

I. A DREAM CALLED BIG DATA

The term Big Data has been used by academics since the early 1990s. Big Data refers to the acquisition, accumulation, and analysis of very large and diverse sets of information. The term was originally concocted to emphasize a new phenomenon: the explosive growth of available data.

This explosive growth of data is essentially the consequence of the digitalisation of our lives. Everything we do, watch, and listen. Our relationships, experiences, hopes and – especially – fears. Our life is the primary source of Big Data.

How is this data generated? Mostly through web search queries, social networks, apps, voice assistants, smartphones, physical beacons, IoT. If it is connected, it is producing Big Data.

Big Data is everywhere, and yet not everyone can harness its power. The technology necessary to acquire, and process Big Data is not yet easily available nor cheap. However, Big Data, and the technological innovations that it has enabled, such as machine learning, are undoubtedly one of the cornerstones of the 21st century.

In 2014, former Working Party 29¹ released a statement on the impact of the development of Big Data. While acknowledging that the challenges of Big Data would require innovative thinking on how key data protection principles were applied, many collective and individual benefits were expected from the development of the field of Big Data.

Six years have passed since that day, and the General Data Protection Regulation has come into effect. What happened to the dream of collective and individual benefits? Well, let us just say things could have gone better. Cambridge Analytica and the 2016 USA elections and Brexit happened. Facebook suffered a 5 billion fine from the FTC. The Privacy Shield fell under an historical decision by the European Court of Justice. Facial recognition systems have been used to suppress Hong Kong protesters. TikTok and its data processing activities have become an international matter.

The world woke up to a world where Big Data is being used and abused by a few public and private actors, for purposes far from the original dream of collective and individual benefits. What seemed like a reality at hand was in fact just a dream.

In the private sector, the consolidation of dominant and monopolistic positions in the digital market in these last few years created a “winners take all” effect on the availability and exploitation of the Big Data gold mine. Google and Facebook share among themselves the biggest piece of the pie, leaving just a few crumbs to everyone else.

More than 92% of the internet searches go through Google, processing more than 40.000 queries every second. More than 3.5 billion searches per year – leaving the remaining 8% of the market to Bing, Yahoo, Baidu, YANDEX and DuckDuckGo.

How do they use all this data? The Internet Health Report 2019² shows that Alphabet, Google’s parent company, earns 85% of their revenue from digital advertising. The remaining 14.5% of their revenue comes from other services and new innovations.

¹ Since 2016 the Working Party 29 assumed the name of European Data Protection Board

² <https://internethealthreport.org/2019/>

Facebook, on the other hand, makes more than 98.5% of its revenue from advertising. In 2018 Google and Facebook alone controlled about 84% of the global digital advertising market outside of China.

These monopolistic players twisted what was once a dream of unprecedented collective benefits and harnessed the Big Data power to create an advertising machine meant to enrich few by exploiting many – the unaware consumers.

II. THE HARSH REALITY

Big Data gave birth to what today we call surveillance capitalism. Surveillance capitalism can be described as a “new economic order that claims human experience as free raw material for hidden commercial practices of extraction, prediction and sales”³.

The idea behind surveillance capitalism is simple: every human action produces raw data. This data can be harnessed and processed, to create “prediction products” – aimed to predict our future actions, thoughts, and even feelings. Basically, behavioural advertising. Consumers in the age of surveillance capitalism are both the raw production materials and the customers of their own products.

However, Big Data and surveillance capitalism are just concepts. The true driving force of the new data-driven economy are mathematical models and algorithms that feed on Big Data.

Cathy o’ Neil⁴ describes an algorithm as “an opinion embedded in math” – what we humans would call a stereotype. These mathematical stereotypes are creating many harmful effects, often concealed as “services” or promises of “optimization”. Almost every human field is ridden with automated decisions made by algorithms that covertly affect millions of people every day: online advertising, job recruiting, politics, education, justice, insurance, credit scoring.

Automated decision making and algorithms create harmful effects because their only feedback is money. If an algorithm is creating profits, then it is working as intended. Corporations and data scientists usually do not care about the receiving end of their operations. As long as the model produces money, nobody cares if it negatively affects some (many) people at critical stages of their life: getting a job, asking for a loan, being sentenced to prison, and so on.

III. OBSCURITY BY DESIGN

Most times, consumers do not know what is happening to their data, or even realise that their data is being processed and shared to third parties. A simple action, like asking the Google assistant or Siri for directions, produces huge quantities of data that are processed in real time.

The General Data Protection Regulation tried to impose new and stronger transparency obligations towards data controllers, with little success. It takes 32 minutes and a College degree to read and understand Google’s privacy policy⁵. The same effort required to read *The Art of War* by Sun Tzu. To this day, Google still declares that they do not sell personal data. Which is true: they do not sell personal data – they just process it and sell its output to the best bidder.

³ The Age of Surveillance Capitalism, Shoshana Zuboff (2019) 4

⁴ Weapons of Math Destruction, Cathy O’Neil (2016)

⁵<https://therealtime.com/wp-content/uploads/2020/09/privacy-policies-tech-giants-small-8-24-scaled.jpg>

The information overload and the tremendous complexity of data processing activities have extinguished any hope for real transparency. Consumers simply cannot understand what is going on and they cannot assess the depth and consequences of their personal data being processed. This (justifiable) lack of understanding is what makes people ignore their privacy. Most consumers gladly accept any kind of data processing operation, if they perceive some sort of reward, such as free services and products. Truth is, only an infinitesimal part of their data is being processed to create useful services and innovations. For the most part, it is just raw material for profiling and behavioural advertising.

The obscurity of data processing creates relevant information asymmetry, and therefore market failure – a situation in which the allocation of goods and services is not “Pareto Efficient”⁶.

Eric Schmidt, CEO of Google from 2001 to 2011, once said that the Internet will disappear. What he meant is that the Internet would just blend into our lives and vanish into the background, until it is undistinguishable. The boundaries between online and offline are already vanishing, as we are almost always connected with one another, and with countless other connected objects. The way we live and interact with each other, and the way we see the world, is increasingly being shaped by the new “Onlife”⁷ paradigm.

Automated data processing and decision making is already strongly embedded in our daily lives. Our life and our actions are affected by automated data processing and algorithms even without ever being directly identified by the corporations who run these algorithms. Advertisers do not care about our identity, as long as they can predict and influence our actions towards a precise goal.

The digitalisation of our lives has increasingly subjected us all to a reality of obscure manipulation and algorithmic repression. People are inevitably profiled into unknown stereotypes that rarely fit their life and personality. Obscure algorithms are used to target people and exploit their fears or ambitions, to sell things they do not need when they are most vulnerable. Predatory ads target the most vulnerable of our society, at their most desperate times, promising a false road to prosperity. In the USA, the for-profit college industry

specifically targets individuals using keywords such as “low self-esteem”, “stuck”, “welfare mom with kids”, “pregnant ladies”, “recent divorce”, “recent death”, “recent incarceration”, “dead end jobs-no future”⁸. These ads are meant to sell overpriced bachelor’s degrees to desperate people, exploiting their data against them. During the 2016 campaign for the Trump administration, 3.5 million black Americans were profiled and categorised as ‘deterrence’ by the Cambridge Analytica algorithms – voters that had to stay home on election day⁹. In 2020 an algorithm determined grades for tens of thousands UK A-level students. Approximately 39% of A-level results were downgraded, and disadvantaged students were the worst affected as the algorithm copied the inequalities that exist in the U.K.’s education system. Long standing structural inequalities are often deeply embedded in algorithms’ datasets and almost never accounted for¹⁰.

⁶ Pareto efficiency or Pareto optimality is a situation where no individual or preference criterion can be better off without making at least one individual or preference criterion worse off

⁷ The Fourth Revolution: How the Infosphere is Reshaping Human Reality, Luciano Floridi, 2014

⁸ Weapons of Math Destruction, Cathy O’Neil (2016)

⁹ <https://www.channel4.com/news/revealed-trump-campaign-strategy-to-deter-millions-of-black-americans-from-voting-in-2016>

¹⁰ https://www.researchgate.net/publication/343472595_The_Ethics_of_Algorithms_Key_Problems_and_Solutions 11

Examples like these are countless. People's personal data is being processed by automated system that only a privileged minority (barely) understand how they work. Citizens and consumers are being exploited and they do not even know it.

In all this obscurity, there is however a light. Digital rights organisations, journalists, and advocates are rising to fight questionable or unlawful practices. Citizens are starting to realise the reality of automated data processing, also thanks to the distribution and popularisation of short movies such as "The Great Hack" or the "Social Dilemma".

STOP DREAMING, AND TAKE BACK REALITY

The status-quo can and should be changed, but we need the right tools.

First, we need to reshape the notion of privacy. The original "right to be left alone" theorised by Samuel Warren and Louis Brandeis does not suffice anymore. This kind of privacy is dead.

In the infosphere, nobody can be left alone. Anything we do produces data; from the moment we are born and in certain ways even before. Therefore, privacy needs to be more than just the right to be left alone.

Most of our interactions take place without conscious reflection; an autonomic behaviour that is the result of individual learning processes. However, autonomous action requires deliberate reflection on our choices¹¹. Algorithms, profiling, targeted advertising, and service customisation deceitfully deprive us from the ability to reflect on our choices, because we lack the information and knowledge necessary to understand how some choices are presented to us. The risk that algorithmic systems can hinder human autonomy by covertly shaping our choices has been also recently discussed by the European Commission's European Group on Ethics in Science and Technologies and by many AI ethics frameworks. Obscure profiling and automated decision making endanger our self-determination freedom and give rise to numerous other negative externalities, akin to environmental harms or pollution¹².

Thus, privacy – or better yet, the "protection of natural persons with regards to processing of personal data"¹³, should act as a link between society and the individual, a precondition to any other fundamental right. A safe space to enable self-determination and freedom to think, communicate and form relationships, free from manipulation and repression. Privacy should not be a private interest, but a public good, just like the natural environment. Privacy is the fundamental filter that lets us breathe in an otherwise polluted and toxic social environment.

PRIVACY, AND US

This new notion of privacy, as a public good and safe space from manipulation and repression, requires a public effort. Consumers alone cannot be expected to protect their privacy – the same way as they cannot be expected to solve climate change or environmental pollution. The public sector and other relevant stakeholders should work together to protect privacy as a public good.

We have the tools at our disposal to change the status-quo, and the recent European Court of Justice decision to invalidate the EU-US Privacy Shield shows that the European Union has great leverage to change the rules of the game. The General Data Protection Regulation (GDPR) is an incredible tool –

¹¹ The Dawn of a Critical Transparency Right for the Profiling Era, Mireille Hildebrandt, 2012

¹² A Taxonomy of privacy, Daniel J. Solove, 2006

¹³ Reg (EU) 2016/679 (General Data Protection Regulation)

and yet just a tool. Even the GDPR does not address all imbalances of power in the information society, and many of its principles, such as privacy by design, transparency, or data portability, have yet to be taken seriously by the relevant stakeholders.

Consumers organisations, industry groups, the public sector, and the supervisor authorities, should work together to push for the enforcement of these principles – to give back control of their data (and their lives) to consumers and citizens.

Fairness, accountability, and transparency are the key elements required to rebuild social trust and to make sure that privacy and fundamental human rights will be embedded into technology. Big Data and algorithms can benefit the society and the individuals, but we need strong governance, vision, and alert watch dogs. Consumers organisation should start tackling important, yet complex and little-known topics such as algorithmic bias and discrimination, behavioural advertising, and surveillance capitalism. They should promote knowledge, especially towards young people, and look deeper into the asymmetries between consumers and organisations that have access to their data and their profiles.

Algorithms should undergo public scrutiny and auditing. Decision making algorithms should also be subject to legibility tests¹⁴, as a tool to promote transparency, comprehensibility, and non-discrimination.

Organisations should be required to provide meaningful and comprehensible information about the contextual use of their algorithms, their consequences and impact towards individuals.

Consumers organisations should work together with digital rights organisations to help consumers exercise their access and portability rights. Article 22 of the GDPR should be strongly enforced and delved into, as the only existing safeguard against the dictatorship of automated decision making.

On the other hand, the public sector should work together with supervisor authorities, consumers organisations and digital rights organisations to determine the limits of monitoring and monetising people. Data is constantly gathered more than what is needed to provide services, to extract monopoly rents¹⁵. In the meanwhile, Apple is promoting its newest iPhone and its iOS as the most privacy-friendly smartphone, while people who cannot afford it are stuck with low-mid tier smartphones ridden with exploitative, pre-installed bloatware and limited privacy settings¹⁶.

Privacy should not be a commodity nor exclusively a private interest. Everyone should be guaranteed the same right to privacy, despite their personal wealth. Privacy is a fundamental right and a public good and should be treated as such.

Likewise, products should not be allowed onto the market if there is a risk of privacy harm – the same as legislators do not allow to market products if there is a risk of environmental or health harm. The right to human dignity demands limits to the degree to which an individual can be scanned, monitored and monetised — irrespective of any claims to putative “consent.”¹⁷ Therefore, consumers organisations should work together with industry groups to enforce true privacy by design services and products, creating ethical data processing frameworks and digital commons, available to everyone.

¹⁴ Why a right to legibility of automated decision making exists, Malgieri, 2017

¹⁵ <https://thebaffler.com/latest/capitalisms-new-clothes-morozov>

¹⁶ <https://privacyinternational.org/advocacy/3320/open-letter-google> ¹⁷ Privacy 2030, A New Vision For Europe, IAPP, 2019

Data security should also be a relevant part of the debate, as data insecurity creates relevant privacy harms and negative externalities, such as identity theft and distrust towards the digital economy.

Citizens should demand fairness, transparency, and accountability. But they cannot be left alone in this fight against surveillance capitalism, data maximisation and exploitation. It is time to act. Time to enforce our fundamental human rights and ensure a safe digital future for everyone, as human beings, not data points to be exploited.

BIBLIOGRAPHY

A Taxonomy of privacy, Daniel J. Solove, 2006

Privacy 2030, A New Vision For Europe, IAPP, 2019

Since 2016 the Working Party 29 assumed the name of European Data Protection Board

The Age of Surveillance Capitalism, Shoshana Zuboff (2019)

The Dawn of a Critical Transparency Right for the Profiling Era, Mireille Hildebrandt, 2012

The Fourth Revolution: How the Infosphere is Reshaping Human Reality, Luciano Floridi, 2014

Reg. (EU) 2016/679 (General Data Protection Regulation)

Weapons of Math Destruction, Cathy O'Neil (2016)

Why a right to legibility of automated decision making exists, Malgieri, Comandé, 2017
<https://internethealthreport.org/2019/>

<https://thebaffler.com/latest/capitalisms-new-clothes-morozov>

<https://privacyinternational.org/advocacy/3320/open-letter-google>

<https://therealtime.com/wp-content/uploads/2020/09/privacy-policies-tech-giants-small-8-24-scaled.jpg>

https://www.researchgate.net/publication/343472595_The_Ethics_of_Algorithms_Key_Problems_and_Solutions

<https://www.channel4.com/news/revealed-trump-campaign-strategy-to-deter-millions-of-black-americans-from-voting-in-2016>