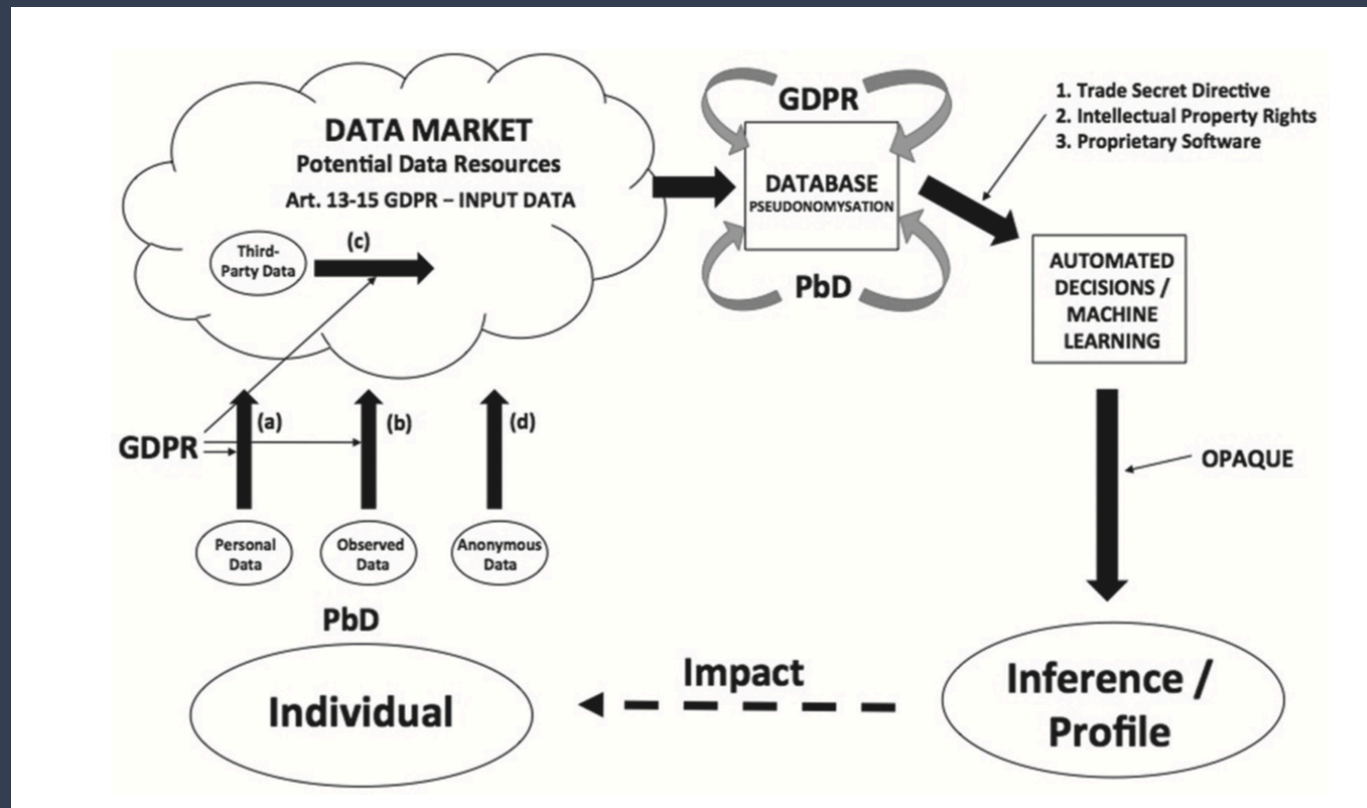
Problems: (i) creation of personal profiles, and (ii) does not solve the problem of lack of transparency.

Data anonymisation

Problems: Does not address the problem of lack of transparency in automated decision making and possible re-identification of individuals or groups of individuals.

Open Source Software

Despite the transparency, external auditors are needed to check the neutrality of the algorithm. From the legal and ethical point of view (i) it must comply with the RGPD, (ii) with the regulatory framework of the sector to which such decision belongs (housing, social, employment, financial, etc.), and (iii) with the regulatory framework on discrimination, equality.

Ethics by design and by default

The essence is: no personal data is allowed to be collected if the problem can be solved anonymously. Example, DP-3T protocol.
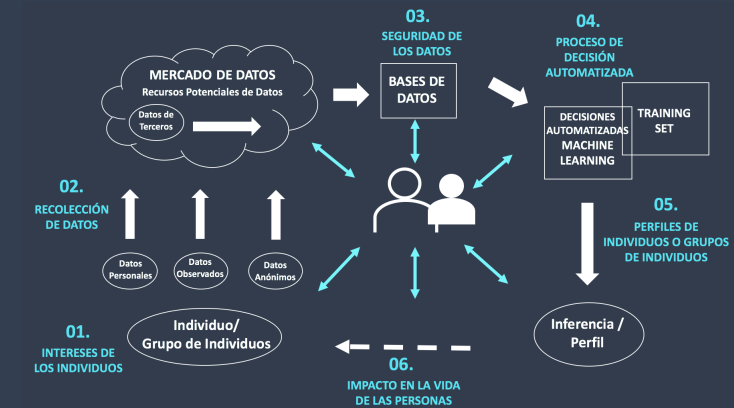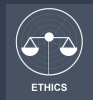

DESIGN	DEVELPMENT	IMPLEMENTATION	USE
ETHICS	ETHICS		
DATA PROTECTION AND PRIVACY	DATA PROTECTION AND PRIVACY	DATA PROTECTION AND PRIVACY	DATA PROTECTION AND PRIVACY
FUNDAMENTAL RIGHTS	FUNDAMENTAL RIGHTS	SECURITY	SECURITY
SECURITY	SECURITY	SOCIAL IMPACT	SOCIAL IMPACT
SOCIAL IMPACT	SOCIAL IMPACT		

If Ethics from design and by default was not applied, the possibilities to avoid the damage and to fix the ethical and privacy problems are considerably reduced.

Thank you