

AI: towards a critical utopia

Nexa Summer School 2019 - 1st edition
26th of June - 2nd of July 2019

AI's impact on human rights in the digital era

Monica A. Senor

Who I am

- I was a practicing criminal lawyer specialising in ICT law, cyber crimes, privacy and data protection until Novembre 2018
- Since December 2018 I've been an officer at the Italian Data Protection Supervisory Authority, Garante per la Protezione dei Dati Personali
- The opinions I'll express are mine and do not reflect the official policy or position of the Authority

Who are you?

Lecture overview

1. What is AI in the context of human rights?
2. What human rights are?
3. AI v. freedom of expression
4. AI v. data protection
5. AI v. non discrimination principle
6. Conclusions



What is AI?

weak AI v. strong AI

Artificial intelligence
Algorithm
Machine Learning

Design stage



Humans **procure** artificial intelligence systems and define their **intentions**.

Deployment stage



Humans **define the input** and **design the algorithm**.



Humans **clean and label the input data**.



In some cases, humans **define the outputs** of artificial intelligence systems.



Artificial intelligence systems **define the model** used and algorithms **learn and adapt independently**.

Implementation stage



Humans **decide how to apply and use the outputs**.

[A page of dense, handwritten text in a historical script, likely Latin or French, with some red ink used for initials or headings. The text is arranged in approximately 25 lines across the page.]

What human rights are?

All human beings are born free
and equal in dignity and rights

Article 1 of the Universal Declaration of Human Rights

Everyone is entitled to all the rights and freedoms set forth in this Declaration, without distinction of any kind, such as race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status

Article 2 of the Universal Declaration of Human Rights

- right to life, liberty and security
- protection against enslavement
- prohibition of genocide
- right to freedom of opinion and expression
- right to freedom of movement and residence
- right to education
- right to privacy
- right to a fair trial
- right to an effective remedy by the competent national tribunals for acts violating the fundamental rights granted by the constitution or by law



AI v. freedom of expression

Freedom of speech and debates
or proceeding in Parliament
ought not to be impeached or
questioned in any court or
place or Parliament

Bill of Rights, 1689

The free communication of ideas and of opinions is one of the most precious rights of man. Any citizen may therefore speak, write and publish freely, except what is tantamount to the abuse of this liberty in the cases determined by Law

Declaration of Human and Civic Rights, 1789

Congress shall make no Law respecting
an establishment of religion, or
prohibiting the free exercise thereof; or
abridging the freedom of speech, or of
the press; or the right of the people
peaceably to assemble, and to petition
the Government for a redress of
grievances

U.S.A. Constitution, I Amendment, 1791

Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers

Universal Declaration of Human Rights, 1948

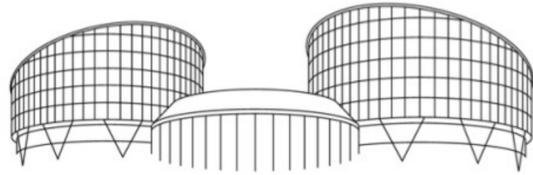
Everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers

The exercise of these freedoms, since it carries with it duties and responsibilities, may be subject to such formalities, conditions, restrictions or penalties as are prescribed by law and are necessary in a democratic society, in the interests of national security, territorial integrity or public safety, for the prevention of disorder or crime, for the protection of health or morals, for the protection of the reputation or rights of others, for preventing the disclosure of information received in confidence, or for maintaining the authority and impartiality of the judiciary

Convention for the Protection of Human Rights
and Fundamental Freedoms, 1950

Everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers

Charter of Fundamental Rights of the
European Union, 2000



EUROPEAN COURT OF HUMAN RIGHTS
COUR EUROPÉENNE DES DROITS DE L'HOMME

Press Unit
Unité de la Presse

Factsheet – Hate speech

March 2017

This factsheet does not bind the Court and is not exhaustive

Hate speech

“**Freedom of expression** constitutes one of the essential foundations of [a democratic] society, one of the basic conditions for its progress and for the development of every man. Subject to paragraph 2 of Article 10 [of the European Convention on Human Rights], it **is applicable not only to ‘information’ or ‘ideas’ that are favourably received or regarded as inoffensive or as a matter of indifference, but also to those that offend, shock or disturb the State or any sector of the population.** Such are the demands of that pluralism, tolerance and broadmindedness without which there is no ‘democratic society’. This means, amongst other things, that every ‘formality’, ‘condition’, ‘restriction’ or ‘penalty’ imposed in this sphere must be proportionate to the legitimate aim pursued.” ([Handyside v. the United Kingdom](#) judgment of 7 December 1976, § 49).

“... [T]olerance and respect for the equal dignity of all human beings constitute the foundations of a democratic, pluralistic society. That being so, as a matter of principle **it may be considered necessary** in certain democratic societies **to sanction or even prevent all forms of expression which spread, incite, promote or justify hatred based on intolerance** ..., provided that any ‘formalities’, ‘conditions’, ‘restrictions’ or ‘penalties’ imposed are proportionate to the legitimate aim pursued.” ([Erbakan v. Turkey](#) judgment of 6 July 2006, § 56).

Recital 46
Directive 2000/31/CE
(e-commerce Directive)

In order to benefit from a limitation of liability, the provider of an information society service, consisting of the storage of information, upon obtaining actual knowledge or awareness of illegal activities has to act expeditiously to remove or to disable access to the information concerned; the removal or disabling of access has to be undertaken in the observance of the principle of freedom of expression and of procedures established for this purpose at national level

"The complexity of decision-making inherent in content moderation may be exacerbated by the introduction of automated processes. Unlike humans, algorithms are today not capable of evaluating cultural context, detecting irony or conducting the critical analysis necessary to accurately identify, for example, "extremist" content or hate speech and are thus more likely to default to content blocking and restriction, undermining the rights of individual users to be heard as well as their right to access information without restriction or censorship"

UN Special Rapporteur's Report 2018



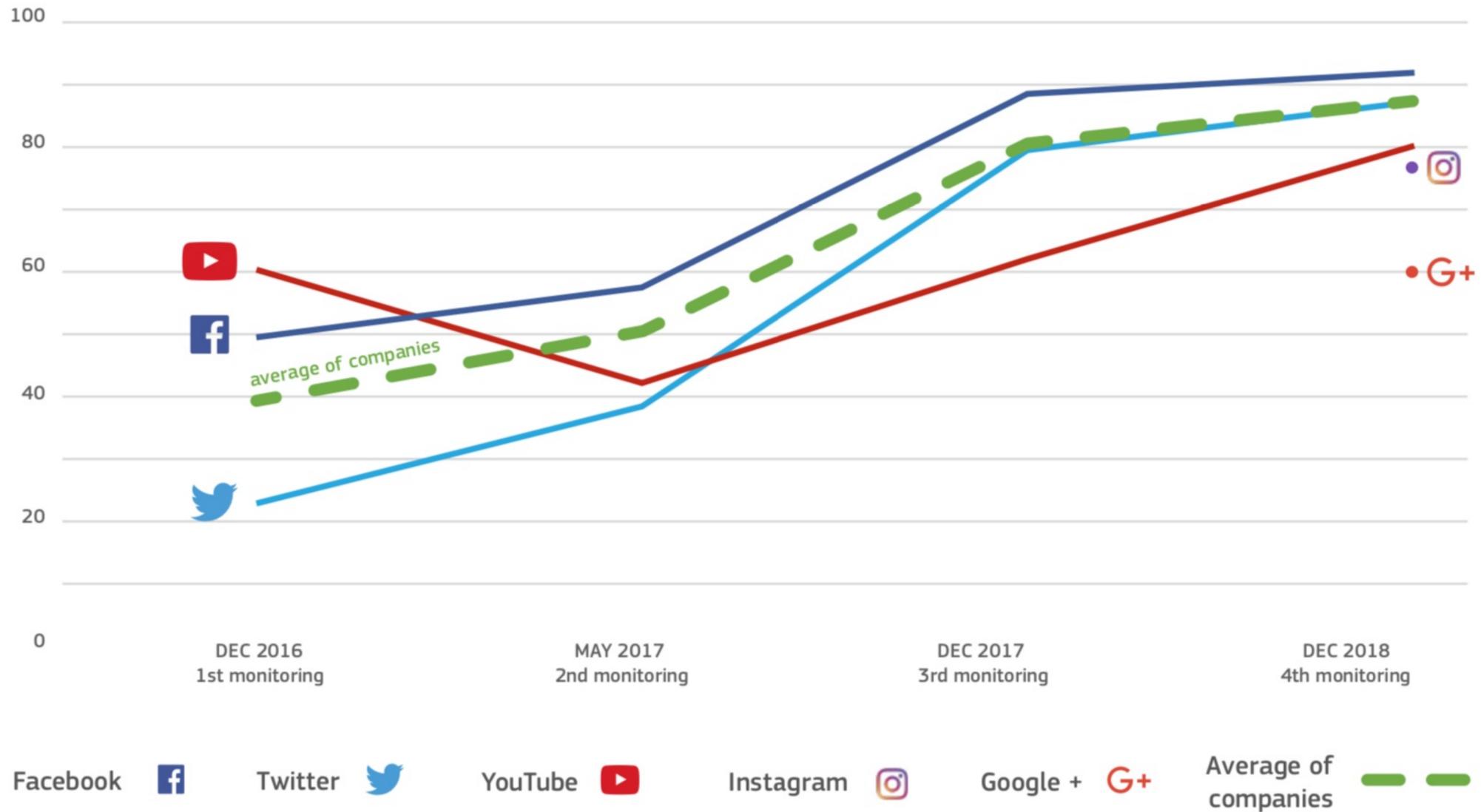
The Napalm Girl case

The EU Code of conduct on countering illegal hate speech online

To prevent and counter the spread of illegal hate speech online, in May 2016, the Commission agreed with Facebook, Microsoft, Twitter and YouTube a "Code of conduct on countering illegal hate speech online"

In the course of 2018, Instagram, Google+, Snapchat and Dailymotion joined the Code of Conduct. Jeuxvideo.com joined in January 2019

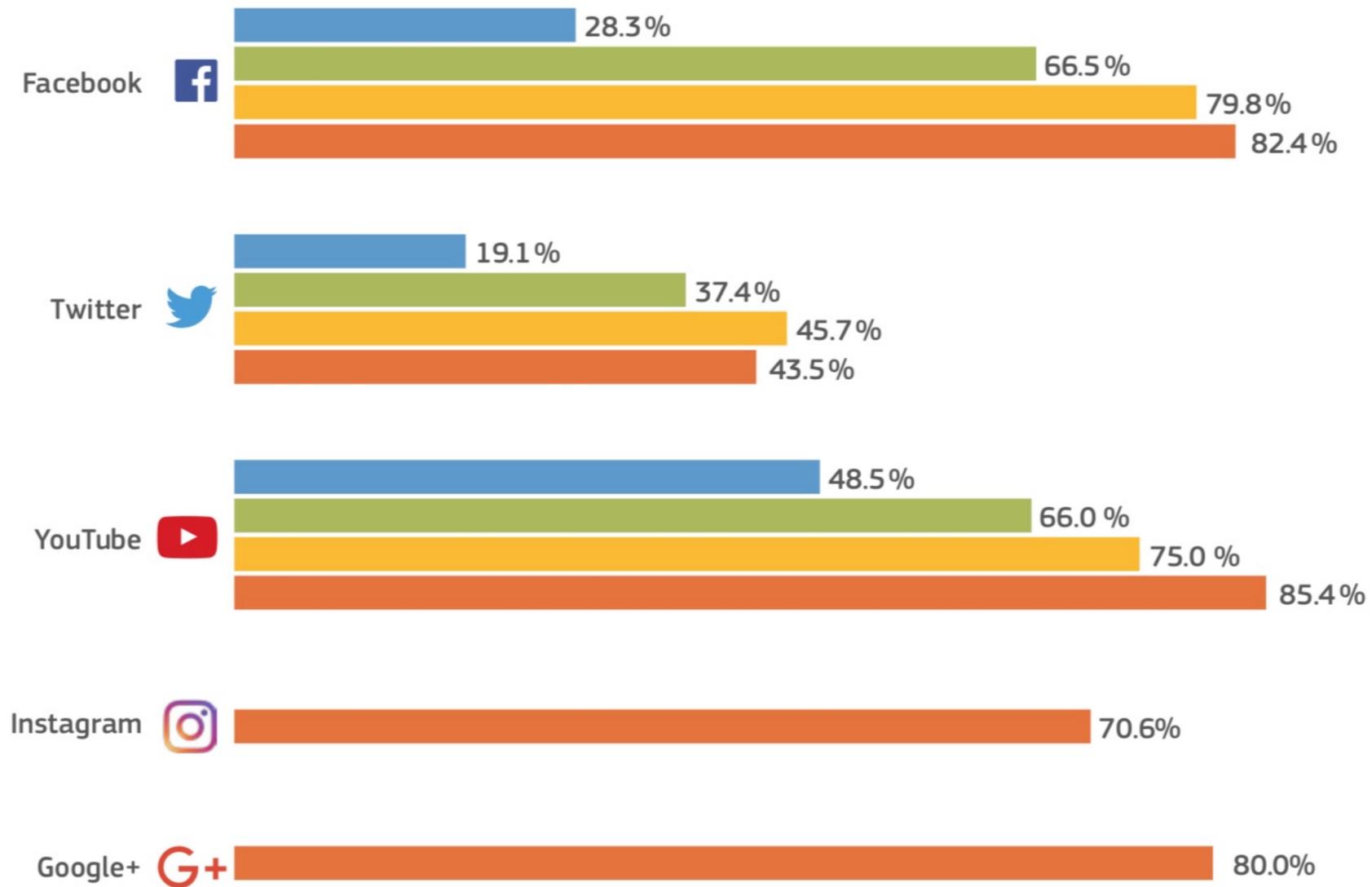
Rate of notifications reviewed within 24 hours since the launch of the Code of Conduct



¹ The figures for Google+ are based on a significantly lower number of cases compared to the other IT companies.

Removals per IT Company

1st Monitoring (Dec. 2016) 2nd Monitoring (May 2017) 3rd Monitoring (Dec. 2017) 4th Monitoring (Dec. 2018)



² YouTube has also limited the features of an additional 23 videos: this implies that while not being removed, a video may not be liked, commented, or shared and does not appear in searches.

The ways in which information is stored, transmitted and secured in the digital age uniquely affect the exercise of the **right to hold opinions** too

Search queries, browsing activities, email, chat, social networking and documents in the cloud are the bricks with which users build their opinions

"Content curation has long informed the capacity of the individual to form opinions. For example, media outlets elevate particular stories to the front page with the intention of shaping and influencing individual knowledge about significant news of the day; commercial advertising has also sought to induce favourable opinions of and cultivate desire for particular products and services"

UN Special Rapporteur's Report 2018

RICHARD H. THALER

WINNER OF THE NOBEL PRIZE IN ECONOMICS

and CASS R. SUNSTEIN

WINNER OF THE HOLBERG PRIZE



Nudge

NEW YORK TIMES Bestseller

Improving Decisions About
Health, Wealth, and Happiness

"One of the few books . . . that fundamentally changes the way I think about the world." —Steven D. Levitt, coauthor of FREAKONOMICS

"A nudge, as we will use the term, is any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives. To count as a mere nudge, the intervention must be easy and cheap to avoid. Nudges are not mandates. Putting fruit at eye level counts as a nudge. Banning junk food does not"

"The use of artificial intelligence extends and enhances the tradition of content curation on the Internet, providing more sophisticated and efficient means of personalising and curating content for the user at a scale beyond the reach of traditional media"

UN Special Rapporteur's Report 2018



AI v. data protection

HARVARD
LAW REVIEW.

VOL. IV.

DECEMBER 15, 1890.

NO. 5.

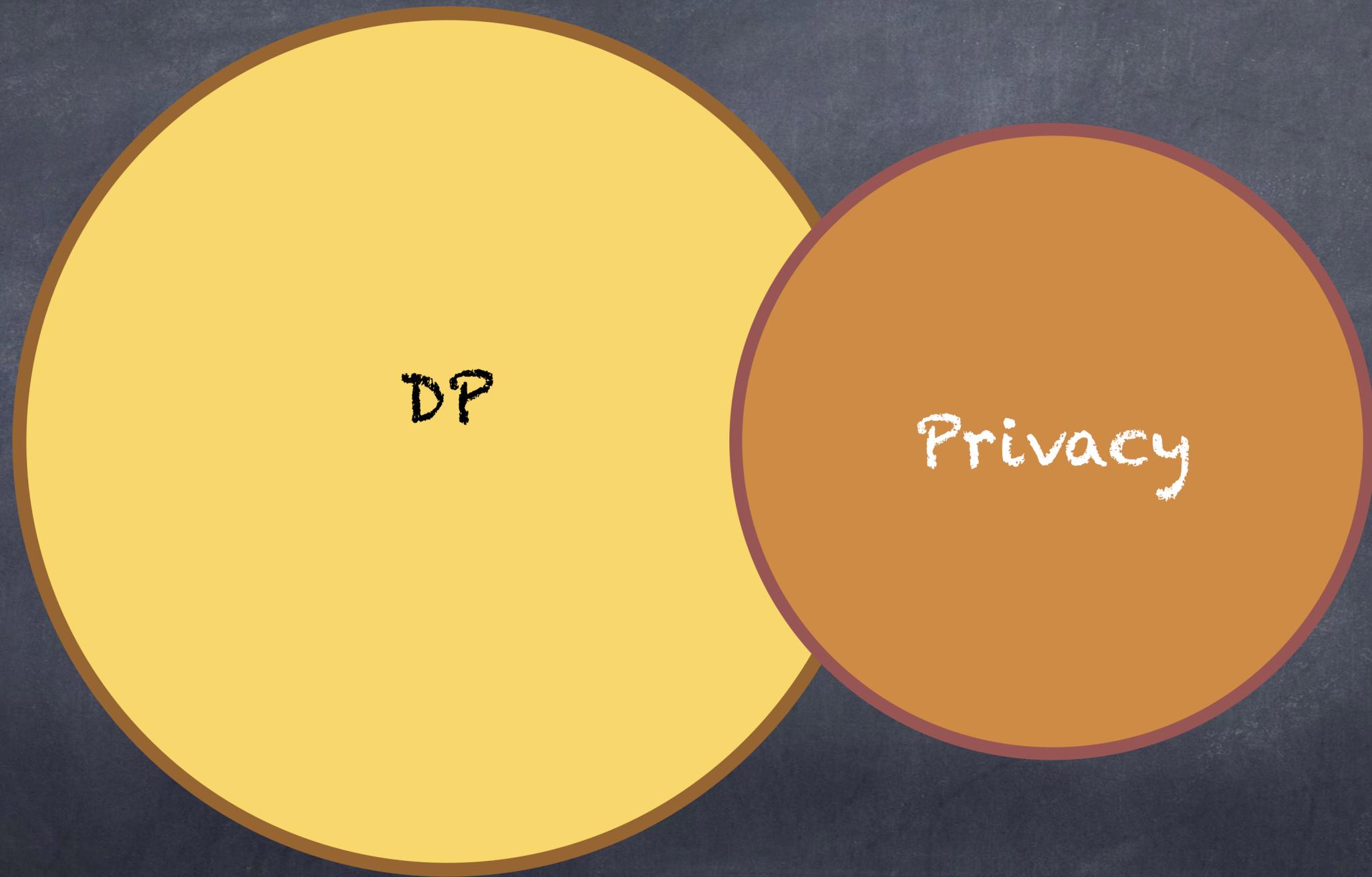
THE RIGHT TO PRIVACY.

"It could be done only on principles of private justice, moral fitness, and public convenience, which, when applied to a new subject, make common law without a precedent; much more when received and approved by usage."

WILLES, J., in *Millar v. Taylor*, 4 Burr. 2303, 2312.

THAT the individual shall have full protection in person and in property is a principle as old as the common law; but

S. Warren & L. Brandeis



DP

Privacy

The data protection right is a (new) fundamental right, pursuing the need to safeguard the integrity and the identity of digital bodies from the modern forms of human control and power (habeas data)

In the beginning there
were physical
bodies ...



The rule of law has been achieved
with the full protection of the
natural person through the integrity
and the inviolability of his/her
physical body from any form of
subjection to human power
(habeas corpus)

... nowadays we talk about
digital bodies

(dematerialization of the
human physical body into
digital information)



We are out information

Stefano Rodotà

European Convention on Human Rights, Art. 8

Right to respect for private and family life

Everyone has the right to respect for his private and family life, his home and his correspondence

There shall be no interference by a public authority with the exercise of this right except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others

Charter of Fundamental Rights of
the European Union, 2000, Art.7

Respect for private and family life

Everyone has the right to respect
for his or her private and family
life, home and communications

Charter of Fundamental Rights of the European Union, 2000, Art. 8

Protection of personal data

Everyone has the right to the protection of personal data concerning him or her.

Such data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law. Everyone has the right of access to data which has been collected concerning him or her, and the right to have it rectified.

Compliance with these rules shall be subject to control by an independent authority.

- CoE Convention 108 is the first and, to date, the only international legally binding instrument dealing with data protection. The Convention underwent a modernisation process
- Under EU law, data protection has been acknowledged as a distinct fundamental right. It is affirmed in Article 16 of the Treaty of the Functioning of the EU, as well as in Article 8 of the EU Charter of Fundamental Rights
- Under EU law, data protection was regulated for the first time by the Data Protection Directive in 1995
- In view of rapid technological developments, the EU adopted new legislation in 2016 to adapt data protection rules to the digital age. The General Data Protection Regulation became applicable in May 2018, repealing the Data Protection Directive

AI

DP (GDPR)

Big data

Minimisation, purpose
Limitation,
transparency, consent

Machine learning

Automated decision
making



AI v. non discrimination

Loomis v. Wisconsin

"Our analysis of Northpointe's tool, called COMPAS (which stands for Correctional Offender Management Profiling for Alternative Sanctions), found that black defendants were far more likely than white defendants to be incorrectly judged to be at a higher risk of recidivism, while white defendants were more likely than black defendants to be incorrectly flagged as low risk"

Pro Publica Analysis on COMPAS bias

Directive (EU) 2016/680

Article 11

Automated individual decision-making

Member States shall provide for a decision based solely on automated processing, including profiling, which produces an adverse legal effect concerning the data subject or significantly affects him or her, to be prohibited unless authorised by Union or Member State law to which the controller is subject and which provides appropriate safeguards for the rights and freedoms of the data subject, at least the right to obtain human intervention on the part of the controller.

Decisions referred to in paragraph 1 of this Article shall not be based on special categories of personal data referred to in Article 10, unless suitable measures to safeguard the data subject's rights and freedoms and legitimate interests are in place.

Profiling that results in discrimination against natural persons on the basis of special categories of personal data referred to in Article 10 shall be prohibited, in accordance with Union law

GDPR, Art. 22

The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her

WP29/EDPB opinion 251/2018

The term "right" in the provision does not mean that Article 22(1) applies only when actively invoked by the data subject. Article 22(1) establishes a general prohibition for decision-making based solely on automated processing. This prohibition applies whether or not the data subject takes an action regarding the processing of their personal data



Conclusions

The Law of the horse

Lots of cases deal with sales of horses; others deal with people kicked by horses; still more deal with the licensing and racing of horses, or with the care veterinarians give to horses, or with prizes at horse shows. Any effort to collect these strands into a course on 'The Law of the Horse' is doomed to be shallow and to miss unifying principles

Frank H. Easterbrook



There are no digital rights but
human rights in the digital era

Aran Balkan

Human rights impact
assessment?

States

Public consultations

External independent audits

Updating existing regulations

Ensuring a competitive field in the AI domain

Companies

TOS based on human rights principles

Transparency

Preventing discrimination

AI code fully auditable

Useful Links

The UN Special Rapporteur's Report: <https://freedex.org/wp-content/blogs.dir/2015/files/2018/10/AI-and-FOE-GA.pdf>

The Code of Conduct on hate speech: https://ec.europa.eu/info/sites/info/files/hatespeech_infographic3_web.pdf

The Loomis case: <https://harvardlawreview.org/2017/03/state-v-loomis/>

The ProPublica Analysis on COMPAS: <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>

WP 29 Guidelines on automated decision making: https://ec.europa.eu/newsroom/article29/document.cfm?action=display&doc_id=49826

Policy and investment recommendations for trustworthy AI: https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60343

The question of whether a computer can think is
no more interesting than the question of whether
a submarine can swim

Edsger W. Dijkstra

Thank you!

Monica A. Senor
@MASenor

ALL slides CC BY-NC-ND 4.0