Berkeley
UNIVERSITY OF CALIFORNIA

BERKELEY ROUNDTABLE ON
THE INTERNATIONAL ECONOMY

BRIE Working Paper
2024-2

Generative AI and the Future of Work: Augmentation or Automation?

John Zysman, Mark Nitzberg

Generative AI and the Future of Work: Augmentation or Automation?

May 2024

[This text is simultaneously published as [Zysman, John & Nitzberg, Mark (2024): Generative AI and the Future of Work:Augmentation or Automation?, Working Paper, Weizenbaum Series 38, Berlin: Weizenbaum Institute. DOI: https://doi.org/10.34669/wi.ws/38 (https://doi.org/10.34669/wi.ws/38)]

This research report was principally funded by the German Federal Ministry of Labour and Social Affairs (684 11 - "Denkfabrik Digitale Arbeitsgesellschaft")

Authors:

John Zysman
Professor Emeritus Department of Political Science Berkeley Roundtable on the International Economy (BRIE) University of California, Berkeley

Mark Nitzberg
Executive Director
Center for Human-Compatible Artificial Intelligence (CHAI) Berkeley Al Research (BAIR) University of California, Berkeley

Generative AI and the Future of Work: Augmentation or Automation?

Abstract [1]

This report examines the potential impact of Generative artificial intelligence (AI) systems, such as ChatGPT, on the future of work and, by implication, on productivity. It argues that although Generative AI is powerful, it has significant limitations and risks that require humans to remain “in the loop” not only to prevent systems from going off the rails, but to capture value. Rather than taking a deterministic view that artificial intelligence (AI) will inevitably destroy jobs, the article suggests that an analysis should start with how firms can strategically deploy these tools to gain an advantage. It asks whether “augmentation” or “simplistic automation” lies ahead. Our objective is to move beyond hype and despair.[2]

The existing digital infrastructure has enabled AI to be adopted quickly. However, projections based solely on automating existing tasks fail to capture the complex reorganizations that are likely to happen. Firms in sectors such as professional services, materials, and pharmaceuticals seem to have particular exposure to the use of Generative AI tools. Adaptations will vary across contexts and depend greatly on who controls the decisions about deployment. Maintaining the centrality of humans is likely to prove crucial—in training systems, curating data, and assessing outputs. One question is which business strategies and public policies encourage that engagement and make it possible.

Although AI regulation debates matter, promoting social prosperity depends heavily on directly shaping the trajectory of the development and use of AI. This requires influencing the constraints and the incentives that firms face, as well as the strategic mindsets of decision makers. Which groups are engaged in the discussions and debates is of vital importance. The article recommends that, beyond the traditional policy proposals, an independent public-interest consultancy needs to be established in order to design creative business strategies that augment workers in a manner that will support, rather than hinder, social prosperity. Ultimately, avoiding a dystopian scenario might hinge on fostering new norms in which human capabilities remain essential.

Contents

1.
Introduction: Automation or Augmentation?

This essay considers the impact of Generative artificial intelligence (hereafter, GenAI) on work, the organization of work, and, by implication, on productivity.[3] Its goal is to frame the debate, so as to go beyond the hype and despair.[4] Hype and despair have accompanied every digital era, from the transistor and microprocessors, to cloud computing and platforms, and now modern AI and GenAI. In writing it, we seek to present a balanced understanding of how the evolution of this latest technology revolution could be steered.

GenAI is "a type of artificial intelligence technology that utilizes deep learning models to create various forms of content, such as text, images, and code, based on the data they were trained on."[5] This represents a significant leap in computing capability. The seeming speed of deployment and adoption in itself generates a sense of urgency that society and the economy have reached a dramatic turning point.

GenAI emerged abruptly in both the public eye and commercial environments not only because of significant innovation, but because the essential infrastructure and technology for its deployment were already in place. Data, computing power, and networks were all at hand when machine learning and transformer-based systems were developed and deployed. The extraordinary pace of experimentation with these new tools, if not yet full or diverse adoption, was possible because the complementary digital tools and infrastructure needed for the deployment of GenAI were at the ready. The several prior decades, perhaps best characterized as the emergence of the "Platform Economy", saw extraordinary expansion in all of the digital capacities required for Generative AI to emerge: storage, computation, and information communication technologies.[6] Then, of course, demand for new chip designs and graphics chips, from the expanding video games sector and then crypto mining, drove semiconductor innovation that could then be directly applied to machine learning and Generative AI. This was particularly true of Nvidia, whose valuation has surged to an extraordinary level. The giant tech companies—Google, Amazon, Microsoft—had the financial resources to pursue the next "thing," and the rocket engines were ignited.

GenAI's remarkable capacities will facilitate, encourage, and perhaps drive significant changes in the organization of work as well as in market competition across diverse aspects of the economy, and perhaps generate real productivity increases. To date, digital automation has meant routine-biased technical change (RBTC), which substitutes electronic capital for human labor in routine tasks. Now, with AI, we see RBTC on steroids, and deployment of the technology has been extended to tasks that are less routine and more cognitively heavy.[7] Will GenAI continue this trajectory of digital automation and represent a renewed threat in the form of displacement of jobs and workers? Or will the new capabilities enable the augmentation of workers' capacities and generate a more intelligent economy? Will work and labor markets experience genuine augmentation, or simple automation?

We speculate that the limits on the capabilities of GenAI, and risks inherent in its use, will constrain the utility of these tools and their deployment. At its core, despite the new sophistication offered by GenAI, the tools that leverage it are, and for now will remain, principally rooted in statistical prediction without any model of how the world works. Consequently, GenAI tools can spin up predictions and presentations that, despite being developed in easy-to-use dialogues and presented with seeming certainty, may bear little relation to any reality or be an invented reality altogether. Second, the nature of large language models (LLM) means that it is impossible to work backward from an outcome to understand how the results were obtained. For example, explicitly implementing or applying results embeds the possibility, perhaps the probability, of bias or persuasive but incorrect statements and processes. One noted AI scholar remarked: "If it is a large language model, it is a beast conjured rather than designed."

Therefore, we postulate that it might be an inherent consequence of the statistical foundation of AI and GenAI that, if GenAI is to be usefully and safely exploited, humans must remain in the loop throughout its development, training, curation, and deployment. Although GenAI is exceptionally powerful, its limits and risks will require doing so. The goal of this essay is to frame debate, to go beyond the hype and despair that has accompanied every era of new digital technology, beginning with the transistor and microprocessors, through cloud computing and platforms, and now modern AI and GenAI.[8] It tries to go beyond the hype of promoters and the despair engendered by the hype in order to obtain a balanced understanding of how the evolution of this technology could be steered.

Obviously, we do not know how GenAI technology will evolve, how its capacities will expand, or whether its risks and limits will remain entrenched or be resolved. Consequently, we can only speculate about the path of deployment. How GenAI tools are actually deployed will depend, among other things, on economic calculations, firm capacities, and regulatory considerations. That said, we have several expectations:

1. Firms will experiment in order to find ways to create market advantage.
2. Some experiments will support worker efforts to augment worker productivity.
3. Some experiments will lead to new approaches to work and the organization of work, and others will fade away or be abandoned altogether.

No single strategy for implementing these tools will be employed, but early successes may well define broader trajectories for a portion of the economy.[9] In the face of the uncertainty about the evolution of GenAI technology as well as its path for deployment, how, then, should we consider its consequences for work and, more widely, for productivity? Responses to this question should take the following steps:

1) Clarify which sectors or competitive environments are the most exposed to GenAI or likely to be influenced by it in the short or medium term.
2) Consider which business strategies and tactics used by firms that employ these tools might create a market advantage.
    - This assumes that firms, not public administrations, will likely be the drivers of development, and this assumption should be stated.
3) In any discussion of strategies and tactics, consider the impact on work.
    - Jobs, and the tasks that comprise them, are an outcome of business and administrative choices, but should not be the initial focus of analysis; the conventional focus on "tasks" is misplaced.

Our discussion proceeds as follows:

- Section 2, intended to create a baseline for our discussion of GenAI, summarizes what we understood in 2022 about the impact of machine learning AI on work and the economy, before ChatGPT and its ilk took center stage. As part of this discussion, we speculate that the enduring concerns about job losses might give way to concerns about labor shortages, as jobs are adapted to create advantage from the new technologies.
- Section 3 considers what the new technology is and what it can and cannot do based on work with collaborators at the Berkeley Center for Human Compatible Artificial Intelligence (CHAI). Some of this is basic, but is worth restating for the sake of clarity.
- Section 4 considers how we might translate the developments noted in Section 3 into at least sensible speculation about the consequences of GenAI for the future of work.
- Section 5 extends the analysis in Section 4 to consider the implications for policy and firm strategy. We ask: If the theme becomes labor shortages, rather than job displacement, how should that influence both policy and strategy?

2.

A Baseline and Beyond

GenAI represents a new phase in the development and deployment of AI. Before considering this digital newcomer's impact on work now and in the future, we should establish a baseline about the effects to date of earlier developments in machine learning AI and other aspects of the digitally driven shifting structure of work. We also note that, over the next few years, concerns in the advanced countries about the displacement of labor might give way to concerns about shortages in skilled labor. That, of course, would turn the policy debate on its head.

Some clear lessons can be learned from the first round of AI innovation that came after the long AI winter: machine learning with neural networks.[10]

- This era of RBTC on steroids, as many have argued, has negatively affected activities with a skill level between those that are low skill and high skill.[11]
- Although the technology affects all economies, its national impact has significant differences, based on training, welfare, and labor market issues.
- Automation has moved from manufacturing to services, and from cognitive routine activities to those that are less routine. Routine services have certainly been automated, and ever broader swaths of services are affected. Two distinctive elements together shift how value is created:

    1) Many products have become service offerings.
    2) Digital platform firms have become a crucial part of the market economy.

The core question in the next sections, then, is as follows: what is new in GenAI, and how might its features drive the next round of work adjustments? But, first, we suggest a baseline.

AI Machine Learning Is Routine-Biased Technological Development on Steroids

As Tyson and Zysman have argued, AI's impact on work thus far is best characterized as "routine-biased technological deployment on steroids," in that we are "adding intelligence to automation tools that substitute for humans in *physical* tasks and substituting for humans in routine, but increasingly nonroutine *cognitive* tasks as well."[12] This last category - nonroutine cognitive tasks - is likely to be most affected by GenAI.

The studies on the impact of this machine learning (ML)–based AI round of innovation on work conclude that, as in prior eras of automation, jobs remain in spite of these shifts. [13] Rather than expecting work tasks to be swept away by AI, a more accurate view is that the production systems themselves are being reconfigured in response to AI.[14] The question has been: what kinds of work tasks will be collaterally impacted by this broader reconfiguration? Employment has been polarized, and, although automation has not entirely swept away jobs, good jobs for middle- and semi-skilled workers are in short supply, and they have experienced stagnant wage growth, both of which contribute to inequality.

The erosion of the middle- and semiskilled routine jobs has bifurcated the labor market, such that low-skilled jobs are available at the bottom, and high-skill, high-wage jobs are available at the top.[15] The technology in itself does not drive the inequality; rather, the inequality is driven by how it is deployed. Although the tendency toward bifurcation is felt across the advanced countries, the impact differs across labor markets, with specific rules and arrangements dampening the extent of the bifurcation of work, and social welfare programs cushioning the consequences.[16]

The discussion on automation has long focused on manufacturing. Over the years, robots in automobile production, with assembly lines and paint shops filled with these mechanical workers, were emblematic of the impact of automation. Strictly speaking, much of that automation has not been about work tasks but, instead, about the technical advantage afforded by robots and, consequently, the reorganization of work activities. Semiconductor fabrication is highly automated because the characteristics of the production processes mean that the chips cannot be manufactured in any other way. Similarly, advances in battery production, which is already automated, will be needed in order to realize the hope that solid-state batteries will facilitate a shift to an electricity-driven green economy.

Of relevance to what we may anticipate from the deployment of GenAI, the national differences in how prior digital automation tools are used are dramatic. The Japanese numerically controlled (NC) machine tool "cells," from earlier years, with multiple simple tools that provide flexibility for the entire line while depending on worker know-how, can be contrasted with Germany's high-end, multipurpose complex tools and America's standard tools, which were often aimed at replacing workers on a one-to-one basis. Studies on the deployment of automation by Helper et al. and later by Acemoglu et al. show that consulting firms and those with MBAs have quite different visions of how to use these tools compared to the companies that conduct the deployment coordination of these tools in-house.[17] Similarly, the German debate over the potential of "Industry 4.0" predominantly hinges on different views about who controls the production process and, inevitably, how the tools are used.[18] The lesson from these prior experiences is that the parties who deploy these tools, for what purpose, and in what context, are crucial for the ultimate consequences.

Also evident but still worth noting is that, as the economy has moved from manufacturing to diverse services, routine service activities are being automated. Nevertheless, an increasing portion of jobs involve cognitive tasks—that is, nonroutine cognitive tasks. Consider the process of checking out at a grocery store: work formerly performed by a clerk has been transferred to the customer, who now acts as cashier and bag packer. Even this "customer as cashier" could be eliminated by sensor and camera-based systems, as seen in the case of Grabango,[19] a company that simply tracks activity and presents a receipt to the customer upon exiting.

One assertion that we need to consider, as part of the move to services, is the extent to which GenAI will be able to *displace* even more sophisticated cognitive tasks and the extent to which the use of GenAI for cognitive tasks results in *work reorganization with new tasks*. Very sophisticated tasks, such as actuarial estimation of insurance risks, have been augmented by computers, and decades ago began to facilitate the construction of customized policies by insurance agents as they work with their clients. That led to more diversified insurance offerings, enabling them to write policies on the fly. Rather than selling rigidly fixed policies, insurance agents had to learn how to work with clients to customize the offerings to fit their needs.

Because traditionally manufactured products are increasingly being sold as services or portals to services, how value is created and where crucial work activities are located in the production/distribution system has also shifted.[20] The transition to products sold as services is certainly facilitated by the ability to collect, store, and analyze large datasets.[21] Consider, for example, tires and aircraft engines: auto and truck fleets can obtain tires as a service, purchasing a set number of miles from the purveyor. Michelin, for instance, ensures the quality, condition, and availability of tires based on sensor-generated data managed by the tire outlet. The condition of the tires is then monitored not by mechanics on the road, but through data analytics. Similarly, aircraft engines can be acquired based on the number of hours of lift. In both cases, the "use" of data enables the seller to effectively manage their products to the benefit of the user and incorporate that data into product improvement/innovation cycles. Certainly, the products, whether tires or engines, are still manufactured, but the sales process and work tasks involved become increasingly entangled with digital statistical analytics. Through the power of AI tools, these statistical analytics subsequently become even more formidable in real-time.

Third, digital platform firms are changing the character of work.[22] Social media sites are hosted on digital platforms, which in turn are based on cloud computing, big data, and the AI-bolstered statistical capabilities that underpin LLM models to target users and link together buyers and sellers. Digital platforms have impacted work in the distribution of goods as well as in social media and digital services. Amazon is the most obvious case in point. A majority of retail and operations now occur online or rely on online processes.[23] Doing so transfers work from a brick-and-mortar main street store to the warehouse and transport sector. To facilitate not only efficiency but effective and rapid delivery, the warehouses are then increasingly automated, with the driver of this automation being predominantly attributable to the digital platform firm. We could say that this represents a move from "brick and mortar" to "warehouses and wheels."

In addition to establishing a baseline with which to consider GenAI, studies about the influence of ML-based AI on work provide guidance on how to look at GenAI. The fear that ML AI, discriminative AI, could eviscerate the world of work came from analyses focused on the tasks that might be displaced by RBTC AI.[24] As discussed above, although RBTC AI shifted the balance of work activities, there are still jobs, although they are different and often differently remunerated. Moreover, the most widely cited analyses of AI and work have focused on the tasks performed in existing jobs and work. Though interesting, these studies can be deeply misleading. A focus on existing tasks means imagining the future principally through the lens of the current organization of production and distribution. This approach is flawed because the tasks of tomorrow may have only a limited relation to the tasks today, and even the tasks that remain may be reorchestrated in radically different ways, as they continually adapt to fit new requirements. Hence, these studies assume path dependence. Recognizing the shortcomings of these AI analyses is the first step towards understanding the difficulties inherent to forecasting the consequences of GenAI. We count at least four shortcomings:

1) First, and most important, jobs, and the tasks embedded in jobs, are the result of a combination of products/service choices, the markets involved, and basic production strategies. If an analysis starts with the tasks, then it is beginning at the end, with no way to infer the decisions that will drive work and work organization.
2) Second, these analyses indicate which tasks might be affected by RBTC AI, but they do not indicate the cost of introducing AI tools. Hence, they do not analyze the economic and financial case for deploying them.[25]
3) Third, jobs consist of many tasks, which are often configured in very different ways by different organizations. Automating some tasks might actually enhance the quality of jobs or increase worker productivity, rather than eliminating jobs. This is related to the broader notion that technology leads to a wide reorganization of work, in which tasks and jobs are shuffled. Adopting new technologies does not automatically mean replacement of workers. The key, then, concerns the choices made about how to use these tools and which skills will be required.
4) Fourth, as has increasingly been recognized and bears repeating, job categories today are very different from those thirty or fifty years ago. What percentage of the tasks today even existed fifty years ago? Or, in other words, what percentage of yesterday's tasks still exist now? Some examples of tasks and jobs today that were unheard of thirty years ago are online influencers, YouTube creators, and data scientists.

What might the future hold?

Will GenAI support a continuation of the polarization of work, which analysts widely agree has been underway to date? Or, as some optimists contend, are we at the beginning of a new era of prosperity and productivity? Analysts at the *Economist* already argue that the demand, demographic, and digital processes that have driven inequality are already being reversed.[26] They note that demand is booming, population growth is slowing in places such as China, and the emerging generation of AI might provide a productivity boost for the "lower performers." They claim that tighter labor markets, a product of these demand and demographic trends, are generating real wage increases. Indeed, in an article amusingly subtitled "Bots and Tots," Hal Varian states that digital tools will be needed to offset demographic trends in advanced countries that are likely to experience labor shortages.[27] (Of course, as an aside, immigration will be part of the economic and policy story.) Florian Butolo makes the case that the finer and more sophisticated division of labor, a basic characteristic of market capitalism and industrialization, means that automation has generated ever more jobs. It is not simply that new products are substituting for old ones. Rather, the process of production and distribution becomes more complex with more phases, jobs, and tasks. For example, in metal work, the simple blacksmith with tongs, a hammer, and an anvil becomes a complex multiphase automated production process, with AI analytics designing the materials.[28]

What is the role of GenAI in these trends? Will work be displaced, and will the polarization of work continue, or will humans need to be kept in the loop of operations, with probability machines that prompt the creation of new categories of high-quality jobs?

In Section 3, we begin to consider these questions, distinguishing between classic Machine Learning (ML) and the new innovations of GenAI. In Section 4, we examine issues about emerging tools that may be deployed by existing firms and entrepreneurs to create competitive advantage and generate productivity. We cannot imagine which businesses will be generated that do not now exist, but we can ask which sectors could be affected by these tools, and what types of competition might be generated by the deployment of these tools. Posing these questions, which are all the more important in the case of GenAI, will help us understand not only which tasks might be affected but which jobs will be required, which jobs might be generated, and how these jobs and the tasks that they encompass will be orchestrated: work organization.

3.

From AI Winter through Machine Learning to Large Language Models

Capabilities, limitations, and the need for humans

Will GenAI represent a new phase in the automation of work or simply an acceleration of the prior trajectory? Will it open possibilities for the augmentation of worker opportunities and capabilities or simply create a wider swath of simplistic automation? How do we understand its impact on work and work organization? To address these questions, in this section we consider how GenAI is built as well as how it operates, distinguishing it from the prior phase of AI development rooted in ML and neural networks (NN). GenAI is characterized by dramatic new capabilities, but has significant limits and technical risks as well. The delineation of the capabilities, limits, and technical risks of GenAI can indicate the enduring role of people in the development, training, curation, and deployment of these tools.

An overview characterization can highlight some of the core issues. The AI systems that emerged in the early 2010s, after a long "AI winter" during which limited progress was made, are all built with statistical tools trained on enormous assemblages of data. Indeed, 90 percent of the world's data was generated just in the past two years.[29] The statistical foundations of current AI have a series of implications for all current AI tools. Four implications have a direct bearing on the world of work and the role of people in the deployment of AI tools.

1. The behavior of these systems is determined by their training data. No one has yet discovered the principles by which they operate. Do they construct a model of the world, or a framework for interpreting particular situations and inputs of data? As

Alison Gopnik has long argued, AI models don't know things that small children know.[30] Consider two examples. In the first one, a ball rolls into the street from between two cars. How would a fully autonomous car respond? Based on real-life experience, and a "Model of the world," we infer that in all likelihood, a person—probably a small person—will emerge to fetch the ball. As such, the car should stop. But if the car has no notion that children play with balls on sidewalks - or worse, has no sense of what a sidewalk is, or what children are - calamity could ensue. In a second example, related to social logic, if Susan is the mother of James, then James's mother is Susan. But the very best current data-driven AI models fail to reach that conclusion.

- In many settings, proper behavior depends on models of human behavior and physical reality, which, when incompatible, can be reconciled by norms and causal logic, respectively. Current data-driven AI systems do not seem to create such models. This is the principal reason that people must be integrated into processes that use today's AI systems.

2. GenAI models can make serious errors, sometimes devising entirely fictional responses altogether: so-called "hallucinations, "an anthropomorphic term for plausible bullshit. Given that this is an inherent feature of large language models, LLMs, and similar systems, people will likely be required to sort things out.
    - An AI system that suggests a shirt of the wrong color in a product search makes an error of little consequence. Suggesting the wrong train in a scheduling task, on the other hand, could have serious consequences. The wrong operating parameter at a chemical plant or nuclear facility could be catastrophic.
    - Plausible bullshit is a very specific and dangerous form of error. Michael Cohen's court submission containing fake legal precedent cases produced by ChatGPT, for example, attracted public attention to this particular threat. We all have colleagues whose résumés suddenly listed new publications as well. Such fabrication is likely inherent in open-ended statistical systems, and must be appropriately contended with.
3. Explainability is a real challenge. The models' behavior is determined by a very large circuit, whose construction depends on the vast data used to train it. As a result, accounting for or explaining outcomes might be impossible.
4. Because AI systems, and the GenAI systems that are now emerging, are trained on existing data—artifacts of our past behavior—they inherently embed the bias and misdirection of the past represented in that data. The resulting bias and misdirection must be continuously sorted out.

The capabilities, limits, and technical risks of GenAI show the enduring role of people in the development, training, curation, and deployment of these tools. We already know that a combination of people and AI tools produces better outcomes than either of them alone in domains as diverse as chess and radiology.

Ongoing assessment of capabilities and risks will be essential. Will the capabilities scale with ever-greater computing power and data, and how can we ensure these greater capabilities expand the limits while containing the risks? AI safety or, rather, the safe deployment of AI, will be an essential question, with implications for social and economic well-being as well as the ultimate risk to human existence from AI.[31]

Now, we turn to the systems themselves.

What is distinctive about GenAI?[32]

AI emerged in the 2010s from its AI winter with breakthroughs in machine learning (ML) and deep neural network NN architectures.[33] The applications developed were tools for specific tasks: labeling and identifying sounds and images, ranking the appeal of news items or products for individuals, and predicting creditworthiness and suitability for jobs. In early 2023, GenAI systems, which excel at numerous creative and cognitive tasks, became commercially viable with the launch of ChatGPT (GPT4), which has "multimodal" capabilities that can seemingly interpret and produce images, video, and complex mixed-media artifacts. The principles of transformer architectures have enabled leaps in wide-ranging task areas, including computer vision, protein folding, and, recently, clinical health care. Here, we distinguish the AI of the 2010s, characterized by machine learning and deep neural network capabilities, from the transformer era of the 2020s. Through assessing their capabilities and limitations, we can conclude that there is a key place for humans in the effective deployment of even the most powerful systems.

David Mumford, renowned mathematician and field medal recipient, on the occasion of receiving the Inaugural Basic Science Lifetime Award in Beijing, presented a crucial idea:

> The basic learning algorithm used in "deep learning" was devised many decades ago but was long considered useful only in highly constrained tasks like reading hand-written zip codes. In the 2010s, however, two things happened: first, the widespread use of computers generated huge datasets of text and images; and second, it was realized that the graphical processing units (GPUs) in all computers could be adapted to hugely speed up the deep learning algorithm. Suddenly, deep learning could be used with data many orders of magnitude larger than previously, and the learning process could be implemented on much faster computers. This, plus the invention of an extension of deep learning called "transformers," led to amazing successes. Speech and face recognition began finally to work at a useful level of accuracy and much more ambitiously, language models learned to generate coherent, contextually relevant responses to questions on virtually any topic. This is best known in the series of GPT (**G**enerative **P**re-training **T**ransformer) programs which stunned the world in their seemingly human-like responses to most queries. These programs take trillions of words as their data and then train billions of parameters embedded in fairly simple computer code.
>
> But GPTs lack any embodiment in the real world. Despite their remarkable linguistic proficiency, these models exhibit a form of cognitive limitation akin to a quadriplegic blind individual. Note that Helen Keller, though blind and deaf, retained touch and muscles and became a brilliant writer. But GPTs lack the ability to directly perceive and interact with the actual external world beyond its vast textual input. As a result, it generates so-called "hallucinations"—simply put, it makes stuff up. For a GPT, Santa Claus might as well as exist as not.[34]

Deep learning: Machine learning using deep neural networks

Although AI algorithms had been in widespread use for years, in 2010, the term AI was still obscure in the research world thanks to a thirty-year academic winter. Hence, it was reserved for film, science fiction, and futurism. Even the University of California at Berkeley, the world's top-rated school in electrical engineering and computer science for many years, had no AI lab per se. Rather, its top AI professors ran research labs focused on ML, computer vision, natural language processing, and other subfields of AI. Even in industry, AI algorithms used in web search, ads, product recommendations, computer games, navigation systems, image processing, and voice recognition were not called AI.

Not until 2012 were breakthroughs made in performance from "deep learning" algorithms trained on millions of data objects instead of thousands. Only then were the ".com" signs along the highways of Silicon Valley taken down, replaced by those with "AI."[35] In this case, "deep" refers to the increase (in 2011, from 3 to 8) in the number of layers in a neural network, a data structure that loosely models interconnected elements of the brain—"neurons"—organized in layers, connected to neighboring neurons in adjacent layers with "weights" (also called parameters), which specify the relative importance of each neuron's current value in combination with the neighbors' values. The process of adjusting the weights based on known data is called "training." After this training, the process of using the neural network is called "inference." During inference, new or unseen input data pass through the network, using the weights to combine the values of neurons along the way to obtain the desired outputs, such as classification labels, regression values, or other types of predictions.

In the 2010s, AI put deep learning center stage, using ingredients from the rise in the prior decade of always-on, ultra-powerful, sensor-rich, ubiquitous, connected devices with highly centralized digital services: vast corpora of data, huge computing resources for training and testing, and distribution of the data, and algorithms and resulting services on fast networks.

Several characteristics of deep networks and deep learning algorithms are relevant to our discussion of GenAI:

- *The algorithms are data-driven.* The algorithms in neural network-based programs do not determine how they will operate. Instead, they are trained on a large corpora of data, such as millions of images of healthy versus diseased tissue, either hand-labeled as such or with innate characteristics.[36] The black box nature of neural network systems leads to the problem of explainability: why did the algorithm produce that particular recommendation or prediction? This central limitation of all data-driven systems has given rise to research on explainable AI, a challenge with no definitive solution (yet).
- *Training requires large amounts of data and computing power* to adjust weights (also called parameters) repeatedly, once per element of training data. This is done for many (often millions, sometimes billions of) weights times many (often millions, sometimes billions of) data elements.
- *Past performance does determine future results.* Deep networks' behavior is determined by the training data. This is seen in a range of applications in which bias in past behavior—for example, by health-care professionals, judges, or employers—is encoded in the AI systems used to automate processes.
- *Parallel processing is key.* This is why GPU (for Graphics Processing Units) chips applied so readily to deep learning systems: they were built to compute live scenes from mathematical descriptions in order to create realistic visuals for immersive computer

games, requiring many parallel, independent calculations that must be done at once. NVidia recognized this as an opportunity, and repackaged its GPUs as AI chips.
- In the 2010s, deep networks, which increase the number of layers in discriminative AI, were used to analyze, estimate, predict, and discriminate but not to create. Creation is the preserve of GenAI, made possible by transformers.

Limitations of deep networks, before the GenAI boom

The wave of AI in the 2010s applied deep learning to perform many analytical and predictive tasks that had eluded AI researchers for decades. The application of these techniques achieved breakthrough results in locating and identifying things in images (e.g. pavement, cars, trees, and buildings for autonomous driving; identifying potential pathologies vs healthy tissue in medical images), playing games such as Go and chess, improving search results, translating text between languages, and automating many decision processes with social implications, such as news and social media ranking, employee performance evaluation, and predictive policing.

Many of the promises made about AI in the 2010s were unrealized because of key limitations in deep learning. By 2020, IBM Watson, an AI-powered digital assistant heralded for helping hospitals and farms as well as offices and factories, had faltered. One major reason was the black box nature of deep learning: even when decisions are correct, but especially when they are wrong—for example, in health care—they need to be justified and explained. Many AI promises did not deliver: Hence, by 2020, AI initiatives at many Fortune 500 companies were shut down.[37]

Which limitations of deep networks mattered in AI during the 2010s?

- *Opacity:* explaining outputs from black-box, data-driven systems is already its own research area, "Explainable AI." Systems that used one AI system to explain the behavior of another were not yet accepted in high-stakes settings.
- *Logical reasoning*: Deep networks do not capture even simple causal relationships.
- *Human context*: Collaboration with humans requires an understanding of norms and interpretation of human intent in behavior. For example, the challenges of driving in the real world, which involves a complex social ballet even at a simple crossroads, made most auto companies' predictions of fully autonomous vehicles by 2020 specious.
- *Planning at multiple scales*: Many types of tasks require creating plans at different time scales and levels of abstraction.
- *Worldview*: Context often determines whether and how a task should be undertaken.

The rise of the Transformer

By the mid-2010s, many challenges in natural language processing, speech processing, and vision had been overcome through variants of deep network architectures and processes. These breakthroughs allowed them to handle sequences of data and add so-called attention mechanisms to find the most salient elements in a sequence being processed.

But with the invention of the transformer architecture at Google, in 2017 such systems became practical to train on large datasets. It had not been practical to train prior architectures (e.g., long short-term networks, or LSTN's, as an example) on exceptionally large datasets. The transformer could be trained in a highly parallel manner and enabled a leap from millions to billions—ultimately, hundreds of billions of parameters and of training data elements.[38] There was a period leading up to the release of GPT3 in 2020 during which models were trained on corpora of increasing orders of magnitude.

The breadth of application of the transformer-based systems extended with other developments into still pictures, sound, voice, video, and more.

Uses and limitations of Transformer-based architectures

Uses: Transformer architectures have now been extended to an astonishing range of applications, well beyond their original domain of tasks in natural language processing. The range of new applications of GenAI systems includes:

- *Language translation:* Transformers enabled a leap in the quality of automated language translation tasks because of their ability to handle sequential data more effectively, with an attention mechanism.
- *Text generation*: They excel at generating coherent and contextually relevant text, not only for writing tasks but enabling applications in chatbots, as well as large-scale content creation: poetry, lesson plans, summarizations, reports, letters, articles, advertising copy, and even book chapters.
- *General problem solving*: Large models are capable of proposing coherent-seeming responses that apply to a staggering range of tasks.
- *Answering questions*: They can answer questions based on a given context by understanding the relationship between different parts of the text.
- *Understanding and creating images*: Combining vision and language, transformers are used in image and video interpretation, as well as the creation of images and videos.
- *Other modalities*: Spoken words, music, biological agents, and many other types of structured data have been used to train transformers for various practical purposes.

Limitations: GenAI systems, and data-driven AI systems more broadly, have significant limitations that the AI research community ardently seeks to overcome. These limitations include:

- *Hallucination*: Large, transformer-based GenAI systems produce outputs based on what is most likely in a given context and based on their training. Even when trained on a carefully curated corpora of writing and illustrations from the best verifiable sources, there is no guarantee of reliable or sensible outputs. There are numerous N partial solutions, for example RAG (for Retrieval-Augmented Generation) that seeks to limit responses to specific citable sources.
- *Computation cost*: Large, transformer-based models famously require vast computational resources for training and inference, both memory for the parameters and training data and processing power for training and for inference.
- *Data needs*: Large models require vast corpora of data for training. There are challenges in curating data in sufficient quantity and quality for these systems. Bias in data that

reflects undesirable patterns is virtually unavoidable, posing a challenge that is the subject of intense research interest as well as regulatory debate.

- *Opacity, again:* Understanding why a transformer model or other data-driven, deep learning–based systems has made a specific prediction or decision is another hot topic in AI research. Because these systems are not engineered in the usual sense, the mystery of their operation also attracts intense research attention. At present, large models are "wrapped" in safety systems that attempt to detect and prevent dangerous or undesirable outputs.
- *Logic gaps*: Despite their apparent ability to reason, large models have been shown to lack basic logical inference.[39]
- *Need for subject-matter experts*: Despite their surprising performance in task areas for which they were not intentionally trained (so-called "emergent behaviors, such as playing chess), large models do not excel at commercially viable levels in most specialized task areas. Their performance can be improved with "fine-tuning," that is, secondary training on specific sub-areas of, for instance, finance or biochemistry. However, to achieve commercially useful performance, subject-matter experts are almost always required to check outputs, just as professional translators must review AI-generated translations of important documents.

GenAI, including its current limitations and the question of whether those limitations are inherent or resolvable, has significant implications for work and the organization of work. "Humans in the loop" play crucial roles in the development and deployment of GenAI in three important ways:

- Training the systems in the first place.[40]
- Curating the data as systems evolve.
- Judging and applying outputs for use and applications.

In the next section, we ask how we could go about assessing the meaning of these powerful tools, with very real limitations for the world of work.

4.

Assessing the impact of emerging GenAI on work: Issues and challenges

The challenge lies in assessing how these new GenAI tools will influence work, organization of work, and, crucially, the place of people in the loop as intelligent systems and tools are deployed. The rapid establishment of LLMs following the development of transformer technology, combined with their actual and potentially extraordinary capabilities, generated widespread hype and fear. For the moment, we are limited to analytic speculation about the impacts. Indeed, how the capabilities and risks of current GenAI will evolve is uncertain. As argued in the prior section, the statistical foundations of AI tools suggest inherent, and likely ongoing, limitations on its capabilities, as well as continuing built-in risks.

This section develops three positions. First, the sudden impact of GenAI rests not only on the power of the tools themselves, but on the infrastructure already in place to launch and then facilitate use. Second, the flood of hype obscures the potential for fundamentally new ways of organizing work roles. It also hides obstacles to deployment of the tools. Third, we must determine which sectors are most exposed to competition using these tools, and assess which sectors will be hit the hardest and the earliest. Analysis should begin with these exposed sectors and how they might respond - not with a catalog of existing tasks. Beginning analysis with the existing tasks amounts to peering at the future through a lens ground in the past. Finally, in considering the broader impact on the economy, we remark on the pace of deployment and the impact on productivity.

Infrastructure at the ready

By the first anniversary of its announcement by Open AI, Chat GPT3 had amassed 180 million users. GenAI suddenly streamed into our consciousness because the infrastructure required—computation, communication, data storage, ML breakthroughs—was already in place to at least begin experimentation and deployment. GenAI is unusual in this regard. Powerful general purpose technologies require the creation of new infrastructure as well as agreements regarding technical standards, to begin widespread use and deployment. This was the case with electricity, in which classic battles over standards were entangled with the basic deployment of generation and distribution. Electricity altered not only streetlights but the configuration of factories. Automobiles, with their internal combustion engines using fossil fuel, substituted for horse-drawn carriages on existing roads until a broader network of roads and gas stations could be built, ultimately changing the organization of cities. Automobile impact depended in part on policy; in the US, for example, public policy often downplayed public transit and led to the construction of the highway system that facilitated the emergence of the suburban world. Indeed, after World War II, Japan built roads and infrastructure in anticipation of, and with a desire to facilitate, the emergence of the automotive sector.

The users of GenAI rely on infrastructure built in previous decades. The existing telephone networks were an initial step in the path toward today's networked economy, with *computer* networking adoption by both consumers and businesses having been influenced directly by the pricing of analog telephone lines. Long before digital networks and optical fiber spread, the pricing system for analog telephones in the US was that of one fixed price for a single line, whatever the usage in a given period. Consequently, in the United States this facilitated home use of systems such as AOL (America Online). Consumers could link their computers to networks for extended periods via existing phone lines in the US at no additional cost per line. Of course, one downside was that, while one person was doing so, no one else on that line could use the phone; nonetheless, use of that line was unrestricted and added no marginal cost. American usage therefore surged not only because the infrastructure built for phone lines already existed, but because regulations were well-suited to adoption for internet use. The adoption of internet standards and the de facto agreement creating the World Wide Web, brought forth the digital world that we know.

First, what will be the pace of deployment?

The practical deployment of GenAI is likely to be slower than the new tool set itself. It will certainly be slower than suggested by advertising claims.[41] Radically new general purpose technological innovation has a powerful impact on firms, markets, and the economy as a whole— but this impact unfolds over time, and GenAI is likely to follow a similar pattern. Given that GenAI tools will be useful in some applications and inappropriate in others, firms and public administrations will need to discover how to use these tools effectively. They must assess the balance of costs and benefits from implementing new processes and generating new products. Experimentation by companies, nongovernmental organizations, and governments will ultimately reveal how these tools will be deployed and distributed given their capabilities and limitations. That experimentation will be about both goals and strategy. It may produce radical change, but also investigations into how existing organizations can achieve incremental improvements in productivity and perhaps quality of work. In that sense, we will see both top-down experimentation, with goals, strategies, and new business models, and bottom-up experimentation that produces incremental changes and tweaks in existing work tasks and jobs.[42] Hence, setting aside the fear of the destruction of most work performed by humans, a more modest observation is that new businesses and business models will be created and, thereby, new jobs and new organization of work will emerge. Existing firms will have to adapt in order to capture market advantage from these tools, and work processes will change. As mentioned in Section 3, the capabilities and limits of these tools suggest that humans and human capabilities will still need to be "in the loop" in order to assess these tools.

Even so, at most firms, particularly those without existing complementary information technology (IT) skills, workers will need to be attracted and trained as an ecosystem of supporting complementary activities emerges. Therefore, the existing pattern likely to continue is one of an elite "best" that quickly deploys advancing IT tools, distantly followed by "the rest."[43]

Moreover, as noted, although AI and GenAI both apply statistical learning to detect patterns that allow plausible recombination and projection of learned elements into new outputs,[44] they have limits and risks that will halt their deployment. If commitments made by an AI bot are binding, as was the case with an Air Canada bot that promised refunds outside the formal Air Canada rules, firms will need to exercise care in using them. If documents are generated with hallucinatory claims or citations, again, caution will be needed.

An emergent supporting ecosystem and people in the loop to offset the limits and risks are likely to lead to even more complex production and finer divisions of labor, as Butollo has argued.[45] The overall balance, as mentioned earlier, can only be the subject of careful speculation.

An approach to assessing the impact on work of GenAI

The debates we are having about what to do about GenAI are due to the ease of experimentation and adoption, combined with its evident capabilities as well as striking flaws and limits. Projecting the impact of GenAI on the world of work is particularly difficult, as it calls for more than assessing existing patterns of deployment, since we are only at its beginning. All we have for now are hints and experiments. Going beyond speculation, we suggest some steps for obtaining a systematic understanding of the potential consequences of GenAI for work.

The conventional approach taken in assessing the prior phase, ML-NN–based AI, focused on existing tasks in jobs. That approach will not suffice for assessing the current situation, as it is, at best, limited and often entirely misplaced. Focusing on existing tasks, a form of bottom-up analysis defined by the existing system, can only reveal incremental possibilities and consequences; because of the nature of the analysis, it cannot contemplate fundamental reconfigurations.[46]

The focus on existing tasks inherently looks at the future through the lens of the past. Starting by examining existing tasks, jobs, and work arrangements inherently requires that we imagine the future as a marginal extension of the present; in other words, focusing on tasks and jobs is akin to starting at the end of the process. Admittedly, that provides some sense of the scale of change that might ensue, but it provides false certainty about what is to be counted.

To start at the beginning, analysis should ask about how tools are used to create competitive advantage or administrative benefit, rather than starting at the end, which entails the structure of tasks, jobs, and work organization that could be prompted by AI. Therefore, the first step, as mentioned earlier, is determining which sectors are most directly affected—and who makes the decisions. Tasks and jobs are always the results of decisions, explicit or implicit, about how to use new technologies.

Thus, the future will be shaped by the ways in which firms use these technologies to create advantage. Exposed *sectors*, rather than tasks, are the place to start. If we regard sectors as speculative use cases, we can use them to develop general, potentially testable, hypotheses about how deployment might affect work. Then, we could develop surveys and case studies to dig deeper and perform some measurements. Generalizations from case studies should be considered in the light of two observations:

- First, many different paths and many different experiments will emerge, reflecting the various regulatory contexts, business communities, and demographics, including skill availability. Different national contexts will induce different sets of experiments and deployments. Firms in each national context and across sectors will make distinct strategic decisions. Hence, a key challenge will be performing analyses and structuring data collection that reflect the broad range of experiments performed and paths taken.
- Second, conversely, consulting firms selling advice on how to capture the gains from these new tools will inevitably take formulaic approaches that narrow the range of paths for clients to explore. Hence, the models proposed by different consulting firms should be studied in order to get a sense of the possible types of experimentation. Do the predominant consulting firms that purport to help clients adapt to GenAI actually propose distinct trajectories of development and strategic options?

Creating an inventory of possibilities, even partial or momentary, requires identifying the differences in competing approaches within sectors, as well as in national variations, as a result of differing contexts.[47] This analysis can be applied to an examination of the various consulting approaches and, where possible, different understandings of the possibilities created by GenAI in various sectors.

We set aside pure speculation about the totally new products and services that could emerge with GenAI. Imagining and generating new businesses and understanding how to capture the gains from technology in radical innovation is the more exciting part of the story. From a historical perspective, this is obvious, as this is how great fortunes are made. The examples in the current digital era are the radical developments by Amazon or Google. However, this is only one part of the story, and perhaps not the dominant part. These dramatic possibilities cannot be ignored, but they should not distract us from the diverse innovations and adjustments that are likely to occur in the existing economy. We consider here how existing firms might act with these new tools to protect or expand existing businesses.

Existing jobs and tasks, the established organization of work, will be relevant in that analysis insofar as they imply the skill sets available in the labor market at the start of the transformation. The new strategies and operations will be shaped, or limited, by the available skills of the existing workforce. An analysis of the impact in this respect could usefully proceed as follows:

Step 1. Consider who is exposed?

First, one could consider which sectors are most exposed to GenAI. A European Central Bank study in 2023 does just that.[48] Such a consideration should try to match the set of capabilities claimed for GenAI, the capabilities that go beyond those of classic ML AI, to the basic product offerings and production processes of sectors or groups of firms. The point is not to examine all tasks but, rather, to consider which undertakings, involving which kinds of jobs and tasks, create or contribute to a distinctive advantageous market position for a particular firm or firms in a specific sector or competitive space. These undertakings are likely to consist of sets of jobs and tasks, but tasks are a result of choices about how to compete and produce, rather than the starting point for analysis. Let us consider some examples, which are illustrative but not comprehensive:

- Professional services, from law to financial advice and strategic consulting, will be directly affected. We already know that GenAI tools speed up the process of designing scenarios for consultants or drafts of legal briefs, as well as computer code itself. Importantly, in all these instances, people will need to review and vet the accuracy of generated materials.
    - Jumping ahead for a moment, we cannot assume that these kinds of jobs will automatically be displaced by implementation of GenAI tools. The cost of producing output may well be reduced, hence, offerings and perhaps employment might be expanded. As suggested in the European Central Bank study, within firms that deployed new technologies, employment levels increased.[49]
- Many sectors will eventually be affected, such as materials and pharmaceuticals. The popular press has already noted the possibilities of new materials suggested by GenAI. These new materials, when matched with application domains, are positioned to change both materials companies and those who deploy these new materials.

Step 2. Consider possible adaptations and strategic moves

Responses to significant competitive challenges involve strategic reorganization, and often relocation, of existing production and distribution. Responses to GenAI should be expected to display the same pattern. Consider two examples of global competition driving change in company strategies and work organization. In the past few years, market pressure in the neoliberal era in the US and the rest of the West arguably came from increasingly global markets. Part of that pressure came from the entry of low-wage countries, in which firms—whether local or from the advanced countries—brought or borrowed, on whatever terms, existing technology. This created immediate pressure on labor costs in the advanced countries and, consequently, led to the relocation of in-house functions, in the form of outsourcing and automation. Benefits accrued to the firms that drove the changes, but not to the domestic workers who were displaced. Automating aspects of existing domestic production was one response. Structuring production so that discrete processes could be transferred abroad easily was another. The result was a reorganization of work, the character of tasks, and their location.

Similarly, product and production innovation abroad forced real changes in American industry—for example, in consumer electronics and automobiles. Japanese product and production innovation in consumer goods and consumer capital goods, for a moment at least, overturned global markets in the 1980s, such as competition with the US in electronics and automobiles. In the electronics industry, Japanese product innovation took professional goods first developed in the US and adapted them for consumer markets, which generated the Walkman and VCR craze. One result was a surge in offshore procurement and production by American firms. In the automobile sector, Japanese production innovation reduced the cost and improved the quality of high-volume goods, while high quality and lean production systems, most popularly associated with Toyota, forced Western firms to reconsider and then reconfigure the very organization of their production lines and supply systems. One aspect of that adaptation in the automotive sector was undoubtedly increased automation and robotics. Moreover, robotics and NC machine tools were, at least initially, deployed quite differently in Japan than in the US, as the Japanese put particular emphasis on flexibility and quality.

Whatever the market opportunities or challenges, our examples suggest that the deployment of the technologies will be influenced by the context in which choices are made, as well as by who makes those choices.[50] In looking ahead to the adaptations that firms might make due to GenAI, context and control are essential matters. We consider, first, the context and, then, aspects of control.

The context—national, sectoral, organization—in which decisions are made is essential. The Japanese context, and the historical trajectory of growth, led to very distinct market conditions, which induced a lean production model. By contrast, American automakers, due to their own context, were entrenched in models relying on mass production, which consequently shaped their initial responses to Japanese competition. When American firms automated existing processes, they tended to view the robots and digitally controlled machine tools as simple alternatives to human labor, rather than as an opportunity for a fundamental rethinking of production processes. In a personal communication some forty years ago with one of the authors, a senior leader at a major American automobile company was reluctant to acknowledge that Japanese firms either had a real cost advantage or had reformulated the production process.[51]

For an analyst hoping to forecast the future impact of a technology, the need to consider the national context complicates matters. It is speculative, but analytically interesting, to consider the evolutionary trajectory of a technology in different contexts. When skilled labor is in short supply, as in Japan now, technology development focuses on digital applications to enable less-skilled workers to perform at the level of more-skilled workers and on worker training to augment their skills.[52] The national context, analytically, involves policy—for example, relative tax rates on capital and labor, and the structure of labor negotiations. It also involves tolerance for venture startups and a willingness to accept firms that break the rules until they can gain a foothold in the market. All of this varies by national context.

The "cookbook" for a new technology has a multitude of recipes (i.e., possibilities and opportunities), each with different consequences for work. GenAI will likely be deployed in diverse ways that reflect these particular contexts, often beginning with the national development trajectories.

Who makes decisions?

Control of decisions is central to our discussion. Who makes the decisions, the purposes of the decisions, and the underlying conceptions of the necessary strategies that drive those decisions, shape how tools are deployed and, in turn, how work is organized. Those who control the technology develop and deploy it for their own purposes and based on their own understanding of opportunities and possibilities. Technology is usually deployed so that those who control it can capture a disproportionate part of the productivity gains from it. This has been true in the West for centuries, from interaction between landlords and peasants to relations between factory bosses and workers.[53]

Differences in who makes the choices about deployment affect the outcomes even within particular national contexts and within specific sectors. A working paper by Acemoglu et al. in 2023 shows that differences in the education of managers affects wages and the share of labor in the US and Denmark.[54] They argue that "changing managerial attitudes and practices towards rent-sharing have been a major contributing factor to the decline in the labor share and slowdown of wage growth" and observe a decline in wages in both countries within five years of the appointment of managers with a business education.

A presentation by Helper et al. in 2004 on the introduction of automation by firms reached conclusions that, not surprisingly, differences in who makes the deployment decisions affects the way the tools are used. They found that two distinct management paradigms influence the deployment of robotic and automation systems.[55] The following is excerpted from that presentation:

- A *pragmatic* approach assumes that the person closest to production has knowledge no one else has, and that there is a big role for learning by doing – that people learn and machines do not, along with the argument that you don't automate until you have simplified the production process.
- A *Taylorist* view sees labor and technology as substitutes, with specialization of tasks permitting the separation of the brain from hand work.

They note that in the Taylorist approach, engineers' ideas can be implemented directly and without worker intervention.

The work by Helper et al. has two implications. First, the difference in paradigms does not seem to influence the decision about *whether* to adopt automation, but it does influence *how* the robots are used. Those employing a Taylorist approach to robots tend to substitute robots for shopfloor workers; in contrast, the "pragmatic" paradigm seems to encourage the use of robots to complement shopfloor workers.

Second, the Taylorist approach, with its labor-displacing implications, is associated with the use of external integrators: that is, external consultants to integrate new technologies.[56] When external integrators are used, less internal investment is made in the analysis and management of data, which reduces the indigenous capacity of a firm to sustain production innovation. With external integrators, data is separated from its context, risking the capture of value by the integrator and limiting the ability of the client firm to imagine innovative next steps.

Let us take this a step further. Tomaso Pardi, Martin Krzywdzinski, and Boy Luethje argue that deployment strategies become political projects. Their analysis of the German auto industry contends that the notion of the fourth industrial revolution is a pretext for the reorganization of work, often to the disadvantage of workers.[57] In a different domain, Judge et al. claim that the crypto craze was a significant political project, derived from an effort to undermine dominant financial institutions.[58] Finally, a classic instance of varied adaptation to common changes in market conditions is the Danish, French, and German response to the flood of American grain onto world markets in the second half of the nineteenth century. The production outcomes in agriculture were a product of the political goals of elite groups, to which Germany and France responded with trade barriers so as to restrict imports. The policies in each country were a result of the famous Iron-Rye political alliances over their shared industrial and agricultural interests. The purpose of the Iron-Rye alliances was to preserve existing production and market arrangements. For Germany, in particular, this entrenched the landed elites in their social as well as economic position. The Danes perceived matters differently, preferring to innovate by feeding cheap imported grain to cows and pigs. They invested in the social and market institutions that supported this redirection of the agriculture sector, subsequently generating an entirely new agricultural category that included dairy farming.[59]

Economic and social contexts, as well as the politics supporting them, modulate and structure the apparent choices available as new technologies unfold. That said, we must consider how the direction of technology development within these choices is influenced by the training, paradigms, ideologies, and resulting conceptions of what is possible held by decision makers. The takeaway is that the ways in which technology is deployed, and who gains advantages from this process of deployment, are unsettled for now.

In sum, speculation on the impact of GenAI must consider the following questions:

- Which sectors are most powerfully exposed?
- What are the different contexts in which decision makers will act, and what problems do they have to solve?
- Who are the decision makers?

5.
Is an AI economy for social prosperity possible?

Will the current trajectory during this era of AI ML, described in Section 2, inevitably continue, with wage polarization and the extraordinary centralization of wealth and control at big tech companies? What might be done to encourage and promote an outcome that augments and improves the quality of work, provides good jobs, and permits social equality? After noting the traditional policy menu intended to create incentives for encouraging investment in people, we consider two less conventional approaches to the challenge.

Forecasting the potential of GenAI has some basic problems. First, we do not know with certainty why and how the GenAI systems produce their answers. The underlying code, as Brian Judge has observed, does not define the system operation. So, code is not law, to paraphrase Larry Lessig.[60] Consequently, we cannot reverse engineer responses - only cross-check them to ensure some sense of fact and reality. Reweighting the elements of a system that itself is opaque makes it difficult to fully assess what is curated and with what consequences. So, we have opaque systems producing uncertain answers, with responses that range from accurate to pure fantasy.

Second, GenAI, however sophisticated the underlying user interface that permits the illusion of conversation may be, remains a system of predictive statistical analysis, generating claims of fact that are not rooted in any actual understanding, sometimes producing outright hallucinations and entirely unreal, alternate worlds. How those limits and risks will evolve is, for now, very uncertain.[61] Validating facts, flagging hallucinations, embedding understanding in outputs, and training and curating the systems are all tasks for which people are well suited. Because the system and its responses are opaque and difficult to shape and control, we believe that it will be essential for humans to remain in the loop at every phase of development:

1. Training and understanding how the system is being developed.
2. Curating training data while building the foundation for an AI firm or sorting out categories of responses that may require rejiggering of the system.
3. Judging by the user whether responses are initially accurate and useful and, then, determining how to apply and use these outputs.

Whether GenAI can generate social prosperity in an abundant economy depends on how these tools are deployed and for what purposes. For now, there is insufficient evidence to determine one way or another whether the effective use of "people," keeping people in the loop, can increase the productive use of labor and capital and avoid dislocations due to the existing risks and limits. An emerging perspective, both in the literature and interestingly emerging in advertising, is that GenAI may have different implications, creating opportunity and productivity for the middle class, rather than displacing it, continuing the bifurcation into very good jobs and poor jobs, with an evisceration of the middle.[62]

As we await the results of that research debate, which will take a long time if there is ever agreement, how can policy encourage competitive strategies that will keep people in the loop and enable firms to develop innovative competitive strategies? The current debate over regulation of AI, though important, is likely to be of secondary importance to how GenAI affects work in the immediate future. Three paths in progress concerning AI regulation are relevant and might shape the trajectory of AI development:[63]

- Establishing guardrails on the use and development of AI is the focus of the EU regulations, the UK debate, and the White House proposals. They do not directly affect the issues about work discussed here.
- Sector-specific rules need to be rethought. The difficulty is that this process will be slow due to struggles over where decisions are made, how they are made, and by

whom. The pace will also be slow because GenAI systems are not only opaque but subject to conflicts of interest about what is appropriate. Moreover, limited real expertise is available to implement actionable rules.[64] Establishing a central government advisory board to support agencies that make decisions about sectors and standards might be useful.

- *Tweaking existing* regulations is a third path. Again, recasting regulations is a very slow process, riddled with conflicting interests and diverse decision-making bodies. International rules will be even slower to develop. Moreover, it is not adequate, and perhaps not even relevant, to simply declare that the existing principles will apply in new conditions. Old settlements will be refought as new rules are formulated, and entirely new questions will emerge.

None of the three paths directly addresses the matter of AI safety, or safe AI.[65] Safe AI is essential to avoid dystopias, and the effort to ensure it is essential. That said, safety alone will not generate the prosperous AI economy that we might hope to achieve.

The crucial question, again, is how the tools will be deployed, to what end, and by whom. Deployment depends on both the "context" in which choices are made and who, with what frameworks, controls the decisions. Policy mostly tries to shape the context in which choices are made, whereas context structures the incentives - both benefits and costly consequences - for decision makers. These incentives are embedded in the character of particular national economies, including demographics, entrenched regulatory and legal systems, and policy choices. Obviously, these contextual constraints can be channeled and tweaked through policy.

Policy goals should focus on making the technology complementary to people, rather than enabling it to displace workers through automation. Daron Acemoglu and Simon Johnson argue:

> The goal should be to deploy generative AI to create and support new occupational tasks and new capabilities for workers. If AI tools can enable teachers, nurse practitioners, nurses, medical technicians, electricians, plumbers, and other modern craft workers to do more expert work, this can reduce inequality, raise productivity, and boost pay by leveling workers up.[66]

Their policy list includes traditional suggestions, such as extending tax breaks for training and employment on par with those for capital goods.[67] They also suggest an increase in "funding for human-complementary technology research, recognizing that this is not currently a private sector priority." In our view, this would be a valuable counterpart to the voluminous funds channeled, usefully and necessarily, toward AI safety.

We suggest a complementary approach that begins with considering the lens through which decision makers who control choices can view their options. Those who control the actual decisions will have strategic maps, models, and worldviews that will likely determine which options, possibilities, and risks they perceive. Therefore, influencing the strategic maps and models of those making the decisions about the deployment of AI will matter greatly, along with the context in which these choices are made. We argue in Section 4 that we need to understand how firms in sectors exposed to GenAI can use these tools to create advantage in the market. The business frameworks, akin to ideologies, that circulate will likely narrow existing views, limit apparent choices, and powerfully influence the decisions about the use of these technologies. As mentioned earlier, the work by Acemoglu et al. and Helper et al. shows that differences in *who* makes the decisions affect the *choices and strategies* adopted. Can those decision frameworks be influenced, or can we only influence the incentives and constraints within these frameworks?

The history of radical general purpose technologies that, like GenAI, have the capacity to influence an entire economy shows that imaginative new approaches and strategies, encompassing previously unfathomed possibilities, are necessary.[68] New strategies and tactics for firms, not just the displacement of workers, are likely to be required, sector by sector, to create distinctive, defensible market advantage. Those who make decisions and the lens through which they view their options when making these decisions will be critical. For example, as previously argued, we know from Acemoglu et al. and Helper et al. that strategies developed by those trained by MBA programs and standard consultancy firms to use cookie-cutter solutions tend to develop programs that depend more heavily on simply displacing workers and substituting capital for labor. The alternative of rethinking strategy and reimaging jobs, tasks, and reformulating work organization thus receives no consideration.

How, then, beyond reshuffling incentives such as taxes, can the strategic and tactical imagination be shaped, or perhaps re-shaped? Conventionally, research studies on existing deployment experiments and strategic analyses of possibilities will be useful. But we can go beyond that. Certainly, significant prizes and publicity for firms that develop innovative strategies for deploying GenAI tools in a manner that augments the possibilities, skills, and well-being of the workforce could establish that effectively using the workforce can be a competitive advantage.

A more radical, third option would be to establish a not-for-profit public service business consultancy to work with firms to generate imaginative options and operational plans to implement them. To have the necessary access to the C-suite in firms, such a consultancy would need to be established with and initially led by widely respected business leaders providing leadership for whom this might be a public service culmination of an already distinguished private sector career, a capstone. Funding, board, and operational teams would need to be tripartite, with funding and participation from public, private sector, and worker organizations. A variety of approaches include working with initial client firm's pro-bono on the basis that traditional fees are due if the exercise proves useful. The purpose would be to demonstrate that creative, and potentially radical, rethinking not only can create business advantages, but can begin a new dialogue about the importance of keeping people in the loop of AI applications given the distinctive, decisive contributions humans make. Certainly, creating such an organization, establishing working teams, and developing clients would be difficult at best. Yet, just as King Arthur's knights of the round table re-educated a warrior community, such a consultancy could generate a new discourse and new norms.[69] Such a consultancy could begin by working with a few willing and visible firms to explore new strategies, developing its own literature based on its experience.

Gen AI will bring change. Let us make the decisions to create social prosperity with competitive firms.

[1] Consistent with the theme of this article, the initial draft, revised and rewritten by Nitzberg and Zysman, was generated by "Claude," the tool created by Anthropic.

[2] John Zysman, Martin Kenney, & Laura Tyson, Beyond Hope and Despair: Developing Healthy Communities in an Era of Intelligent Tools, Innovation Policy Lab Working Paper Series 2019-01, https://munkschool.utoronto.ca/media/2642/download?inline=/download/.

[3] Set aside the differing definitions and analytics of the notion of productivity and simply note that it highlights the link between the inputs into the productive economy and its outputs as measured, albeit very imperfectly, in the market economy. See, e.g., Diane Coyle, *The Measure of Progress* (Princeton University Press: forthcoming in 2025).

[4] Zysman, J., Kenney, M., & Tyson, L. Beyond Hype and Despair: Developing Healthy Communities in the Era of Intelligent Tools (January, 2019), available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3414691

[5] Defined by Perplexity, itself a GenAI tool, on March 21, 2024.

[6] M. Kenney & J. Zysman, The Rise of the Platform Economy, *Issues in Science and Technology,* 32(3) (2016), retrieved from https://issues.org/rise-platform-economy-big-data-work; J. Zysman & M. Kenney, The Next Phase in the Digital Revolution: Intelligence Tools, Platforms, Growth, Employment, *Communications of the ACM,* 61(2) (2018), 54-63, retrieved from https://cacm.acm.org/magazines/2018/2/224635-the-next-phase-in-the-digital-revolution/abstract/; A. Garcia Calvo, M. Kenney, & J. Zysman, Understanding Work in the Online Platform Economy: The Narrow, the Broad, and the Systemic Perspectives, Berkeley Round Table on the International Economy (2022, July 15), available at SSRN: https://ssrn.com/abstract=4164068 or http://dx.doi.org/10.2139/ssrn.4164068/.

[7] Laura D. Tyson & John Zysman, Automation, AI, and Work, *Journal of the American Academy of Arts & Sciences,* 151(2) (2022), retrieved from https://www.amacad.org/sites/default/files/daedalus/downloads/Daedalus_Sp22_AI-%26-Society_0.pdf.

[8] Zysman, J., Kenney, M., & Tyson, L. Beyond Hype and Despair: Developing Healthy Communities in the Era of Intelligent Tools (January, 2019), available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3414691

[9] P. A. David, Narrow Windows, Blind Giants and Angry Orphans, Working Paper no. 10, Center for Economic Policy Research, Stanford University (1986); a point hammered in the literature on the example of the Qwerty keyboard for typing; it is not only the answer that becomes embedded but a solution. See also D. Noble, Social Choice in Machine Design: The Case of Automatically Controlled Machine Tools, and a Challenge for Labor, *Politics & Society,* 8(3-4) (1978), 313-347. Noble argues that technology "bears the social imprint of its authors," in the sense that there is always a range of possibilities or alternatives for its development delimited over time by the social choices of those with the power to choose.

[10] https://en.wikipedia.org/wiki/AI_winter AI winter is, as reported by Wikipedia, a period of reduced funding and interest in AI research. There have been several doubt times between the hype times. - Wikipedia has an article on them. When research failed to deliver on a promise, there would be reduced interest and funding.

[11] Laura Tyson and I have presented our views on this. The article notes those on whose work we drew. Tyson & Zysman, Automation, AI & Work. Much of the discussion here, including the language in some cases, is drawn from that article.

[12] Laura D. Tyson, John Zysman; Automation, AI & Work. Daedalus 2022; 151 (2): 256–271. doi: https://doi.org/10.1162/daed_a_01914

[13] S. Albanesi et al., Reports of AI Ending Human Labour May Be Greatly Exaggerated, European Central Bank (2023, November 28), retrieved from https://www.ecb.europa.eu/pub/economic-research/resbull/2023/html/ecb.rb231128~0a16e73d87.en.html

[14] F. Butolo, The Rebound Effects of Automation, Weizenbaum Series #27, Berlin (2024); Butolo argued that ever finer and more sophisticated divisions of labor balance the labor markets and drive demand for workers, shifting the required skills. https://www.weizenbaum-library.de/items/9d438f0a-f0e4-41d2-b933-1bcdf8a0d2bf Also, as argued by Hal Varian, demographic shifts are crucial to the condition of labor markets and influence the adoption of digital technologies; Automation versus procreation (aka bots versus tots), VOXeu, CEPR, March 30, 2020, https://cepr.org/voxeu/columns/automation-versus-procreation-aka-bots-versus-tots/.

[15] Tyson & Zysman, Automation, AI & Work; H. J. Holzer, Understanding the Impact of Automation on Workers, Jobs, and Wages, Brookings Institution (2022, January 19), retrieved from https://www.brookings.edu/articles/understanding-the-impact-of-automation-on-workers-jobs-and-wages/; D. Acemoglu & P. Restrepo, Tasks, Automation, and the Rise in U.S. Wage Inequality, *Econometrica,* 90(5) (2022), 1973-2016, retrieved from https://economics.mit.edu/sites/default/files/2022-10/Tasks%20Automation%20and%20the%20Rise%20in%20US%20Wage%20Inequality.pdf; D. Acemoglu & P. Restrepo, Unpacking Skill Bias: Automation and New Tasks, *AEA Papers and Proceedings,* 110 (2020), 356-361, DOI: https://doi.org/10.1257/pandp.20201063; A. Di Battista, S. Grayling, E. Hasselaar, T. Leopold, R. Li, M. Rayner, & S. Zahidi, Future of Jobs Report 2023, World Economic Forum (2023, May), retrieved from https://www3.weforum.org/docs/WEF_Future_of_Jobs_2023.pdf.

[16] While RBTC had driven market forces of inequality, the extreme inequality that sets apart the .1% from the 1% and indeed the .01% from the 1% is a product of tax policy and financial market regulations.

[17] D. Acemoglu, A. He, & D. le Maire, Eclipse of Rent-Sharing: The Effects of Managers' Business Education on Wages and the Labor Share in the US and Denmark, National Bureau of Economic Research, Working Paper 29874 (2022), retrieved from http://www.nber.org/papers/w29874/.

[18] T. Pardi, M. Krzywdzinski, & B. Luethje, Digital Manufacturing Revolutions as Political Projects and Hypes: Evidences from the Auto Sector, International Labour Organization, Working Paper 3 (2020, April), retrieved from https://www.ilo.org/wcmsp5/groups/public/---dgreports/---inst/documents/publication/wcms_742905.pdf.; S. Helper & J. Kiehl, Developing supplier capabilities: Market and non-market approaches, *Industry and Innovation* 11(1-2) (2004): 89-107

[19] Grabango, Checkout-Free Technology. (n.d.), https://www.grabango.com. Note that Grabango’s approach is quite distinct from Amazon’s now abandoned technology.

[20] How far this will extend is speculation. The tendency is most evident in, for example, fleet sales where the buyer can fix predictable costs, since the purchase is by land miles by trucks or air miles for airplane engines. While the purveyor can learn and internalize the learning, keeping their costs of provision below the sales price.

[21] J. Zysman, The Algorithmic Revolution? The Fourth Service Transformation, *Communications of the ACM,* 49(7) (2006, July), retrieved from https://cacm.acm.org/magazines/2006/7/5869-the-algorithmic-revolutionthe-fourth-service-transformation/abstract/; J. Zysman et al., Services with Everything: The ICT-Enabled Digital Transformation of Services, in *The Third Globalization: Can Wealthy Nations Stay Rich in the Twenty-First Century?,* ed. D. Breznitz, & J. Zysman, 99-129, Oxford University Press (2013, March).

[22] M. Kenney & J. Zysman, The Rise of the Platform Economy, *Issues in Science and Technology*, 32(3) (2016), retrieved from https://issues.org/rise-platform-economy-big-data-work/ (https://issues.org/rise-platform-economy-big-data-work/); D. Bearson, M. Kenney, & J. Zysman, Measuring the impacts of labor in the platform economy: new work created, old work reorganized, and value creation reconfigured, *Industrial and Corporate Change*, 30(3) (2020, December), 536–563, DOI: https://doi.org/10.1093/icc/dtaa046 (https://doi.org/10.1093/icc/dtaa046).

[23] M. Keutel, G. Lunawat, & M. Schmid, Future of Retail Operations: Winning in a Digital Era, McKinsey & Company (2020, January), 4-5, 14*,* retrieved from https://www.mckinsey.com/~/media/mckinsey/industries/retail/our%20insights/future%20of%20retail%20operations%20winning%20in%20a%20digital%20era/mck_retail-ops-2020_fullissue-rgb-hyperlinks-011620.pdf.

[24] C. B. Frey & M. Osborne, The Future of Employment: How Susceptible Are Jobs to Computerisation? Oxford Martin Programme on the Future of Work (2013, September), retrieved from https://www.oxfordmartin.ox.ac.uk/downloads/academic/future-of-employment.pdf.

[25] Nor do they consider that apart from the financial costs, these indeed, these tools may be adopted to address labor power. All too often part of the objective of introducing automation is eliminating workers to limit political or union pressures from the union movement.

[26] See *Economist*, December 2, 2023. There are several relevant articles: Why Economists Are at War over Inequality (2023, November), retrieved from https://www.economist.com/finance-and-economics/2023/11/30/income-gaps-are-growing-inexorably-arent-they; Welcome to a Golden Age for Workers (2023, November), retrieved from https://www.economist.com/finance-and-economics/2023/11/28/welcome-to-a-golden-age-for-workers; and Why Economists Are at War over Inequality (2023, November), retrieved from https://www.economist.com/finance-and-economics/2023/11/30/income-gaps-are-growing-inexorably-arent-they/

[27] Varian, H. Automation versus procreation (aka bots versus tots) (2020, March), retrieved from https://cepr.org/voxeu/columns/automation-versus-procreation-aka-bots-versus-tots/.

[28] Butollo, The Rebound Effects of Automation.

[29] Marr, B. (2018, May 21). How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read. Bernard Marr & Co, retrieved from https://bernardmarr.com/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/

[30] Many of Alison Gopnik's works support this argument. For example: https://www.psychologicalscience.org/observer/children-creativity-intelligence/. See her homepage for the rich collection, http://alisongopnik.com/Alison_Gopnik_Books.htm.

[31] We distinguish existential threats into two categories: Threats generated by the systems acting with, or acting as if they have, actual intent: Systems where a mistaken instruction sets off a debacle.

[32] This discussion of machine learning based AI draws from, quotes, and rephrases two articles by Mark Nitzberg and John Zysman:
Zysman, J. & Nitzberg, M. (2021, March 15). Algorithms, Data, and Platforms: The Diverse Challenges of Governing AI. Journal of European Public Policy, 29(11);
Zysman, J. & Niztberg, M. (2020, October). Governing AI: Understanding the Limits, Possibility, and Risks of AI in an Era of Intelligent Tools and Systems. BRIE Working Paper No., 2020-5, available at SSRN:https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3681088

[33]Most estimates of the AI winter, 1985 to 2010. The winter was a period where the general public and the sponsors like, DARPA, saw that the paths being followed at the time were not likely to yield results. Knowledge representation systems, expert systems, were not delivering on their promise.

[34] D. Mumford, Consciousness, Robots, and DNA, forthcoming in *Proceedings of the First International Congress of Basic Science,* Beijing, July 2023, preprint accessed Apr. 7, 2024, https://www.dam.brown.edu/people/mumford/blog/2024/Cons.html.

[35] A. Krizhevsky, I. Sutskever, & G. E. Hinton, Imagenet Classification with Deep Convolutional Neural Networks, *Advances in Neural Information Processing Systems*, 25(2) (2012), DOI:10.1145/3065386

[36] In the case of deep reinforcement learning, systems are trained by programming an exploration of the relevant world, such as making legal moves on a chessboard, where a scoring mechanism called a reward function gives a rating of proximity to a desired outcome.

[37] J. Varian, Why Most Companies Are Failing at Artificial Intelligence: Eye on A.I. Fortune (2019, October), retrieved from https://fortune.com/2019/10/15/why-most-companies-are-failing-at-artificial-intelligence-eye-on-a-i/.

[38] A. Vaswani et al., Attention Is All You Need, *Advances in Neural Information Processing Systems,* 30(1) (2017, June), retrieved from https://papers.nips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf.

[39] L. Berglund et al., The Reversal Curse: LLMs Trained on "A Is B" Fail to Learn "B Is A," forthcoming in *Proceedings of ICLR 2024*, https://arxiv.org/pdf/2309.12288.pdf, accessed April 7, 2024.

[40] Training and curating have overlapping meanings: choosing training data, sometimes analyzing the data and pruning out the unwanted stuff. Training requires data. Selecting the data to use for training may be considered part of the whole training process but for the computer programmer, training data is just assumed to be a given. Teaching folks to code is no longer enough. Programmers must understand statistics and experimental design. Hence, "data science" vs "computer science

[41] Erik Brynjolfsson and Laura Tyson have provided an excellent overview of the productivity and deployment problem: Laura Tyson and Erik Brynjolfsson (co-chairs), John Haltiwanger, Larry Katz, Michael Strain, NASEM Committee on AI and the Workforce Productivity Effects of AI.

[42] Thanks to Florian Butolo for emphasizing this point.

[43] D. Andrews, C. Criscuolo, & P. N. Gal., The Best versus the Rest; The Global Productivity Slowdown, Divergence across Firms and the Role of Public Policy, OECD Productivity Working Papers, no. 5, November 2016.

[44] The phrasing in this sentence was "generated," in part, with the help of Claude.

[45] Op. Cit. Butollo, The Rebound Effects of Automation.

[46] Certainly, we could test that proposition by carefully calculating tasks in era 1 and setting them against the realities that emerged. That though seems a good deal of work to demonstrate what would seem to be evident. The worlds of work that emerged were not logical extensions of their predecessors. S. Lund et al., The Future of Work after COVID-19, McKinsey Global Institute (2019, February), retrieved from https://www.mckinsey.com/featured-insights/future-of-work/the-future-of-work-after-covid-19#/; J. Manyika et al., Jobs Lost, Jobs Gained: Workforce Transitions in a Time of Automation, McKinsey Global Institute (2017, December), retrieved from https://www.mckinsey.com/~/media/mckinsey/industries/public%20and%20social%20sector/our%20insights/what%20the%20future%20of%20work%20will%20mean%20for%20jobs%20skills%20and%20wages/mgi-jobs-lost-jobs-gained-executive-summary-december-6-2017.pdf; V. Ratcheva, T.A. Leopold, & S. Zahidi, S., Jobs of Tomorrow: Mapping Opportunity in the New Economy, World Economic Forum (2020, January), retrieved from https://www3.weforum.org/docs/WEF_Jobs_of_Tomorrow_2020.pdf.

[47] S. Berger, How We Compete: What Companies Around the World Are Doing to Make It in Today's Global Economy, Crown, 2005.

[48] Albanesi et al., Reports of AI Ending Human Labour May Be Greatly Exaggerated.

[49] Ibid.

[50] D. Acemoglu & S. Johnson, *Power and Progress,* PublicAffairs, New York (2023); D. Acemoglu & S. Johnson, Choosing AI's Impact on the Future of Work, *Stanford Social Innovation Review* (2023, October), retrieved from https://ssir.org/articles/entry/ai-impact-on-jobs-and-work; D. Acemoglu, D. Autor, & S. Johnson, Can We Have Pro-Worker AI? Choosing a Path of Machines in Service of Minds, MIT Shaping the Future of Work Initiative (2023, September), retrieved from https://shapingwork.mit.edu/wp-content/uploads/2023/09/Pro-Worker-AI-Policy-Memo.pdf. Acemoglu and Johnson have brought the notion of who controls into the mainstream debate. See above. The arguments about context are a core part of comparative political economy.

[51] John Zysman, personal conversation.

[52] K. Kushida, Japan's Aging Society as Technological Opportunity: How Japan's Extreme Demographics are Shaping Innovation Trajectories in Automation, Artificial Intelligence and Intelligence Augmentation, Carnegie Endowment for International Peace (forthcoming 2024); R. Bernstein, Google Chief Economist Weighs Bots vs Tots, Edhat (2020, March 13), retrieved from https://www.edhat.com/news/google-chief-economist-weighs-bots-vs-tots/

[53] Acemoglu & Johnson, *Power and Progress*; Acemoglu & Johnson, Choosing AI's Impact on the Future of Work; Acemoglu et al., Can We Have Pro-Worker AI?

[54] D. Acemoglu, A. He, & D. le Maire, Eclipse of Rent-Sharing: The Effects of Managers' Business Education on Wages and the Labor Share in the US and Denmark, National Bureau of Economic Research, Working Paper 29874 (2022, March), retrieved from https://www.nber.org/system/files/working_papers/w29874/w29874.pdf. Their sweeping analysis develops this point.

[55] See an NBER presentation by Susan Helper, Raphael Martins, and Robert Seamans based on their 2019 work Complements or Substitutes: Firm Level Management of Labor and Technology. See also Developing Supplier Capabilities: Market and Non-market Approaches (tandfonline.com); S. Helper & J. Kiehl, Developing supplier capabilities: Market and non-market approaches, *Industry and Innovation* 11(1-2) (2004): 89-107.

[56] This makes sense because an internal analysis would turn to the existing operations teams and workers, and external integrators would take a more formal abstracted analysis without attention to existing knowledge.

[57] Pardi et al., Digital Manufacturing Revolutions as Political Projects and Hypes.

[58] B. Judge, B. Eichengreen, & J. Zysman, The Mirage of Decentralized Finance, Berkeley Roundtable on International Economies (2023, May), available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4459315/.

[59] A. Gerschenkron, Economic Backwardness in Historical Perspective, Harvard University Press, 1962.

[60] S. Russell; Mark Nitzberg; and Brian Judge "When Code Isn't Law: Rethinking Regulation for Artificial Intelligence", In draft form April 2024;
This reference reverses Larry Lessig's infamous comment that Code is law in cyberspace. See: L. Lessig, *Code and Other Laws of Cyberspace,* Basic Books (1999).

[61] A lively discussion of these issues is in Gary Marcus's blog. See garymarcus@substack.com

[62] How One Tech Skeptic Decided A.I. Might Benefit the Middle Class. Steve Lohr. New York Times April 2, 2024. A version of this article appears in print on April 2, 2024, Section B, Page 1 of the New York edition with the headline: Tech Skeptic Finds Benefit In A.I. for the Middle Class.

[63] M. Nitzberg & J. Zysman, Algorithms, Data, and Platforms: The Diverse Challenges of Governing AI, *Journal of European Public Policy,* 29(11) (2022, July), retrieved from https://www.tandfonline.com/doi/abs/10.1080/13501763.2022.2096668; J. Zysman & M. Nitzberg, Governing AI: Understanding the Limits, Possibility, and Risks of AI in an Era of Intelligent Tools and Systems, Woodrow Wilson Center, Science and Technology Innovation Program (2020, December), retrieved from https://www.wilsoncenter.org/sites/default/files/media/uploads/documents/WWICS%20Governing%20AI.pdf. Judge, Nitzberg, and Zysman have made earlier statements on these issues, separately and together.

[64] It might be useful to establish a central government consulting agency to be able to advise sectoral and standards decision-making bodies.

[65] Stuart Russell makes the distinction between making AI safe and making safe AI. The first is, taking these systems whose behavior cannot be entirely understood or predicted, and trying to surround them with guard rails. The second is to make a new generation of AI built from the ground up from components that can be proven to adhere to their prescribed designs, and to behave in predictable ways. Stuart Russell, *Human Compatible: Artificial Intelligence and the Problem of Control*, Viking, 2019.

[66] Acemoglu et al., Can We Have Pro-Worker AI?

[67] Ibid., 1. As they write:

> 1. Equalize tax rates on employing workers and on owning equipment/ algorithms to level the playing field between people and machines.
> 2. Update Occupational Safety and Health Administration rules to create safeguards (i.e., limitations) on the surveillance of workers. Finding ways to elevate worker voice on the direction of development could also be helpful.
> 3. Increase funding for human-complementary technology research, recognizing that this is not currently a private sector priority.
> 4. Create an AI center of expertise within the government, to help share knowledge among regulators and other officials.
> 5. Use that federal expertise to advise on whether purported human complementary technology is appropriate to adopt in publicly provided education and healthcare programs, including at the state and local level.

[68] M. Chui et al., The Economic Potential of Generative AI: The Next Productivity Frontier, McKinsey & Company (2023, June), retrieved from https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#introduction/.

[69] Thanks to Steve Cohen who used this analogy to show that the French Planning Commission in the immediate post war years did more than just allocate capital but led by Jean Monnet sought to re-educate the business community. At times Cohen's position was treated by scholars as a cute analogy but only that. However, the memoirs of François Bloch-Laine whose role as head of the French Tresor, through whom the French plan had to operate, stated explicitly that the plan was a form of adult education not just a source of capital. Cohen was right.