NBER WORKING PAPER SERIES

A RESEARCH AGENDA FOR THE ECONOMICS OF TRANSFORMATIVE AI

Erik Brynjolfsson
Anton Korinek
Ajay K. Agrawal

Working Paper 34256
http://www.nber.org/papers/w34256

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2025

We thank Basil Halperin and Philip Trammell for their excellent research assistance on this draft, Matty Smith for editorial assistance, and Guy Ben-Ishai, Diane Coyle, Sarah Friar, Christie Ko, Sam Manning, James Manyika, Albert Wenger and numerous participants in the March 2024 Asilomar Workshop on the Economics of Transformative AI for valuable discussions and suggestions. The Stanford Digital Economy Lab, CSH Vienna, and the Centre for Innovation and Entrepreneurship at the University of Toronto's Rotman School of Management provided funding and support. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w34256

A Research Agenda for the Economics of Transformative AI
Erik Brynjolfsson, Anton Korinek, and Ajay K. Agrawal
NBER Working Paper No. 34256
September 2025
JEL No. A11, O33, O40

ABSTRACT

As we approach Transformative Artificial Intelligence (TAI), there is an urgent need to advance our understanding of how it could reshape our economic models, institutions and policies. We propose a research agenda for the economics of TAI by identifying nine Grand Challenges: economic growth, innovation, income distribution, decision-making power, geoeconomics, information flows, safety risks, human well-being, and transition dynamics. By accelerating work in these areas, researchers can develop insights and tools to help fulfill the economic potential of TAI.

Erik Brynjolfsson
Stanford University
and NBER
erik.brynjolfsson@gmail.com

Anton Korinek
University of Virginia
Department of Economics and EconTAI
and NBER
akorinek@virginia.edu

Ajay K. Agrawal
University of Toronto
Rotman School of Management
and NBER
ajay.agrawal@rotman.utoronto.ca

1. Motivation

AI capabilities have improved radically in recent years. Our institutions, organizations, skills, and economic models are struggling to keep pace. In this growing gap lie the greatest risks of the coming decade, as well as the greatest opportunities. We urgently need to improve our understanding of AI's economic implications.

Before we can formulate effective solutions, we must define the fundamental questions, laying the groundwork for a productive research agenda. By doing so, we seek to catalyze the insights and analytical tools necessary to navigate AI's evolving landscape while fostering shared prosperity. A question well-posed is half answered.

As rapid as improvements in AI have been, from mastering Ph.D.-level exams to writing essays and creating art, there are reasons to believe bigger advances will occur within the next few years. Timelines for achieving future AI milestones have accelerated. In 2020, the median prediction on Metaculus for the arrival date of a "general AI system" that could outperform most humans was 2062. As of this writing, the median prediction is 2033, while a quarter of participants expect it to be achieved by 2028 (Barnett, 2020). Dario Amodei recently wrote that AI that exceeds almost all humans in almost all cognitive tasks might appear as soon as 2026 (Amodei, 2024), and Demis Hassabis said he thought that this was likely within the decade (Webber, 2025). While it is inherently difficult to predict dates for future inventions, we can't dismiss the possibility that extremely powerful AI systems will be available soon.

Such powerful AI systems would transform society. Even current AI technologies have the potential to affect large sectors of the economy. One analysis estimates that current generative AI systems will significantly affect at least half of the tasks of 19% of workers in the US economy, while 80% of jobs will have 10% or more of their tasks affected (Elondou et al., 2024).[1] Looking forward, a report from the National Academies on AI and the Future of Work estimates that productivity growth rates could double in the coming years as a result of these technologies (National Academies of Sciences, Engineering, and Medicine, 2024).

More capable forms of AI would have broader and deeper effects. For example, interactions among AI agents could reshape the foundations of economic activity as they negotiate contracts, make investment decisions, and drive consumer choices. This dynamic interplay between AI and human agents is poised to redefine the economic landscape in ways we are only beginning to grasp.

The economic and societal changes are likely to unfold slowly at first, then rapidly (Brynjolfsson et al., 2021). We urgently need a proactive research agenda to guide the redesign

[1] What's more, recent literature finds that these technologies are already significantly affecting productivity and work in specific applications like contact centers (e.g., Brynjolfsson et al., 2025) and freelance work (e.g., Hui et al., 2024).

of our economic institutions, norms, and policies. We need to understand how to realize TAI's potential while minimizing its risks.[2]

We present a research agenda here to motivate attention to critical questions about how TAI will impact the economy and societal well-being. Among the key concerns are how TAI will influence productivity and invention, as well as the extent to which it may exacerbate inequality or concentrate decision-making power. Other pressing issues include the interaction between TAI and global governance, implications for truth and misinformation, catastrophic risks, and ensuring that economic incentives align with broader social welfare.

Our agenda is relevant to all researchers and policymakers interested in the broader effects of AI on society. Economics provides powerful frameworks for analyzing how transformative technologies affect society through production processes, resource allocation, and market mechanisms. Economists study how technologies like AI change production functions—the relationship between inputs (labor, capital, data) and outputs—and how these changes ripple through markets, affecting productivity, wages, prices, and growth. Unlike technical analyses that focus on capabilities, economic analysis emphasizes societal outcomes: who benefits, what trade-offs emerge, and how institutions might adapt to technological change.

Research on the economic implications of TAI can help policymakers craft better policies. This includes redesigning labor laws, tax systems, social insurance programs, antitrust and market regulations, as well as updating macroeconomic policies, and global governance structures. A well-informed policy response will be essential to ensuring that TAI creates broad-based prosperity rather than exacerbating economic division.

We systematically analyze how TAI will affect economic processes at multiple levels, examining three interconnected dimensions: (1) how TAI affects innovation processes and the creation of new ideas; (2) how TAI reshapes the organization of production factors through new business models, market structures, and institutional arrangements; and (3) how TAI substitutes for or augments individual factors of production, particularly human labor and cognitive capabilities. This three-level approach helps structure our research questions and connects them to established economic methodologies.

[2] Not all economists agree that AI will be transformative. Acemoglu (2025) argues that AI will only be able to automate a limited fraction of tasks with little impact on growth, while Narayanan and Kapoor (2025) suggest AI may be a 'normal technology' whose diffusion will follow traditional patterns, with modest economic impact. These perspectives are valuable but they may underestimate the potential for AI to transform innovation processes themselves and create feedback loops that accelerate capability development. Moreover, even if the chances of AI fundamentally transforming the economy were small, the implications are large enough that it would still be wise to do scenario planning, just as we should be prepared for other important outcomes even if they are unlikely.

2. What TAI Means for the Economy

From an economic perspective, we define Transformative AI as artificial intelligence that enables a sustained increase in total factor productivity growth of at least 3 - 5x historical averages. Total factor productivity growth measures the rate at which an economy's output increases beyond what can be explained by additional inputs of labor and capital. Such growth may occur because AI facilitates a radical new set of goods, services, or production processes, because AI changes the relative scarcity of inputs, particularly by making cognitive labor significantly more abundant relative to other factors, or because AI creates novel economic organizations and institutions.

While transformative AI could take many forms, we describe a potential manifestation that Dario Amodei calls "powerful AI," which he summarizes as "a country of geniuses in a datacenter" (Amodei, 2024): an advanced AI model with intelligence surpassing the most capable humans across multiple fields, able to perform complex tasks like solving mathematical theorems, creating novel works, or directing experiments autonomously.[3]

Despite its capabilities, powerful AI could be constrained by several bottlenecks, including a lack of physical embodiment, limits on the response times of physical systems or external software, requirements for computational infrastructure to support its large-scale operation, and policies. Such bottlenecks could hinder the speed and scale of its economic impact. TAI's actual form may differ from the specific instantiation we describe here.

It is useful to monitor both technological and economic indicators that capture complementary aspects of the transition. We list both sets of indicators in the appendix. The most significant economic indicators of TAI reflect how AI affects economic activity, represented by measures such as productivity, real gross domestic product (GDP), or real GDP/capita. Some observers define TAI via threshold indicators for annual output growth or productivity growth—for example, one metric suggests that TAI would be marked by annual growth of real output of at least 30%, amounting to roughly a ten-fold increase in the growth rate from the average of the Industrial Age (Davidson, 2023). Mustafa Suleyman (2023) defines an "Artificially Capable Intelligence" as an AI system that can transform $10,000 into $1 million within a specific time frame.

3. Grand Challenges and Key Questions

To effectively prepare for the development of TAI, we must first understand the key economic questions and challenges it presents. We identify nine Grand Challenges where economic analysis can shape TAI's trajectory and impact. For each challenge, we define several key research questions that need to be addressed, listed in the following box:

---

[3] The related concepts of Artificial General Intelligence (AGI), Human-Level AI (HLAI), and Strong AI differ somewhat but are often used interchangeably. All of them are likely to yield transformative economic implications.

Economic Growth

1. *How can TAI change the rate and determinants of economic growth?*
2. *What will be the main bottlenecks for growth?*
3. *How can TAI affect the relative scarcity of inputs including labor, capital and compute?*
4. *How will the role of knowledge and human capital change?*
5. *What new types of business processes and organizational capital will emerge?*

Invention, Discovery and Innovation

6. *For what processes and techniques will TAI boost the rate and direction of invention, discovery, and innovation?*
7. *Which fields of innovation and discovery will be most affected and what breakthroughs could be achieved?*

Income Distribution

8. *How could TAI exacerbate or reduce income and wealth inequality?*
9. *How could TAI affect labor markets, wages and employment?*
10. *How might TAI interact with social safety nets?*

Concentration of Decision-making and Power

11. *What are the risks of AI-driven economic power becoming concentrated in the hands of a few companies, countries or other entities?*
12. *How might AI shift political power dynamics?*

Geoeconomics

13. *How could AI redefine the structure of international relations, including trade, global security, economic power and inequality, political stability, and global governance?*

Information, Communication, and Knowledge

14. *How can truth vs. misinformation, cooperation vs. polarization, and insight vs. confusion be amplified or dampened?*
15. *How can TAI affect the spread of information and knowledge?*

AI Safety & Alignment

16. *How can we balance the economic benefits of TAI with its risks, including catastrophic and existential risks?*
17. *What can economists contribute to help align TAI with social preferences and welfare?*

Meaning and Well-being

18. *How can people retain their sense of meaning and worth if "the economic problem is solved" as Keynes predicted?*
19. *What objectives should we direct TAI to help us maximize?*

Transition Dynamics

20. *How does the speed mismatch between TAI and complementary factors affect the rollout of TAI and how can adjustment costs be minimized?*
21. *How can societies prepare for and respond to potential transition crises, e.g., sudden mass unemployment, system failures, or conflicts triggered by TAI developments?*

We expand on each of these Grand Challenges in the ensuing nine subsections and pose additional research questions. Our goal is to inspire the research community to tackle these challenges and questions.

3.1 Economic Growth

Improved technological capabilities are the key drivers of growth in standard economic models (e.g., Aghion et al., 2019). How can economists detect early signs of an AI-driven "growth explosion" improving upon approaches like those of Nordhaus (2021)?

The nature of growth constraints is likely to change significantly. Current bottlenecks, such as limited supply of workers, human capital, and physical capital, may evolve or be supplanted by new limiting factors. Which bottlenecks will gain prominence in a TAI-driven economy, and how will these differ from traditional constraints? Will energy availability, computational resources, or raw materials become key constraints? Current frontier models require significant energy resources, rare earth minerals, and water for cooling. Research is needed to quantify these ecological constraints and model how they might affect the diffusion and economic impact of TAI.

The role of human capital is likely to undergo significant changes. How will the value and relevance of human capital evolve as TAI is developed? Which human skills will become redundant and which are likely to remain in demand as TAI capabilities grow? How should education and training systems adapt to prepare workers?

The diffusion process will be crucial in determining how the benefits of TAI are realized and distributed. What factors will influence the rate of TAI adoption, and how will these affect economic growth? What policies could promote optimal diffusion of TAI?

3.2 Invention, Discovery and Innovation

Since innovation is the primary driver of economic growth, it is important to understand how TAI may transform the nature and extent of innovation.

How and where will TAI automate scientific discovery? Traditional innovation processes are costly and constrained by the time required to formulate hypotheses, conduct experiments, and iterate on solutions. TAI systems may dramatically reduce costs and time involved. How will the ability to automate experimentation and problem-solving at scale influence the rate of technological progress? What are the likely bottlenecks? What will be the impact on the frequency and quality of innovations, and, consequently, the rate of economic growth?

TAI may also affect the extent of disciplinary and non-disciplinary innovation. Many of the most transformative innovations emerge from the integration of knowledge across distinct fields, yet human cognitive limitations often prevent experts from identifying novel connections between disciplines. Many discovery-based institutions (e.g., universities and institutes) today are organized around disciplines, including training, promotion, and dissemination. To what extent will TAI, equipped with access to cross-disciplinary data and capable of synthesizing disparate knowledge, identify previously unrecognized complementarities? By generating insights that bridge domains, how might TAI impact not only the rate of innovation but the direction by shifting the innovation process from discovering local maxima (disciplinary discoveries) to more global discoveries (not constrained by disciplines)? What are the implications for economic growth?

Today, innovation is concentrated in firms and institutions with substantial resources and expertise. Will TAI democratize the innovation process? Will it allow small firms, entrepreneurs, and individuals to perform sophisticated innovation activities, including R&D, previously out of reach? Although humans may still need to articulate desired outcomes, will TAI be able to do most of the rest of the work of the innovation process? What will be the new bottlenecks?

To what extent will TAI shift the distribution of the agents that engage in the innovation process? Or will the production of innovation shift from human brains to compute? How will this impact the rate and direction of innovation?

What will be the marginal value of intelligence itself? In what domains will frontier models be most valuable compared to less sophisticated types of AI?

3.3 Income Distribution

Labor is the main source of income for the majority of the population, and labor markets therefore play a crucial role in income distribution. Technological progress has traditionally gone hand in hand with both job displacement and the creation of new work, although not always in balance (Acemoglu & Restrepo, 2019; Autor et al., 2022). Advances in AI may be far more rapid and transformative than earlier technologies (Korinek & Stiglitz, 2019). Simple economics suggests that if a machine is a close substitute for a worker in a job, the worker's market wage will tend to fall to the cost of having a machine do the same tasks.

Will the capabilities of TAI largely displace workers, or will there be areas with rising labor demand? How might changing bottlenecks affect the distribution of economic gains across sectors and populations? How will this depend on economic policy?

If machines can perform essentially all work tasks, the remaining jobs may be either transitional or may involve demand for labor for intrinsic human-centered reasons (Korinek, 2024). What will those jobs look like, including the balance between cognitive and physical work? What are the implications for equilibrium wages, employment levels, and unemployment rates?

If compute and robots become close substitutes for labor, will they remain sufficiently scarce and expensive to preserve a significant role for labor?

The purveyors of TAI may earn significant gains from their inventions. Will this lead to an increase in income inequality in relative terms? Or will a growing economy lift all boats and ensure that the remaining labor demand is sufficient to spread around the surplus equitably? And how will these developments affect the concentration and inequality of wealth? How will these questions depend on whether AI systems are proprietary or open source?

Our existing social safety nets are designed around the notion that labor is the most important source of income. Social security and health benefits depend on people's jobs or labor income. There is disability insurance for people who can't work, and unemployment insurance for people who lose their jobs. Will our social safety nets still perform their role effectively in a TAI world? If not, then how can they be adapted? What mechanisms of social insurance and income distribution can we design to ensure that TAI's benefits are shared broadly across society?

3.4 Concentration of Decision-making and Power

There's an ongoing debate about concentration in the AI industry itself. The success of ever-larger models suggests the possibility that the AI industry may become increasingly concentrated, while the success of low-cost models of nearly-equivalent performance and the success of open source models could support increased competition (Korinek & Vipra, 2025). Will TAI be dominated by a single AI system, a small number of comparable systems or a plethora of systems with varying capabilities and strengths?

Perhaps more importantly, how will TAI affect concentration in the rest of the economy? TAI and digital technologies change the cost of information processing, and as machine learning enables rapid analyses and application of previously tacit knowledge, economic organizations and institutions will also change. For instance, will larger retailers and manufacturers gain an increased competitive advantage over small single-unit shops and plants? Will centrally-planned economies find new life? Or will AI democratize expertise and lead to flourishing competition? More subtly, which types of information and decisions, if any, will become more centralized, and which types will become more decentralized, and under what conditions?

Relatedly, if TAI eviscerates demand for labor and drives down wages, as is contemplated in the preceding section, what will be the effects on distribution on economic bargaining power? Will capital owners, who tend to be more concentrated than labor, become more powerful? Will governments or technologists or another group ascend, and if so, how will they be governed?

As TAI enables more autonomous digital agents to participate in economic activities, questions of human agency and control become economically significant. Research is needed on the efficiency and welfare implications of delegating economic decisions to AI systems, the optimal division of control rights between humans and machines, and the design of mechanisms that preserve human autonomy while leveraging AI capabilities.

Furthermore, economic concentration often leads to concentration of political power. So changes in the locus of economic decision-making may ripple through to other parts of the economy. In particular, even if a highly centralized system optimizes more types of economic decision-making, then will it at the same time erode individual autonomy and liberty?

The political economy of AI regulation warrants economic analysis. Key questions include how concentrated AI development affects regulatory capture risks, how different regulatory approaches might impact innovation rates and diffusion patterns, and what governance structures could balance innovation incentives with societal risk management. Game-theoretic approaches could illuminate the strategic interactions between developers, regulators, and competing nations

3.5 Geoeconomics

Geoeconomics is an emerging field which examines the use of a country's economic strength to exert influence on foreign entities to achieve geopolitical or economic goals. It involves leveraging economic instruments like trade policy, investment, and sanctions to advance national interests.

TAI may have profound implications for geoeconomics, including (1) the military applications and their effect on global security, (2) shifts in economic power due to productivity gains that are not uniformly distributed across states, (3) international governance and regulatory frameworks, and (4) challenges to political stability and information control within countries.

How will TAI reshape the economics of deterrence and the balance of power among states? How will TAI impact the stability of economic and military alliances and rivalries? Could TAI challenge traditional notions of nation-state power and enable small actors to wield significant influence, for instance, by creating an autonomous "army”?

How will TAI change the economics of cyber warfare and the defense of critical infrastructure? Can regulatory frameworks manage the dual-use nature of AI technologies without hindering economic growth?

Differential adoption of TAI could also exacerbate global economic inequalities between nations and affect trade flows. Will existing industrial capacity, research infrastructure, or policy frameworks determine which countries capture the most value? How might TAI reshape global trade patterns, particularly the competitive position of emerging economies? Will TAI alter the geopolitical importance of industries such as manufacturing, agriculture, and services, and how might trade policies evolve in response?

What policies can countries with limited technological capacity adopt to avoid marginalization in a TAI-driven global economy? How might global inequality affect the economic and political influence of technologically lagging states? What role can foreign aid or technology-sharing initiatives play in closing the TAI gap between countries? How will TAI-induced inequality affect migration patterns, regional stability, and global cooperation? How will TAI influence cooperation on shared challenges like climate change or public health crises?

International governance frameworks for TAI could address geopolitical risks and ensure that the technology evolves peacefully and benefits all countries. How can international institutions balance national sovereignty with coordinated oversight? How can the global community prevent regulatory arbitrage, where countries adopt divergent rules to gain competitive advantages? What factors influence whether major technology companies succeed in regulatory capture?

These geopolitical questions connect to broader discussions of digital sovereignty, including debates around digital tokens, data governance, and platform regulation. Economic research on TAI should build on existing work examining how digital technologies reshape political and economic power across borders.

3.6 Information, Communication, and Knowledge

A key determinant of a society's economic success is how it manages information, communication, and knowledge. Laws, institutions, incentives, and norms that promote the creation and transmission of accurate information tend to boost economic growth. For instance, the U.S. Constitution gives Congress the power to promote the progress of science and useful arts by granting authors and inventors exclusive rights to their writings and discoveries for a limited time. More broadly, the jury system, blind scientific reviews, libel laws, and the scientific method itself are systems intended to favor truthful information.

Digital information systems, communications, and social media are increasingly intermediated by AI-based recommender systems and populated by AI agents. In some contexts, these systems have been found to disproportionately promote misinformation, perhaps unintentionally (Vosoughi et al., 2018). How will TAI affect the quality of information flows? How can we ensure that the collection of information, performed for example by the media, is sufficiently compensated to create incentives for the production of high-quality information?

Relatedly, AI systems are creating content by themselves by processing information. How useful and innovative will this content be? Will it provide deeper and broader insights, even novelty? Or will it be misleading and destructive? AI-based deepfakes can misrepresent people and events. Could TAI simply overwhelm human-produced content with the sheer quantity of content it produces?

In the spirit of the preceding section, TAI may also affect political stability by enhancing the ability of states and non-state actors to manipulate information. What are its implications for authoritarian regimes, democratic governance, and social cohesion? How might TAI shape public opinion, disrupt political processes, or suppress dissent? Will TAI influence the economics of state surveillance practices and civil liberties? Understanding these dynamics will help policymakers harness TAI's benefits while mitigating its political risks.

Relatedly, TAI will likely transform both individual and collective intelligence in ways that require economic analysis. Research is needed on how AI affects information markets and knowledge production. These questions connect to foundational economic concepts of information asymmetry, public goods, and knowledge spillovers.

3.7 AI Safety & Alignment

AI safety and alignment refer to the challenge of ensuring that AI systems behave consistently with human values and intentions (Bengio et al., 2025). As AI grows more powerful and autonomous, the economic implications of their safety and alignment become crucial. How do the costs of AI safety and alignment compare to the economic benefits? What are the economic incentives for developing safe and aligned AI systems? And with whom should they be aligned?

Economists can develop frameworks and methodologies to align AI systems with social preferences and welfare (Korinek & Balwit, 2024). We have experience designing mechanisms for harmonizing the interests of agents with their principals. How can social welfare functions be adapted to capture the complexities of AI alignment? What economic mechanisms can be designed to internalize the positive and negative externalities generated by AI systems? How can we design incentive structures that encourage developers and users to prioritize alignment with broader societal goals? How can market mechanisms promote the development of safe and aligned AI systems? What tools can address the challenges of preference aggregation and value learning in AI systems?

One particular question concerns the trade-off between growth and safety (Jones, 2024). TAI systems may significantly increase economic growth but also affect the potential for catastrophic and even existential risks to humanity. What methods can economists develop to assess potential existential risks from TAI? If TAI also mitigates other risks or extends longevity, then how can we assess the trade-offs? Under what conditions is it rational to continue rapid AI progress or to slow or halt it?

An important factor to consider is the risk of AI race dynamics, whereby actors who benefit from being the first to develop higher capabilities prioritize speed over safety. For example, within labs, individual researchers may perceive career benefits to advancing more rapidly. Within nations, labs may race against each other to be the first to ship more powerful capabilities. And at the geopolitical scale, individual countries may race to outdo each other. In a worst-case scenario, race dynamics create existential risks. This is a classic externality problem, and economists have experience both in analyzing externalities and designing measures to internalize them. Another relevant question is how to address the tension between the economic benefits and the safety risks of open-source systems.

3.8 Meaning and Well-being

Keynes' (1931) prediction about solving the "economic problem" raises fundamental questions about human purpose and fulfillment in a TAI world. What can economics contribute to our understanding of meaning and well-being in a world without work? How can we analyze the production and distribution of non-monetary sources of fulfillment? Importantly, what is our ultimate objective in a world where machines can perform essentially all work? Is it desirable for work to maintain its current societal importance if we achieve TAI?

The psychological and social impacts of widespread labor displacement may present a complex challenge. Interestingly, some studies show that retirees experience increased happiness and life satisfaction, while the involuntarily unemployed tend to suffer decreased well-being. This raises several questions: What factors contribute to the positive experience of retirees versus the negative experience of the unemployed? How do societal expectations, financial security, and the voluntary nature of retirement influence these outcomes? Can we design economic policies that mimic the positive aspects of retirement for those displaced by TAI?

Will individuals make welfare-maximizing choices on how to spend their time if labor market opportunities decline? Or are there some aspects of work that are subject to externalities or internalities and that justify policy interventions to encourage work even if its economic value declines (Korinek & Juelfs, 2023)? For example, externalities may arise when an individual's work affects others in society beyond just producing marketable output, such as by fostering social connections or political stability. Internalities may occur when individuals don't fully internalize the effects of their work choices on their own welfare, such as effects on mental health and loneliness. In hindsight, it would have been useful to measure mental health in teens before, during, and after social media were introduced with a keener eye. How can we accurately measure and model these externalities and internalities? What policy interventions might be justified to address these market failures, and how would they differ from current labor market policies?

As we potentially transition to a post-work society, economics will play a crucial role in shaping new institutions to support human flourishing. What frameworks can help us analyze what provides meaning when TAI can handle most economic tasks? How might the distribution of meaning-generating activities be optimized for social welfare? Can TAI itself help create or facilitate new sources of meaning and fulfillment? How can we equitably distribute the benefits of TAI, not just in terms of material wealth, but also in terms of access to fulfilling activities and meaning?

3.9 Transition Dynamics

Optimizing policies and institutions for a world of TAI is not enough. We must also successfully navigate the transition from our current economic institutions, organizations, and processes. As technology advances, bottlenecks are likely to emerge. Furthermore, some complementary factors (e.g. human skills, organizational structures, regulatory frameworks) will remain fixed in the short term and only adjust over time. The success of the transition will depend in large part on identifying and managing these bottlenecks and complementary factors.

Transitional dynamics deserve particular attention as the speed of adjustment following TAI could determine the effects on labor markets, including the prevalence of technological unemployment. The mismatch between rapidly evolving AI capabilities and slower-moving complementary factors could create significant adjustment costs. In addition, how do policy interventions—such as targeted retraining subsidies, adaptive regulatory sandboxes, and organizational-innovation grants—differentially accelerate the adaptation of complementary factors and thereby minimize aggregate adjustment costs during the rollout of TAI? There are many pathways from our current pre-TAI equilibrium to the new post-TAI economy. Research that informs policy during this transition phase could significantly enhance social welfare.

4. Methodologies for the Economics of TAI

Economics provides a rich toolkit to analyze social science questions. Moreover, TAI may itself create new tools for economists. We lay out the main methodologies that we envision to make progress on the TAI research agenda that we outline above.

4.1 Theoretical Approaches

Economic theory plays a crucial role in understanding the potential economic impacts of TAI. Since TAI may represent a radical break from the past, one of the primary difficulties empirical researchers face is the scarcity of relevant historical data to extrapolate from. A significant amount of work in this field involves predicting a highly uncertain future. In this context, theoretical approaches that leverage higher-level regularities from the past—such as fundamental laws of economics—become especially valuable.

These approaches encompass a range of modeling techniques to understand economic shifts. Growth models, including those focusing on AI take-off dynamics, are at the forefront of this research. These models aim to capture non-linear and potentially explosive growth patterns that TAI might induce, helping to forecast scenarios of rapid technological advancement and its economic consequences. Complementing these are micro-to-macro approaches that bridge the gap between individual-level impacts of AI adoption and their aggregate effects on the broader economy.

A special focus within theoretical approaches is on normative frameworks, particularly those drawing from welfare theorems, social choice, and public finance principles. These frameworks address critical questions about the societal implications of TAI, such as how to equitably share the benefits and how to reform taxation systems in a world where traditional labor may be significantly diminished. How to steer AI development in socially beneficial directions (Korinek & Stiglitz, 2020; Brynjolfsson, 2022)? By tackling these normative questions, theoretical approaches not only help predict the future economic landscape but also provide valuable insights for policymakers and society at large in navigating the transformative effects of AI.

4.2 A Transformative AI Dashboard

Extensive work is already underway to track *current* AI capabilities directly (see, e.g., Maslej et al., 2024, Ch. 2). Economic indicators of a coming AI-driven boom in economic growth are less well-developed, but would also be valuable because the relationships between benchmark task scores and economic impacts are not always intuitive or deterministic. Nordhaus (2021) takes a first step in this direction, looking for signs that computing is substituting for labor well enough to drive a near-term growth explosion. As of the date of his work, he largely does not find them. But the evidence may change as TAI comes closer and as other metrics are examined.[4]

Analysis on this topic could extend to many other economic indicators, such as the substitution of labor for capital across various links in the AI and robotics supply chains. Further insights could arise by studying trends in a broader range of measures of substitutability: e.g., the elasticity of substitution among different types of labor and capital instead of their aggregates. These data could fit a variety of growth models, reflecting differing assumptions about hard-to-measure variables, such as the relationship between research breakthroughs and economic growth on a given time horizon. Finally, by keeping the associated data and model-based forecasts current with an easily accessible "economic transformation dashboard," researchers and policymakers could stay up to date about whether a period of AI-driven explosive growth appears to be approaching, and if so, what shape it is taking, as described in further depth in the appendix.

4.3 New Metrics for Welfare

New metrics are also needed for assessing economic welfare and consumption. As AI and other digital technologies become responsible for a larger share of production and distribution, the marginal costs of many goods and services will fall to nearly zero. When goods have zero price, they often have zero weight in GDP as it is conventionally measured. Likewise, labor productivity is typically measured as GDP per hour worked, so any mismeasurement of GDP will ripple through to productivity measures as well.

New methods, such as massive online choice experiments, can help assess the valuations that consumers have for goods and services that are poorly captured by traditional

[4] It's worth noting that Nordhaus's model focuses heavily on aggregate macroeconomic variables (such as the capital share). More "microfounded" work of this kind includes that of Besiroglu et al. (2023), who find a rising capital share in AI R&D in particular, suggesting that, in effect, machines may soon improve machine capabilities without being bottlenecked by a lack of human research capacity. On the other hand, Acemoglu (2025) argues that automating only the particular tasks AI is most clearly on track to automate would have little impact on growth.

measures (Brynjolfsson & Collis, 2020). As these methods are extended, refined, and scaled, they can create an updated measurement toolkit to better track AI's contributions in the coming years.

4.4 Task-level Assessments of Potential Impact

It has proven fruitful to analyze the effects of AI and related technologies at the task level rather than the level of entire occupations, firms, or industries. This method has been applied to understand the potential effects of machine learning (Brynjolfsson et al., 2018; Brynjolfsson & Mitchell, 2017) and generative AI (Elondou et al., 2024) on work and employment. While prior work used the Bureau of Labor Statistics O-Net taxonomy of about 18,000 tasks, future work could apply natural language processing to classify hundreds of millions of job postings and resumés, creating a much more fine-grained and dynamic task taxonomy.

4.5 Simulating Economies Using AI Agents

A fascinating new approach in economic research leverages artificial intelligence itself, particularly through simulations and agent-based modeling. By creating LLM-based agents, researchers can simulate human behavior with increasing accuracy and at scale (Manning & Horton, 2025; Anthis et al., 2025). Imagine the potential of modeling thousands, even millions, of these agents interacting in a simulated economy. The simulations could potentially produce similar outputs to randomized controlled trials (RCTs) to assess alternative policy options and interventions but at much greater speed. This may offer new insights into the economy and how AI might reshape labor, consumer behavior, industrial growth, and unforeseen economic bottlenecks.

Simulations are also increasingly used to develop AI systems, especially for robotics. What are the economic implications of developing simulation-based AI systems? What are the externalities associated with simulations and thus the welfare implications of creating these as a public good (e.g., open source) or establishing interoperability standards? How do the costs and benefits of using simulations to train AI compare with traditional data collection methods, particularly in industries with high physical-world uncertainty? What investment strategies best reduce the reality gap and improve real-world robot performance? How will early adopters gain competitive advantages, and what risks will late adopters face if simulation technologies remain imperfect?

4.6 Scenario Planning

Scenario planning offers a structured approach to preparing for multiple possible futures in the context of transformative AI. Unlike forecasting, which typically attempts to predict the most likely future state, scenario planning develops a set of distinct but plausible futures to help researchers and organizations understand different possible outcomes (Korinek, 2023).

This methodology is particularly valuable for studying TAI given the high degree of uncertainty around both the development timeline and the eventual capabilities. By systematically exploring different combinations of key variables – such as the speed of AI development, the degree of AI capabilities, the distribution of productivity gains, and the evolution of labor markets – researchers can identify robust economic patterns and outcomes.

The scenario planning process can be enhanced by incorporating insights from other methodological approaches discussed above. Theoretical models can help define the parameters and relationships that shape different scenarios, while simulations can test the internal consistency of various narratives. This integrated approach helps identify critical indicators and economic mechanisms across multiple scenarios. Moreover, scenario planning bridges technical, economic research, and policy communities, as well as the broader public, through a common framework for analyzing TAI futures.

5. Conclusion

The transition to an economy shaped by TAI will not follow a predetermined path. Some scenarios offer the promise of vastly enhanced wealth, where TAI drives unprecedented productivity, improves social welfare, and distributes benefits fairly. However, without thoughtful management, the outcome could be dystopian, with increased inequality, mass unemployment, social instability, and even catastrophe, leaving many people worse off.

This research agenda highlights the key economic questions and encourages researchers to develop the tools necessary to inform policies that maximize positive outcomes. By identifying key economic indicators, anticipating challenges, and advancing this research agenda, we hope to increase the likelihood that TAI will lead to shared prosperity and a sustainable future for humanity.

As the research agenda outlined in this paper progresses, it will also yield insights relevant to the economic analysis of AI-related policy domains. These domains include: labor laws and their relationship to TAI-driven employment changes; taxation systems and their effects on distribution of TAI-generated wealth; education and skill development programs in a TAI-driven economy; social insurance and income distribution mechanisms; factors affecting social and political stability during technological change; macroeconomic frameworks and their accounting for TAI's impact on productivity and growth; antitrust and market regulations in relation to TAI-driven market concentration; intellectual property frameworks and their effects on innovation incentives and broader economic outcomes; environmental and resource policies as they relate to TAI's ecological footprint; and international coordination mechanisms regarding TAI development and deployment. Economic analysis of these areas could contribute to a better understanding of TAI's effects across society and the resulting trade-offs.

By addressing these questions proactively, economic researchers can inform policymakers on how to harness the potential of TAI while mitigating its downsides.

There is a critical need for interdisciplinary research that integrates economics, computer science, public policy, and other social sciences. While significant resources are being allocated to the *technical* development of TAI, there has been comparatively little investment in understanding its *economic* implications and in research to inform policy on how to steer TAI in desired directions and prepare for the impacts. Further steps in this research area could include developing a more comprehensive research agenda, forming working groups, creating partnerships across sectors, and building institutional capacity for these research efforts. Such approaches could contribute to understanding how TAI might affect human welfare.

As economists, we have a unique responsibility to develop robust analytical frameworks that can illuminate the potential disruptions and opportunities presented by TAI. Organizations and processes evolve far more slowly than technology itself—which makes our work all the more urgent. We should not wait until powerful TAI systems are fully realized before undertaking this critical research. By directing our collective attention to these questions now, economists can contribute vital insights about how different approaches might affect the distribution of benefits and risks associated with TAI. Our work can help society understand pathways toward broadly shared prosperity and human flourishing as this transformative technology unfolds.

Appendix: Economic Indicators for TAI

We outline a set of technological and economic indicators to help us better comprehend the nature and rate of transformation.

Indicators of factor inputs help us track the resources devoted to AI development and deployment, which reflect the scale and pace of advancement. They also provide insights into potential resource constraints or environmental impacts of AI progress.

a) **Compute:** This set of indicators measures the total computational power used in frontier AI training and/or deployment, often quantified in floating-point operations per second (FLOPS) or petaflop/s-days. Investments in the data center and associated infrastructure are another useful metric. We should also track improvements in compute efficiency, such as the reduction in compute needed to achieve specific AI benchmarks over time.
b) **Labor:** This could measure the share of workers or total compensation paid to personnel working on producing cutting-edge AI and robotics systems, and other research inputs.
c) **Energy:** This set of indicators measures the total energy used by AI systems, including both training and inference phases. It could be quantified in kilowatt-hours (kWh) or joules, and might be normalized per unit of compute or per AI task performed. Improvements in energy efficiency need to be tracked in parallel.
d) **Other raw materials:** This could include the consumption of other important input factors, such as materials crucial for AI and robotic hardware. Metrics might include the volume or value of these materials used in manufacturing.
e) **Data:** One of the most important inputs is data used to train, refine, and test the systems.

Technological indicators, reflecting the economic definition of the term "production technology" as reflecting how we turn inputs into outputs, capture how AI is transforming the economy's production process, for example, by replacing or augmenting human labor. They provide crucial insights into the potential for AI to transform the economy and the changing nature of work.

a) **Advances in AI capabilities:** This has traditionally been measured through performance on standardized benchmarks across a wide range of domains (e.g., MMLU, BigBench, HumanEval, MATH). We must track both the absolute performance and the rate of improvement over time. Importantly, part of our role as economists is to examine to what extent these technical indicators map into economic usefulness.
b) **Advances in robot capabilities:** This includes metrics on dexterity, mobility, and task completion rates for physical robots. We might measure the percentage of human physical tasks that robots can perform or the speed and accuracy with which they complete standardized tasks. Our focus is on economic usefulness.
c) **Substitutability of labor with capital:** This could be quantified through the elasticity of substitution measures between AI/robotic systems and human labor across different job categories, together with the elasticity of substitution of the resulting outputs in

consumers' consumption baskets. This also includes the percentage of tasks within occupations that can be automated (i.e., perfectly substituted).

Production/output indicators gauge the economic impact of AI on overall production, efficiency, and the labor market. They are crucial for understanding how the benefits of AI are distributed across the economy and society.

a) **Productivity (including Total Factor Productivity (TFP) and labor productivity):** We should measure changes in output per unit of input, both for the economy as a whole (TFP) and specifically for labor including areas not well-measured by traditional GDP and productivity metrics such as household production, healthspan, and environmental quality. This could be tracked at the firm, industry, and economy-wide levels.
b) **Output growth:** This would measure the overall increase in economic output (e.g., GDP growth) attributable to AI adoption. We might develop methods to isolate AI's contribution to growth from other factors.
c) **Effects on labor demand:** have both a price and a quantity dimension:
    i) On the price side, labor demand is reflected in wages. We should track changes in wage levels across different skill categories and occupations, paying particular attention to those most affected by AI.
    ii) On the quantity side, labor demand is reflected in job numbers and flows, including job displacement and creation. Moreover, it may also be reflected in labor force participation rates or unemployment numbers. At the sectoral level, we can observe shifts in employment across sectors.
    iii) Relatedly, new business formation, as well as the creation of new occupations and tasks should be tracked.
    iv) Time use can provide another useful set of metrics, for production, administrative and in household work. AI will trigger process innovation just as prior technologies did.

Financial market indicators often reflect expectations about future technological impacts, making these indicators valuable for anticipating economic shifts. They can provide early signals of how investors and businesses are valuing AI's potential.

a) **Equity markets:** We could track the stock performance of AI-focused companies and the adoption of AI in various sectors. This might include specialized AI stock indices or the AI-related revenue of major tech companies.
b) **Venture Capital and Private Equity Investments:** Many relevant AI-related investments are not publicly traded and should also be tracked.
c) **Energy prices:** Given the energy-intensive nature of AI, tracking energy prices (especially electricity) could provide insights into the costs and constraints of AI deployment.
d) **Interest rates:** Changes in interest rates might reflect expectations about AI-driven productivity growth. We should analyze the relationship between AI advancement and long-term interest rate trends.

Industry-level phenomena reveal how AI is transforming the structure of the economy and creating new opportunities. They provide insights into the dynamism of the economy and the pace of creative destruction driven by AI.

a) **Emergence of new industries:** We should track the number and growth of new AI-enabled industries. In principle, this could include measures such as the number of new NAICS codes related to AI, but that is likely to be a severely lagging indicator. Other metrics include the revenue and employment in entirely new categories of business.
b) **Rapid industry reshuffling:** This could be measured through changes in market concentration (e.g., Herfindahl-Hirschman Index) within industries, the rate of company formations and bankruptcies, and shifts in industry compositions of major stock indices.

Income distribution and inequality indicators reflect the societal impacts as AI potentially reshapes the distribution of economic gains—these indicators become crucial for understanding societal impacts. They can help policymakers identify and address potential increases in inequality resulting from AI adoption.

a) **Labor share vs. capital share:** We should track the proportion of national income going to workers versus capital owners. This could be measured economy-wide and within specific sectors, especially those heavily impacted by AI.
b) **Gini coefficient:** This standard measure of income inequality should be monitored at various levels (within countries, between countries, and globally) to assess how AI affects overall income distribution.

International indicators help us understand how AI might affect global economic relationships and potentially exacerbate or reduce international inequalities. They are crucial for anticipating geopolitical implications and informing international economic policies.

a) **Global terms of trade:** We should monitor how AI affects the relative prices of exports and imports for different countries. This could include tracking changes in the value of knowledge-intensive exports relative to raw materials or manufacturing.
b) **Cross-country gaps in GDP/capita:** This would involve measuring how AI adoption affects economic convergence or divergence between countries at different levels of development.

It is important to reiterate that our final objective is ultimately social welfare so the purpose of all these indicators is that they reflect different dimensions of AI's effects on welfare. Utilitarian welfare depends not only on individual utility derived from material consumption of goods and services but also on non-material goods such as health, happiness, and meaning. As we track these economic indicators, we must keep in mind the broader implications of AI-triggered transformation for human well-being.

References

Acemoglu, D. (2025). The simple macroeconomics of AI. *Economic Policy, 40*(121), 13-58. https://doi.org/10.1093/epolic/eiae042

Acemoglu, D., & Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives*, *33*(2), 3-30. https://doi.org/10.1257/jep.33.2.3

Aghion, P., Jones, B. F., & Jones, C. I. (2019). Artificial intelligence and economic growth. In A. Agrawal, J. Gans, & A. Goldfarb (Eds.), *The economics of artificial intelligence: An agenda* (pp. 237-282). University of Chicago Press. https://www.nber.org/system/files/chapters/c14015/c14015.pdf

Amodei, D. (2024, October). Machines of loving grace. *DarioAmodei.com*. https://darioamodei.com/machines-of-loving-grace

Anthis, J. R., Liu, R., Richardson, S. M., Kozlowski, A. C., Koch, B., Evans, J., Brynjolfsson, E., & Bernstein, M. (2025). LLM social simulations are a promising research method. *arXiv*. https://arxiv.org/abs/2504.02234

Autor, D., Chin, C., Salomons, A., & Seegmiller, B. (2022, August 14). New frontiers: The origins and content of new work, 1940-2018. https://economics.mit.edu/sites/default/files/2022-11/ACSS-NewFrontiers-20220814.pdf

Barnett, M. (2020). When will the first general AI system be devised, tested, and publicly announced? [In progress]. *Metaculus*. https://www.metaculus.com/questions/5121/date-of-artificial-general-intelligence/

Bengio, Y., Mindermann, S., Privitera, D., Besiroglu, T., Bommasani, R., Casper, S., Choi, Y., Fox, P., Garfinkel, B., Goldfarb, D., Heidari, H., Ho, A., Kapoor, S., Khalatbari, L., Longpre, S., Manning, S., Mavroudis, V., Mazeika, M., Michael, J., ... Zeng, Y. (2025). I*nternational AI Safety Report* (Report No. DSIT 2025/001). https://www.gov.uk/government/publications/international-ai-safety-report-2025

Besiroglu, T., Emery-Xu, N., & Thompson, N. (2023). Economic impacts of AI-augmented R&D. *arXiv*. https://arxiv.org/abs/2212.08198

Brynjolfsson, E. (2022). The Turing trap: The promise & peril of human-like artificial intelligence. *arXiv*. https://arxiv.org/abs/2201.04200

Brynjolfsson, E., & Collis, A. (2020). How should we measure the digital economy? (Hutchins Center Working Paper No. 57). Hutchins Center on Fiscal & Monetary Policy at Brookings. https://www.brookings.edu/wp-content/uploads/2020/01/WP57-Collis_Brynjolfsson_updated.pdf

Brynjolfsson, E., Li, D., & Raymond, L. (2025). Generative AI at work. *The Quarterly Journal of Economics, 140*(2), 889-942. https://doi.org/10.1093/qje/qjae044

Brynjolfsson, E., & Mitchell, T. (2017). What can machine learning do? Workforce implications. *Science*, *358*(6370), 1530-1534. https://doi.org/10.1126/science.aap8062

Brynjolfsson, E., Mitchell, T., & Rock, D. (2018). What can machines learn, and what does it mean for occupations and the economy? *AEA Papers and Proceedings*, *108*, 43-47. https://doi.org/10.1257/pandp.20181019

Brynjolfsson, E., Rock, D., & Syverson, C. (2021). The productivity J-curve: How intangibles complement general purpose technologies. *American Economy Journal: Macroeconomics, 12*(1), 333-372. https://doi.org/10.1257/mac.20180386

Davidson, T. (2023, August 9). Report on whether AI could drive explosive economic growth. *Open Philanthropy*. https://www.openphilanthropy.org/research/report-on-whether-ai-could-drive-explosive-economic-growth/

Elondou, T., Manning, S., Mishkin, P., & Rock, D. (2024). GPTs are GPTs: Labor market impact potential of LLMs. *Science*, *384*(6702), 1306-1308. https://doi.org/10.1126/science.adj0998

Hui, X., Reshef, O., & Zhou, L. (2024). The short-term effects of generative artificial intelligence on employment: Evidence from an online labor market. *Organization Science*, *35*(6), 1977-1989. https://doi.org/10.1287/orsc.2023.18441

Jones, C. I. (2024). The AI dilemma: Growth versus existential risk. *American Economic Review: Insights*, *6*(4), 575-590. https://doi.org/10.1257/aeri.20230570

Keynes, J. M. (1931). Economic possibilities for our grandchildren. In *Essays in persuasion* (pp. 358-373). Macmillan.

Korinek, A. (2023). Scenario planning for an A(G)I future. *IMF Finance & Development Magazine*, *60*(4). https://www.imf.org/en/Publications/fandd/issues/2023/12/Scenario-Planning-for-an-AGI-future-Anton-korinek

Korinek, A. (2024). Economic policy challenges for the age of AI (NBER Working Paper No. 32980). National Bureau of Economic Research. https://doi.org/10.3386/w32980

Korinek, A., & Balwit, A. (2024). Aligned with whom? Direct and social goals for AI systems. In J. B. Bullock et al. (Eds.), *The Oxford handbook of AI governance*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780197579329.013.4

Korinek, A., & Juelfs, M. (2023). Preparing for the (non-existent?) future of work. In J. B. Bullock et al. (Eds.), *The Oxford handbook of AI governance*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780197579329.013.44

Korinek, A., & Stiglitz, J. E. (2019). Artificial intelligence and its implications for income distribution and unemployment. In A. Agrawal, J. Gans, & A. Goldfarb (Eds.), *The* *economics of artificial intelligence: An agenda* (pp. 349-390). University of Chicago Press. https://www.nber.org/system/files/chapters/c14018/c14018.pdf

Korinek, A., & Stiglitz, J. E. (2020). Steering technological progress. National Bureau of Economic Research. https://conference.nber.org/conf_papers/f143989.pdf

Korinek, A., & Vipra, J. (2025, January). Concentrating intelligence: Scaling and market structure in artificial intelligence. *Economic Policy*, *40*(121), 225-256.

Manning, B., & Horton, J. (2025). General social agents [Working paper]. https://benjaminmanning.io/files/optimize.pdf

Maslej, N., Fattorini, L., Perrault, R., Parli, V., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J. C., Shoham, Y., Wald, R., & Clark, J. (2024). Chapter 2: Technical performance. In *The AI Index 2024 annual report*. AI Index Steering Committee, Institute for Human-Centered AI, Stanford University. https://aiindex.stanford.edu/wp-content/uploads/2024/04/HAI_AI-Index-Report-2024_Chapter2.pdf

Narayanan, A., & Kapoor, S. (2025, April 14). AI as normal technology. *Knight First Amendment Institute*. https://knightcolumbia.org/content/ai-as-normal-technology

National Academies of Sciences, Engineering, and Medicine. (2024). *Artificial intelligence and the future of work.* The National Academies Press. https://doi.org/10.17226/27644

Nordhaus, W. D. (2021). Are we approaching an economic singularity? Information technology and the future of economic growth. *American Economic Journal: Macroeconomics*, *13*(1), 299-332. https://doi.org/10.1257/mac.20170105

Suleyman, M. (2023, July 14). Mustafa Suleyman: My new Turing test would see if AI can make $1 million. *MIT Technology Review*. https://www.technologyreview.com/2023/07/14/1076296/mustafa-suleyman-my-new-turing-test-would-see-if-ai-can-make-1-million/

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146-1151. https://doi.org/10.1126/science.aap9559

Webber, A. (2025, April 16). Demis Hassabis is preparing for AI's endgame. *TIME*. https://time.com/7277608/demis-hassabis-interview-time100-2025/