

Ethics Washing in AI

Moshe Y. Vardi

Rice University

vardi@cs.rice.edu



Background

AI: Early Optimism

1958

H. A. Simon and A. Newell:

"within ten years a digital computer will be the world's chess champion"

1967

M. Minsky:

"Within a generation ... the problem of creating 'artificial intelligence' will substantially be solved."

"AI Winters"

**First AI Winter:
1974-1980**

Slow progress and dearth
of research funding

**Second AI Winter:
1987-1993**

The "Japanese Fifth-
Generation bust" and
dearth of research funding

AI Breakthroughs

A horizontal timeline with three circular markers at the top labeled 1997, 2011, and 2016. Below each marker is a rectangular box containing text. The timeline is represented by a horizontal line with dots at the ends and vertical lines connecting the markers to their respective boxes.

1997

IBM's Deep Blue beats Kasparov

2011

IBM's Watson defeats the two greatest Jeopardy! champions, Brad Rutter and Ken Jennings, by a significant margin.

2016

AlphaGo beats Lee Se-dol to take the DeepMind Challenge series!

AI Breakthroughs, I

1997

IBM's Deep Blue beats Kasparov



by a significant margin.

AI Breakthroughs, II

1997

IBM's Deep Blue beats Kasparov



by a significant margin.

AI Breakthroughs, III



2016

AlphaGo beats Lee
Se-dol to take the
DeepMind
Challenge series!

AI Develops Intuition

“AlphaGo combines tree-search techniques with search-space-reduction techniques that use machine learning.”

- In other words, AlphaGo develops “intuition” for Go playing.

Machine learning provides a response to Polanyi’s Paradox (1966):

"We can know more than we can tell... The skill of a driver cannot be replaced by a thorough schooling in the theory of the motorcar"

A Transportation Revolution?



Tesla Car Hits a Concrete Barrier

Collision between an SUV operating with Partial Driving Automation and a Crash Attenuator in Mountain View, California, March 2018

The US National Transportation Safety Board (NTSB) said the car had been driving semi-autonomously using Tesla's Autopilot software.

Feb. 2020: NTSB said more crashes were foreseeable if Tesla did not implement changes to its Autopilot system.

Uber AV Hits a Pedestrian

- Uber AV, with a safety driver on board, killed a pedestrian in Tempeh, AZ, March 2018.
- NTSB:** Uber did not have a formal safety plan in place.
- NTSB:** Uber had previously disabled the AV's factory-set automatic emergency braking system in order to prevent erratic driving.
- Sept. 2020:** Safety driver charged with homicide. Uber will not face charges.

Machine Bias

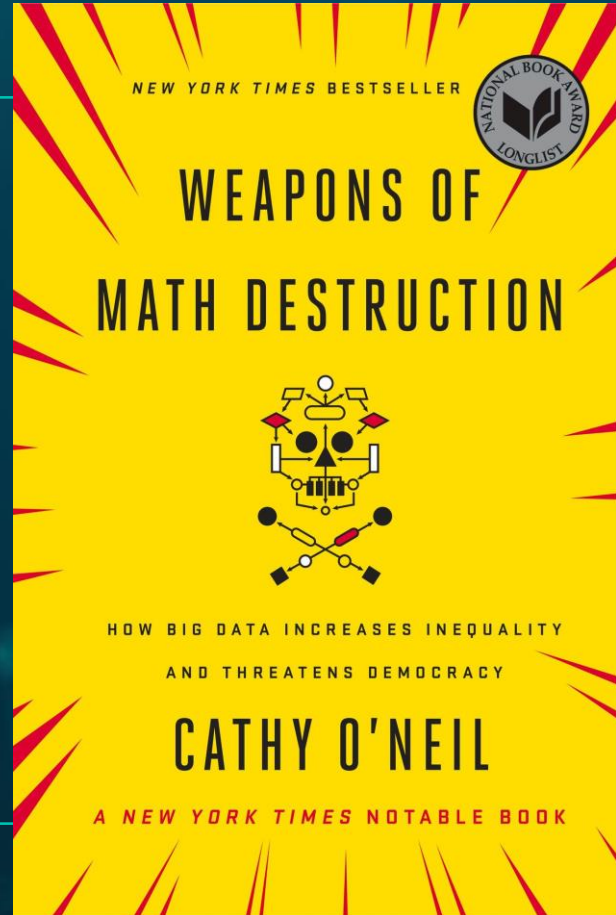
ProPublica, May 2016: There's software used across the country to predict future criminals. And it's biased against blacks!



Weapons of Math Destruction, '17

“How Big Data Increases Inequality and Threatens Democracy”

NYT: “A frightening look at how algorithms are increasingly regulating people”



The “Techlash”

Peggy Noonan, WSJ, Oct. 2017:

“Because all of their personal and financial information got hacked in the latest breach, because our country’s real overlords are in Silicon Valley and appear to be moral Martians who operate on some weird new postmodern ethical wavelength. And they’ll be the ones programming the robots that’ll soon take all the jobs!”



Ethics

An Ethics Crisis!

NYTimes, Feb. 2018

“Tech’s Ethical ‘Dark Side’: Harvard, Stanford and Others Want to Address It.”

Boston Globe, March 2018

“Computer science faces an ethics crisis. The Cambridge Analytica scandal proves it!”

NYTimes, Oct. 2018

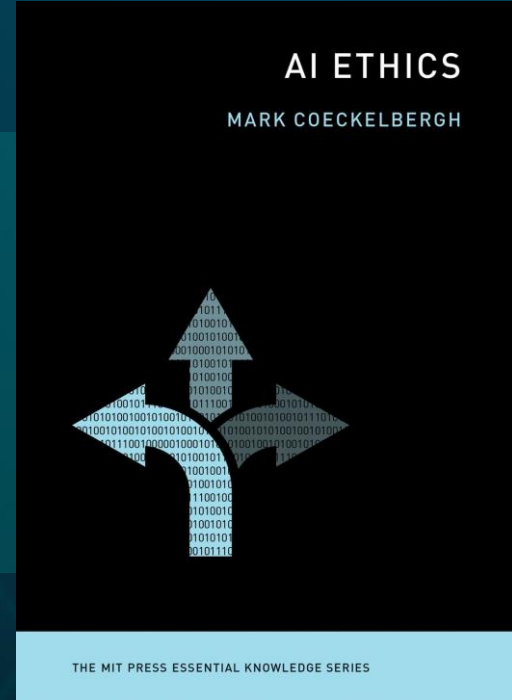
“Some think **chief ethics officers** could help technology companies navigate political and social questions.”

VentureBeat, Jan. 2019

“Facebook backs Institute for Ethics in Artificial Intelligence with \$7.5 million.”

AI Ethics - April 2020

“Artificial intelligence powers Google's search engine, enables Facebook to target advertising, and allows Alexa and Siri to do their jobs. AI is also behind self-driving cars, predictive policing, and autonomous weapons that can kill without human intervention. These and other AI applications raise complex ethical issues that are the subject of ongoing debate.”



Rome Call for AI Ethics

Rome, Feb. 28, 2020: The Pontifical Academy for Life, Microsoft, IBM, FAO, and the Italian Government, signed as first the “Call for an AI Ethics”.

“The values and principles that we are able to instill in AI will help to establish a framework that regulates and acts as a point of reference for digital ethics.”

Google's Responsible AI

Responsible AI Practices:

“The development of AI is creating new opportunities to improve the lives of people around the world, from business to healthcare to education. It is also raising new questions about the best way to build fairness, interpretability, privacy, and security into these systems.”

Institute for Ethics in AI

Jan. 20, 2019: “Facebook and the Technical University of Munich Announce a New Independent TUM Institute for Ethics in Artificial Intelligence.”

“At Facebook, ensuring the responsible and thoughtful use of AI is foundational to everything we do — from the data labels we use, to the individual algorithms we build, to the systems they are a part of.”

Ethics as an Escape from Regulation

Ben Wagner, 2018: *Ethics as an escape from regulation.
From “**ethics washing**” to ethics-shopping?*

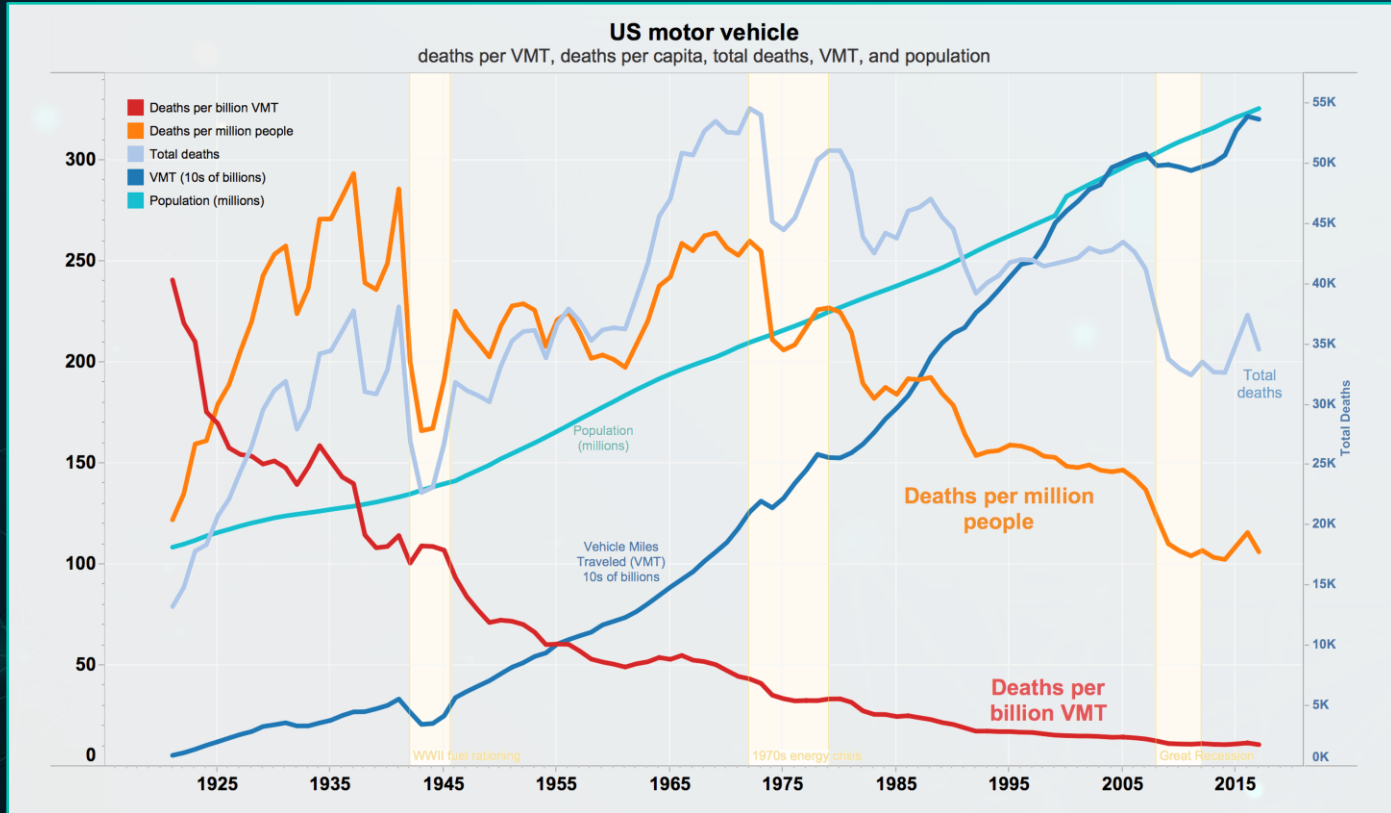
“Much of the debate about ethics seems to provide an **easy alternative** to government regulation.”

The Automobile



1908: *Ford Model T*

US Motor Vehicle Death



Traffic Safety Measures

- Crash avoidance, e.g., mirrors
- Driver assistance, e.g., anti-lock brakes
- Crashworthiness, e.g., airbags
- Pedestrian safety, e.g., crosswalks
- Car-to-car safety, e.g., traffic lights
- Laws, e.g., DWI, safety training

NOT: ethics training for drivers!

YES: public policy!

ML and the Justice System, I



Shirley A. Jackson, President, RPI, April 2017:

“Can you foresee a day when smart machines, driven with artificial intelligences, will assist with courtroom fact-finding or, more controversially even, judicial decision-making?”

John G. Roberts, Chief Justice:

“It’s a day that’s here, and it’s putting a significant strain on how the judiciary goes about doing things.”

NY Times: “An Algorithm that Grants Freedom or Takes it Away.”



ML and the Justice System, II

Many applications:

- Bail
- Sentencing
- Parole
- Separating children from parents
- ...

Proprietary Algorithms



Advancing Justice.



Embracing Community.

"The New Jim Crow"

Michelle Alexander, NYT, 11/'18:

Recent criminal justice reforms contain the seeds of a frightening system of "e-carceration."

"These advanced mathematical models appear colorblind on the surface . . . but are also significantly influenced by pervasive bias in the criminal justice system."

“My Dog is a Very Well-trained Dog”



- My dog Fluffy has been trained to detect risk of recidivism.
- Its accuracy has even been tested!
- But it is a “black box” and does not provide explanations.

Question: Would you allow Fluffy to make parole decisions?
Should you give Fluffy ethics training?

Regulating Decisions Systems

James Larus and
Chris Hankin,
CACM, Aug. 2018:

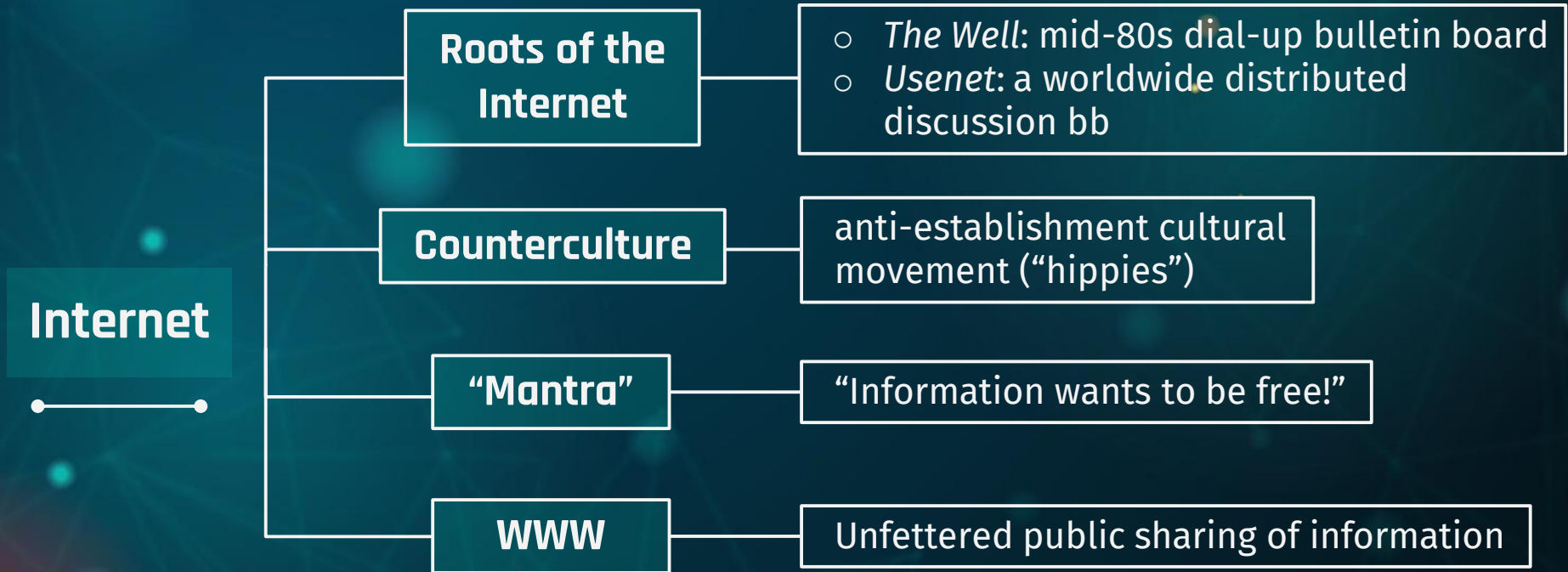
“Disdain for regulation is pervasive throughout the tech industry. In the case of automated decision making, this attitude is mistaken.”

“The widespread adoption of AD systems will be economically disruptive and will raise new and complex societal challenges, such as worker displacement; autonomous accidents; and, perhaps most fundamentally, confusion and debate over what it means to be human.”



The Internet

Example: Internet



Problem 1: How can you find information on the WWW?

Internet (continued)

		Result
SOLUTION 1	Directory (Yahoo)	Does not scale!
SOLUTION 2	Search Engine (Altavista)	Stank!
SOLUTION 3	Good Search Engine (Google)	Good 👍

Problem 2: How do you monetize free information?

Brilliant Solution: *advertising* (copied from newspapers and TV)

AI-Driven Internet

Observation

Online advertising effectiveness is very low!

Solution

Micro-targeted advertising
– match ads with individual preferences.

Requirement


Personal data!

Example: “A secret partnership with Mastercard illuminates the massive amounts of user data Google scrapes and sells—at times without their knowledge.”

ML on Steroids: targeted advertising based on personal data.


Micro-Targeting Advertising


What is Exxon?

 ExxonMobil

- Sponsored -

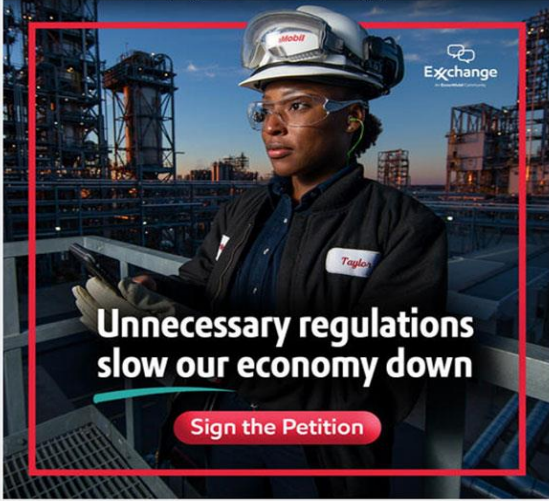
Subscribe to learn how Global Thermostat is capturing CO2 directly from the air.



 ExxonMobil

Sponsored · Paid for by EXXON MOBIL CORPORATION

The oil and gas industry is THE engine that powers America's economy. Help us make sure unnecessary regulations don't slow energy growth!



**Unnecessary regulations
slow our economy down**

[Sign the Petition](#)

exxchange.com
Show Your Support Today
[Sign Up](#)

“If you’re not paying for it, you’re the product.”

Epistemic Inequality:

“The combination of state surveillance and its capitalist counterpart means that digital technology is separating the citizens in all societies into two groups: the watchers (invisible, unknown, unaccountable) and the watched – ***epistemic inequality!***” (2019)

THE AGE OF SURVEILLANCE CAPITALISM

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

—
**SHOSHANA
ZUBOFF**

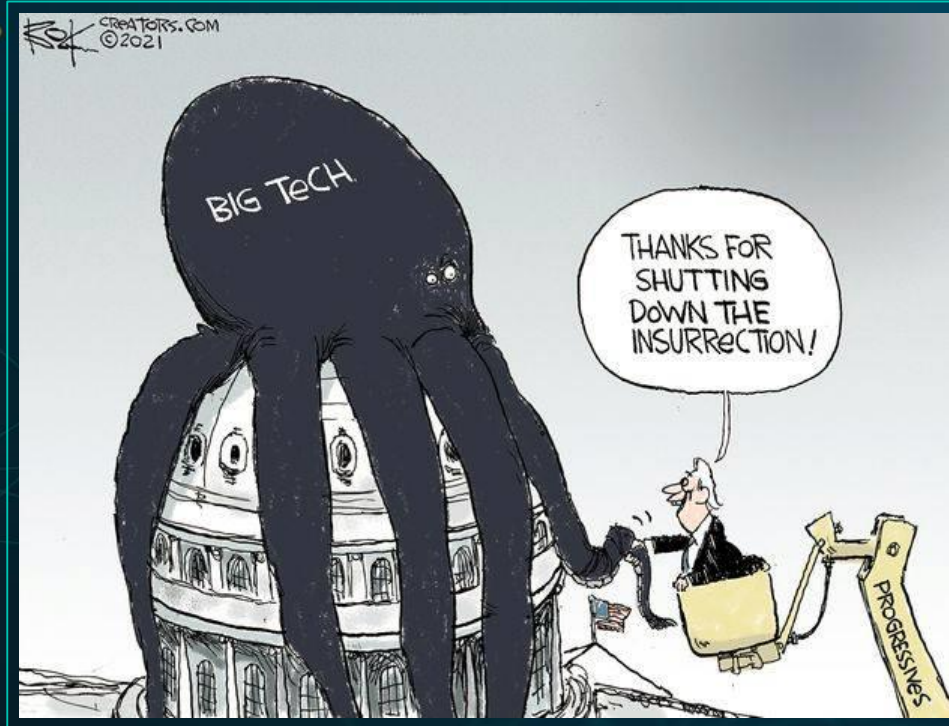
The Coup We Are Not Talking About

Shoshana Zuboff, NYT, 1/2021:

“We can have democracy, or we can have a surveillance society, but we cannot have both.”

“In an information civilization, societies are defined by questions of knowledge . . . Who knows? Who decides who knows? Who decides who decides who knows? Surveillance capitalists now hold the answers to each question, though we never elected them to govern.”

Are We Living in a Tech Dystopia?



Facebook's Engagement Engine

ML on Steroids: Content moderation –
maximize user engagement.

- **Nano-targeting** Users Based Solely on
Their Perceived Interests

Social-media – addiction by design:

- **A.C. Madrigal, The Atlantic, 2013:** “What an anthropologist's examination of Vegas slot machines reveals about the hours we spend on social networks.”
- **NYU, 2/21:** “Social Media Use Driven by Search for Reward, Akin to Animals Seeking Food, New Study Shows”

The Tech Industry's War on Kids

Richard Freed, Medium, March 2018:

“What none of these parents understand is that their **children's and teens' destructive obsession with technology is the predictable consequence of a virtually unrecognized merger between the tech industry and psychology.**

This alliance pairs the consumer tech industry's immense wealth with the most sophisticated psychological research, making it possible to develop social media, video games, and phones with **drug-like power** to seduce young users.”

Death Rates Due to Suicide

Center for Disease Control, Oct. 2019:

2000-2007	Between 2007-2017
The suicide rate among persons aged 10–24 was stable	The suicide rate among persons aged 10–24 increased 56%
	The suicide rate among children aged 10-14 nearly tripled
	The suicide rate among teenagers (14-19) increased by 76 percent

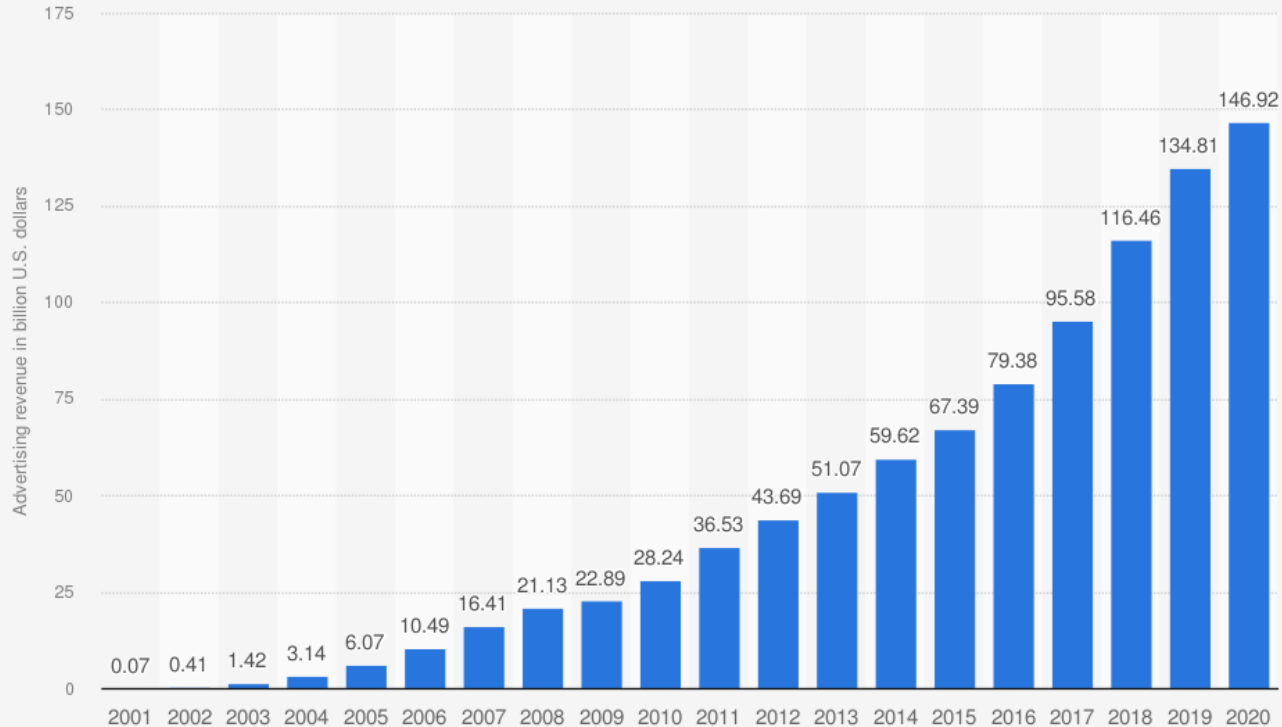
Death Rates Due to Suicide

Center for Disease Control, Oct. 2019:

2000-2007	Between 2007-2017
The suicide rate among persons aged 10-24 was stable	June 2007: iPhone The suicide rate among persons aged 10-24 increased by 20 percent
	July 2008: Facebook iPhone app The suicide rate among children aged 10-14 nearly tripled
	The suicide rate among teenagers (14-19) increased by 76 percent

Google's Business

Advertising revenue of Google from 2001 to 2020 (in billion U.S. dollars)



Business vs Ethics

Imagine: We send Sergei Brin and Larry Page to an “ethics boot-camp”. Will they ditch Google’s targeted-advertising business model?

MYV, Jan. 2019: “If society finds the surveillance business model offensive, then the remedy is public policy, in the form of laws and regulations, rather than an ethics outrage.”

But Google is stuck between its “ethics façade” and its unethical business model.

Google's Ethics Scandal



What Really Happened When Google Ousted **Timnit Gebru**?

Wired, 6/'21:

“She was a star engineer who warned that messy AI can spread racism. Google brought her in. Then it forced her out. Can Big Tech take criticism from within?”

Google's Ethics Contradiction

"The withering email that got an ethical AI researcher fired at Google"

Dec. 2020

Feb. 2021

"Google fires Margaret Mitchell, another top researcher on its AI ethics team"

"This is bigger than just Timnit': How Google tried to silence a critic and ignited a movement."

"AI ethics research conference suspends Google sponsorship."

Mar. 2021

"Citing privacy concerns, Google says it won't use technologies that track individuals across multiple websites."

WSJ's Bombshell Facebook Investigation

9/21: Several scathing articles about Facebook

"We make body image issues worse for one in three teen girls."

"We are not actually doing what we say we do publicly"

"Misinformation, toxicity and violent content are inordinately prevalent among reshares."

"An unfortunate side effect is that harmful and misinformative content can go viral, often before we can catch it and mitigate its effects."

There is No Free Lunch!

QUESTION 1 • Are Google's service really free to consumers?

QUESTION 2 • How does Google make almost \$150B from advertising?

Alex Webb, Bloomberg, 3/'21:

"The Whole Web Pays For Google And Facebook To Be Free"

ANSWER • Consumers pay for Google services, but in a totally **opaque** way!

OUTCOME • Mass surveillance!

Opacity

Standard argument: People willingly give their data in exchange for “free” services.

Recent Pew Survey:



74% of US Facebook users did not know that FB maintains list of their traits and interests.



51% are not comfortable with FB compiling this information.



27% disagree with FB's characterization.

Question: Have you read Gmail's “terms of service”?

The Destruction of the Information Market

Markets: prices for goods and services are determined by sellers and buyers

- A supremely efficient mechanism for value discovery!

Internet: “How Hippies destroyed the Internet”

Question: Is the issue lack of *ethics*, or is it lack of *public policy*?



Policy

Why No IT Public Policy?

Big Tech, 01/2022: ~\$10T in market cap!

- At least 1/3 of the SP500
- Estimated 2021 revenue: \$1.4T

The IT industry has successfully lobbied against any attempt to legislate/regulate IT public policy under the mantra “*regulation stifles innovation*”.

- *Wired*, 09/'18: “The case against Elon Musk will chill innovation!”

A Techno-Utopian Culture

John Perry Barlow, 1996

A Declaration of the Independence of Cyberspace:

“Governments of the Industrial World, you weary giants of flesh and steel, I come from Cyberspace, the new home of Mind. On behalf of the future, I ask you of the past to leave us alone.”

A Techno-Utopian Culture

But: What happens in cyberspace does not stay in cyberspace!

Imagine: A cyberattack on the US power grid

- **WSJ:** “America’s Electric Grid Has a Vulnerable Back Door—and Russia Walked Through It”
-

The Disruption Culture

C. Christensen, 1995:

Disruptive Innovation

— a process by which a product or service takes root initially in simple applications at the bottom of a market and then relentlessly moves up market, eventually displacing established competitors.

- Example: Disk-drive technology

“Disruption” has become a Silicon Valley mantra!

- V. Heffernan, NYT, 10/'20: “Ennobling destruction and sabotage makes the most brutal forms of capitalism seem like God’s will!”

From Disrupting to Breaking

Facebook's motto (until 2014):

"Move fast and break things.
Unless you are breaking stuff,
you are not moving fast enough."

We now know:

- FB's "frictionless sharing" enables massive dissemination of fake news, misinformation, disinformation, and extreme content.

WaPo, 10/'21: "The case against Mark Zuckerberg: Insiders say Facebook's CEO chose growth over safety."

Almost-Breaking Democracy



The Capitol Insurrection, **Jan. 6, 2021**

Corporate Responsibility

Corporate Social Responsibility (since 1960s):

- A form of self-regulation
- Sacrificing profits for social goals
- **Corporate stakeholders:** shareholders, customers, employees, community

Shareholder Value (since 1980s): *the ultimate measure of a company's success is the extent to which it enriches shareholders.*

- **Milton Friedman, 1970** : “The social responsibility of a business is to increase its profits.”
- Business Roundtable, 2019: Back to stakeholders!

Technology is Driving the Future

Question:

Who is doing the steering?



Answer: Tech corporations!

Technology has moved very fast, but public policy has lagged behind.

Innovation is not a goal; it is a means for societal progress!

50 Years of Blaming Milton Friedman

Binyamin Appelbaum, NYTime, 9/20:

“Critics have been fighting ever since to get corporations to acknowledge broader responsibilities. **It’s the wrong battle.** *Instead of redefining the role of the corporation, we need to redefine the role of the state.*”

“**Government remains the most powerful means to express our collective will.** The necessary solution is to create stronger incentives for good behavior and laws against bad behavior.”

Ethics vs Public Policy

Is ethics important? Of course!

- **Rome Call:** *“All human beings are born free and equal in dignity and rights.”*

Ethics should *inform* public policy.

- **Chief Justice Earl Warren:** *“In civilized life, the law floats in a sea of ethics.”*

But: The law is institutionalized; ethics is not!

- Without laws and regulations, ethics is a subject for philosophy books!

Note: I teach a course on “Computing, Ethics, and Society”.

Regulation is Coming!

Tim Cook, Apple's CEO, 1/'19:

"We all deserve control over our digital lives. That's why we must rein in the data brokers."

Mark Zuckerberg, 3/'19, WaPo:

"The Internet needs new rules. Let's start in these four areas: harmful content, elections, privacy, and data portability."

Reuters, 1/'20:

"Alphabet CEO backs temporary ban on facial recognition tech."

WaPo, 1/'22:

"Why 2022 could be a 'watershed year' for tech regulation".

How to Regulate? |

Caveat: Smart regulation is hard!

- Good governance is a necessary condition.

Different Approaches:

- US: Technology specific
- Europe: Societal-impact specific
- What about international treaties?

The Big Question: Is Big Tech too big?

- Tim Wu, NYT, 11/'18: "Be Afraid of Economic Bigness. Be Very Afraid".
- Should FB be forced to spin-off Instagram and WhatsApp?

The Social Trilemma

Emily Bazelon, NYT, 1/'21:

“Why Is Big Tech Policing Speech?
Because the Government Isn’t!”

“We are uncomfortable with government doing it; we are uncomfortable with Silicon Valley doing it. But we are also uncomfortable with nobody doing it at all!”

WaPo, 3/22: “Facebook breaks its own rules to allow for some calls to violence against Russian invaders”

“Facebook will allow some calls for violence against Russian invaders as the war in Ukraine drags on, creating an unusual exception to its long-standing hate speech rules that prohibit such language.”

Who Drives Technology?

Tristan Greene, Next Web, 2/20: When we discuss the ethics of AI, we're asking two simple questions:

- Is it ethical to build an AI for this specific purpose?
- Is it ethical to build an AI with these capabilities?

MYV: Who benefits? Who decides?

Kate Crawford, 2021: "Stop talking about AI ethics. It's time to talk about power."

Don't be driven by technology. Drive it!