

Auditing Machine Learning Algorithms

Jan Roar Beckstrom
Chief Data Scientist, The Innovation Lab
Office of the Auditor General of Norway

A European approach to Artificial intelligence

The European Commission's approach to AI centres on excellence and trust. It aims to boost research and industrial capacity and ensure fundamental rights.

First: AI is a good thing

- Can contribute to large-scale efficiency improvements in the public sector
- Can enable significantly better public services
- Not inconceivable that e.g. “case processing time” will become an almost irrelevant concept

Then: There are major risks using AI...

Automating poverty

A series exploring how our governments use AI to target the vulnerable

Digital dystopia / How algorithms punish the poor



Comment Benefits of 'welfare robots' and the need for human oversight

UN Big tech allowed to target and surveil the poor creating 'digital welfare state'

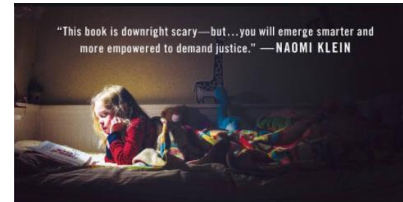


US / Zombie debts are hounding struggling Americans. Will you be next?



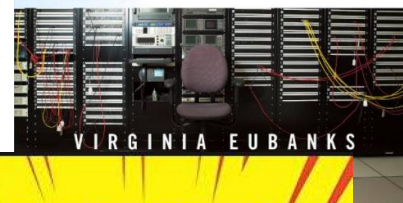
Computer says no / The people trapped in universal credit's 'black hole'

How Bristol assesses citizens' risk of harm - using an algorithm

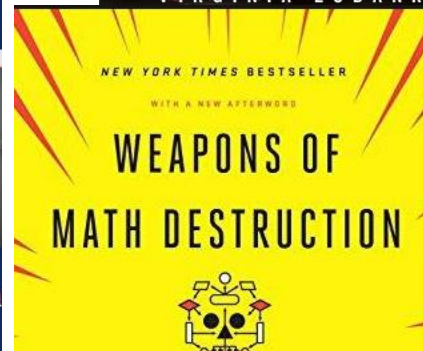


AUTOMATING INEQUALITY

HOW HIGH-TECH TOOLS PROFILE, POLICE, AND PUNISH THE POOR



VIRGINIA EUBANKS



NEW YORK TIMES BESTSELLER

WITH A NEW AFTERWORD

WEAPONS OF MATH DESTRUCTION

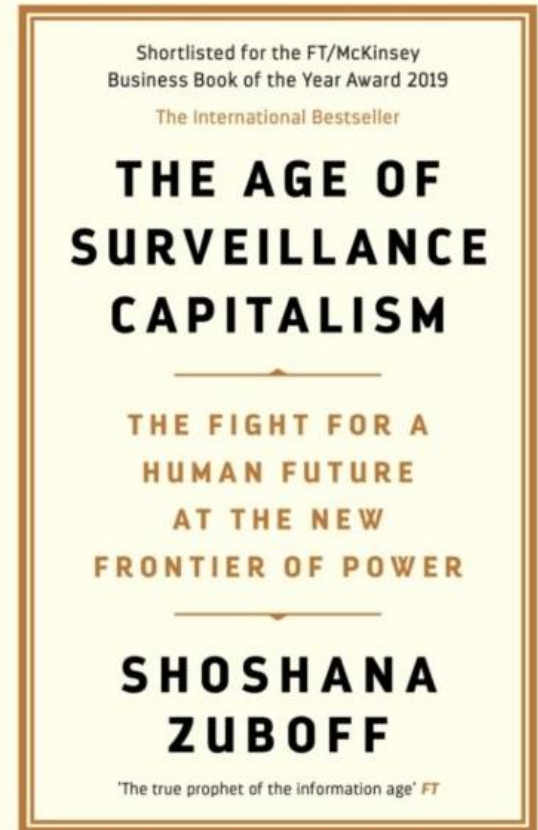


HOW BIG DATA INCREASES INEQUALITY AND THREATENS DEMOCRACY

CATHY O'NEIL

"This is a manual for the 21st-century citizen, and it succeeds where other big data accounts have failed—it is accessible, refreshingly critical, and feels relevant and urgent."

—FINANCIAL TIMES



Shortlisted for the FT/McKinsey Business Book of the Year Award 2019

The International Bestseller

THE AGE OF SURVEILLANCE CAPITALISM

THE FIGHT FOR A HUMAN FUTURE AT THE NEW FRONTIER OF POWER

SHOSHANA ZUBOFF

'The true prophet of the information age' FT

Around the world



India / How a glitch in biometric welfare system can be lethal



Australia / The automated system leaving welfare recipients cut off with nowhere to turn



UK / Benefits system automation could plunge claimants deeper into poverty



The Guardian view / On automating poverty: OK computers?



Pasco's sheriff created a futuristic program to stop crime before it happens.

It monitors and harasses families across the county.

[nature](#) > [news](#) > [article](#)NEWS | 24 October 2019 | [Update 26 October 2019](#)

Millions of black people affected by racial bias in health-care algorithms

Study reveals rampant racism in decision-making software used by US hospitals – and highlights ways to correct it.

Heidi Ledford



You have full access to this article via your institution.

Download PDF



Related Articles

A fairer way forward for AI in health care



Bias detectives: the researchers striving to make algorithms fair



Why audit algorithms?

Problematic issues with ML models – two examples

* **Black box:** Machine learning algorithms usually operate as black boxes and it is unclear how they derived a certain decision

* **Bias:** By default, machine learning models pick up biases from the training data. This can turn your machine learning models into racists.

Source:

<https://christophm.github.io/interpretable-ml-book/>

ISSAI 3000

The International Standard for Performance Auditing

Quote:

“Performance auditing aims to contribute to improved economy, efficiency and effectiveness in the public sector.

It also aims to contribute to good governance, accountability and transparency.”

In Europe we have the General Data Protection Regulation (GDPR):

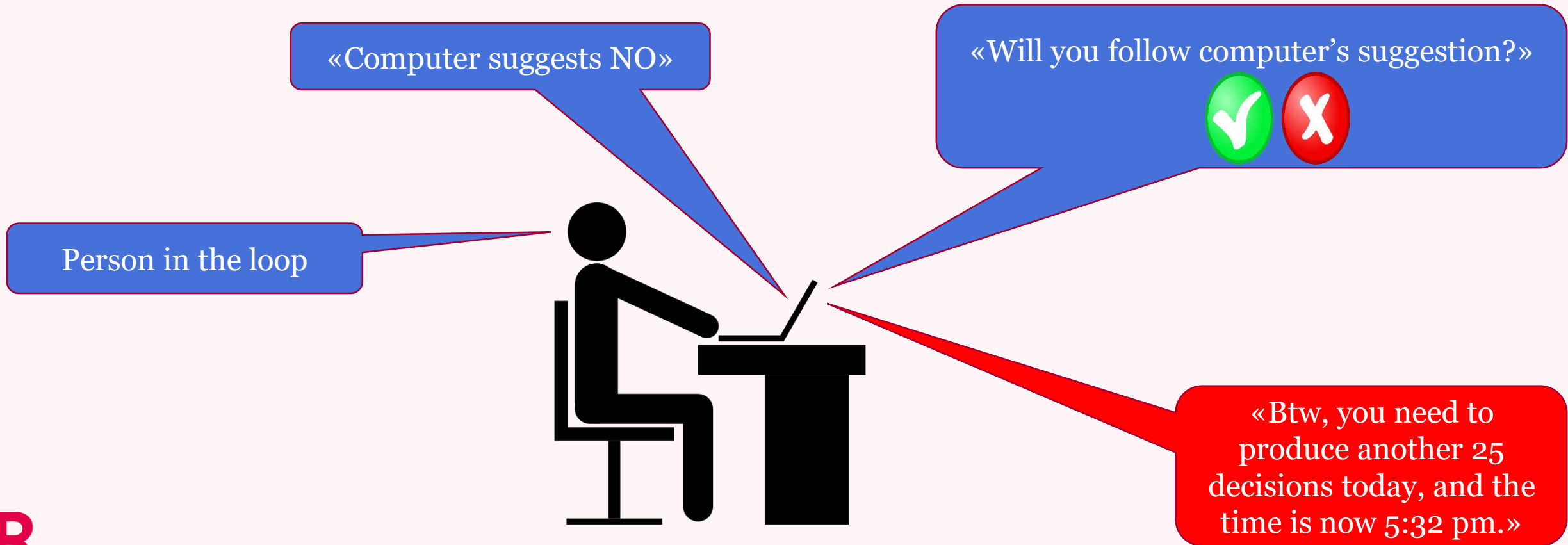
GDPR Article 22 “Automated individual decision-making, including profiling“:

“The data subject shall have the right not to be subject to a decision based solely on automated processing”

In short: There should be **“a person in the loop”**

But...

(Automated, dystopian case work...)

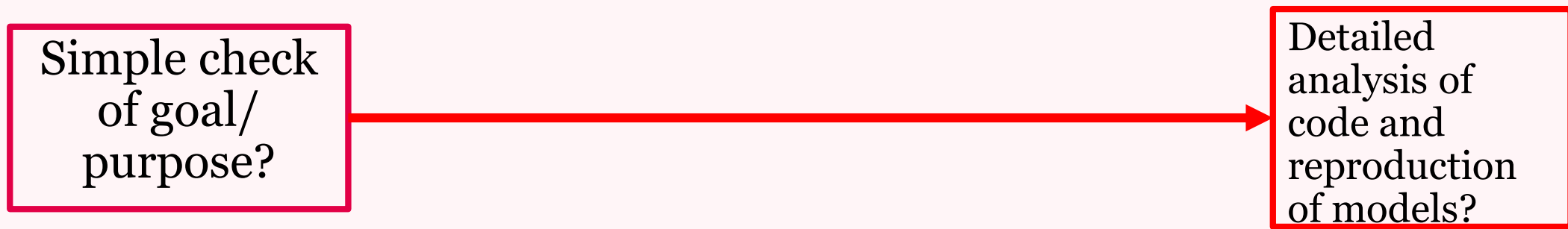


A note of caution:

It may still be the case that machines can be made less biased than humans?

“We are still using these algorithms called humans that are really biased. We’ve tested them and known that they’re horrible, but we still use them to make really important decisions every day.” (Rayid Ghani, computer scientist, Carnegie Mellon University)

Then, how to audit ML?



But where to start?

www.auditingalgorithms.net

- A guide written by auditors (us), for auditors (you)
- An international collaboration (SAIs of Germany, UK, Netherlands, Finland, Norway)
- Includes an “auditability checklist” and a simple “audit helper tool” in Excel
- Largely non-technical

An interesting question:

To what extent should auditors be «democracy activists»?

What if an ML-system reproduces existing inequalities present in society?

Shall we limit ourselves to «technical» aspects of fairness and bias?

Or should we «contribute in changing the world if the world is unfair»?

See also «*Bias Preservation in Machine Learning: The Legality of Fairness Metrics Under EU Non-Discrimination Law*»

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3792772

My view:

“Now, it is not clear how far the responsibility of the auditor reaches in protecting democracy against unwise use of AI and algorithms. This is uncharted territory for the profession.

Still, I believe we have a very clear role in contributing to AI being implemented in the public sector in ways that are **fair, just, explainable and transparent.**”

To conclude:

Auditors are indispensable
as governments start using
machine learning & Ai