# *Pause Giant AI Experiments?*

Avi Goldfarb
University of Toronto

FOURTH ESTATE

# Opinion | Why I Welcome Our Future AI Overlords

Who's afraid of ChatGPT?





AI/MACHINE LEARNING

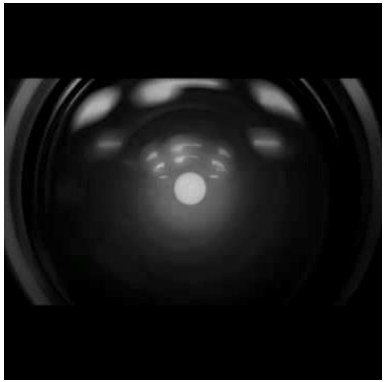
## Generative AI is about to become a $23 trillion industry – and that's not counting its dark side of scams, deepfakes and romance bots

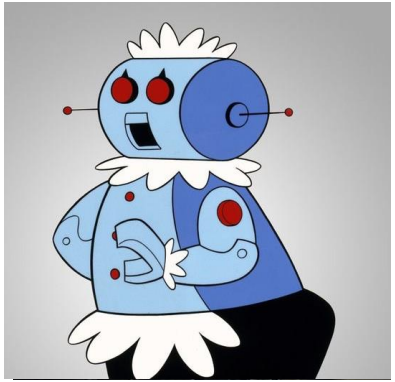
Essay | Browsers, the printing press, Freud and AI

## How AI could change computing, culture and the course of history

Expect changes in the way people access knowledge, relate to knowledge and think about themselves

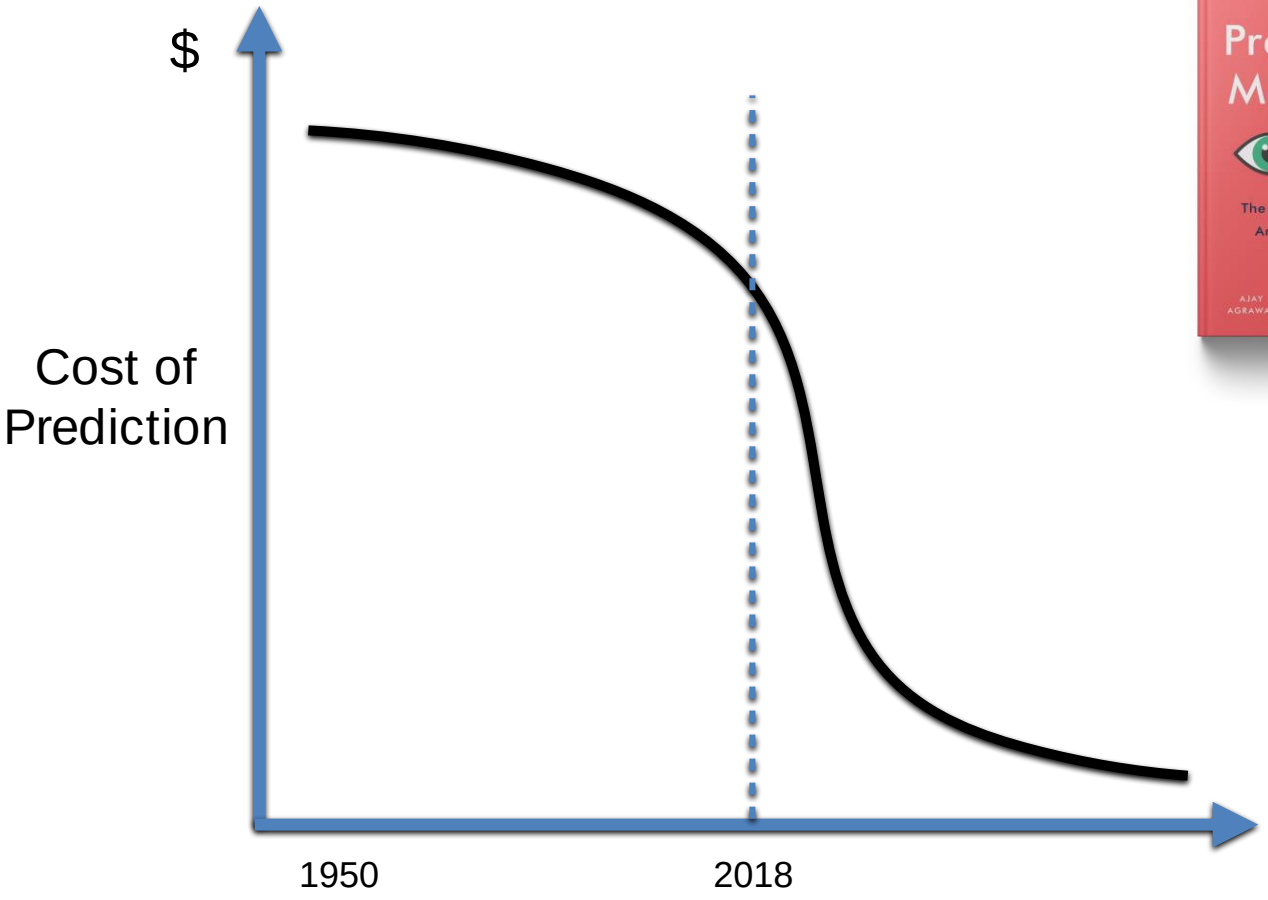
# What is Artificial Intelligence?

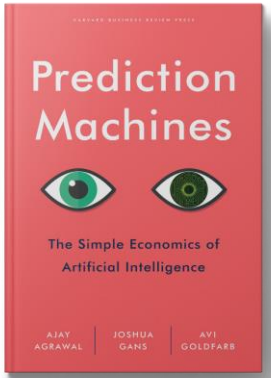


**PREDICTION:**
**Using information that you *do* have to generate information that you *don't* have**

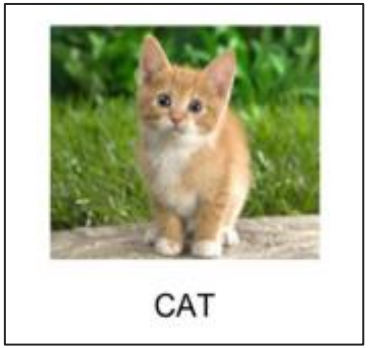
# Artificial Intelligence




Mortgages
APPROVED
Your Mortgag


RISK?
YES
NO


CAT

OpenAI

A How will AI affect productivity and inequality?

Artificial Intelligence (AI) has the potential to significantly affect productivity and inequality in multiple ways, and many of these effects are interconnected.

**Productivity**

1. **Automation of Routine Tasks:** AI can automate routine tasks, which can greatly enhance productivity. This is true in various sectors like manufacturing, where robots can work 24/7, and services, where AI can handle tasks such as customer inquiries or data analysis.
2. **Intelligent Assistance:** AI can provide intelligent assistance to human workers, helping them to be more efficient and effective. For example, AI can help doctors diagnose diseases, architects design buildings, or teachers personalize learning.
3. **Innovation:** AI can help drive innovation, creating new products, services, and ways of doing things. This can lead to increased productivity across the economy.

**Inequality**

1. **Job Displacement:** While AI can increase productivity, it can also lead to job

FOURTH ESTATE

# Opinion | Why I Welcome Our Future AI Overlords

Who's afraid of ChatGPT?





AI/MACHINE LEARNING

## Generative AI is about to become a $23 trillion industry – and that's not counting its dark side of scams, deepfakes and romance bots

Essay | Browsers, the printing press, Freud and AI

## How AI could change computing, culture and the course of history

Expect changes in the way people access knowledge, relate to knowledge and think about themselves

future of life INSTITUTE

# Pause Giant AI Experiments: An Open Letter

We call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4.

*View this open letter online*.

| Published | PDF created | Signatures |
| --- | --- | --- |
| March 22, 2023 | May 5, 2023 | 27565 |

AI systems with human-competitive intelligence can pose profound risks to society and humanity, as shown by extensive research[1] and acknowledged by top AI labs.[2] As stated in the widely-endorsed Asilomar AI Principles, *Advanced AI could represent a profound change in the history of life on Earth, and should be planned for and managed with commensurate care and resources.* Unfortunately, this level of planning and management is not happening, even though recent months have seen AI labs locked in an out-of-control race to develop and deploy ever more powerful digital minds that no one – not even their creators – can understand, predict, or reliably control.

"Should we risk loss of control of our civilization? Should we develop nonhuman minds that might eventually outnumber, outsmart, obsolete and replace us?"

"Should we let machines flood our information channels with propaganda and untruth?"

"Should we automate away all the jobs, including the fulfilling ones?"

National Bureau of Economic Research

# THE ECONOMICS OF ARTIFICIAL INTELLIGENCE

## *An Agenda*

Edited by Ajay Agrawal, Joshua Gans, and Avi Goldfarb

# 1

# Artificial Intelligence and the Modern Productivity Paradox
## A Clash of Expectations and Statistics

Erik Brynjolfsson, Daniel Rock, and Chad Syverson

The discussion around the recent patterns in aggregate productivity growth highlights a seeming contradiction. On the one hand, there are astonishing examples of potentially transformative new technologies that could greatly increase productivity and economic welfare (see Brynjolfsson and McAfee 2014). There are some early concrete signs of these technologies' promise, recent leaps in artificial intelligence (AI) performance being the most prominent example. However, at the same time, measured productivity growth over the past decade has slowed significantly. This deceleration is large, cutting productivity growth by half or more in the decade preceding the slowdown. It is also widespread, having occurred throughout the Organisation for Economic Co-operation and Development (OECD) and, more recently, among many large emerging economies as well (Syverson 2017).[1]

# Potential for a Productivity Boom?

9

## Artificial Intelligence and Economic Growth

Philippe Aghion, Benjamin F. Jones, and Charles I. Jones

### 9.1 Introduction

This chapter considers the implications of artificial intelligence for economic growth. Artificial intelligence (AI) can be defined as "the capability of a machine to imitate intelligent human behavior" or "an agent's ability to

RESEARCH - REPORT

## Machines of mind: The case for an AI-powered productivity boom

Martin Neil Baily, Erik Brynjolfsson, and Anton Korinek · Wednesday, May 10, 2023

There is an emerging literature that estimates the productivity effects of AI on specific occupations or tasks. Kalliamvakou (2022) finds that software engineers can code up to twice as fast using a tool called Codex, based on the previous version of the large language model GPT-3. That's a transformative effect. Noy and Zhang (2023) find that many writing tasks can also be completed twice as fast and Korinek (2023) estimates, based on 25 use cases for language models, that economists can be 10-20% more productive using large language models.

future of life INSTITUTE

# Pause Giant AI Experiments: An Open Letter

We call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4.

*View this open letter online*.

| Published | PDF created | Signatures |
| --- | --- | --- |
| March 22, 2023 | May 5, 2023 | 27565 |

AI systems with human-competitive intelligence can pose profound risks to society and humanity, as shown by extensive research[1] and acknowledged by top AI labs.[2] As stated in the widely-endorsed Asilomar AI Principles, *Advanced AI could represent a profound change in the history of life on Earth, and should be planned for and managed with commensurate care and resources.* Unfortunately, this level of planning and management is not happening, even though recent months have seen AI labs locked in an out-of-control race to develop and deploy ever more powerful digital minds that no one – not even their creators – can understand, predict, or reliably control.

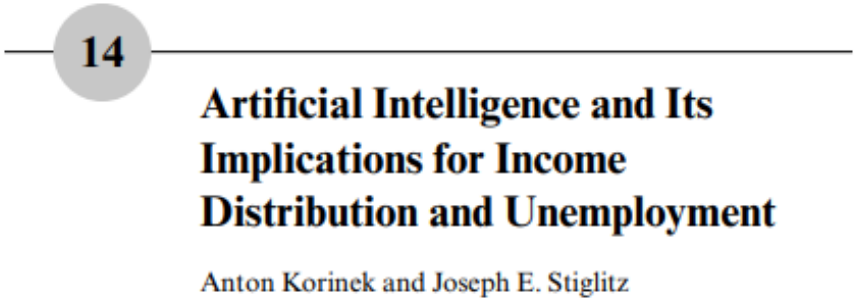
"Should we risk loss of control of our civilization?" Should we develop nonhuman minds that might eventually outnumber, outsmart, obsolete and replace us?"

"Should we let machines flood our information channels with propaganda and untruth?"

"Should we automate away all the jobs, including the fulfilling ones?"

**“Should we risk loss of control of our civilization? Should we develop nonhuman minds that might eventually outnumber, outsmart, obsolete and replace us?”**

# “Should we develop nonhuman minds that might eventually outnumber, outsmart, obsolete and replace us?”


14

Artificial Intelligence and Its Implications for Income Distribution and Unemployment

Anton Korinek and Joseph E. Stiglitz

“If progress in AI cannot be halted, our description above suggests mechanisms that may ensure that humans can afford a separate living space and remain viable: because humans start out owning some of the factors that are in limited supply, if they are prohibited from transferring these factors, they could continue to consume them without suffering from their price appreciation.”

# “Should we develop nonhuman minds that might eventually outnumber, outsmart, obsolete and replace us?”

9

## Artificial Intelligence and Economic Growth

Philippe Aghion, Benjamin F. Jones, and Charles I. Jones

### 9.1 Introduction

This chapter considers the implications of artificial intelligence for economic growth. Artificial intelligence (AI) can be defined as “the capability of a machine to imitate intelligent human behavior” or “an agent’s ability to

# “Should we develop nonhuman minds that might eventually outnumber, outsmart, obsolete and replace us?”

*American Economic Journal: Macroeconomics 2021, 13(1): 299–332*
*https://doi.org/10.1257/mac.20170105*

## Are We Approaching an Economic Singularity? Information Technology and the Future of Economic Growth†

*By* William D. Nordhaus*

### III. Rapid Technological Change through Superintelligent Innovation

A first possible source of extremely rising economic growth comes from rapid improvements in technology generated by superintelligent agents. This approach can be seen easily using a Cobb-Douglas production function of the form $Y_t = K_t^{\alpha}(A_t L_t)^{1-\alpha}$. Here and below, assume that $Y$ is output, $K$ is capital, $L$ is labor, $A$ is labor-augmenting technology, $s$ is the savings rate, and $t$ is time. For most of the discussion, I assume the savings rate is constant. For a given rate of labor-augmenting technological change of $h$, the growth of output will be $g \rightarrow n + h$. Singularity quite naturally arises if technological change becomes extremely rapid.

# The A.I. Dilemma: Growth versus Existential Risk

Charles I. Jones*
Stanford GSB and NBER

September 12, 2023 — Version 0.7
*Preliminary, comments appreciated*

**Abstract**

Advances in artificial intelligence (A.I.) are a double-edged sword. On the one hand, they may increase economic growth as A.I. augments our ability to innovate or even itself learns to discover new ideas. On the other hand, many experts note that these advances entail existential risk: creating a superintelligent entity misaligned with human values could lead to catastrophic outcomes, including human extinction. This paper considers the optimal use of A.I. technology in the presence of these opportunities and risks. Under what conditions should we continue the rapid progress of A.I. and under what conditions should we stop?

## Simple Model

*Key Point 4 (Mortality improvements):* *Mortaility risk and existential risk are in the same units and do not run into the diminishing marginal utility of consumption. If A.I. improves life expectancy, the existential risk cutoffs are much higher, on the order of 25–30% for* $\gamma = 2$.

**Economic**

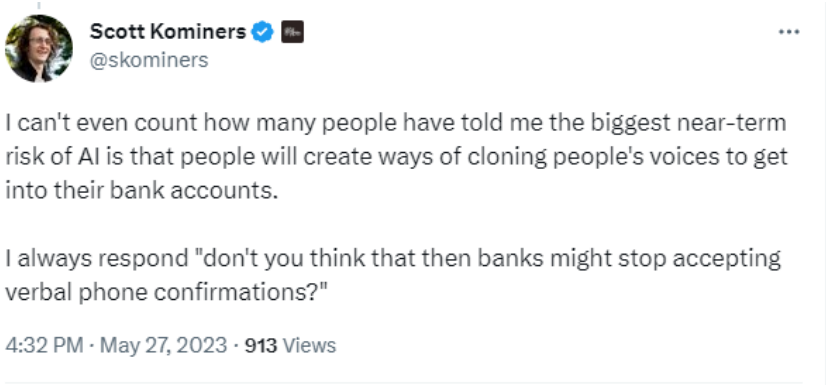
- Choose
  - Consumption: $c = c_0 e^{gt}$ — growth at exogenous rate $g$, e.g. 10% per year
  - Existential risk: Probability of survival is $S(T) \equiv e^{-\delta T}$.
- Simplify so the model is essentially static:
  - All growth and x-risk occurs immediately
  - If survive, consume constant $c_T$ forever
- $N$ people $\Rightarrow$ social welfare

$$U = \int_0^\infty e^{-\rho t} N u(c) dt = \frac{1}{\rho} N u(c)$$

**“Should we let machines flood our information channels with propaganda and untruth?”**

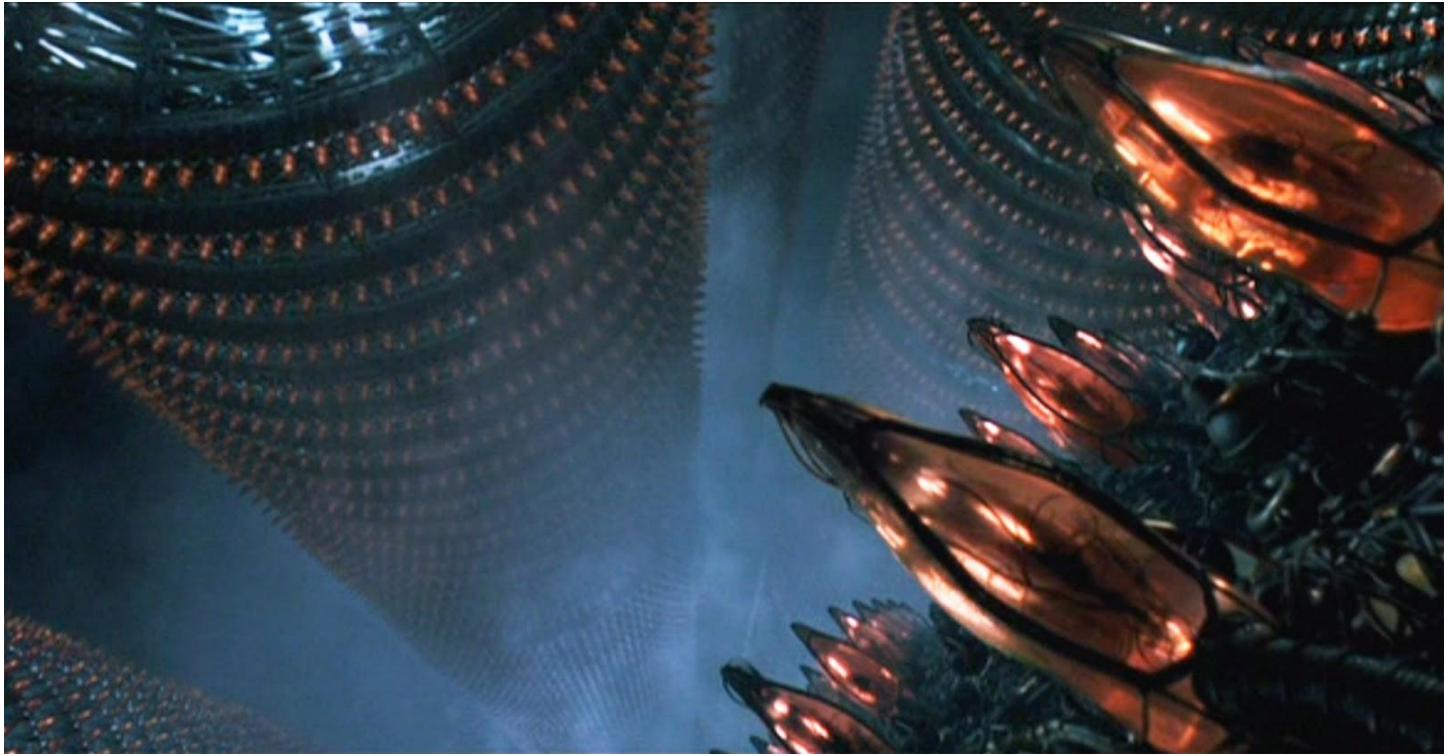
# “Should we let machines flood our information channels with propaganda and untruth?”

- The economics are more complicated than popular discourse suggests. Lots of big open questions. Little research to date.



**“Should we automate away all the jobs, including the fulfilling ones?”**

# Wrong question!



9

# Artificial Intelligence and Economic Growth

Philippe Aghion, Benjamin F. Jones, and Charles I. Jones

## 9.1 Introduction

This chapter considers the implications of artificial intelligence for economic growth. Artificial intelligence (AI) can be defined as "the capability of a machine to imitate intelligent human behavior" or "an agent's ability to

"Baumol (1967) observed that sectors with rapid productivity growth, such as agriculture and even manufacturing today, often see their share of gross domestic product (GDP) decline while those sectors with relatively slow productivity growth—perhaps including many services—experience increases. As a consequence, economic growth may be constrained not by what we do well but rather by what is essential and yet hard to improve."

7

# Artificial Intelligence, Income, Employment, and Meaning

Betsey Stevenson

The evolution of artificial intelligence (AI) evokes strong emotions in people. Some imagine a dystopia in which people are replaced by machines. Machines will develop the content we read, and the entertainment we enjoy. Artificial intelligence will pick our friends and our politicians, and ultimately take away any sense of human agency. And worst of all, those machines

"There are really two separate questions: there is an employment question, in which the fundamental question is, can we find fulfilling ways to spend our time if robots take our jobs? And there is an income question, can we find a stable and fair distribution of income?."

# WHY MIGHT INEQUALITY INCREASE?

# The Global Decline of the Labor Share*

Loukas Karabarbounis, Brent Neiman

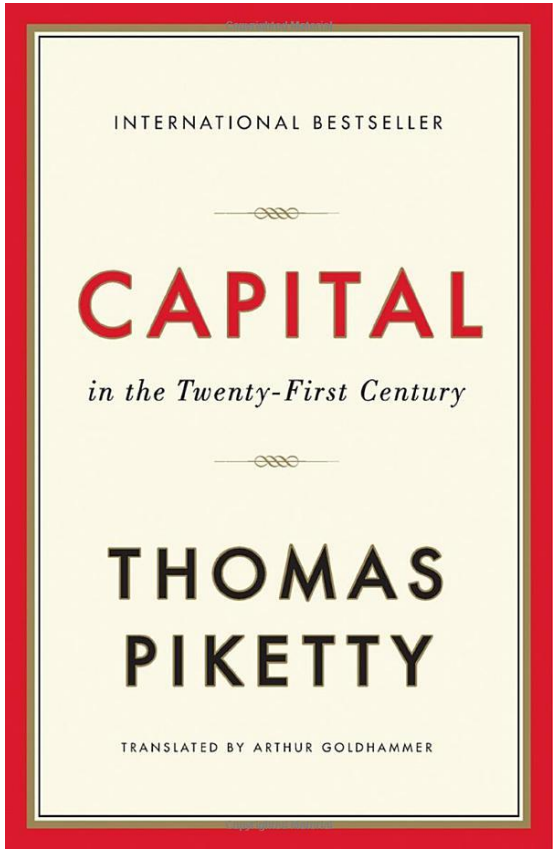
*The Quarterly Journal of Economics*, Volume 129, Issue 1, February 2014, Pages 61–103,
https://doi.org/10.1093/qje/qjt032
**Published:** 24 October 2013



**Abstract**

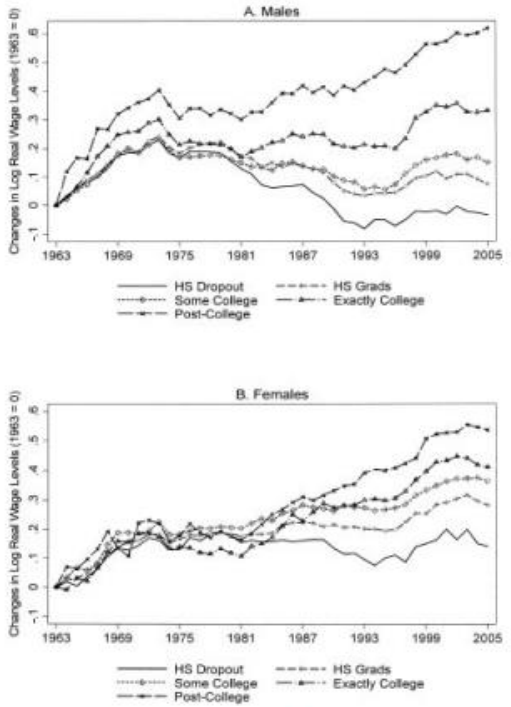
The stability of the labor share of income is a key foundation in macroeconomic models. We document, however, that the global labor share has significantly declined since the early 1980s, with the decline occurring within the large majority of countries and industries. We show that the decrease in the relative price of investment goods, often attributed to advances in information technology and the computer age, induced firms to shift away from labor and toward capital. The lower price of investment goods explains roughly half of the observed decline in the labor share, even when we allow for other mechanisms influencing factor shares, such as increasing profits, capital-augmenting technology growth, and the changing skill composition of the labor force. We highlight the implications of this explanation for welfare and macroeconomic dynamics.

**JEL:** E21 - Consumption; Saving; Wealth, E22 - Investment; Capital; Intangible Capital; Capacity, E25 - Aggregate Factor Income Distribution



# Computing and the internet increased inequality

FIGURE 5.—TRENDS IN COMPOSITION-ADJUSTED REAL LOG WEEKLY FULL-TIME WAGES BY GENDER AND EDUCATION, 1963–2005 (MARCH CPS)




See notes to table 1 for details on samples and data processing.

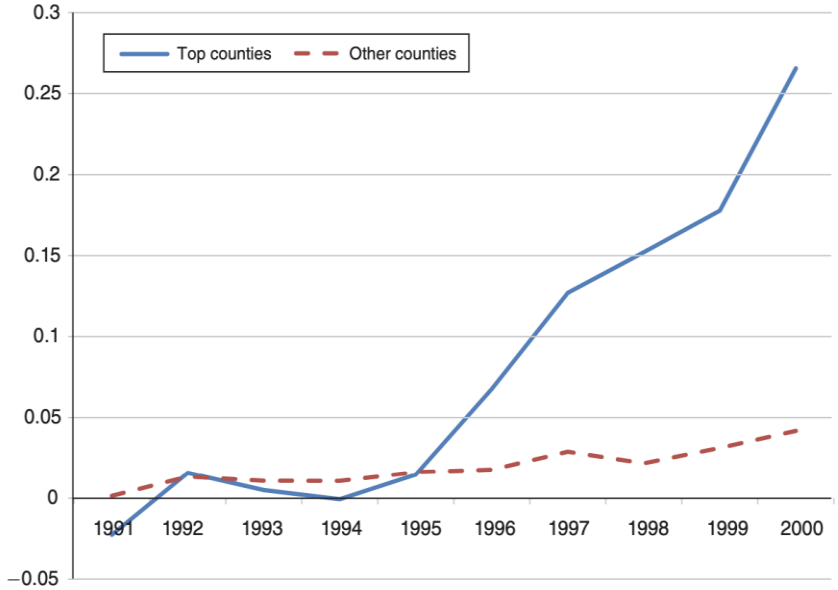
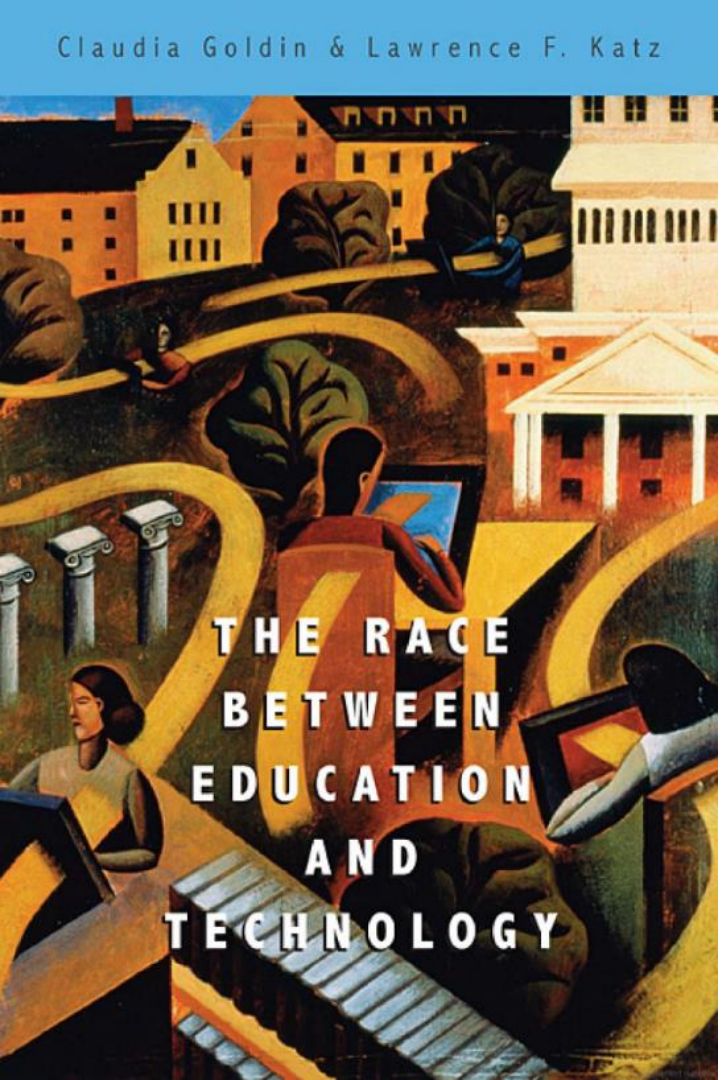



FIGURE 2. MARGINAL EFFECT OF ADVANCED INTERNET YEAR-BY-YEAR IN TOP COUNTIES




# TRENDS IN U.S. WAGE INEQUALITY: REVISING THE REVISIONISTS

David H. Autor, Lawrence F. Katz, and Melissa S. Kearney*

*Abstract*—A recent "revisionist" literature characterizes the pronounced rise in U.S. wage inequality since 1980 as an "episodic" event of the first half of the 1980s driven by nonmarket factors (particularly a falling real minimum wage) and concludes that continued increases in wage inequality since the late 1980s substantially reflect the mechanical confounding effects of changes in labor force composition. Analyzing data from the Current Population Survey for 1963 to 2005, we find limited support for these claims. The slowing of the growth of overall wage inequality in the 1990s hides a divergence in the paths of upper-tail (90/50) inequality—which has increased steadily since 1980, even adjusting for changes in labor force composition—and lower-tail (50/10) inequality, which rose sharply in the first half of the 1980s and plateaued or contracted thereafter. Fluctuations in the real minimum wage are not a plausible explanation for these trends since the bulk of inequality growth occurs above the median of the wage distribution. Models emphasizing rapid secular growth in the relative demand for skills—attributable to skill-biased technical change—and a sharp deceleration in the relative supply of college workers in the 1980s do an excellent job of capturing the evolution of the college/high school wage premium over four decades. But these models also imply a puzzling deceleration in relative demand growth for college workers in the early 1990s, also visible in a recent "polarization" of skill demands in which employment has expanded in high-wage and low-wage work at the

This literature reaches two broad conclusions. First, much of the rise in U.S. earnings inequality during the 1980s appears to be explained by shifts in the supply of and demand for skills combined with the erosion of labor market institutions—including labor unions and the minimum wage—that protected the earnings of low- and middle-wage workers.[2] Second, a number of influential studies argue that the surge of inequality evident in the 1980s reflected an ongoing, secular rise in the demand for skill that commenced decades earlier and perhaps accelerated during the 1980s with the onset of the computer revolution. When this secular demand shift met with an abrupt slowdown in the growth of the relative supply of college-equivalent workers during the 1980s—itself a consequence of slowing educational attainment for cohorts born after 1949 and of smaller entering labor force cohorts—wage differentials expanded



SPECIAL ISSUE REVIEW

# Skills, education, and the rise of earnings inequality among the "other 99 percent"

DAVID H. AUTOR Authors Info & Affiliations

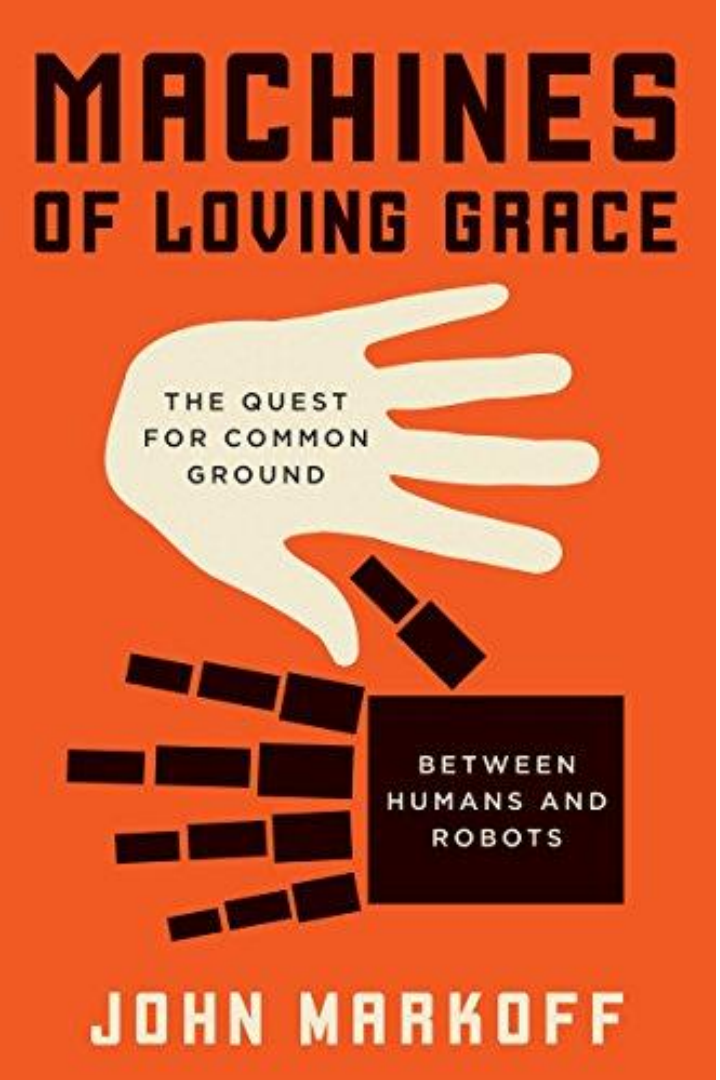
SCIENCE · 23 May 2014 · Vol 344, Issue 6186 · pp. 843-851 · DOI: 10.1126/science.1251868



## Abstract

The singular focus of public debate on the "top 1 percent" of households overlooks the component of earnings inequality that is arguably most consequential for the "other 99 percent" of citizens: the dramatic growth in the wage premium associated with higher education and cognitive ability. This Review documents the central role of both the supply and demand for skills in shaping inequality, discusses why skill demands have persistently risen in industrialized countries, and considers the economic value of inequality alongside its potential social costs. I conclude by highlighting the constructive role for public policy in fostering skills formation and preserving economic mobility.

# MUST INEQUALITY INCREASE WITHOUT REDISTRIBUTION?




Introduction:
One group designed powerful machines that allow humans to perform previously unthinkable tasks, like programming robots for space exploration, while the other works to replace humans with machines, like the developers of artificial intelligence robots to perform the work of doctors and lawyers.

Conclusion:
The solution to the contradiction inherent in AI versus IA lies in the very human decisions of engineers and scientists…who all have intentionally chosen human-centered design.

# The Turing Trap: The Promise & Peril of Human-Like Artificial Intelligence

*Erik Brynjolfsson*

*In 1950, Alan Turing proposed a test of whether a machine was intelligent: could a machine imitate a human so well that its answers to questions were indistinguishable from a human's? Ever since, creating intelligence that matches human intelligence has implicitly or explicitly been the goal of thousands of researchers, engineers, and entrepreneurs. The benefits of human-like artificial intelligence (HLAI) include soaring productivity, increased leisure, and perhaps most profoundly a better understanding of our own minds. But not all types of AI are human-like – in fact, many of the most powerful systems are very different from humans – and an excessive focus on developing and deploying HLAI can lead us into a trap. As machines become better substitutes for human labor, workers lose economic and political bargaining power and become increasingly dependent on those who control the technology. In contrast, when AI is focused on augmenting humans rather than mimicking them, humans retain the power to insist on a share of the value created. What is more, augmentation creates new capabilities and new products and services, ultimately generating far more value than merely human-like AI. While both types of AI can be enormously beneficial, there are currently excess incentives for automation rather than augmentation among technologists, business executives, and policy-makers.*

A good start would be to replace the Turing Test, and the mindset it embodies, with a new set of practical benchmarks that steer progress toward AI-powered systems that exceed anything that could be done by humans alone.

SCIENCE

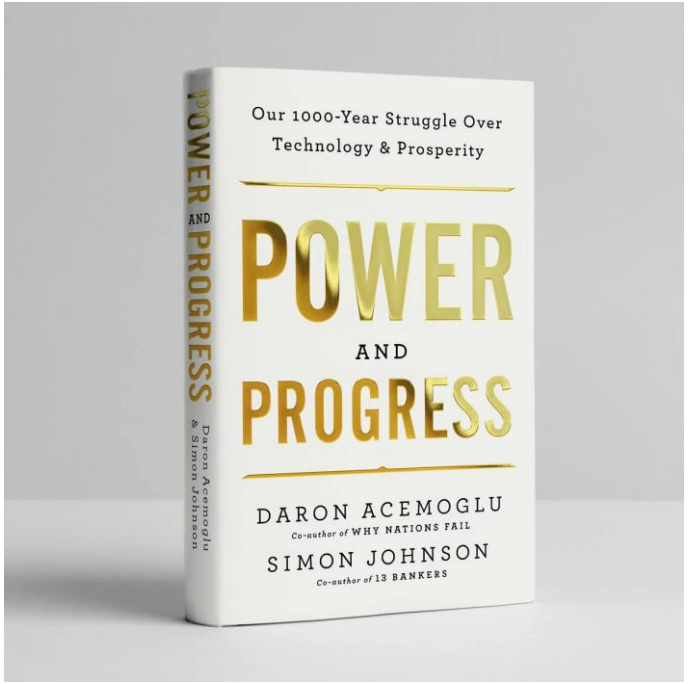
Forum BOSTON REVIEW

# AI's Future Doesn't Have to Be Dystopian

AI can be used to increase human productivity, create jobs and shared prosperity, and protect and bolster democratic freedoms—but only if we modify our approach.

Daron Acemoglu

Democracy, Economy, Politics, Redesigning AI, Science and Technology




Current developments, such as they are, go in the direction of automating teachers–for example, by implementing automated grading or online resources to replace core teaching tasks. But **AI** could also **revolutionize** education by **empowering** teachers to adapt their material to the needs and attitudes of diverse students in real time. We **already know** that what works for one individual in the classroom may not work for another; different students find different elements of learning challenging. AI in the classroom can make teaching more adaptive and student-centered, generate distinct new teaching tasks, and, in the process, increase the productivity of—and the demand for—teachers.



- Some of their writing suggests that they want to change the objectives and philosophy of the entire research field.

- The underlying hypothesis is that if the technical objectives of AI research are changed, then this will steer the economy away from potential loss of jobs, devaluation of skills, inequality, and social discord following from this.

- In this way, society can avoid what Brynjolfsson calls the "Turing Trap", where AI-enabled automation leads to a concentration of wealth and power.

POLICY FORUM

# Do we want less automation?

## AI may provide a path to decrease inequality

AJAY AGRAWAL, JOSHUA S. GANS, AND AVI GOLDFARB Authors Info & Affiliations

*SCIENCE* • 13 Jul 2023 • Vol 381, Issue 6654 • pp. 155-158 • DOI: 10.1126/science.adh9429



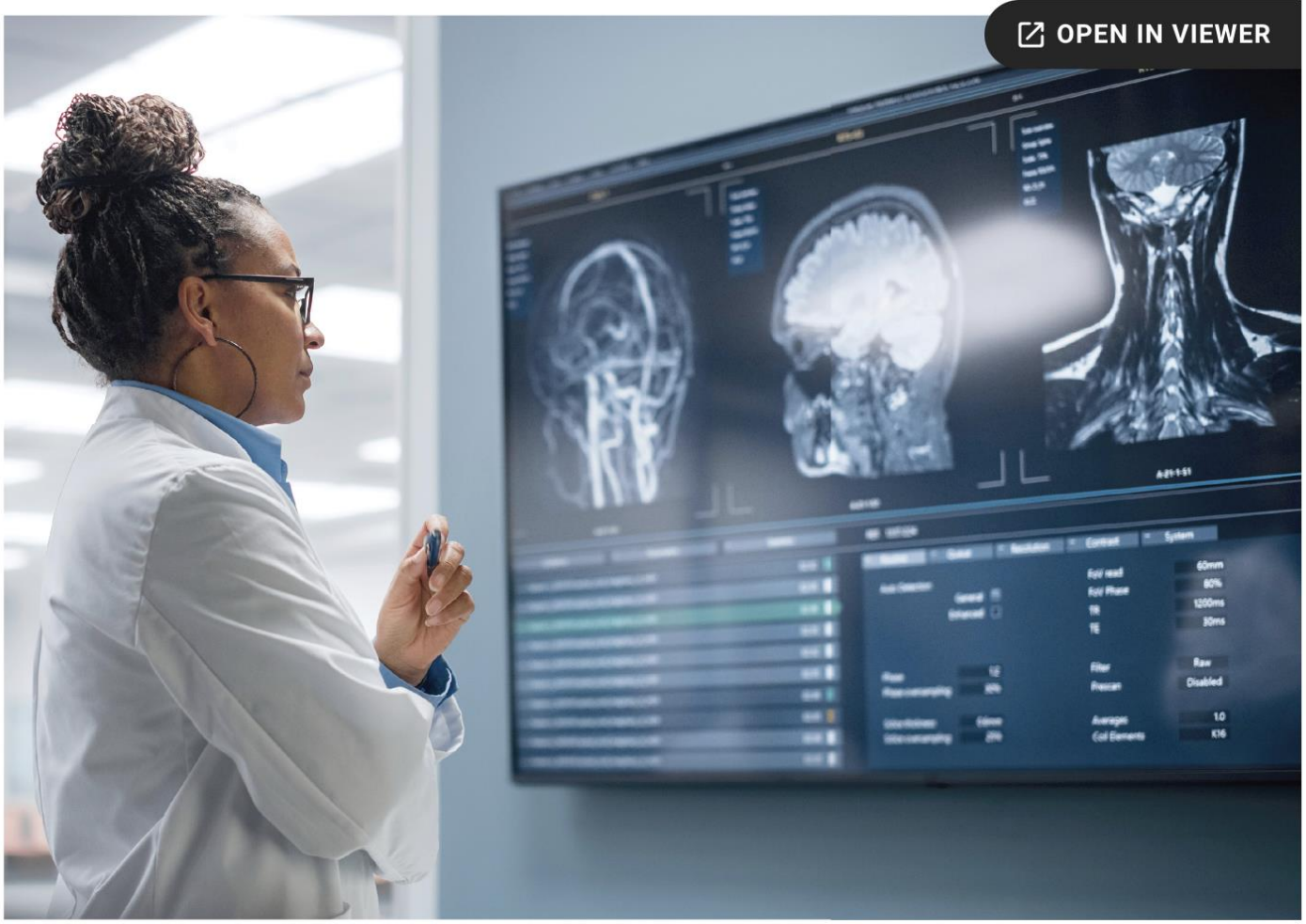
Impressive achievements made through artificial intelligence (AI) innovations in automating the tasks required in many jobs have reinforced concerns about labor market disruption and increased income inequality. This has motivated calls for change in the direction of AI innovation from being guided by task automation to instead focusing on labor augmentation (*1*). But task automation and labor augmentation are not polar opposites. Instead, automation of some tasks can lead to augmentation of labor elsewhere. Furthermore, AI automation may provide a path to reversing the trend of increasing income inequality by enabling disproportionate productivity improvements for lower-wage workers, allowing them to perform at levels that would previously require years of education and experience.

# GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models

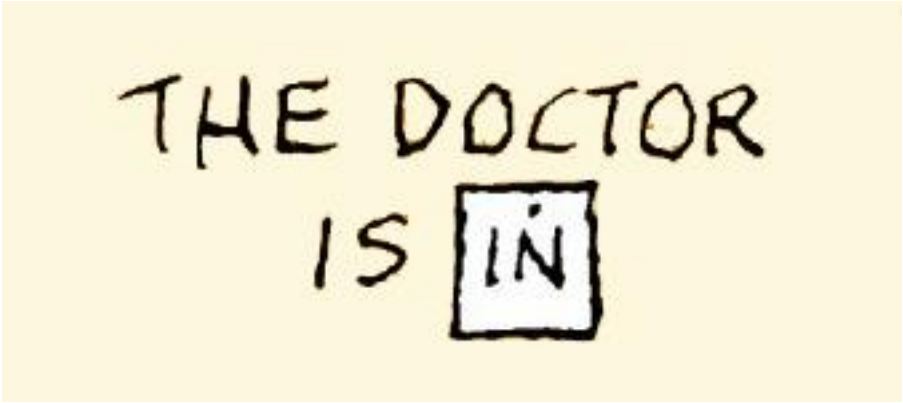
Tyna Eloundou, Sam Manning, Pamela Mishkin, Daniel Rock

We investigate the potential implications of large language models (LLMs), such as Generative Pre-trained Transformers (GPTs), on the U.S. labor market, focusing on the increased capabilities arising from LLM-powered software compared to LLMs on their own. Using a new rubric, we assess occupations based on their alignment with LLM capabilities, integrating both human expertise and GPT-4 classifications. Our findings reveal that around 80% of the U.S. workforce could have at least 10% of their work tasks affected by the introduction of LLMs, while approximately 19% of workers may see at least 50% of their tasks impacted. We do not make predictions about the development or adoption timeline of such LLMs. The projected effects span all wage levels, with higher-income jobs potentially facing greater exposure to LLM capabilities and LLM-powered software. Significantly, these impacts are not restricted to industries with higher recent productivity growth. Our analysis suggests that, with access to an LLM, about 15% of all worker tasks in the US could be completed significantly faster at the same level of quality. When incorporating software and tooling built on top of LLMs, this share increases to between 47 and 56% of all tasks. This finding implies that LLM-powered software will have a substantial effect on scaling the economic impacts of the underlying models. We conclude that LLMs such as GPTs exhibit traits of general-purpose technologies, indicating that they could have considerable economic, social, and policy implications.

Subjects: **General Economics (econ.GN)**; Artificial Intelligence (cs.AI); Computers and Society (cs.CY)
Cite as: arXiv:2303.10130 **[econ.GN]**
(or arXiv:2303.10130v4 **[econ.GN]** for this version)
https://doi.org/10.48550/arXiv.2303.10130




AI-driven diagnostic tools, with appropriate policy support, hold potential to expand opportunities and boost wages for nonphysician health care workers. PHOTO: GORODENKOFF/ISTOCK




760,000 jobs for physicians and surgeons in the US in 2021
Earning a median income of $200,000/year

3 million jobs for registered nurses, with median income $77,600
Millions more pharmacists, physician assistants, and paramedics

Center on Regulation and Markets at BROOKINGS

POLICY BRIEF | JUNE 2023

# THE TURING TRANSFORMATION:

## ARTIFICIAL INTELLIGENCE, INTELLIGENCE AUGMENTATION, AND SKILL PREMIUMS

AJAY AGRAWAL, JOSHUA GANS, AND AVI GOLDFARB

Almon Brown Strowager, an American undertaker from the 19th century, allegedly angry that a local switch operator (and wife of a competing undertaker) was redirecting his customer calls to her husband, sought to take all switch operators to their employment graves. He conceived of and, with family members, invented the Strowager switch that auto-

including a Bar exam, the SAT, and various AP-level courses. AI pioneer and Turing Award winner Geoff Hinton remarked in 2016 that time was up for radiologists and that no one should continue training in that field. Whether that will hold true or not, it is hardly surprising that recent developments in AI have reinforced the widespread view that the intent of AI research is to re-

"One worker's automation is another's augmentation. Automation of rare high value skills can mean augmentation for everyone else. Similarly, augmentation that complements the lucky humans with rare high value skills can mean increased inequality and a hollowing out of the middle class."

- The first 50 years of computing contain many technologies that appear to be intelligence augmenting, creating new capabilities and new products and services.

- The last 10 years have seen a rise in artificial intelligence applications, whose inventors directly aspire to automate tasks currently performed by humans.

- The apparently augmenting technologies appear to have increased inequality.

- But one person’s automation is another’s augmentation.

- Perhaps the automating technologies will decrease inequality, depending on whose work gets automated and whose gets augmented.

future of life INSTITUTE

# Pause Giant AI Experiments: An Open Letter

We call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4.

*View this open letter online*.

| Published | PDF created | Signatures |
| --- | --- | --- |
| March 22, 2023 | May 5, 2023 | 27565 |

AI systems with human-competitive intelligence can pose profound risks to society and humanity, as shown by extensive research[1] and acknowledged by top AI labs.[2] As stated in the widely-endorsed Asilomar AI Principles, *Advanced AI could represent a profound change in the history of life on Earth, and should be planned for and managed with commensurate care and resources.* Unfortunately, this level of planning and management is not happening, even though recent months have seen AI labs locked in an out-of-control race to develop and deploy ever more powerful digital minds that no one – not even their creators – can understand, predict, or reliably control.

"Should we risk loss of control of our civilization?" Should we develop nonhuman minds that might eventually outnumber, outsmart, obsolete and replace us?"

"Should we let machines flood our information channels with propaganda and untruth?"

"Should we automate away all the jobs, including the fulfilling ones?"

# Open questions

- Will AI lead to a large improvement in productivity?
- If it does, which forces dominate with respect to inequality?
- What does equilibrium look like when fake images, sounds, and videos are easy to create?
- How soon, and under what circumstances, should we be concerned about market power?