POLITECNICO
MILANO 1863

Advanced Deep Learning Models and Methods

- Privacy Preserving Learning –

18th February 2022

Prof. Matteo Matteucci – *matteo.matteucci@polimi.it*
Alberto Archetti – *alberto.archetti@polito.it*
Eugenio Lomurno – *eugenio.lomurno@polimi.it*

«Me, Myself, and I»

What might go wrong?!?!?!

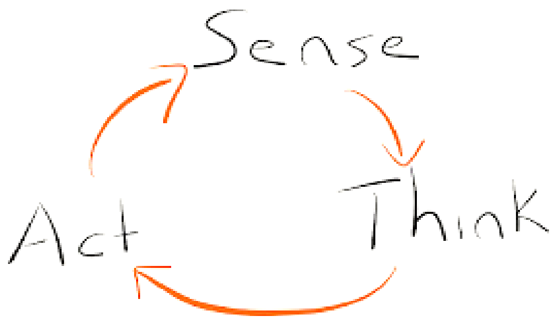
Matteo Matteucci, PhD
Full Professor
Dept. of Electronics, Information & Bioengineering
Politecnico di Milano
matteo.matteucci@polimi.it

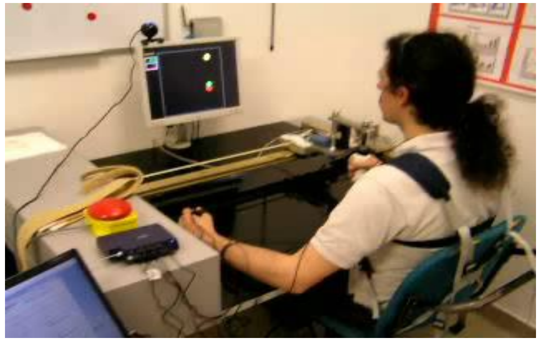
My research interests

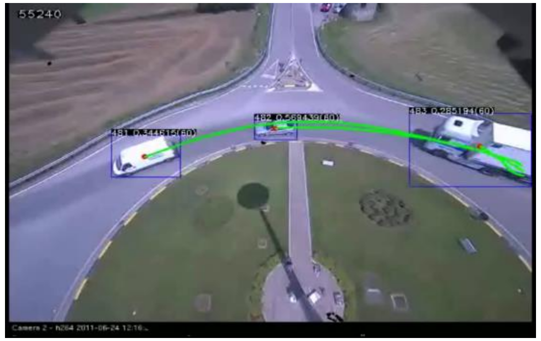
- Robotics & Autonomous Systems
- Machine Learning
- Pattern Recognition
- Computer Vision & Perception

Courses I teach

- Robotics (BS+MS)
- Machine Learning (MS)
- Deep Learning (MS+PhD)
- Cognitive Robotics (MS)

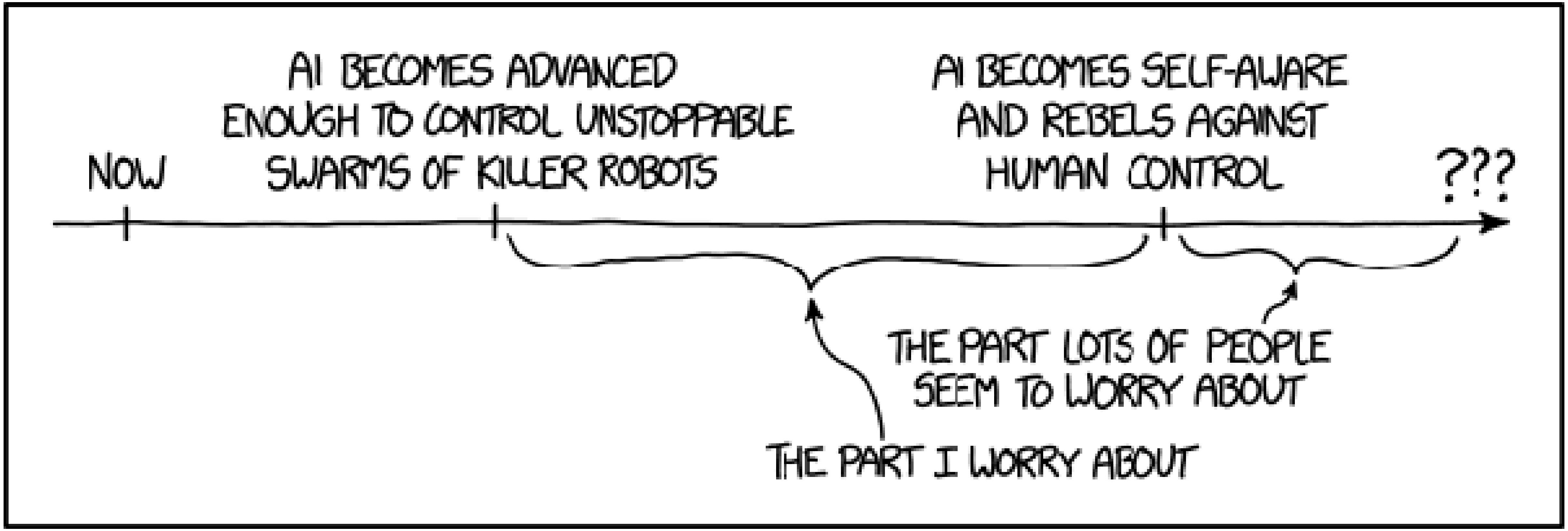






Enable physical and software autonomous systems to perceive, plan, and act without human intervention in the real world

What might go wrong ?!?!?!

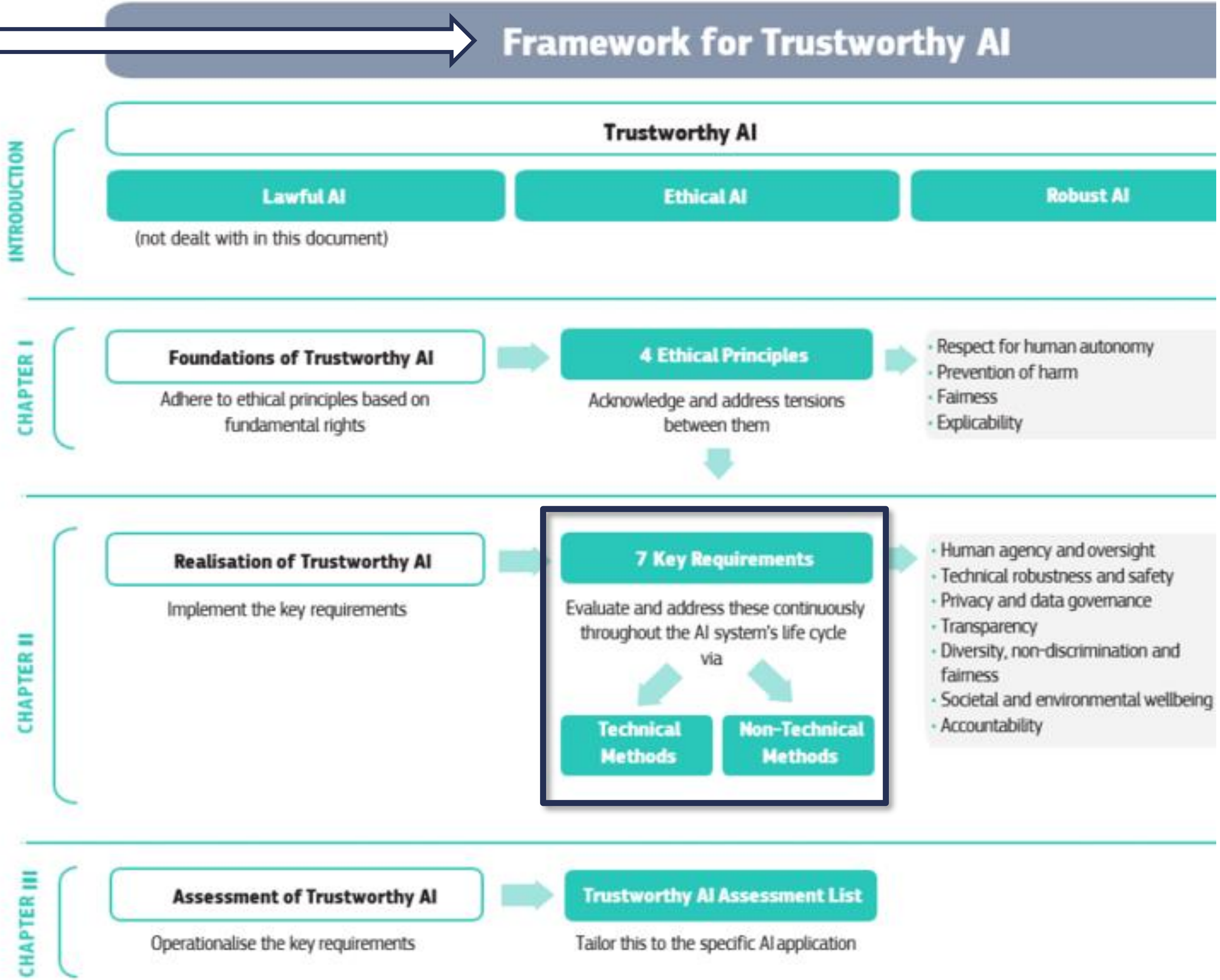

NOW
AI BECOMES ADVANCED ENOUGH TO CONTROL UNSTOPPABLE SWARMS OF KILLER ROBOTS
AI BECOMES SELF-AWARE AND REBELS AGAINST HUMAN CONTROL
???
THE PART LOTS OF PEOPLE SEEM TO WORRY ABOUT
THE PART I WORRY ABOUT

https://imgs.xkcd.com/comics/robot_future.png

EU worries too !!!

Independent High-Level Expert Group on Artificial Intelligence set up by the European Commission

Ethics Guidelines for Trustworthy AI



Framework for Trustworthy AI

Trustworthy AI

- Lawful AI (not dealt with in this document)
- Ethical AI
- Robust AI

Foundations of Trustworthy AI — Adhere to ethical principles based on fundamental rights → **4 Ethical Principles** — Acknowledge and address tensions between them →

- Respect for human autonomy
- Prevention of harm
- Fairness
- Explicability

Realisation of Trustworthy AI — Implement the key requirements → **7 Key Requirements** — Evaluate and address these continuously throughout the AI system's life cycle via **Technical Methods** / **Non-Technical Methods** →

- Human agency and oversight
- Technical robustness and safety
- Privacy and data governance
- Transparency
- Diversity, non-discrimination and fairness
- Societal and environmental wellbeing
- Accountability

Assessment of Trustworthy AI — Operationalise the key requirements → **Trustworthy AI Assessment List** — Tailor this to the specific AI application

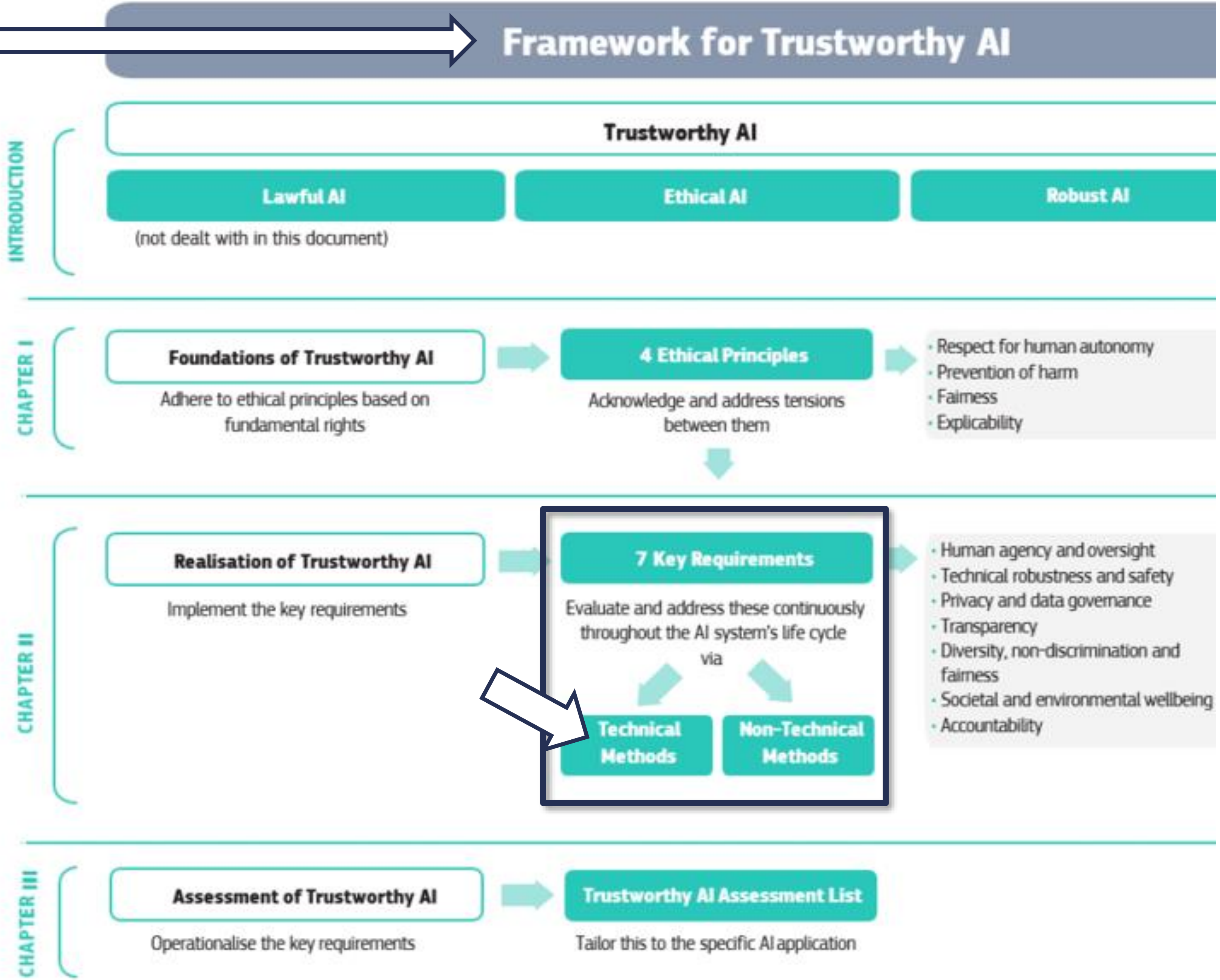
EU worries too !!!

7 key requirements for **ethical AI:**

- Human agency and oversight
- Technically robustness & safe
- Privacy and data governance
- Transparency
- Diversity, non-discrimination and fairness
- Societal and environmental wellbeing
- Accountable

Will your algorithms pass the test? Create AI humans can trust.

#AI #ArtificialIntelligence



Framework for Trustworthy AI

INTRODUCTION

Trustworthy AI

- Lawful AI (not dealt with in this document)
- Ethical AI
- Robust AI

CHAPTER I

Foundations of Trustworthy AI: Adhere to ethical principles based on fundamental rights → **4 Ethical Principles**: Acknowledge and address tensions between them →

- Respect for human autonomy
- Prevention of harm
- Fairness
- Explicability

CHAPTER II

Realisation of Trustworthy AI: Implement the key requirements → **7 Key Requirements**: Evaluate and address these continuously throughout the AI system's life cycle via Technical Methods, Non-Technical Methods →

- Human agency and oversight
- Technical robustness and safety
- Privacy and data governance
- Transparency
- Diversity, non-discrimination and fairness
- Societal and environmental wellbeing
- Accountability

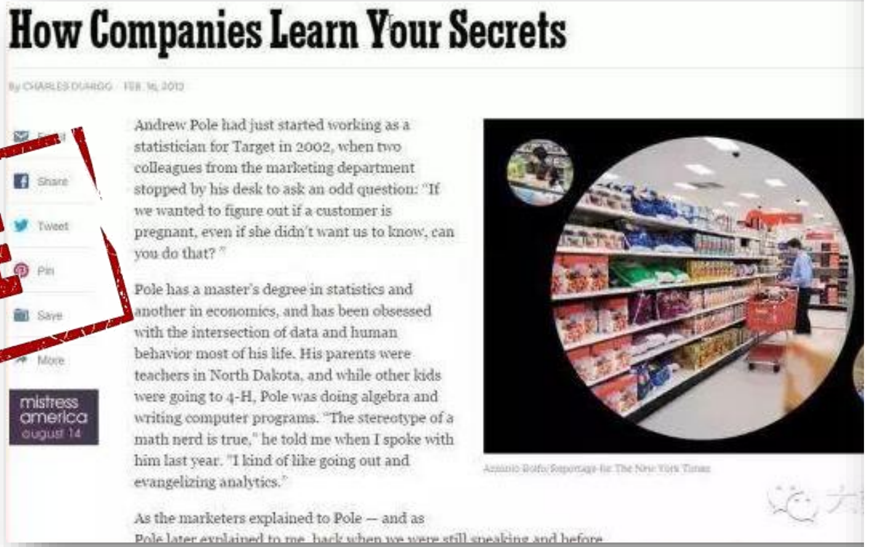
CHAPTER III

Assessment of Trustworthy AI: Operationalise the key requirements → **Trustworthy AI Assessment List**: Tailor this to the specific AI application

Why is this a big concern?

“The enormous data that companies feed into AI-driven algorithms are susceptible to data breaches as well.“

FAKE



How Companies Learn Your Secrets

By CHARLES DUHIGG FEB. 16, 2012

Email
Share
Tweet
Pin
Save
More

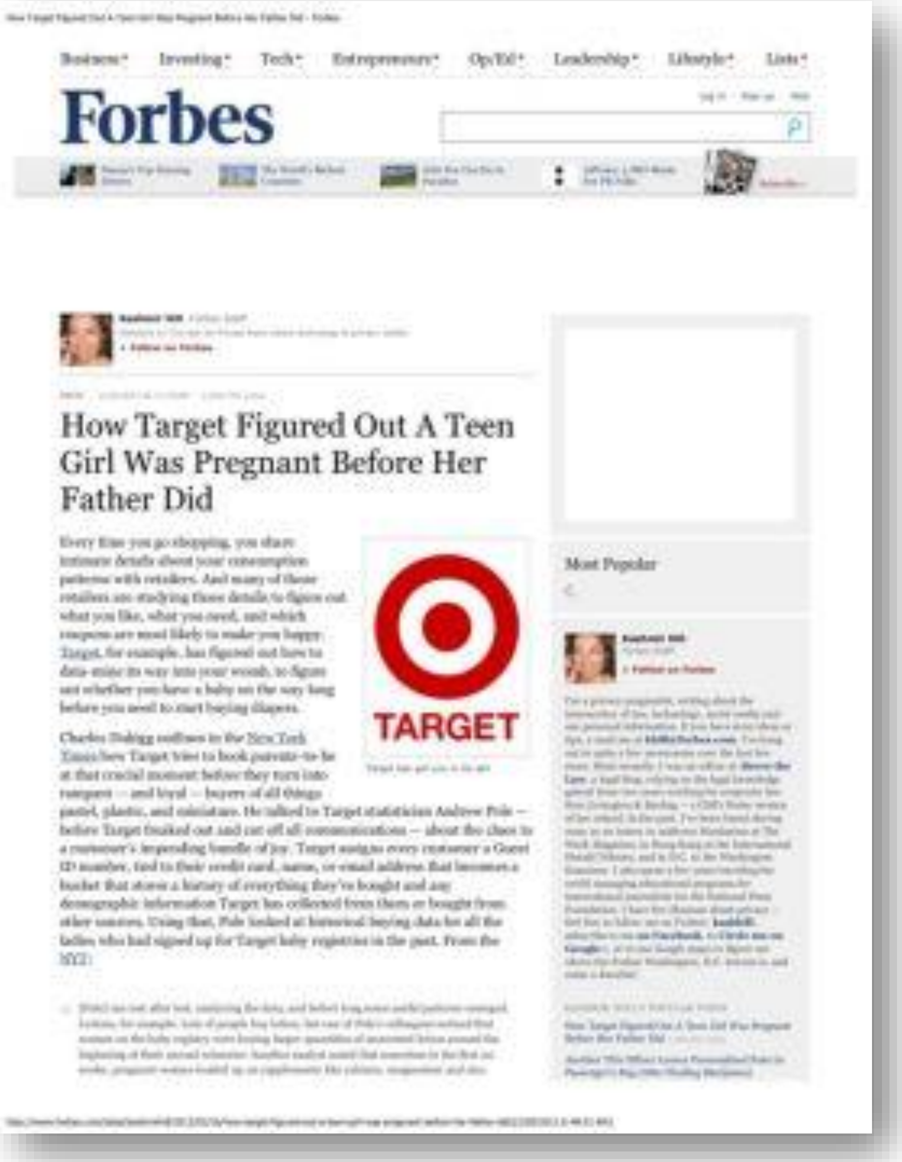
mistress america august 14

Andrew Pole had just started working as a statistician for Target in 2002, when two colleagues from the marketing department stopped by his desk to ask an odd question: “If we wanted to figure out if a customer is pregnant, even if she didn’t want us to know, can you do that? ”

Pole has a master’s degree in statistics and another in economics, and has been obsessed with the intersection of data and human behavior most of his life. His parents were teachers in North Dakota, and while other kids were going to 4-H, Pole was doing algebra and writing computer programs. “The stereotype of a math nerd is true,” he told me when I spoke with him last year. “I kind of like going out and evangelizing analytics.”

Antonio Bolfo/Reportage for The New York Times

As the marketers explained to Pole — and as Pole later explained to me, back when we were still speaking and before



Business▾ Investing▾ Tech▾ Entrepreneurs▾ Op/Ed▾ Leadership▾ Lifestyle▾ Lists▾

Forbes

How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did

Every time you go shopping, you share intimate details about your consumption patterns with retailers. And many of those retailers are studying those details to figure out what you like, what you need, and which coupons are most likely to make you happy. Target, for example, has figured out how to data-mine its way into your womb, to figure out whether you have a baby on the way long before you need to start buying diapers.

TARGET

Charles Duhigg outlines in the New York Times how Target tries to hook parents-to-be at that crucial moment before they turn into rampant — and loyal — buyers of all things pastel, plastic, and miniature. He talked to Target statistician Andrew Pole — before Target freaked out and cut off all communications — about the clues to a customer's impending bundle of joy. Target assigns every customer a Guest ID number, tied to their credit card, name, or email address that becomes a bucket that stores a history of everything they've bought and any demographic information Target has collected from them or bought from other sources. Using that, Pole looked at historical buying data for all the ladies who had signed up for Target baby registries in the past. From the NYT:

[illegible]

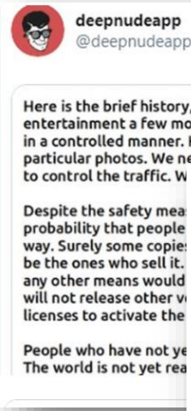
Most Popular

“AI may generate personal data […] created without the permission of the individual.”

https://thinkml.ai/is-artificial-intelligence-a-threat-to-privacy/

Why is this a big concern?

"The enormous data that companies feed into AI-driven algorithms are susceptible to data breaches as well."

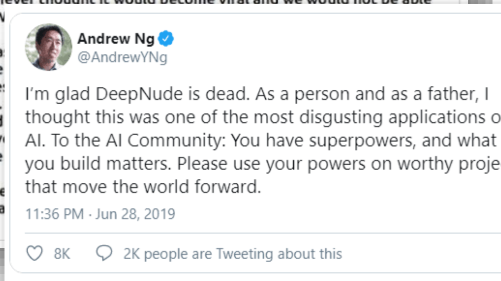


deepnudeapp
@deepnudeapp

Here is the brief history, and the end of DeepNude. We created this project for user's entertainment a few months ago. We thought we were selling a few sales every month in a controlled manner. Honestly, the app is not that great, it only works with particular photos. We never thought it would become viral and we would not be able to control the traffic. W

Despite the safety mea
probability that people
way. Surely some copie
be the ones who sell it.
any other means would
will not release other v
licenses to activate the

People who have not ye
The world is not yet rea

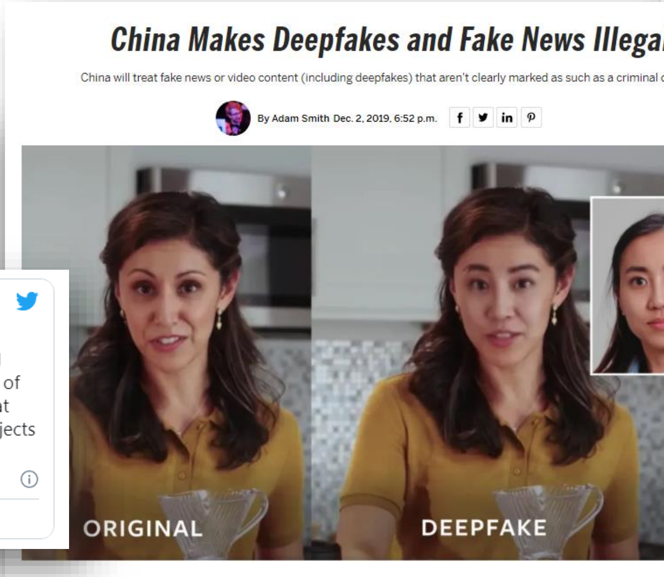


Andrew Ng
@AndrewYNg

I'm glad DeepNude is dead. As a person and as a father, I thought this was one of the most disgusting applications of AI. To the AI Community: You have superpowers, and what you build matters. Please use your powers on worthy projects that move the world forward.

11:36 PM · Jun 28, 2019

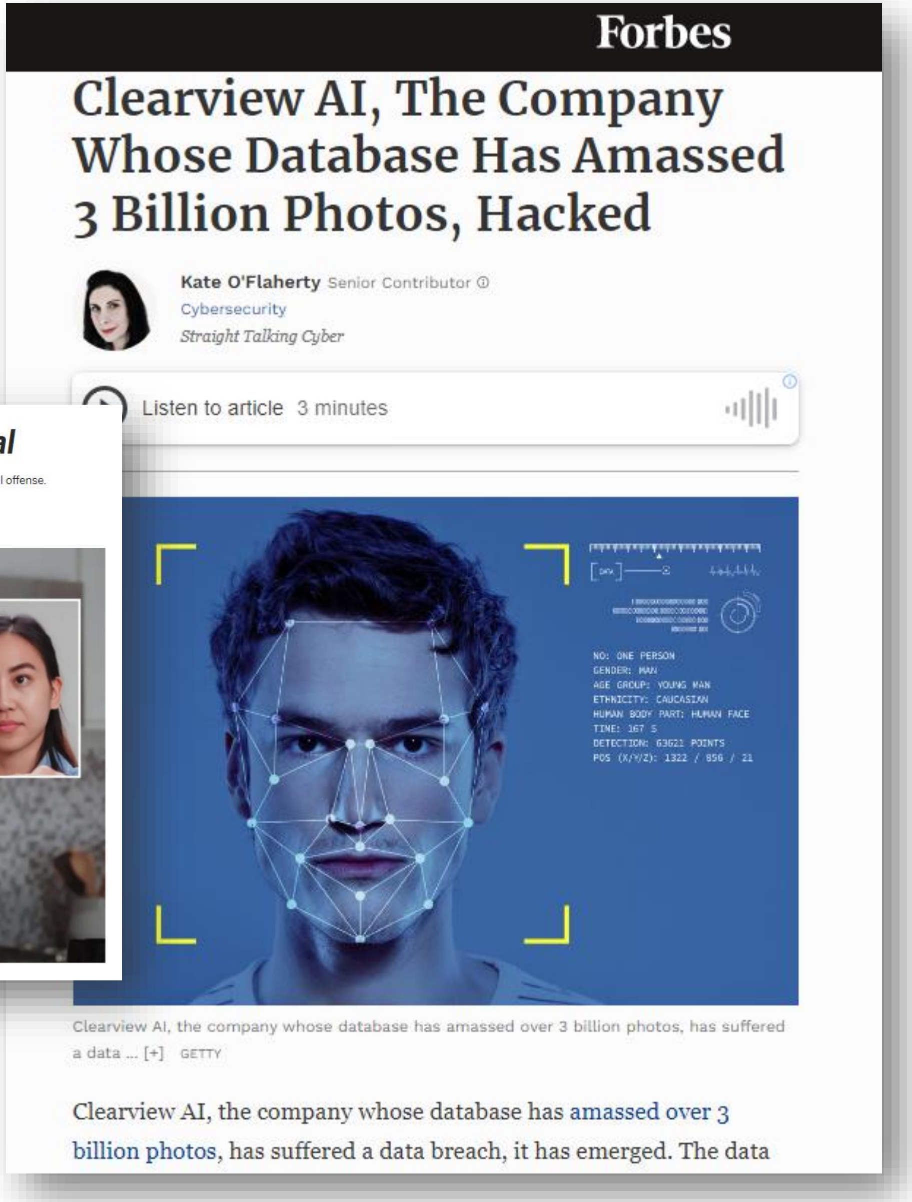
♡ 8K 2K people are Tweeting about this



China Makes Deepfakes and Fake News Illegal

China will treat fake news or video content (including deepfakes) that aren't clearly marked as such as a criminal offense.

By Adam Smith Dec. 2, 2019, 6:52 p.m.



Forbes

Clearview AI, The Company Whose Database Has Amassed 3 Billion Photos, Hacked

Kate O'Flaherty Senior Contributor
Cybersecurity
Straight Talking Cyber

Listen to article 3 minutes

Clearview AI, the company whose database has amassed over 3 billion photos, has suffered a data ... [+] GETTY

Clearview AI, the company whose database has amassed over 3 billion photos, has suffered a data breach, it has emerged. The data

"AI may generate personal data [...] created without the permission of the individual."

https://thinkml.ai/is-artificial-intelligence-a-threat-to-privacy/

The Social Impact of Artificial Intelligence and Data Privacy Issues by Shree Das, 08 September 2020

Why is this a big concern?

"Modern technologies like surveillance cameras, smartphones, and the internet have made our private data collection easier than ever."

Terms & Conditions
1 rules and requirements that one
agree to abide by in order to use a service; 2
special arrangements, provi
ments, rules, specifications, an
at form an integral part of
contract

The Economist
Theresa May v Brussels
Ten years on: banking after the crisis
South Korea's unfinished revolution
Biology, but without the cells
The world's most valuable resource
amazon
Google
Data and the new rules of competition

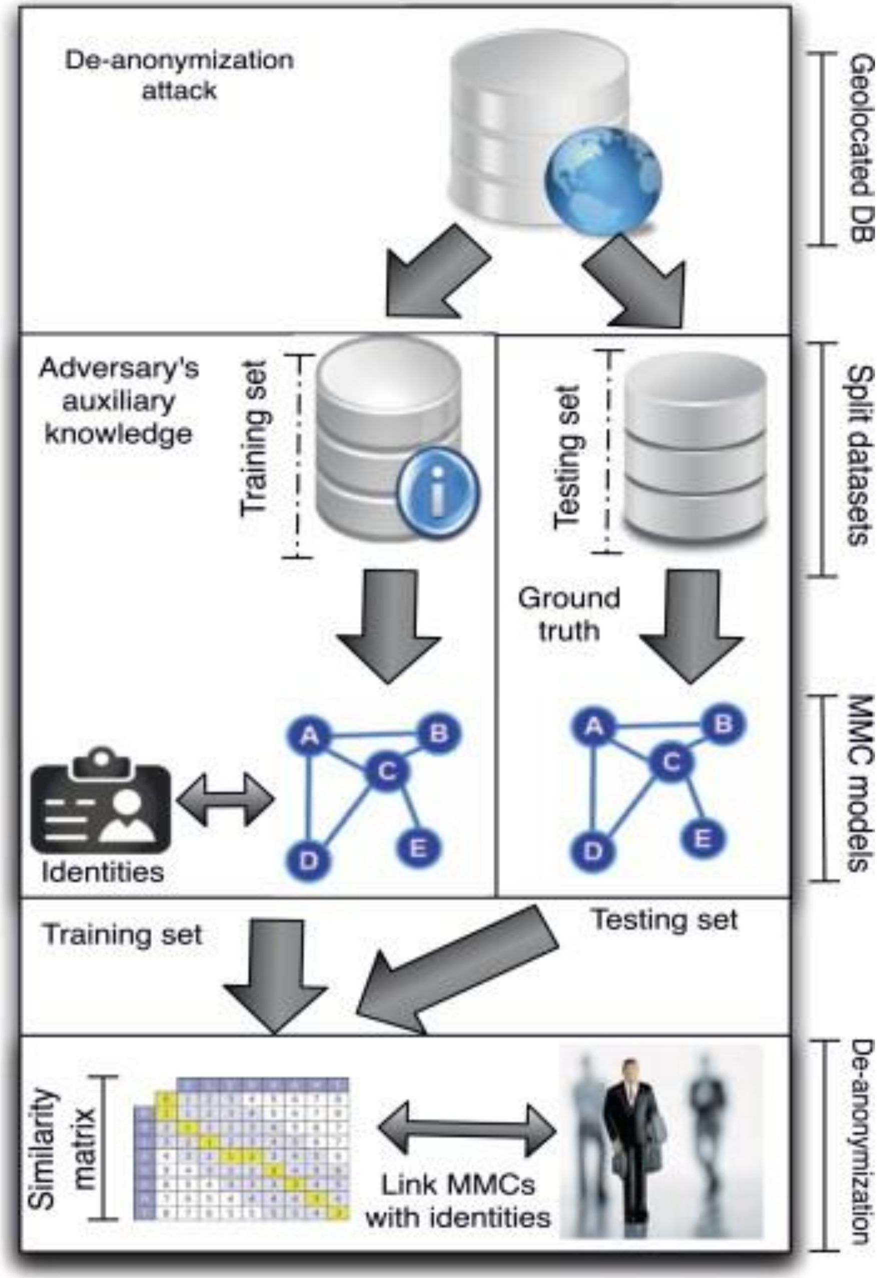
https://thinkml.ai/is-artificial-intelligence-a-threat-to-privacy/

Why is this a big concern?

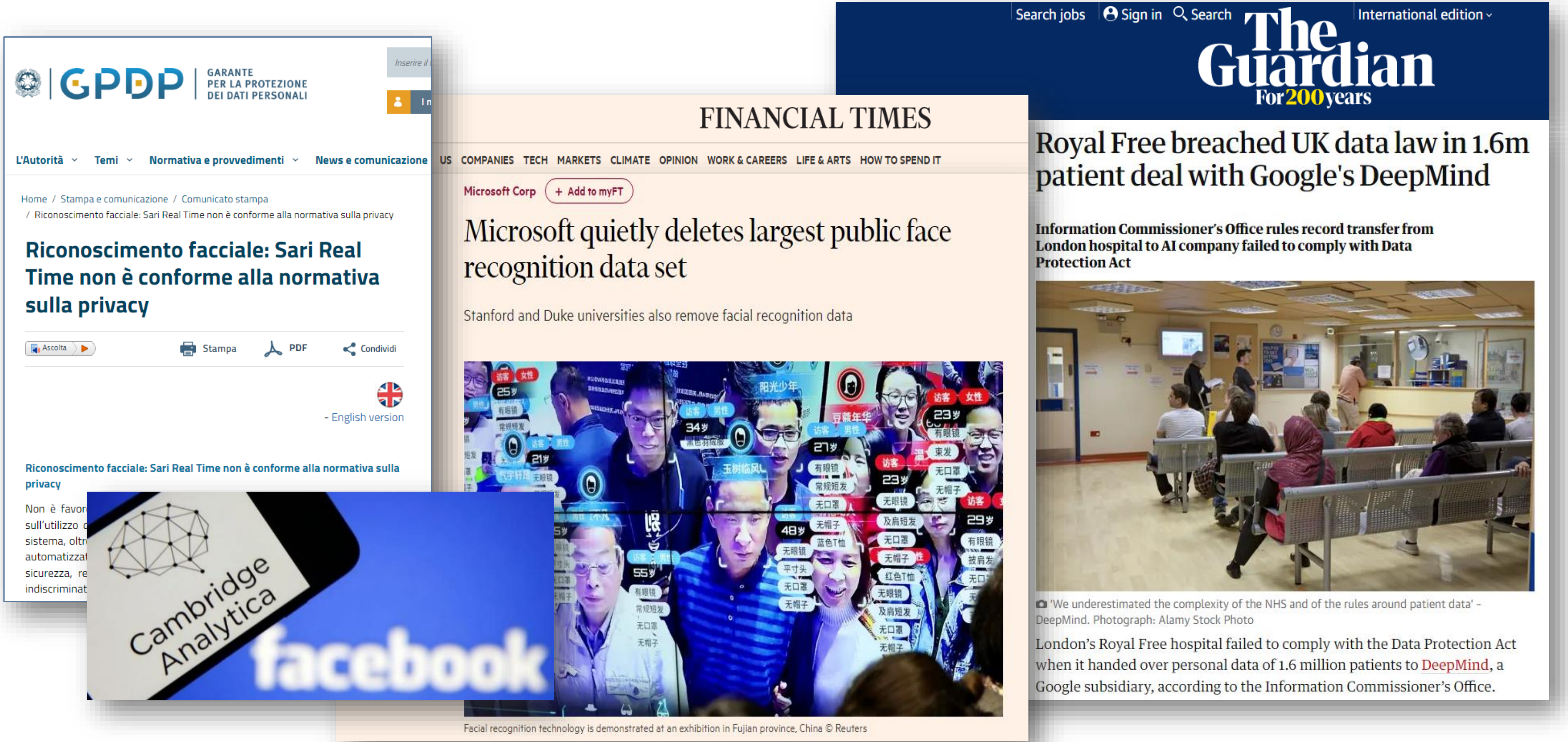
"Modern technologies like surveillance cameras, smartphones, and the internet have made our private data collection easier than ever."

Terms & Conditions
1 rules and requirements that one
agree to abide by in order to use a service; 2
special arrangements, provi
ents, rules, specifications, an
at form an integral part of
r contract

https://thinkml.ai/is-artificial-intelligence-a-threat-to-privacy/



Are you entitled to use those data?



GPDP | GARANTE PER LA PROTEZIONE DEI DATI PERSONALI

L'Autorità · Temi · Normativa e provvedimenti · News e comunicazione

Home / Stampa e comunicazione / Comunicato stampa / Riconoscimento facciale: Sari Real Time non è conforme alla normativa sulla privacy

Riconoscimento facciale: Sari Real Time non è conforme alla normativa sulla privacy

Ascolta · Stampa · PDF · Condividi

- English version

Riconoscimento facciale: Sari Real Time non è conforme alla normativa sulla privacy

Non è favor
sull'utilizzo c
sistema, oltr
automatizzat
sicurezza, re
indiscriminat

FINANCIAL TIMES

US · COMPANIES · TECH · MARKETS · CLIMATE · OPINION · WORK & CAREERS · LIFE & ARTS · HOW TO SPEND IT

Microsoft Corp · + Add to myFT

Microsoft quietly deletes largest public face recognition data set

Stanford and Duke universities also remove facial recognition data

Facial recognition technology is demonstrated at an exhibition in Fujian province, China © Reuters

Search jobs · Sign in · Search · International edition

The Guardian For 200 years

Royal Free breached UK data law in 1.6m patient deal with Google's DeepMind

Information Commissioner's Office rules record transfer from London hospital to AI company failed to comply with Data Protection Act

'We underestimated the complexity of the NHS and of the rules around patient data' - DeepMind. Photograph: Alamy Stock Photo

London's Royal Free hospital failed to comply with the Data Protection Act when it handed over personal data of 1.6 million patients to DeepMind, a Google subsidiary, according to the Information Commissioner's Office.

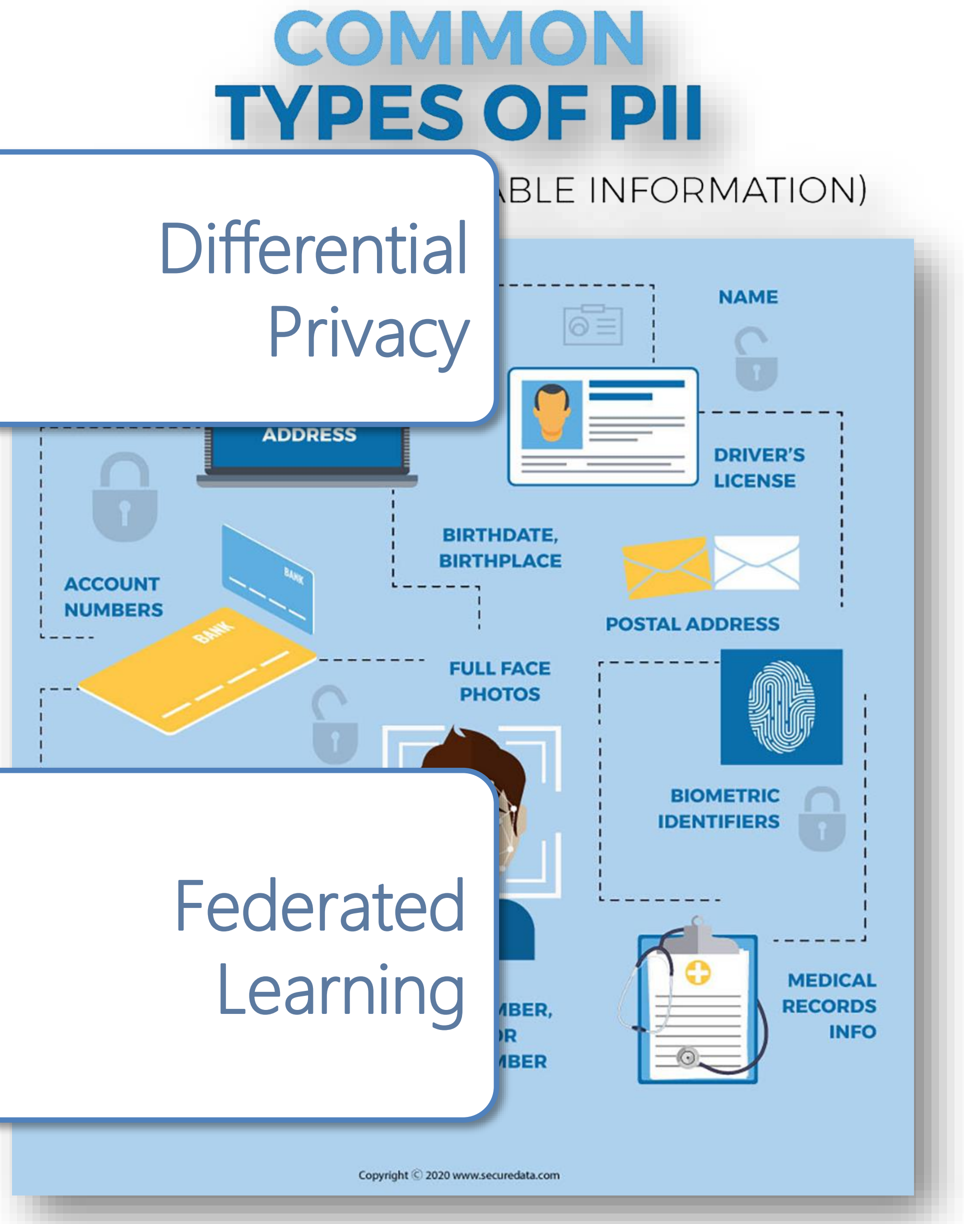
The European approach to trustworthy AI

- Privacy and data protection. AI systems must guarantee privacy and data protection throughout a system's entire lifecycle.

Differential Privacy

- Quality and integrity of data. The quality of the data sets used is paramount […] it may contain socially constructed biases, inaccuracies, errors and mistakes.
- Access to data. In any given organization that handles individuals' data […] data protocols governing […] who can access data and under which circumstances.

Federated Learning



https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines/1.html#privacy

Next on Stage …

13:45 – 15:15: Federated Learning

Alberto Archetti (alberto.archetti@polito.it)

PhD Candidate, Politecnico di Milano

15:30 – 17:00: Differential Privacy

Eugenio Lomurno (eugenio.lomurno@polimi.it)

PhD Candidate, Politecnico di Milano