Introduction

By Ajay Agrawal, Joshua Gans, and Avi Goldfarb

Artificial intelligence (AI) technologies have advanced rapidly over the last several years. As the technology continues to improve, it may have a substantial impact on the economy with respect to productivity, growth, inequality, market power, innovation, and employment. In 2016, the White House put out several reports emphasizing this potential impact. Despite its importance, there is little economics research on the topic. The research that exists is derived from past technologies (such as factory robots) that capture only part of the economic reach of AI. Without a better understanding of how AI might impact the economy, we cannot design policy to prepare for these changes.

To address these challenges, the National Bureau of Economic Research held its first conference on the Economics of Artificial Intelligence in September 2017 in Toronto, with support from the NBER Economics Digitization Initiative, the Sloan Foundation, the Canadian Institute for Advanced Research, and the University of Toronto's Creative Destruction Lab. The purpose of the conference was to set the research agenda for economists working on AI. The invitation emphasized these points as follows:

> "The context is this: imagine back to 1995 when the internet was about to begin transforming industries. What would have happened to economic research into that revolution had the leading economists gathered to scope out a research agenda at that time? Today, we are facing the same opportunity with regard to AI. This time around we are convening a group of 30 leading economists to scope out the research agenda for the next 20 years into the economics of AI.""

Scholars who accepted the invitation were asked to write up and present ideas around a specific topic related to their expertise. For each paper, a discussant was assigned. Throughout the conference, in presentations, discussions, and debates, participants weighed in with their ideas for what the key questions will be, what research has already shown, and where the challenges will lie. Pioneering AI researchers Geoffrey Hinton, Yann LeCun, and Russ Salakhutdinov attended, providing useful context and detail about the current and expected future capabilities of the technology. The conference was unique because it emphasized the work that still needs to be done, rather than the presentation of standard research papers. Participants had the freedom to engage in informed speculation and healthy debate about the most important areas of inquiry.

This volume contains a summary of the proceedings of the conference. We provided authors with few constraints. This meant diversity in topics and chapter style. Many of the papers contained herein are updated versions of the original papers and presentations at the conference. Some discussants commented directly on the papers while others went further afield, emphasizing concepts that did not make it into the formal presentations but instead arose as part of debate and discussion. The volume also contains a small number of papers that were not presented at the conference, but nevertheless represent ideas that came up in the general discussion and that warranted inclusion in a volume describing the proceedings of the conference.

We categorize the papers into four broad themes. First, several papers emphasize the role of AI as a general purpose technology, building on the existing literature on general purpose technologies from the steam engine to the internet. Second, many papers highlight the impact of AI on growth, jobs, and inequality, focusing on research and tools from macro and labor economics. Third, five chapters discuss machine learning and economic regulation, with an emphasis on microeconomic consequences and industrial organization. The final set of chapters explore how AI will affect research in economics.

Of course, these themes are not mutually exclusive. Discussion of AI as a GPT naturally leads to discussions of economic growth. Regulation can enhance or reduce inequality. And AI's impact on economics is a consequence of it being a general purpose technology for scientific discovery (as emphasized in chapter 4 by Cockburn, Henderson, and Stern). Furthermore, a handful of concepts cut across the various parts, most notably the role of humans as AI improves and the interaction between technological advance and political economy.

Below, we summarize these four broad themes in detail. Before doing so, we provide a definition of the technology that brings together the various themes.

What is artificial intelligence?

The Oxford English Dictionary defines artificial intelligence as “the theory and development of computer systems able to perform tasks normally requiring human intelligence.” This definition is both broad and fluid. There is an old joke among computer scientists that artificial intelligence defines what machines cannot yet do. Before a machine could beat a human expert at chess, such a win would mean artificial intelligence. After the famed match between IBM’s Deep Blue and Gary Kasparov, playing chess was called computer science and other challenges became artificial intelligence.

The chapters in this volume discuss three related, but distinct, concepts of artificial intelligence. First, there is the technology that has driven the recent excitement around artificial intelligence: Machine learning. Machine learning is a branch of computational statistics. It is a tool of prediction in the statistical sense, taking information you have and using it to fill in information you do not have. Since 2012, the uses of machine learning as a prediction technology have grown substantially. One set of machine learning algorithms in particular, called “deep learning”, has been shown to be useful and commercially viable for a variety of prediction tasks from search engine design to image recognition to language translation. The chapter in the book authored by us—Agrawal, Gans, and Goldfarb—emphasizes that rapid improvements in prediction technology can have a profound impact on organizations and policy (chapter 3). The chapter by Taddy (chapter 2) defines prediction with machine learning as one component of a true artificial intelligence and provides detail on the various machine learning technologies.

While the recent interest in AI is driven by machine learning, computer scientists and philosophers have emphasized the feasibility of a true artificial general intelligence that equals or exceeds human intelligence (Bostrom 2014, Kaplan 2016). The closing sentence of this volume summarizes this possibility bluntly. Daniel Kahneman writes, “I do not think that there is very much that we can do that computers will not eventually be programmed to do.” The economic and societal impact of machines that surpass human intelligence would be extraordinary. Therefore—whether such an event occurs imminently, in a few decades, in a millennium, or never—it is worth exploring the economic consequences of such an event. While not a focal aspect of any chapter, several of the chapters in this volume touch on the economic consequences of such superintelligent machines.

A third type of technology that is often labeled “artificial intelligence” is better-seen as a process: automation. Much of the existing empirical work on the impact of artificial intelligence uses data on factory automation through robotics. Daron Acemoglu and Pascual Restrepo use data on factory robots to explore the impact of AI and automation on work (chapter 8). Automation is a potential consequence of artificial intelligence, rather than artificial intelligence per se. Nevertheless, discussions of the consequences of artificial intelligence and automation are tightly connected.

While most chapters in the book focus on the first definition—artificial intelligence as machine learning—a prediction technology, the economic implications of artificial general intelligence and automation receive serious attention.

AI as a GPT

A general purpose technology (GPT) is characterized by pervasive use in a wide range of sectors combined with technological dynamism (Bresnahan and Trajtenberg 1995). General purpose technologies are enabling technologies that open up new opportunities. While electric motors did reduce energy costs, the productivity impact was largely driven by increased flexibility in the design and location of factories (David 1990). Much of the interest in artificial intelligence and its impact on the economy stems from its potential as a GPT. Human intelligence is a general purpose tool. Artificial intelligence, whether defined as prediction technology, general intelligence, or automation, similarly has potential to apply across a broad range of sectors.

Brynjolfsson, Rock, and Syverson (chapter 1) argue the case for AI as a GPT. They focus on machine learning and identify a variety of sectors in which machine learning is likely to have a broad impact. They note expected continual technological progress in machine learning and a number of complementary innovations that have appeared along with machine learning. By establishing AI as a GPT, they can turn to the general lessons of the productivity literature on GPTs with respect to initially low rates of productivity growth, organizational challenges, and adjustment costs. They propose four potential explanations for the surprisingly low measured productivity growth given rapid innovation in AI and related technologies – false hopes, mismeasurement, redistribution, and implementation lags – and conclude that lags due to missing complementary innovations are most likely the primary source of missing productivity growth: "… an underrated area of research involves the complements to the new AI technologies, not only in areas of human capital and skills, but also new processes and business models. The intangible assets associated with the last wave of computerization were about ten times as large as the direct investments in computer hardware itself."

Henderson's comment emphasizes the impact of a GPT on employment and the distribution of income, directly linking the discussion of AI as a GPT to questions addressed in the section on Growth, Jobs, and Inequality. She agrees with the central thesis "One of the reasons I like the paper so much is that it takes seriously an idea that economists long resisted – namely that things as nebulous as "culture" and "organizational capabilities" might be a) very important, b) expensive, and c) hard to change." At the same time, she adds emphasis on additional implications: "… I think that the authors may be underestimating the implications of this dynamic in important ways… I'm worried about the transition problem at the societal level quite as much as I'm worried about it at the organizational level."

The next chapters provide micro-level detail on the nature of AI as a technology. Taddy (chapter 2) provides a broad overview of the meaning of intelligence in computer science. He then provides some technical detail on two key machine learning techniques, deep learning and reinforcement learning. He explains the technology in a manner intuitive to economists: "Machine Learning is a field that thinks about how to automatically build robust predictions from complex data. It is closely related to modern Statistics, and indeed many of the best ideas in ML have come from Statisticians (the lasso, trees, forests, etc.). But whereas statisticians have often focused on *model inference* – on understanding the parameters of their models (e.g., testing on individual coefficients in a regression) – the ML community has been more focused on the single goal of maximizing predictive performance. The entire field of ML is calibrated against 'out-of-sample' experiments that evaluate how well a model trained on one dataset will predict new data."

Building on ideas in Agrawal, Gans, and Goldfarb (2018), we argue in chapter 3 that the current excitement around AI is driven by advances in prediction technology. We then show that modeling AI as a drop in the cost of prediction provides useful insight into the microeconomic impact of AI on organizations. We emphasize that AI is likely to substitute for human prediction, but complement other skills such as human judgment—defined as knowing the utility or valuation function. "… a key departure from the usual assumptions of rational decision-making is that the decision-maker does not know the payoff from the risky action in each state and must apply *judgment* to determine the payoff… Judgment does not come for free."

Prat's comment emphasizes that economists typically assume that the valuation function is given, and that loosening that assumption will lead to a deeper understanding of the impact of AI on organizations. He offers an example to illustrate: "Admissions offices of many universities are turning to AI to choose which applicants to make offers to. Algorithms can be trained on past admissions data. We observe the characteristics of applicants and the grades of past and present students… The obvious problem is that we do not know how admitting someone who is likely to get high grades is going to affect the long-term payoff of our university… Progress in AI should induce our university leaders to ask deeper questions about the relationship between student quality and the long-term goals of our higher-learning institutions. These questions cannot be answered with AI but rather with more theory-driven retrospective approaches or perhaps more qualitative methodologies."

The next chapters explore AI as a GPT that will enhance science and innovation. After reviewing the history of artificial intelligence, Cockburn, Henderson, and Stern (chapter 4) provide empirical support for the widespread application of machine learning in general, and deep learning in particular, in scientific fields outside of computer science. "… we develop what we believe is the first systematic database that captures the corpus of scientific paper and patenting activity in artificial intelligence… we find striking evidence for a rapid and meaningful shift in the application orientation of learning-oriented publications, particularly after 2009." The authors make a compelling case for AI as a general purpose tool in the method of invention. The paper concludes by discussing the implications for innovation policy and innovation management: "the potential commercial reward from mastering this mode of research is likely to usher in a period of racing, driven by powerful incentives for individual companies to acquire and control critical large datasets and application-specific algorithms."

Mitchell's comment emphasizes the regulatory effects of AI as a GPT for science and innovation—in terms of intellectual property, privacy, and competition policy: "It is not obvious whether AI is a general purpose technology for innovation or a very efficient method of imitation. The answer has a direct relevance for policy. A technology that made innovation cheaper would often (but not always) imply less need for strong IP protection, since the balance would swing toward limiting monopoly power and away from compensating innovation costs. To the extent that a technology reduces the cost of imitation, however, it typically necessitates greater protection." Several later chapters detail these and other regulatory issues.

Agrawal, McHale, and Oettl (chapter 5) provide a recombinant growth model that explores how a general purpose technology for innovation could affect the rate of scientific discovery. "… instead of emphasising the potential substitution of machines for workers in existing tasks, we emphasise the importance of AI in overcoming a specific problem that impedes human researchers – finding useful combinations in complex discovery spaces… we develop a relatively simple combinatorial-based knowledge production function that converges in the limit to the Romer/Jones function… If the curse of dimensionality is both the blessing and curse of discovery, then advances in AI offer renewed hope of breaking the curse while helping to deliver on the blessing." This idea of AI as an input into innovation is a key component of Cockburn, Henderson, and Stern (chapter 4), as well as in several later chapters. It is an important element of Aghion, Jones, and Jones' model of the impact of AI on economic growth (chapter 9), emphasizing endogenous growth through AI (self-)improvements. It also underlies the papers focused on how AI will impact that way economics research is conducted (chapters 21 through 24).

The section on AI as a general purpose technology concludes with Manuel Trajtenberg's discussion of political and societal consequences (chapter 6). At the conference, Trajtenberg discussed Joel Mokyr's paper "The Past and Future of Innovation: some lessons from economic history", which will be published elsewhere. The paper therefore sits between a stand-alone chapter and a discussion. Trajtenberg's chapter does not comment directly on Mokyr, but uses Mokyr's paper as a jumping off point to discuss how technology creates winners and losers, and the policy challenges associated with the political consequences of the diffusion of a GPT. “The sharp split between winners and losers, if left to its own, may have serious consequences far beyond the costs for the individuals involved: when it coincides with the political divide, it may threaten the very fabric of democracy, as we have seen recently both in America and in Europe. Thus, if AI bursts onto the scene and triggers mass displacement of workers, and demography plays out its fateful hand, the economy will be faced with a formidable dual challenge, that may require a serious reassessment of policy options… we need to anticipate the required institutional changes, to experiment in the design of new policies, particularly in education and skills development, in the professionalization of service occupations, and in affecting the direction of technical advance. Furthermore, economists possess a vast methodological arsenal that may prove very useful for that purpose – we should not shy away from stepping into this area, since its importance for the economy cannot be overstated.” The next set of chapters also emphasize the distributional challenges of economic growth driven by rapid technological change.

Growth, jobs, and inequality

Much of the popular discussion around AI focuses on the impact on jobs. If machines can do what humans do, then will there still be work for humans in the future? The chapters in this section dig into the consequences of AI for jobs, economic growth, and inequality. Almost all chapters emphasize that technological change means an increase in wealth for society. As Jason Furman puts it in chapter 12, "We need more artificial intelligence." At the same time, it is clear that the impact of AI on society will depend on how the increased income from AI is distributed. The most recent GPTs to diffuse, computers and the internet, likely led to increased inequality due to skill-bias (e.g. Autor, Katz, and Krueger 1998; Akerman, Gaarder, and Mogstad 2015) and to an increased capital share (e.g. Autor et al 2017). This section brings together those papers that emphasized (largely macroeconomic) ideas related to growth, inequality, and jobs. If the impact of AI will be like these other technologies, then what will the consequences look like for inequality, political economy, economic growth, jobs, and the meaning of work?

Stevenson (chapter 7) outlines many of the key issues. She emphasizes that economists generally agree that in the long run society will be wealthier. She highlights issues with respect to the short run and income distribution. Summarizing both the tension in the public debate and the key themes in several other chapters, she notes, "In the end, there's really two separate questions: there's an employment question, in which the fundamental question is can we find fulfilling ways to spend our time if robots take our jobs? And there's an income question, can we find a stable and fair distribution of income?"

Acemoglu and Restrepo (chapter 8) examine how AI and automation might change the nature of work. They suggest a task-based approach to understanding automation, emphasizing the relative roles of labor and capital in the economy. "At the heart of our framework is the idea that automation and thus AI and robotics replace workers in tasks that they previously performed, and via this channel, create a powerful *displacement effect*.”. This will lead to a lower labor share of economic output. At the same time, productivity will increase and capital will accumulate, thereby increasing the demand for labor. More importantly, “we argue that there is a more powerful countervailing force that increases the demand for labor as well as the share of labor in the national income: the *creation of new tasks*, functions and activities in which labor has a comparative advantage relative to machines. The creation of new tasks generates a *reinstatement effect* directly counterbalancing the *displacement effect*.” Like Stevenson, the long run message is optimistic; however, a key point is that adjustment costs may be high. New skills are a necessary condition of the long run optimistic forecast, and there is likely to be a short and medium term mismatch between skills and technologies. They conclude with a discussion of open questions about which skills are needed, the political economy of technological change (reinforcing ideas highlighted in the earlier chapter by Trajtenberg), and the interaction between inequality and the type of innovation enabled by automation going forward.

Aghion, Jones, and Jones (chapter 9) build on the task-based model, focusing on the impact on economic growth. They emphasize Baumol’s cost disease: “Baumol (1967) observed that sectors with rapid productivity growth, such as agriculture and even manufacturing today, often see their share of GDP decline while those sectors with relatively slow productivity growth – perhaps including many services – experience increases. As a consequence, economic growth may be constrained not by what we do well but rather by what is essential and yet hard to improve. We suggest that combining this feature of growth with automation can yield a rich description of the growth process, including consequences for future growth and income distribution.” Thus, even in the limit where there is an artificial general intelligence that creates a singularity or intelligence explosion with a self-improving AI, cost disease forces may constrain growth. This link between technological advance and Baumol’s cost disease provides a fundamental limit to the most optimistic and the most pessimistic views. Scarcity limits both growth and the downside risk. The paper also explores how AI might reduce economic growth if it makes it easier to imitate rival's innovations, returning to issues of intellectual property highlighted in Mitchell's comment. Finally, they discuss inequality within and across firms. They note that AI will increase wages of the least skilled employees of technologically advanced firms but also increasingly outsource the tasks undertaken by such employees.

Francois's comment takes this emphasis on cost disease as a starting point, asking what those tasks will be that humans are left to do. "But it is when we turn to thinking about what are the products or services where humans will remain essential in production that we start to run into problems. What if humans can't do anything better than machines? Many discussions at the conference centered around this very possibility. And I must admit that I found the scientists' views compelling on this… The point I wish to make is that even in such a world where machines are better at all tasks, there will still be an important role for human 'work'. And that work will become the almost political task of managing the machines." He argues that humans must tell the machines what to optimize. Bostrom (2014) describes this as the value-loading problem. Francois emphasizes that this is largely a political problem, and links the challenges in identifying values with Arrow's (1951) impossibility theorem. He identifies key questions around ownership of the machines, length of time that rents should accrue to those owners, and the political structure of decision-making. In raising these questions, he provides a different perspective on issues highlighted by Stevenson on the meaning of work and Trajtenberg on the political economy of technological change.

The discussion of the meaning of work is a direct consequence of concerns about the impact of AI on jobs. Jobs have been the key focus of public discussion on AI and the economy. If human tasks get automated, what is left for humans to do? Bessen (chapter 10) explores this question, using data about other technological advances to support his arguments. He emphasizes that technological change can lead to an increase in demand and so the impact of automation on jobs is ambiguous, even within a sector. “The reason automation in textiles, steel, and automotive manufacturing led to strong job growth has to do with the effect of technology on demand… New technologies do not just replace labor with machines, but in a competitive market, automation will reduce prices. In addition, technology may improve product quality, customization, or speed of delivery. All of these things can increase demand. If demand increases sufficiently, employment will grow even though the labor required per unit of output declines.”

Like Bessen, Goolsbee (chapter 11) notes that much of the popular discussion around AI relates to the labor market consequences. Recognizing that those consequences matter, his chapter mostly emphasizes the positive: Growth and productivity are good. AI has potential to increase our standard of living. Like Acemoglu and Restrepo, he notes that the short term displacement effects could be substantial. One frequently-cited solution to the displacement effects of AI is a universal basic income, in which all members of society receive a cash transfer from the government. He then discusses the economics of such a policy and the numerous challenges to making it work. “First… in a world where AI induced unemployment is already high, separating work and income is an advantage. In a world like the one we are in now, offering a basic income will likely cause a sizable drop in the labor market participation by low wage groups… Second, for a given amount of money to be used on redistribution, UBI likely shifts money away from the very poor… Third, …converting things to a UBI and getting rid of the in-kind safety net will lead to a situation in which, even if among a small share of UBI recipients, SOME people will blow their money in unsympathetic ways – gambling, drugs, junk food, Ponzi schemes, whatever. And now those people will come to the emergency room or their kids will be hungry and by the rules, they will be out of luck. That’s what they were supposed to have used their UBI for.” Before concluding, he touches on a variety of regulatory issues that receive more detailed discussion in chapters 16 through 20. His conclusion mirrors that of Francois, emphasizing the importance of humans in determining policy direction, even if AI improves to the point where it surpasses human intelligence.

Furman (chapter 12) is similarly optimistic, emphasizing that we need more, not less AI. "AI is a critical area of innovation in the U.S. economy right now. At least to date, AI has not had a large impact on the aggregate performance of the macroeconomy or the labor market. But it will likely become more important in the years to come, bringing substantial opportunities – and our first impulse should be to embrace it fully." Referencing data on productivity growth and on the diffusion of industrial robots, he then discusses potential negative effects on the economy as AI diffuses, particularly with respect to inequality and reduced labor force participation. The issues around labor force participation highlight the importance of Stevenson's questions on the meaning of work. Like Goolsbee, Furman notes several challenges to implementing a universal basic income as a solution to these negative effects. He concludes that policy has an important role to play in enabling society to fully reap the benefits of technological change while minimizing the disruptive effects.

Returning to the question of labor share highlighted by Acemoglu and Restrepo, Sachs (chapter 13) emphasizes that the income share going to capital grows with automation: "Rather than Solow-era stylized facts, I would therefore propose the following alternative stylized facts: 1) the share of national income accruing to capital rises over time in sectors experiencing automation, especially when capital is measured to include human capital; 2) the share of national income accruing to low-skilled labor drops while the share accruing to high-skilled labor rises; 3) the dynamics across sectors vary according to the differential timing of automation, with automation spreading from low-skilled and predictable tasks towards high-skilled and less-predictable tasks; 4) automation reflects the rising intensity of science and technology throughout the economy…, and 5) future technological changes associated with AI are likely to shift national income from medium-skilled and high-skilled towards owners of business capital…" The chapter concludes with a list of key open questions about the dynamics of automation, the role of monopoly rents, and the consequences for income distribution and labor force participation.

Stiglitz and Korinek (chapter 14) also emphasize income distribution, discussing the implications of AI-related innovation for inequality. They show that, in a first-best economy, contracts can be specified in advance that make innovation Pareto-improving. However, imperfect markets and costly redistribution can imply move away from the first-best. Innovation may then drive inequality directly, by giving innovators a surplus, or indirectly by changing the demand for different types of labor and capital. They discuss policies that could help reduce the increase in inequality, emphasizing different taxation tools. Related to the ideas introduced in Mitchell's comment, they also explore IP policies: "If outright redistribution is infeasible, there may be other institutional changes which result in market distributions that are more favorable to workers. For example, intervention to steer technological progress may act as a 2nd best device… we provide an example in which a change in intellectual property rights – a shortening of the term of patent protection – effectively redistributes some of the innovators' surplus to workers (consumers) to mitigate the pecuniary externalities on wages that they experience, with the ultimate goal that the benefits of the innovation are more widely shared." Stiglitz and Korinek conclude with a more speculative discussion of artificial general intelligence ("super-human artificial intelligence"), emphasizing that such a technological development will likely further increase inequality.

The final chapter in the section on growth, jobs, and inequality calls for a different emphasis. Cowen (chapter 15) emphasizes consumer surplus, international effects, and political economy. With respect to consumer surplus, he writes, "Imagine education and manufactured goods being much cheaper because we produced them using a greater dose of smart software. The upshot is that even if a robot puts you out of a job or lowers your pay, there will be some recompense on the consumer side." Cowen also speculates that AI might hurt developing countries much more than developed, as automation means that labor cost reasons to offshore decline. Finally, like Trajtenberg and Francois, he emphasizes the political economy of AI, highlighting questions related to income distribution.

Taken together, the chapters in this section highlight several key issues, and provide models that identify challenges related to growth, jobs, inequality, and politics. These models set up a number of theoretical and empirical questions about how AI will impact economic outcomes within and across countries.

The discussions are necessarily speculative because AI has not yet diffused widely, so research must either be entirely theoretical or it must use related technologies (such as factory robots) as a proxy for AI. The discussions are also speculative because of the challenges in measuring the relevant variables. In order to determine the impact of AI on the economy, we need consistent measures of AI, productivity, intangible capital, and growth across sectors, regions, and contexts. Going forward, to the extent that progress occurs against the proposed research agenda, it will depend on advances in measurement.

Machine Learning and Regulation

Industry will be a key innovator and adopter of artificial intelligence. A number of regulatory issues arise. The regulatory issues related to truly intelligent machines are touched on by Trajtenberg, Francois, Goolsbee, and Cowen. Mitchell's comment of Cockburn, Henderson, and Stern emphasizes intellectual property regulation. This section focuses on other regulatory challenges with respect to advances in machine learning.

Varian (chapter 16) sets up the issues by describing the key models from industrial organization that are relevant to understanding the impact of machine learning on firms. He highlights the importance of data as a scarce resource, and discusses the economics of data as an input: it is non-rival and it exhibits decreasing returns to scale in a technical sense (because prediction accuracy increases in the square root of N). He discusses the structure of ML-using industries, including vertical integration, economies of scale, and the potential for price discrimination. He emphasizes the difference between learning-by-doing and data network effects: “There is a concept that is circulating among lawyers and regulators called ‘data network effects’. The model is that a firm with more customers can collect more data and use this data to improve its product. This is often true – the prospect of improving operations is what makes ML attractive – but it is hardly novel. And it is certainly not a network effect! This is essentially a supply-side effect known as “learning by doing”… A company can have huge amounts of data, but if it does nothing with the data, it produces no value. In my experience, the problem is not lack of resources, but is lack of skills. A company that has data but no one to analyze it is in a poor position to take advantage of that data.” He concludes by highlighting policy questions related to algorithmic collusion (which was discussed at the conference as “economist catnip,” interesting and fun but unlikely to be of first-order importance), security, privacy, and transparency.

Chevalier’s comment builds on Varian’s emphasis on the importance of data, exploring the potential of antitrust policy aimed at companies that use machine learning. Legal scholars and policymakers have asked whether antitrust essential facilities doctrine should be applied to data ownership. She emphasizes the trade-off between static and dynamic considerations for such a policy: “In evaluating antitrust policies in innovative industries, it is important to recognize that consumer benefits from new technologies arise not just from obtaining goods and services at competitive prices, but also from the flow of new and improved products and services that arise from innovation. Thus, antitrust policy should be evaluated not just in terms of its effect on prices and outputs, but also on its effect on the speed of innovation. Indeed, in the high technology industries, it seems likely that these dynamic efficiency considerations dwarf the static efficiency considerations.” She also explores several practical challenges.

Another regulatory issue that arises from the importance of data is privacy. Tucker (chapter 17) notes that machine learning uses data to make predictions about what individuals may desire, be influenced by, or do. She emphasizes that privacy is challenging for three reasons: cheap storage means that data may persist longer than the person who generated the data intended, non-rivalry means that data may be repurposed for uses other than originally intended, and externalities caused by data created by one individual that contains information about others. “For example, in the case of genetics, the decision to create genetic data has immediate consequences for family members, since one individual’s genetic data is significantly similar to the genetic data of their family members… There may also be spillovers across a person’s decision to keep some information secret, if such secrecy predicts other aspects of that individual’s behavior that AI might be able to project from.” She discusses potential negative impacts of these three challenges, concluding with some key open questions.

Jin (chapter 18) also focuses on the importance of data as an input into machine learning. She emphasizes that reduced privacy creates security challenges, such as identify theft, ransomware, and misleading algorithms (such as Russian-sponsored posts in the 2016 U.S. election). “In my opinion, the leading concern is that firms are not fully accountable for the risk they bring to consumer privacy and data security. To restore full accountability, one needs to overcome three obstacles, namely (1) the difficulty to observe firms’ actual action in data collection, data storage, and data use; (2) the difficulty to quantify the consequence of data practice, especially before low-probability adverse events realize themselves; and (3) the difficulty to draw a causal link between a firm’s data practice and its consequence.” Combined, Tucker and Jin’s chapters emphasize that any discussion of growth and impact of AI requires an understanding of the privacy framework. Access to data drives innovation, underlies the potential for economic growth, and frames the antitrust debate.

The economics of data also create challenges with respect to the rules governing international trade. Goldfarb and Trefler (chapter 19) argue that economies of scale in data through feedback loops, along with economies of scope and knowledge externalities in AI innovation, could create the opportunity for country-level rents and strategic trade policy. At the same time, they emphasize that the geographic constraints on data and knowledge would have to be high for such a policy to be optimal at the country level. They highlight the rise of China: "China has become the focal point for much of the international discussion. The U.S. narrative has it that Chinese protection has reduced the ability of dynamic U.S. firms such as Google and Amazon to penetrate Chinese markets. This protection has allowed China to develop significant commercial AI capabilities, as evidenced by companies such as Baidu (a search engine like Google), Alibaba (an e-commerce web portal like Amazon), and Tencent (the developer of WeChat, which can be seen as combining the functions of Skype, Facebook, and Apple Pay)… we collected time-series data on the institutional affiliation of all authors of papers presented at a major AI research conference… we compare the 2012 and 2017 conferences… While these countries all increased their absolute number of participants, in relative terms they all lost ground to China, which leapt from 10 percent in 2012 to 24 percent in 2017." The authors discuss the international dimensions of domestic regulation related to privacy, access to government data, and industrial standards.

The final regulatory issue highlighted in this section is tort liability. Galasso and Luo (chapter 20) review prior literature on the relationship between liability and innovation. They emphasize the importance of getting the balance right between consumer protection and innovation incentives. "A central question in designing a liability system for AI technologies is how liability risk should be allocated between producers and consumers, and how this allocation might affect innovation… A key promise of AI technologies is to achieve autonomy. With less room for consumers to take precautions, the relative liability burden is likely to shift towards producers, especially in situations in which producers are in a better position than individual users to control risk… On the other hand, during the transitional period of an AI technology, substantial human supervision may still be required… In many of these situations, it may be impractical or too costly for producers to monitor individual users and to intervene. Therefore, it would be important to maintain consumer liability to the extent that users of AI technologies have sufficient incentives to take precautions and invest in training, thus internalizing potential harm to others."

Broadly, regulation will affect the speed at which AI diffuses. Too much regulation, and industry will not have incentives to invest. Too little regulation, and consumers will not trust the products that result. In this way, getting the regulatory balance right is key to understanding when and how any impact of AI on economic growth and inequality will arise.

Impact on the Practice of Economics

Cockburn, Henderson, and Stern emphasize that machine learning is a general purpose technology for science and innovation. As such, it is likely to have an impact on research in a variety of disciplines, including economics. Athey (chapter 21) provides an overview of the various ways in which machine learning is likely to affect the practice of economics. For example: "I believe that machine learning (ML) will have a dramatic impact on the field of economics within a short time frame… ML does not add much to questions about identification, which concern when the object of interest, e.g., a causal effect, can be estimated with infinite data, but rather yields great improvements when the goal is semi-parametric estimation or when there are a large number of covariates relative to the number of observations… a key advantage of ML is that ML views empirical analysis as "algorithms" that estimate and compare many alternative models… 'outsourcing' model selection to algorithms works very well when the problem is 'simple' – for example, prediction and classification tasks, where performance of a model can be evaluated by looking at goodness of fit in a held-out test set." She emphasizes the usefulness of machine learning techniques for policy problems related to prediction (as in Kleinberg et al 2015). The paper then details recent advances in using machine learning techniques in causal inference, which she views as a fundamental new toolkit for empirical economists. She concludes with a list of sixteen predictions of how machine learning will impact economics, emphasizing new econometric tools, new data sets and measurement techniques, increased engagement of economists as engineers (and plumbers), and of course increased study of the economic impact of machine learning on the economy as a whole.

Lederman's comment emphasizes the usefulness of machine learning to create new variables for economic analysis, and how the use of machine learning by organizations creates a new kind of endogeneity problem. "We develop theoretical models to help us understand the data-generation process which, in turn, informs both our concerns about causality as well as the identification strategies we develop… Overall, as applied researchers working with real-world datasets, we need to recognize that increasingly the data we are analyzing is going to be the result of decisions that are made by algorithms in which the decision-making process may or may not resemble the decision-making processes we model as social scientists."

If the study of AI is going to be a key question for economists going forward, Raj and Seamans (chapter 22) emphasize that we need better data. "While there is generally a paucity of data examining the adoption, use, and effects of both AI and robotics, there is currently less information available regarding AI. There are no public datasets on the utilization or adoption of AI at either the macro or micro level. The most complete source of information, the McKinsey Global Institute study, is proprietary and inaccessible to the general public or the academic community. The most comprehensive and widely-used dataset examining the diffusion of robotics is the International Federation of Robotics (IFR) Robot Shipment Data… the IFR does not collect any information on dedicated industrial robots that serve one purpose. Furthermore, some of the robots are not classified by industry, detailed data is only available for industrial robots (and not robots in service, transportation, warehousing, or other sectors), and geographical information is often aggregated…" They provide a detailed discussion of data collection opportunities by government and by academic researchers. If the agenda set up in the other chapters is to be answered, it is important to have a reliable data set that defines AI, measures its quality, and tracks its diffusion.

Related to Athey's emphasis of increased engagement of economists as engineering, Milgrom and Tadelis (chapter 23) describe how machine learning is already affecting market design decisions. Using specific examples from online marketplaces and telecommunications auctions, they emphasize the potential of AI to improve efficiency by predicting demand and supply, overcoming computational barriers, and reducing search frictions. "AI and Machine Learning are emerging as important tools for market design. Retailers and marketplaces such as eBay, TaoBao, Amazon, Uber, and many others are mining their vast amounts of data to identify patterns that help them create better experiences for their customers and increase the efficiency of their markets… two-sided markets such as Google, which match advertisers with consumers, are not only using AI to set reserve prices and segment consumers into finer categories for ad targeting, but they also develop AI-based tools to help advertisers bid on ads… Another important application of AI's strength in improving forecasting to help markets operate more efficiently is in electricity markets. To operate efficiently, electricity market makers… must engage in demand and supply forecasting." The authors argue that AI will play a substantial role in the design and implementation of markets over a wide range of applications.

Camerer (chapter 24) also emphasizes the role of AI as a tool for predicting choice. "Behavioral economics can be defined as the study of natural limits on computation, willpower, and self-interest, and the implications of those limits for economic analysis (market equilibrium, IO, public finance, etc.). A different approach is to define behavioral economics more generally, as simply being open-minded about what variables are likely to influence economic choices… In a general ML approach, predictive features could be - and *should* be – any variables that predict… If behavioral economics is recast as open-mindedness about what variables might predict, then ML is an ideal way to do behavioral economics because it can make use of a wide set of variables and select which ones predict." He argues that firms, policymakers, and market designers can implement AI as either a "bionic patch" that improves human decision-making or "malware" that exploits human weaknesses. In this way, AI could reduce or exacerbate the political economy and inequality issues highlighted in earlier chapters. In addition, Camerer explores two other ways in which AI and behavioral economics will interact. He hypothesizes that machine learning could help predict human behavior in a variety of settings including bargaining, risky choice, and games, helping to verify or reject theory. He also emphasizes that (poor) implementation of AI might provide insight into new ways to model biases in human decision-making.

The book concludes with Kahneman's brief and insightful comment. Kahneman begins with a discussion of Camerer's idea of using prediction to verify theory, but continues with a broader discussion of a variety of themes that arose over the course of the conference. With an optimistic tone, he emphasizes that there are no obvious limits to what artificial intelligence may be able to do. "Wisdom is breadth. Wisdom is not having too narrow a view. That is the essence of wisdom; it's broad framing. A robot will be endowed with broad framing. When it has learned enough, it will be wiser than we people because we do not have broad framing. We are narrow thinkers, we are noisy thinkers, and it is very easy to improve upon us. I do not think that there is very much that we can do that computers will not eventually be programmed to do."

The future of research on the economics of artificial intelligence

The chapters in this book are the beginning. They highlight key questions, recognize the usefulness of several economic models, and identify areas for further development. We can leverage what we know about GPTs to anticipate the impact of AI as it diffuses, recognizing that no two GPTs are identical. If AI is a general purpose technology, it is likely to lead to increased economic growth. A common theme in these chapters is that slowing down scientific progress—even if it were possible—would come at a significant cost. At the same time, many attendees emphasized that the distribution of the benefits of AI might not be even. It depends on who owns the AI, the effect on jobs, and the speed of diffusion.

The task given to the conference presenters was to scope out the research agenda. Perhaps more than anything, this volume highlights all that we don't know. It emphasizes questions around growth, inequality, privacy, trade, innovation, political economy, etc. We don't have answers yet. Of course, the lack of answers is a consequence of the early stage of AI's diffusion. We cannot measure the impact until AI is widespread.

With the current state of measurement, however, we may never get answers. As highlighted in the chapter by Raj and Seamans, we do not have good measures of AI. We also do not have a good measure of improvement to AI. What is the AI equivalent to the computational speed of a microchip or the horsepower of an internal combustion engine that will allow for quality-adjusted prices and measurement? We also do not have good measures of productivity growth when that growth is primarily driven by intangible capital. To answer these questions, the GDP measurement apparatus needs to focus on adjusting for intangible capital, software, and changes to the innovation process (Haskel and Westlake 2017). Furthermore, to the extent that the benefits of AI generate heterogeneous benefits to people as consumers and as workers, measurement of the benefit of AI will be tricky. For example, if AI enables more leisure and people choose to take more leisure, should that be accounted for in measures of inequality? If so, how?

While each chapter has its own take on the agenda, several themes cut across the volume as key aspects of the research agenda going forward. To the extent there is consensus on the questions, the consensus focuses on the potential of AI as a GPT, and the associated potential consequences on growth and inequality. A second consistent theme is the role of regulation in accelerating or constraining the diffusion of the technology. A third theme is that AI will change the way we do our work as economists. Finally, a number of issues appear in many chapters that are somewhat outside the standard economic models of technology's impact. How do people find meaning if AI replaces work with leisure? How can economists inform the policy debate on solutions proposed by technologists in the popular press such as taxing robots or a universal basic income? How does a technology's diffusion affect the political environment, and vice versa?

This book highlights the questions and provides direction. We hope readers of this book take it as a starting point for their own research into this new and exciting area of study.

References

Agrawal, Ajay, Joshua Gans, and Avi Goldfarb. 2018. *Prediction Machines: The Simple Economics of Artificial Intelligence*. Harvard Business Review Press, Boston MA.

Akerman, Anders, Ingvil Gaarder, and Magne Mogstad. 2015 The Skill Complementarity of Broadband Internet. *Quarterly Journal of Economics* 130(4), 1781–1824.

Arrow, Kenneth. 1951. 2nd edition 1963. Social Choice and Individual Values. John Wiley and Sons, New York.

Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen. 2017. The Fall of the Labor Share and the Rise of Superstar Firms. Working paper, MIT.

Autor, David H., Lawrence F. Katz, and Alan B. Krueger. 1998. Computing Inequality: Have Computers Changed the Labor Market? *Quarterly Journal of Economics*, 113(4), 1169–1213.

Bostrom, Nick. 2014. *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press, Oxford UK.

Bresnahan, Timothy F., and M. Trajtenberg. 1995. General purpose technologies 'Engines of growth'? *Journal of Econometrics* 65, 83-108.

David, Paul A. 1990. The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox. *American Economic Review Papers and Proceedings* 80(2), 355-361.

Haskel, Jonathan, and Stian Westlake. 2017. *Capitalism without Capital: The Rise of the Intangible Economy*. Princeton University Press, Princeton NJ.

Kaplan, Jerry. 2016. *Artificial Intelligence: What Everyone Needs to Know*. Oxford University Press, Oxford UK.

Kleinberg, Jon, Jens Ludwig, Sendhil Mullainathan, and Ziad Obermeyer. 2015. "Prediction Policy Problems." *American Economic Review*, 105(5): 491-95.