When It Is Fundamentally Dubious

Stephen Downes

November 2, 2021
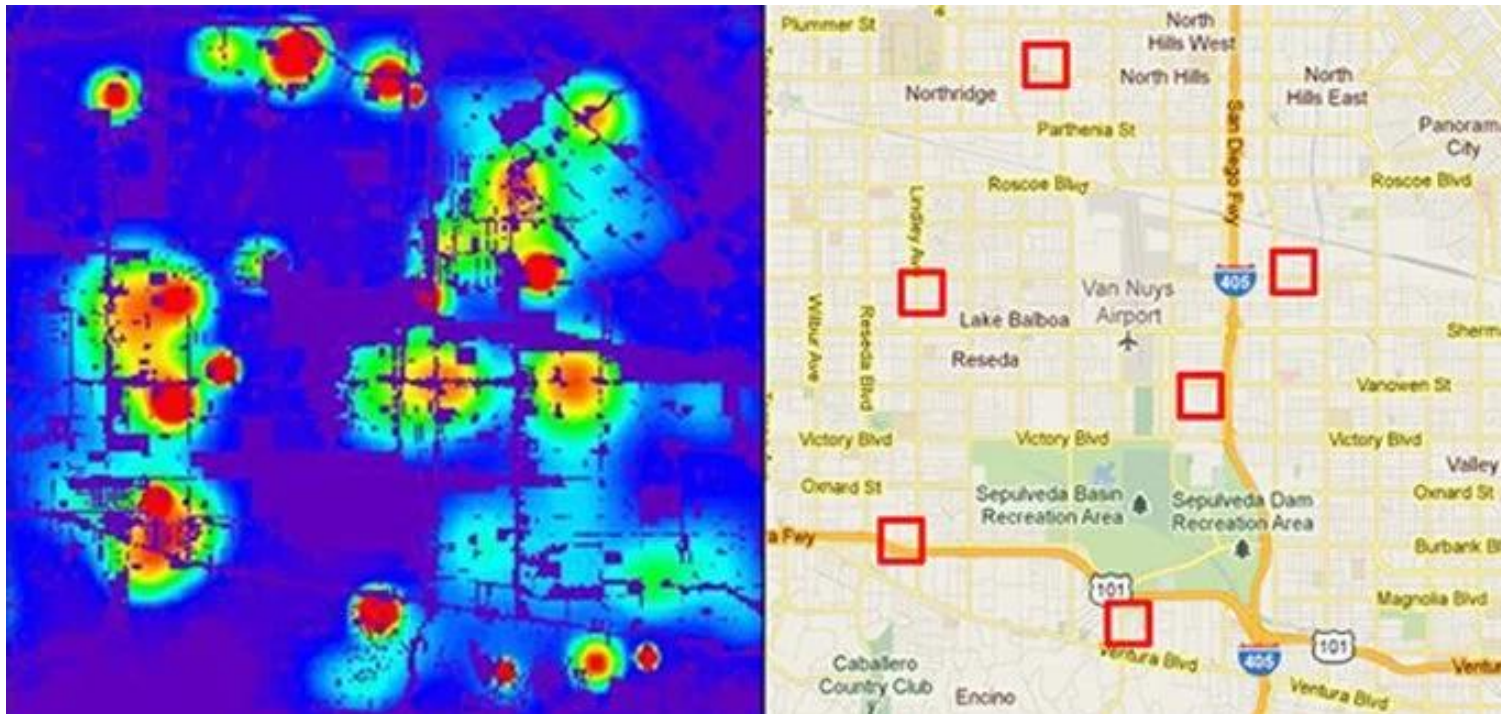
When It Is Fundamentally Dubious

- Some uses of AI and analytics are fundamentally dubious

- In such cases, the ethical question of such use *at all* arises

Incomplete & crude but useful breakdown

Genuine, rapid progress	Imperfect but improving	Fundamentally dubious
Shazam, reverse img search	Spam detection	Predicting recidivism
Face recognition	Copyright violation	Predicting job success
Med. diagnosis from scans	Automated essay grading	Predictive policing
Speech to text	Hate speech detection	Predicting terrorist risk
Deepfakes	Content recommendation	Predicting at-risk kids
Perception	Automating judgment	Predicting social outcomes

https://www.cs.princeton.edu/~arvindn/talks/MIT-STS-AI-snakeoil.pdf

Predicting Criminals

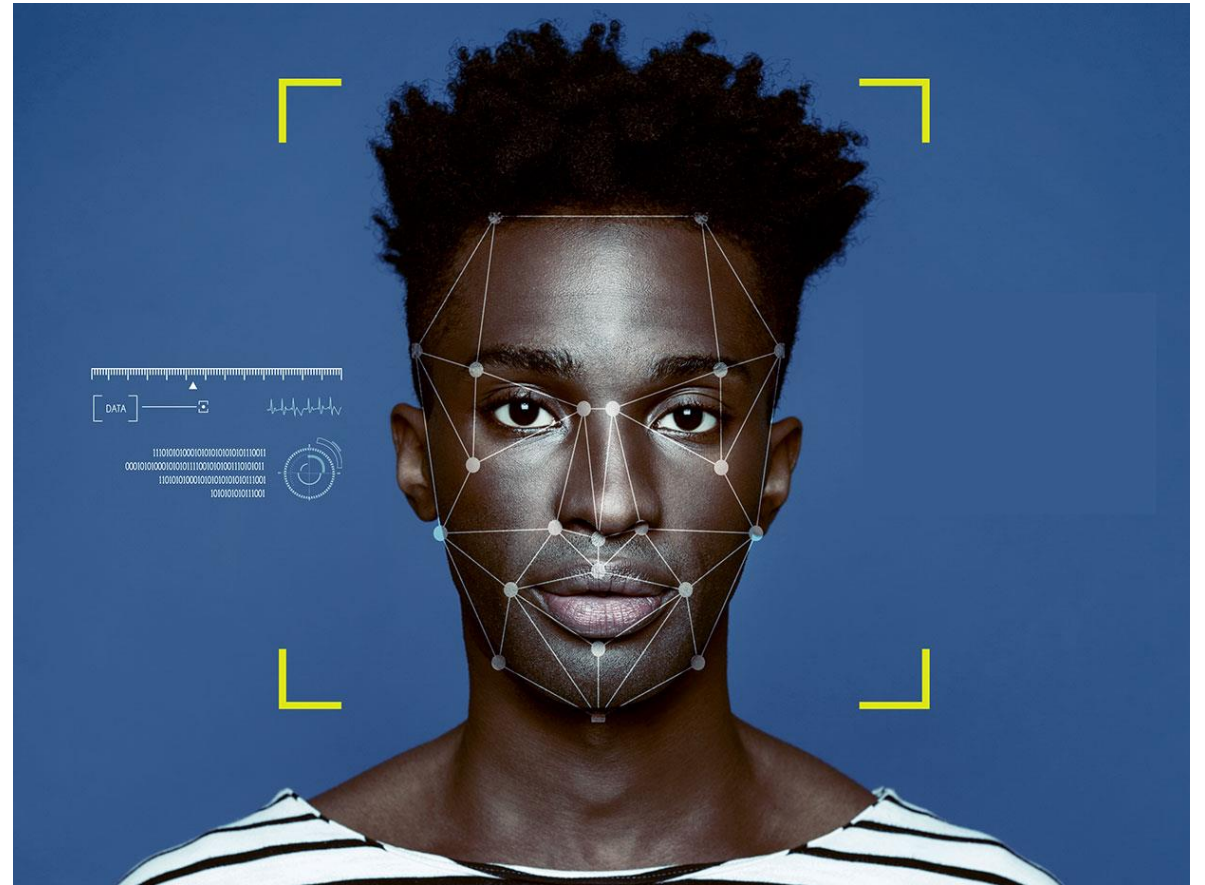"…puts students and families at risk of being unfairly targeted before they've done anything wrong"

https://www.smithsonianmag.com/innovation/artificial-intelligence-is-now-used-predict-crime-is-it-biased-180968337/ "Historical data from police practices, critics contend, can create a feedback loop through which algorithms make decisions that both reflect and reinforce attitudes about which neighborhoods are 'bad' and which are 'good.'"

Racial Profiling

It is arguable that there is no ethical application of analytics to identify specific races for special treatment.

https://spectrum.ieee.org/do-you-have-the-right-complexion-for-facial-recognition#toggle-gdpr
Biased facial recognition system disproportionately labels minority UCLA students and faculty as criminals

Identity Graphs

https://www.data-axle.com/resources/blog/solving-the-identity-graph-gap-through-robust-identity-resolution/
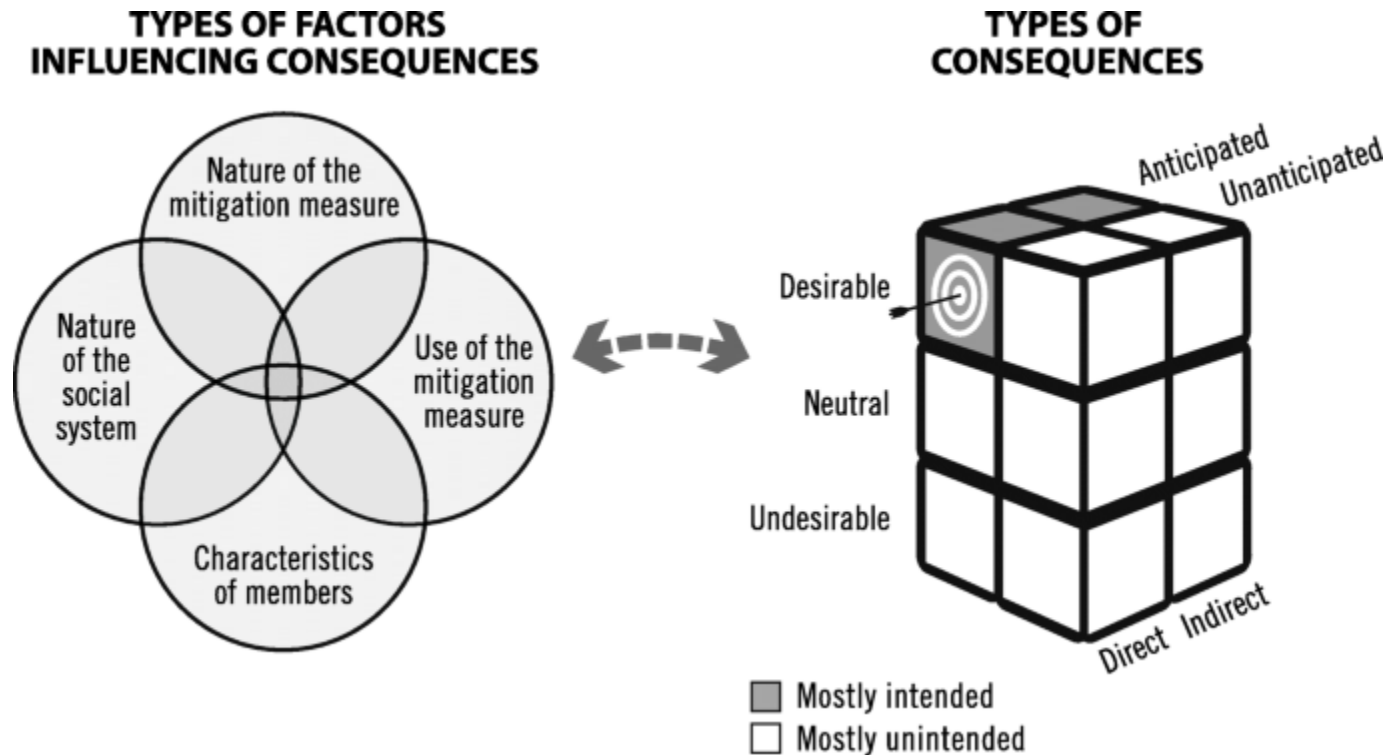
Autonomous Weapons

Examples:

- Armed robot 'dogs' being used as security guards

- Autonomous drones being used in armed conflict

https://www.thedefensepost.com/2021/10/13/robot-dog-assault-rifle/

https://www.popularmechanics.com/military/weapons/a36559508/drones-autonomously-attacked-humans-libya-united-nations-report/

When We Don't Know What the Consequences Will Be

TYPES OF FACTORS INFLUENCING CONSEQUENCES

Nature of the mitigation measure

Nature of the social system

Use of the mitigation measure

Characteristics of members

TYPES OF CONSEQUENCES

Anticipated   Unanticipated

Desirable

Neutral

Undesirable

Direct   Indirect

■ Mostly intended
□ Mostly unintended

For example, the suggestion that colleges should put smart speakers in student dormitories – "we simply have no idea what long-term effect having conversations recorded and kept by Amazon might have on their futures."