



Ethics, Public Policy, and Technological Change

Rob Reich
Mehran Sahami
Head TA: Roberta Fischli

Housekeeping

- Final sections on Friday
- Policy memo due on Thursday
- Final reflection assignment distributed Thursday
- Final class on Wednesday: come prepared with questions & criticisms

Frontier AI: Policy Choices

1. When policy windows open
2. Incremental policy change
3. Structural policy change
4. Radical policy change

Frontier AI: Policy Choices

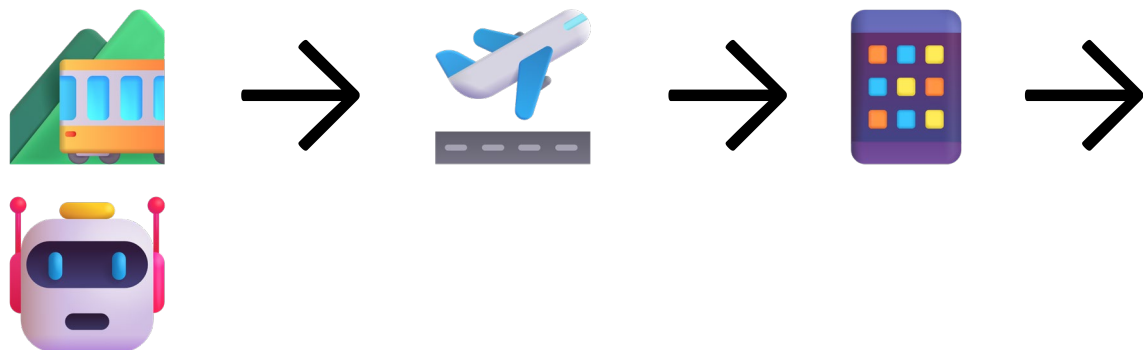
1. When policy windows open
2. Incremental policy change
3. Structural policy change
4. Radical policy change

Delayed regulation is a feature, not a bug



In the historical context of emerging technologies, it is a regular feature that innovation and market dissemination precedes democratic regulation.

We see this occur repeatedly in history.





Source: *Progressive Railroading*

Railroads

1804: the world's first steam-powered locomotive journey

1827: first regular carrier of passengers + freight via
Baltimore + Ohio railroad

1887: Congress passed the Interstate Commerce Act, making
the railroads the first industry subject to federal regulation



Source: *Science Museum*

Telegraph + Telephone

1830s: The telegraph was invented

1851: Western Union Telegraph Company was established

1876: Alexander Graham Bell patents the telephone

1885: American Telegraph and Telephone Company (AT&T) was formed

1934: The Communication Act of 1934 was announced, which established the FCC



Airplane

Source: Pinterest, William Doubledee

1903: Wilbur and Orville Wright invented and flew the Wright Flyer

1908: First commercial flight took off

1920s: Commercial flights become more regular

1935: Qantas opens its first international flight

1958: The Federal Aviation Administration (FAA) was established

Shipping Containers



By Bernhard Fuchs - NYK Virgo, CC BY 2.0,
<https://commons.wikimedia.org/w/index.php?curid=27657317>

1956: Malcolm McLean launches first intermodal container; shipping industry is moving cargo, not sailing ships

1960s: First purpose-built container ships

1970s: Adoption of International Organization for Standardization (ISO) container standards

1984: US passes Shipping Act, replacing 1916 legislation

Generative AI, LLMs, GPTs

2022: OpenAI releases ChatGPT

2023: Claude, Gemini, Llama, DeepSeek

Dec 2025: 800M weekly users of ChatGPT

A policy window has opened for AI

2016: The EU General Data Protection Regulation (GDPR). This marked the strongest privacy and security law in the world.

2023: Biden White House releases Executive Order on Artificial Intelligence

2024: Creation of the AI Safety Institutes

2025: EU AI office's General Purpose AI Code of Practice

2025: Trump White House releases AI Action Plan, era of AI accelerationism and AI sovereignty

What's Next? (India AI summit at end of February 2026)

Value Trade-Offs

- How might AI undermine human agency/liberty?
 - Experience machine (freedom in the metaverse?)
 - Workplace surveillance
 - Autonomous weapons
 - How might AI be in tension with human well-being?
 - Technological displacement of labor
 - Distributional consequences
 - CSAM + NCII
 - AI companions as substitutes for human intimacy
 - Favors tyranny
-

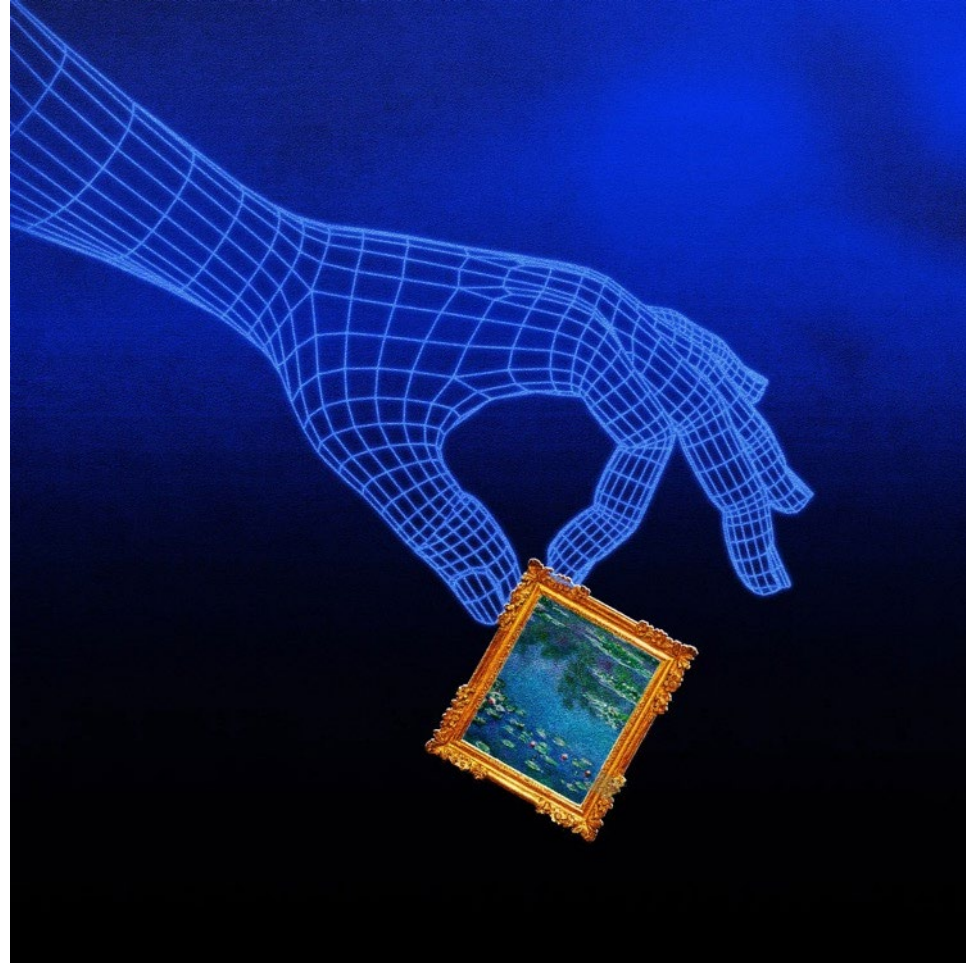
Frontier AI: Policy Choices

1. When policy windows open
2. Incremental policy change
3. Structural policy change
4. Radical policy change

(1) Existing Laws: Copyright

THE NEW YORKER

IS A.I. ART STEALING
FROM ARTISTS?





Fair Learning

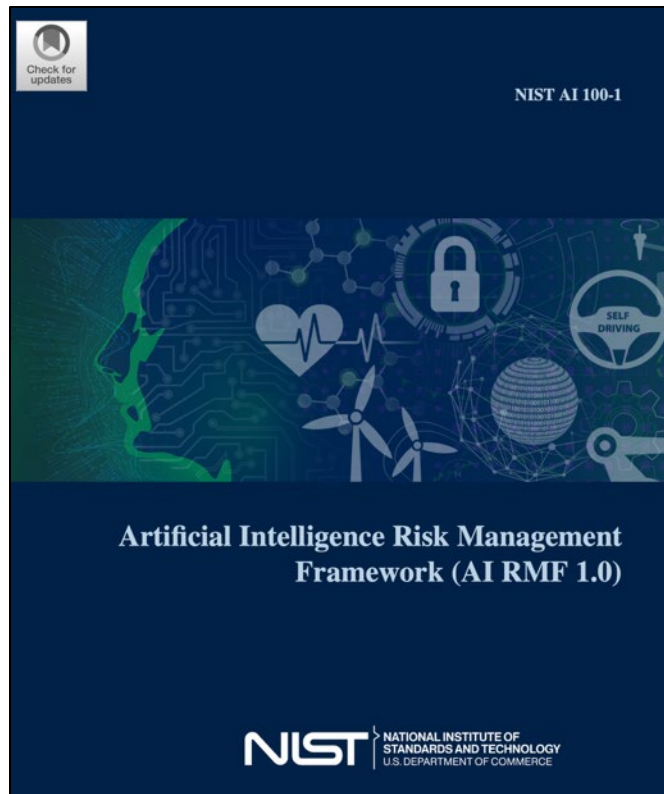
Mark A. Lemley* & Bryan Casey**

“We argue that **ML systems should generally be able to use databases for training**, whether or not the contents of that database are copyrighted.”

Fair Use

- (1) **Purpose and character** of use (e.g., educational? transformative?)
- (2) **Nature** of the copyrighted work (e.g., creative?)
- (3) **Amount and substantiality** of work taken
- (4) **Effect on potential market**

(2) Risk Management and AI Bill of Rights



AI Risk Management Framework



Legal Disclaimer:

The Blueprint for an AI Bill of Rights is **non-binding** and **does not constitute U.S. government policy**.

It **does not supersede, modify, or direct** an interpretation of any existing statute, regulation, policy, or international instrument.

It does not constitute binding guidance for the public or Federal agencies and therefore **does not require compliance** with the principles described herein.

[Principles might depend on **context**, such as law enforcement and national security.]

The Blueprint for an AI Bill of Rights is not intended to, and **does not, create any legal right, benefit, or defense . . .**

(3) Red-Teaming



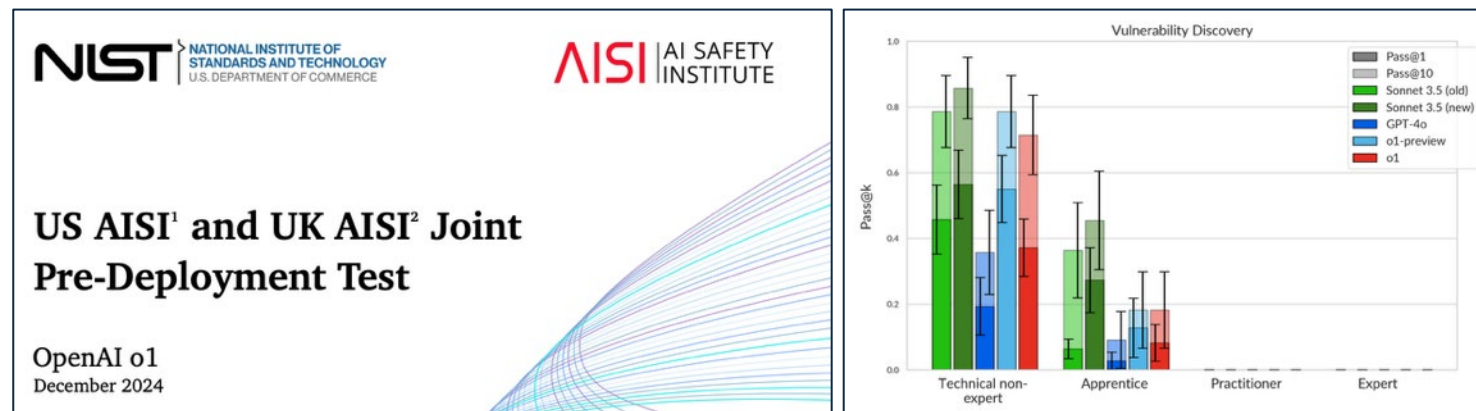


POLITICO Rishi Sunak promised to make AI safe. Big Tech's not playing ball.

US AI Safety Institute “has **no direct testing mandate.**”

“We cannot rely on the **goodwill of AI companies.**”

U.S. AI Safety Institute Signs Agreements Regarding AI Safety Research, Testing and Evaluation With Anthropic and OpenAI



The image shows the logo for the AI Security Institute (AISI). The logo consists of the letters 'AISI' in a large, white, sans-serif font on the left. To the right of 'AISI' is a vertical line, followed by the words 'AI SAFETY' in a smaller, white, sans-serif font. A thick red horizontal line is drawn across the word 'SAFETY'. Below 'AI SAFETY' is the word 'INSTITUTE' in the same white, sans-serif font. The background is a gradient from dark blue at the top to red at the bottom.

SECURITY
AISI | AI ~~SAFETY~~
INSTITUTE

Britain's AI Security Institute "will drop its focus on **bias and free speech** to concentrate on crime and national security issues"

AI Red-Teaming Is Not a One-Stop Solution to AI Harms:

Recommendations for Using
Red-Teaming for AI Accountability

Sorelle Friedler, Ranjit Singh, Borhane Blili-Hamelin,
Jacob Metcalf, and Brian J. Chen

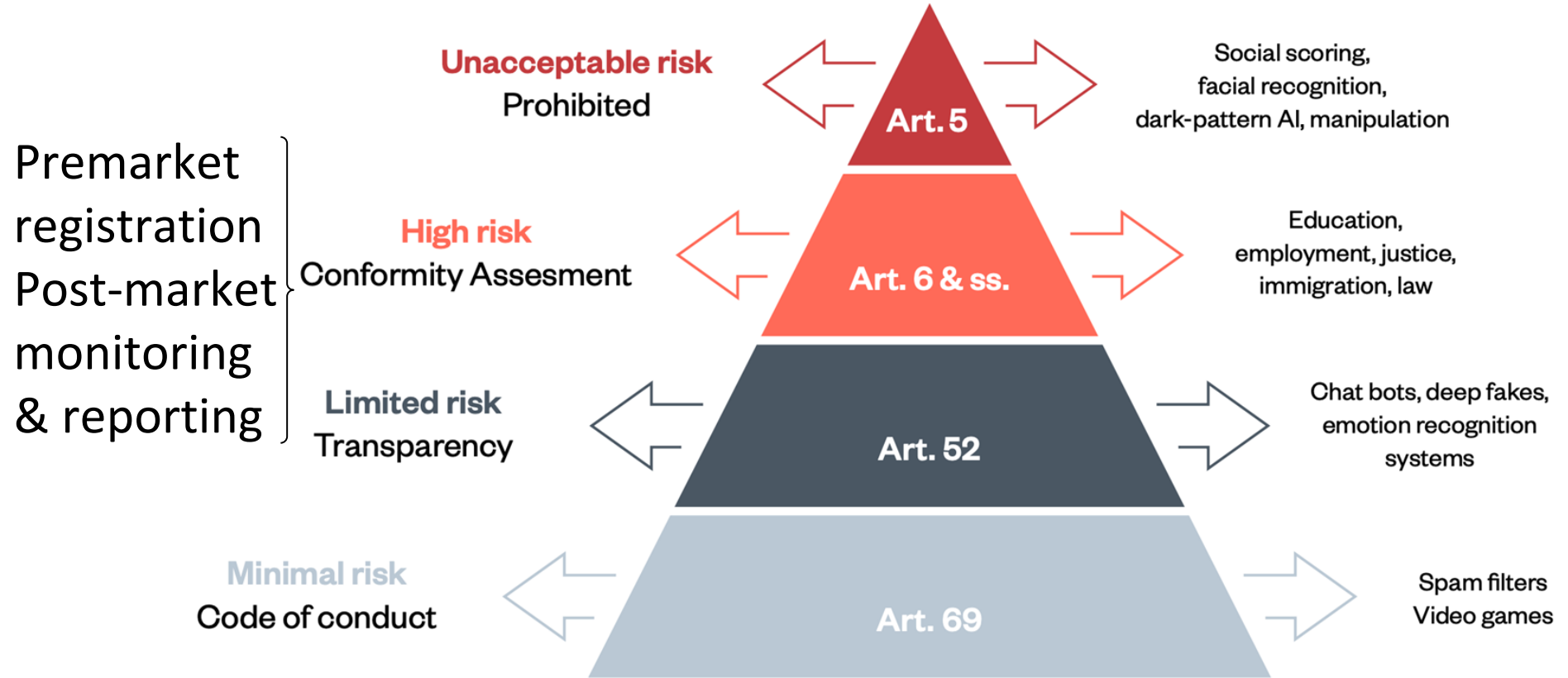
“[R]ed-teaming traditionally prioritizes people who have **advanced technical skills** and excludes the many people who do not.”

Jack Clark: “Pre-deployment testing is a nice idea but very **difficult to implement**”

Frontier AI: Policy Choices

1. When policy windows open
2. Incremental policy change
3. Structural policy change
4. Radical policy change

(1) EU AI Act



(2) A New AI Agency



Digital
Consumer
Protection
Commission

AN FDA FOR ALGORITHMS

ANDREW TUTT*

The rise of increasingly complex algorithms calls for critical thought about how best to prevent, deter, and compensate for the harms that they cause. This Article argues that the criminal law and tort regulatory systems will prove no match for the difficult regulatory puzzles algorithms pose. Algorithmic regulation will require federal uniformity, expert judgment, political independence, and pre-market review to prevent—without stifling innovation—the introduction of unacceptably dangerous algorithms into the market. This Article proposes that certain classes of new algorithms should not be permitted to be distributed or sold without approval from a government agency designed along the lines of the FDA. This “FDA for Algorithms” would approve certain complex and dangerous algorithms when it could be shown that they would be safe and effective for their intended use and that satisfactory measures would be taken to prevent their harmful misuse. Lastly, this Article proposes that the agency should serve as a centralized expert regulator that develops guidance, standards, and expertise in partnership with industry to strike a balance between innovation and safety.

(3) A Public Option

**The
New York
Times**

Universities and Tech Giants Back National Cloud Computing Project

A proposal to give scientists access to huge data sets and powerful computers.



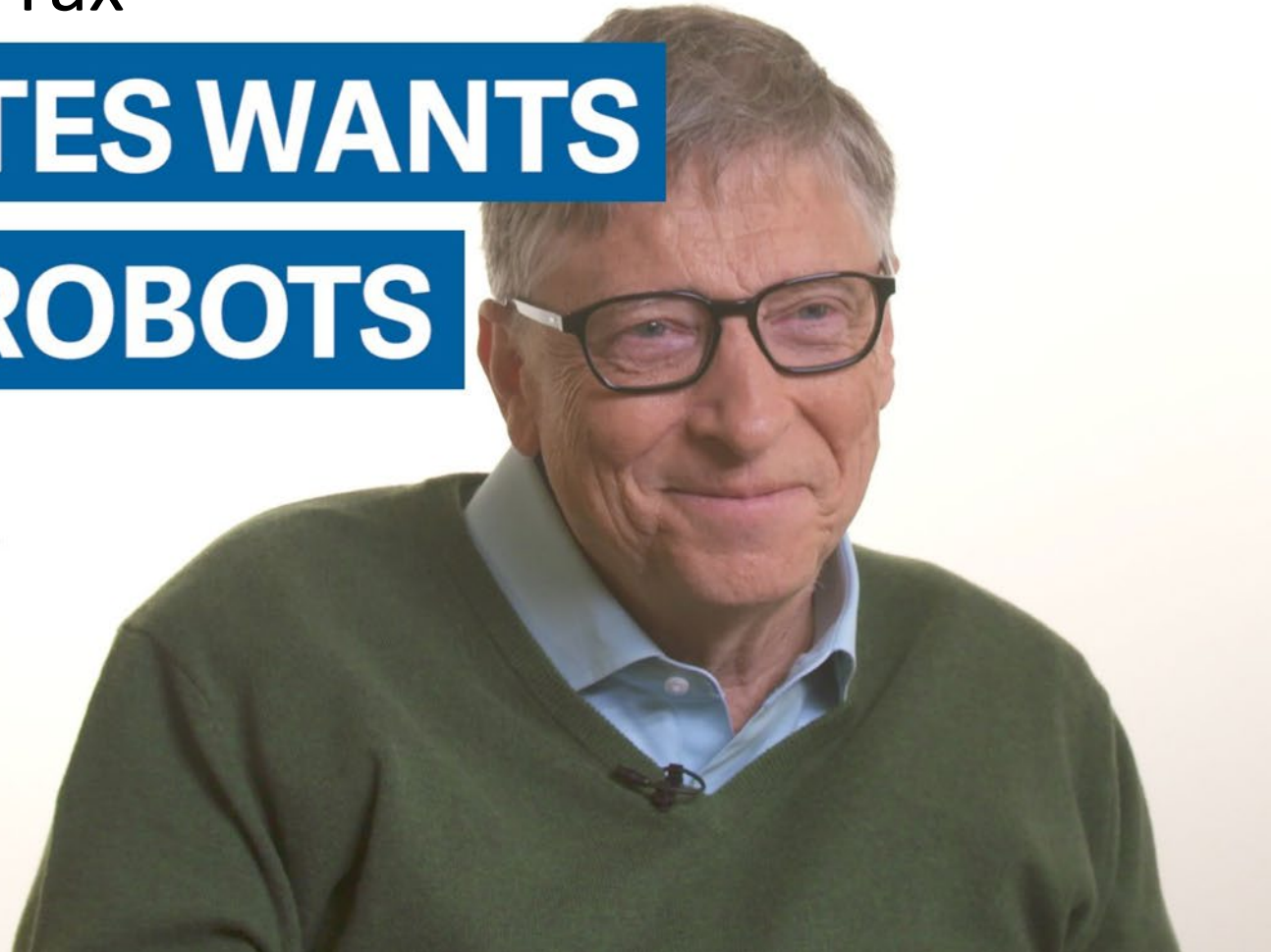
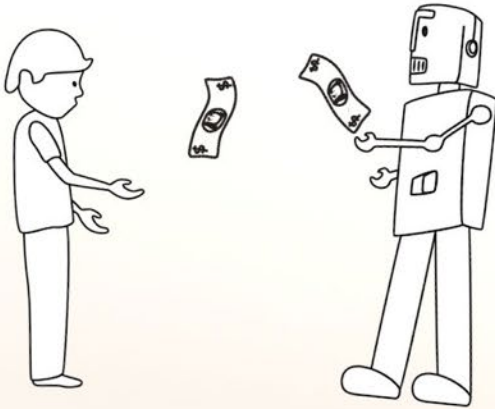
Frontier AI: Policy Choices

1. When policy windows open
2. Incremental policy change
3. Structural policy change
4. Radical policy change

(1) Automation Tax

BILL GATES WANTS

TO TAX ROBOTS



Labor is taxed, on average, at **25%**, while software is taxed at **5%**

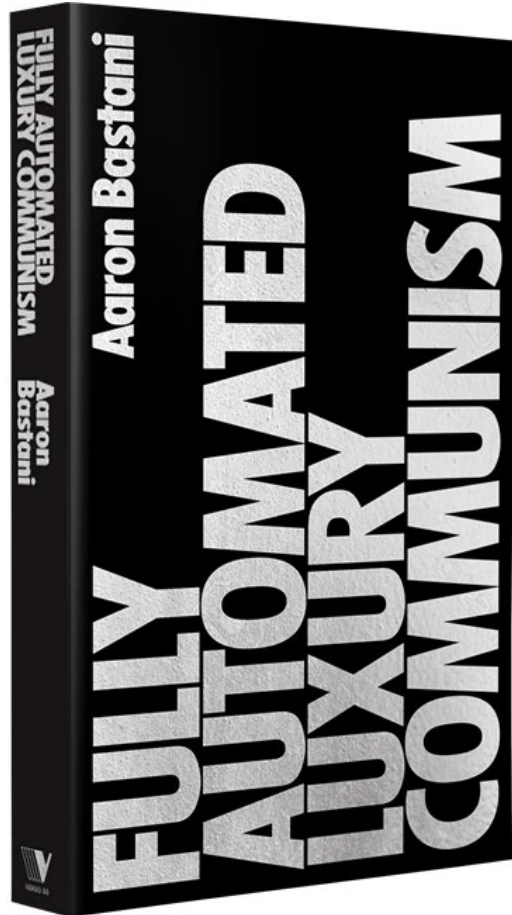
“In essence, the U.S. tax system encourages companies to buy machines will discouraging them from adding workers.” (Acemoglu et al 2020).

(2) Universal Basic Income



Regular, guaranteed, unconditional payment
from government

Fully Automated Luxury Communism?



The New York Times

OPINION

The World Is a Mess. We Need Fully Automated Luxury Communism.

Asteroid mining. Gene editing. Synthetic meat. We could provide for the needs of everyone, in style. It just takes some imagination.

June 11, 2019

(3) Something Else

BOSTON REVIEW



CLASS & INEQUALITY

Forum

The American Corporation is in Crisis —Let's Rethink It

For decades, shareholder primacy has obscured the fact that employees should do well when businesses do well.

Lenore Palladino

“Even the most generous UBI proposals amount to a transfer payment that leaves the **basic logic of capitalism** and widening power asymmetries between workers and employers unchallenged.” (Kelly, 2023)



Unthinkable
Radical
Acceptable
Sensible
Popular
Policy
Popular
Sensible
Acceptable
Radical
Unthinkable



Waiting for an AI Catastrophe?



In the
meantime,
what's
happening
inside
companies
?



Trust & Safety



Product policy



Red-teaming



AI Safety



Research teams

Guest Experts