



AI SUPER- POWERS

Utopia, Dystopia, and the Real AI Crisis

A White Paper Adapted from

AI Superpowers

China, Silicon Valley, and the New World Order

By Kai-Fu Lee

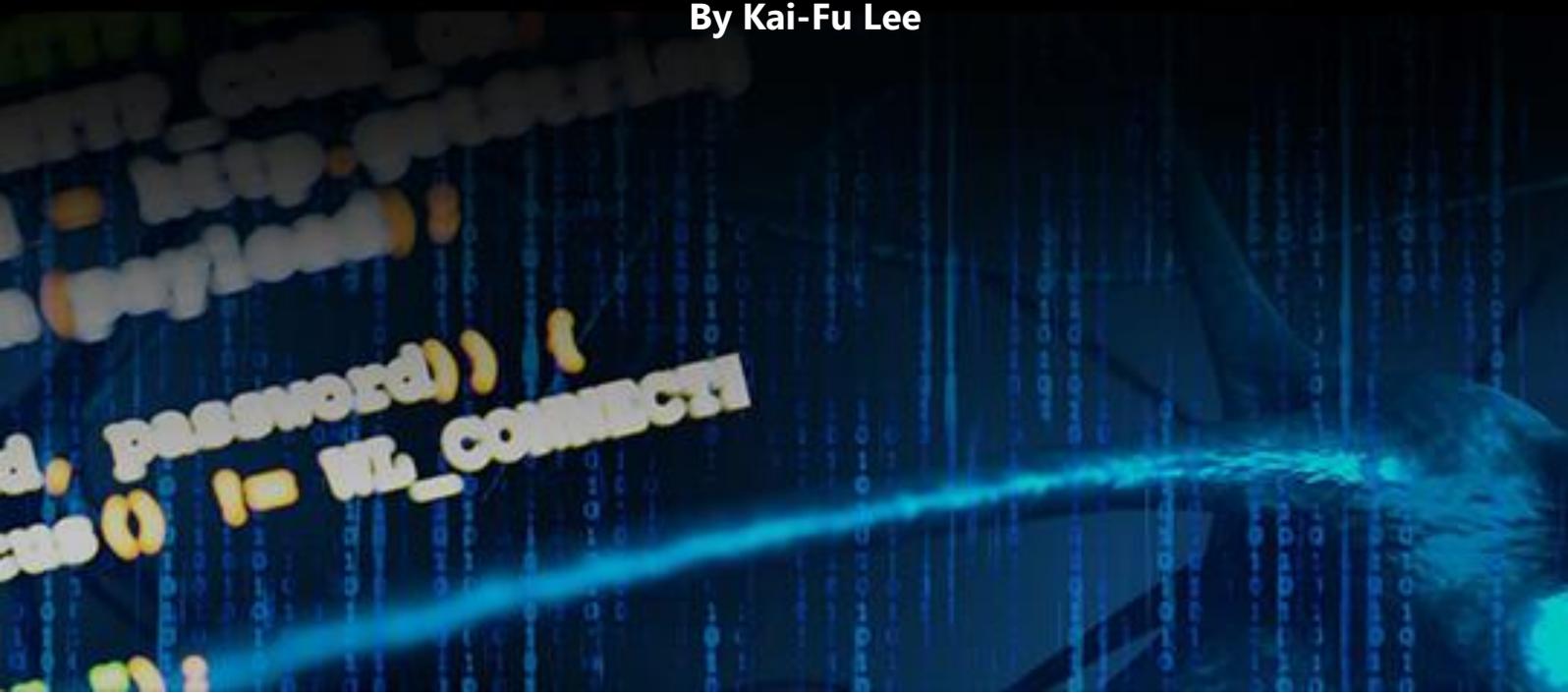


Table of Contents

Introduction	3
Utopia, Dystopia, and the Real AI Crisis	4
Reality Check	7
Folding Beijing: Science-Fiction Visions and AI Economics	
11	
A Real AI Crisis	
21	
Google versus the Rest	24
Power Grids versus AI Batteries	27
The Chip on China’s Shoulder	29
A Tale of Two AI Plans	
31	
Betting on AI	33
Self-Driving Dilemmas	
36	
About the Author	41
About the Book	43
What People are Saying About the Book	
46	

Introduction

AI will bring with it great economies of scale, new efficiencies and a potentially more powerful way of doing business. At the same time, AI can eventually lead to a loss of meaning and purpose as people's jobs disappear – and those suddenly without work find themselves unable to find employment. This leads to depression, substance abuse and general malaise.

AI-induced unemployment will lead to psychological damage. People will face the prospect of not just being temporarily out of work but of being permanently excluded from the functioning of the economy. They will watch as algorithms and robots easily outperform them at tasks and skills they spent their whole lives mastering. It will lead to a crushing feeling of futility, a sense of having become obsolete in one's own skin.

The winners of this AI economy will marvel at the awesome power of these machines. But the rest of humankind will be left to grapple with a far deeper question: when machines can do everything that we can, what does it mean to be human?

That's a question that I found myself grappling with in the depths of my own personal crisis of mortality and meaning. That is something I explore here – alongside the wealth of opportunities AI brings come a score of potential problems for people.

This white paper is based on my book, *AI Superpowers: China, Silicon Valley, and the New World Order*, which explores how AI drives change – in China, the U.S. and the world – but also how it can lead to suffering as well as success. It's important to note that amid the many reasons to believe in the power of AI, we need to be aware of the pitfalls so we can navigate the course ahead. We want to know which people and countries will benefit from this tremendous technology. We wonder whether AI can vault us to lives of material abundance, and whether there is space for humanity in a world run by intelligent machines.

No one has a crystal ball that can reveal the answers to these questions for us. But that core uncertainty makes it all the more important that we ask these questions and, to the best of our abilities, explore the answers. I attempt to do that in my book – and in this adaptation for you.

Utopia, Dystopia, and the Real AI Crisis

All of the AI products and services that I outlined in my previous white paper, on the four waves of AI, are within reach based on current technologies. Bringing them to market requires no major new breakthroughs in AI research, just the nuts-and-bolts work of everyday implementation: gathering data, tweaking formulas, iterating algorithms in experiments and different combinations, prototyping products, and experimenting with business models.

But the age of implementation has done more than make these practical products possible. It has also set ablaze the popular imagination when it comes to AI. It has fed a belief that we're on the verge of achieving what some consider the Holy Grail of AI research, artificial general intelligence (AGI)—thinking machines with the ability to perform any intellectual task that a human can—and much more.

Some predict that with the dawn of AGI, machines that can improve themselves will trigger runaway growth in computer intelligence. Often called “the singularity,” or artificial superintelligence, this future involves computers whose ability to understand and manipulate the world dwarfs our own, comparable to the intelligence gap between human beings and, say, insects. Such dizzying predictions have divided much of the intellectual community into two camps: utopians and dystopians.

The utopians see the dawn of AGI and subsequent singularity as the final frontier in human flourishing, an opportunity to expand our own consciousness and conquer mortality. Ray Kurzweil—the eccentric inventor, futurist, and guru-in-residence at Google—envisions a radical future in which humans and machines have fully merged. We will upload our minds to the cloud, he predicts, and constantly renew our bodies through intelligent nanobots released into our bloodstream. Kurzweil predicts that we are just one decade away from AGI, and that we will reach the singularity by 2045.

Other utopian thinkers see AGI as something that will enable us to rapidly decode the mysteries of the physical universe. DeepMind founder Demis Hassabis predicts that the creation of superintelligence will allow human civilization to solve intractable problems, producing inconceivably brilliant solutions to global warming and previously incurable diseases. With super intelligent computers that understand the universe on levels that humans cannot even conceive of, these machines become not just tools for lightening the burdens of humanity; they approach the omniscience and omnipotence of a God.

Not everyone, however, is so optimistic. Elon Musk has called superintelligence “the biggest risk we face as a civilization,” comparing the creation of it to “summoning the demon.” Intellectual celebrities such as cosmologist Stephen Hawking have joined Musk in the dystopian camp, many of them inspired by the work of Oxford philosopher Nick Bostrom, whose 2014 book *Superintelligence* captured the imagination of many futurists.

For the most part, members of the dystopian camp aren't worried about the AI takeover as imagined in films like the Terminator series, with human-like robots "turning evil" and hunting down people in a power-hungry conquest of humanity. Superintelligence would be the product of human creation, not natural evolution, and thus wouldn't have the same instincts for survival, reproduction, or domination that motivate humans or animals. Instead, it would likely just seek to achieve the goals given to it in the most efficient way possible.

The fear is that if human beings presented an obstacle to achieving one of those goals— reverse global warming, for example—a super intelligent agent could easily, even accidentally, wipe us off the face of the earth.

The fear is that if human beings presented an obstacle to achieving one of those goals— reverse global warming, for example—a super intelligent agent could easily, even accidentally, wipe us off the face of the earth. For a computer program whose intellectual imagination so dwarfed our own, this wouldn't require anything as crude as gun-toting robots. Superintelligence's profound understanding of chemistry, physics, and nanotechnology would allow for far more ingenious ways to instantly accomplish its goals. Researchers refer to this as the "control problem" or "value alignment problem," and it's something that worries even AGI optimists.

Although timelines for these capabilities vary widely, Bostrom's book presents surveys of AI researchers, giving a median prediction of 2040 for the creation of AGI, with superintelligence likely to follow within three decades of that. But read on.

Reality Check

Utopian and dystopian visions of the super intelligent future inspire both awe and a sense of dread in audiences. Those all-consuming emotions then blur the lines in our mind separating these fantastical futures from our current age of AI implementation. The result is widespread popular confusion over where we truly stand today, and where things are headed.

To be clear, none of the scenarios described above—the immortal digital minds or omnipotent superintelligences—are possible based on present-day technologies. Today, there remain no known algorithms for AGI or a clear engineering route to get there. The singularity is not something that can occur spontaneously, with autonomous vehicles running on deep learning suddenly “waking up” and realizing that they can band together to form a super intelligent network.

There remain no known algorithms for artificial genuine intelligence or a clear engineering route to get there. The singularity is not something that can occur spontaneously.

Getting to AGI would require a series of foundational scientific breakthroughs in artificial intelligence, a string of advances on the scale of, or greater, than deep learning. These breakthroughs would need to remove key constraints on the “narrow AI” programs that we run today and empower them with a wide array of new abilities: multidomain learning; domain-independent learning; natural-language understanding; commonsense reasoning, planning, and learning from a small number of examples. Taking the next step to emotionally intelligent robots may require self-awareness, humor, love, empathy, and appreciation for beauty. These are the key hurdles that separate what AI does today—spotting correlations in data and

making predictions—and artificial general intelligence. Any one of these new abilities may require multiple major breakthroughs; AGI implies solving all of them.

The mistake of many artificial general intelligence forecasts is to simply take the rapid rate of advance from the past decade and extrapolate it outward or launch it exponentially upward in an unstoppable snowballing of computer intelligence. Deep learning represents a major leveling up in machine learning, a movement onto a new plateau with a variety of real-world uses: the age of implementation. But there is no proof that this level change represents the beginning of exponential growth that will inevitably race toward AGI and then superintelligence at an ever-increasing pace.

Science is difficult, and fundamental scientific breakthroughs are even harder. Discoveries like deep learning that truly raise the bar for machine intelligence are rare and often separated by many decades, if not longer. Implementations and improvements on these breakthroughs abound, and researchers at places like DeepMind have demonstrated powerful new approaches to things like reinforcement learning. But in the twelve years since Geoffrey Hinton and his colleagues' landmark paper on deep learning, I haven't seen anything that represents a similar sea change in machine intelligence. Yes, the AI scientists surveyed by Bostrom predicted a median date of 2040 for AGI, but I believe scientists tend to overestimate when an academic demonstration will become a real-world product. To wit, in the late 1980s, I was the world's leading researcher on AI speech recognition, and I joined Apple because I believed the technology would go mainstream within five years. It turned out I was off by twenty years.

I cannot guarantee that scientists definitely will not make the breakthroughs that would bring about artificial general intelligence and then superintelligence. In fact, I believe we should expect continual improvements to the existing state of the art. But I believe we are still many decades, if not centuries, away from the real thing. There is also a real possibility that AGI is something humans will never achieve. Artificial general intelligence would be a major turning point in the relationship between

humans and machines—what many predict would be the most significant single event in the history of the human race. It's a milestone that I believe we should not cross unless we have first definitively solved all problems of control and safety. But given the relatively slow rate of progress on fundamental scientific breakthroughs, I and other AI experts like Andrew Ng and Rodney Brooks believe AGI remains farther away than often imagined.

Does that mean I see nothing but steady material progress and glorious human flourishing in our AI future? Not at all. Instead, I believe that human civilization will soon face a different kind of AI-induced crisis. This crisis will lack the apocalyptic drama of a Hollywood blockbuster, but it will disrupt our economic and political systems all the same, and even cut to the core of what it means to be human in the twenty-first century.

In short, this is the coming crisis of jobs and inequality. Our present AI capabilities can't create a superintelligence that destroys our civilization. But my fear is that we humans may prove more than up to that task ourselves.

Folding Beijing: Science-Fiction Visions and AI Economics

When the clock strikes 6 a.m., the city devours itself. Densely packed buildings of concrete and steel bend at the hip and twist at their spines. External balconies and awnings are turned inward, creating smooth and perfectly sealed exteriors. Skyscrapers break down into component parts, shuffling and consolidating into Rubik's Cubes of industrial proportions. Inside those blocks are the residents of Beijing's Third Space, the economic underclass that toils during the night hours and sleeps during the day. As the cityscape folds in on itself, a patchwork of squares on the earth's surface begin their 180-degree rotation, flipping over to tuck these consolidated structures underground.

When the other side of these squares turn skyward, they reveal a separate city. The first rays of dawn creep over the horizon as this new city emerges from its crouch.

Tree-lined streets, vast public parks, and beautiful single-family homes begin to unfold, spreading outward until they have covered the surface entirely. The residents of First Space stir from their slumber, stretching their limbs and looking out on a world all their own.

These are visions of Hao Jingfang, a Chinese science-fiction writer and economics researcher. Hao's novelette *Folding Beijing* won the prestigious Hugo Award for its arresting depiction of a city in which economic classes are separated into different worlds.

In a futuristic Beijing, the city is divided into three economic castes that split time on the city's surface. Five million residents of the elite First Space enjoy a twenty-four-hour cycle beginning at 6 a.m., a full day and night in a clean, hyper-modern, uncluttered city. When First Space folds up and flips over, the 20 million residents of Second Space get sixteen hours to work across a somewhat less glamorous cityscape. Finally, the denizens of Third Space—50 million sanitation workers, food vendors, and menial laborers—emerge for an eight-hour shift from 10 p.m. to 6 a.m., toiling in the dark among the skyscrapers and trash pits.

The trash-sorting jobs that are a pillar of the Third Space could be entirely automated but are instead done manually to provide employment for the unfortunate denizens condemned to a life there. Travel between the different spaces is forbidden, creating a society in which the privileged residents of First Space can live free of worry that the unwashed masses will contaminate their techno-utopia.

The Real AI Crisis

This dystopian story is a work of science fiction but one rooted in real fears about economic stratification and unemployment in our automated future. Hao holds a Ph.D. in economics and management from prestigious Tsinghua University. For her day job, she conducts economics research at a think tank reporting to the Chinese central government, including investigating the impact of AI on jobs in China.

As the four waves of AI spread across the global economy, they have the potential to wrench open ever greater economic divides between the haves and have-nots, leading to widespread technological unemployment.

It's a subject that deeply worries many economists, technologists, and futurists, myself included. I believe that as the four waves of AI spread across the global economy, they have the potential to wrench open ever greater economic divides between the haves and have-nots, leading to widespread technological unemployment. As Hao's story so vividly illustrates, these chasms in wealth and class can morph into something much deeper: economic divisions that tear at the fabric of our society and challenge our sense of human dignity and purpose.

Massive productivity gains will come from the automation of profit-generating tasks, but they will also eliminate jobs for huge numbers of workers. These layoffs won't discriminate by the color of one's collar, hitting highly educated white-collar workers just as hard as many manual laborers. A college degree—even a highly specialized professional degree—is no guarantee of job security when competing against machines that can spot patterns and make decisions on levels the human brain simply can't fathom.

Beyond direct job losses, artificial intelligence will exacerbate global economic inequality. By giving robots the power of sight and the ability to move autonomously, AI will revolutionize manufacturing, putting third-world sweatshops stocked with armies of low-wage workers out of business. In doing so, it will cut away the bottom rungs on the ladder of economic development. It will deprive poor countries of the opportunity to kick-start economic growth through low-cost exports,

the one proven route that has lifted countries like Korea, China, and Singapore out of poverty. The large populations of young workers that once comprised the greatest advantage of poor countries will turn into a net liability, and a potentially destabilizing one. With no way to begin the development process, poor countries will stagnate while the AI superpowers take off.

But even within those rich and technologically advanced countries, AI will further cleave open the divide between the haves and the have-nots. The positive-feedback loop generated by increasing amounts of data means that AI-driven industries naturally tend toward monopoly, simultaneously driving down prices while also eliminating competition among firms. While small businesses will ultimately be forced to close their doors, the industry juggernauts of the AI age will see profits soar to previously unimaginable levels. This concentration of economic power in the hands of a few will rub salt in the open wounds of social inequality.

In an age of intelligent machines, what does it mean to be human?

In most developed countries, economic inequality and class-based resentment rank among the most dangerous and potentially explosive problems. The past few years have shown us how a cauldron of long-simmering inequality can boil over into radical political upheaval. I believe that, if left unchecked, AI will throw gasoline on the socioeconomic fires.

Lurking beneath this social and economic turmoil will be a psychological struggle, one that won't make the headlines but that could make all the difference. As more and more humans see themselves displaced by machines, they will be forced to answer a far deeper question: in an age of intelligent machines, what does it mean to be human?

The Techno-Optimists and the “Luddite Fallacy”

Like the utopian and dystopian forecasts for AGI, this prediction of a jobs and inequality crisis is not without controversy. A large contingent of economists and techno-optimists believe that fears about technology-induced job losses are fundamentally unfounded.

Members of this camp dismiss dire predictions of unemployment as the product of a “Luddite fallacy.” The term is derived from the Luddites, a group of nineteenth-century British weavers who smashed the new industrial textile looms that they blamed for destroying their livelihood. Despite the best efforts and protests of the Luddites, industrialization plowed full-steam ahead, and both the number of jobs and quality of life in England rose steadily for much of the next two centuries. The Luddites may have failed in their bid to protect their craft from automation—and many of those directly impacted by automation did in fact suffer stagnant wages for some time—but their children and grandchildren were ultimately far better off for the change.

This, the techno-optimists assert, is the real story of technological change and economic development. Technology improves human productivity and lowers the price of goods or services. Those lower prices mean consumers have greater spending power, and they either buy more of the original goods or spend that money on something else. Both of these outcomes increase the demand for human labor and thus jobs. Yes, shifts in technology might lead to some short-term displacement. But just as millions of farmers became factory workers, those laid-off factory workers can become yoga teachers and software programmers. Over the long term, technological progress never truly leads to an actual reduction in jobs or increase in unemployment.

It’s a simple and elegant explanation of the ever-increasing material wealth and relatively stable job markets in the industrialized world. It also serves as a lucid

rebuttal to a series of “boy who cried wolf” moments around technological unemployment. Ever since the Industrial Revolution, people have feared that everything from weaving looms to tractors to ATMs will lead to massive job losses. But each time, increasing productivity has paired up with the magic of the market to smooth things out.

Not all inventions are created equal. Some perform a single task (typewriters), some eliminate the need for one kind of labor (calculators), and some disrupt a whole industry (the cotton gin).

Economists who look to history—and the corporate juggernauts who will profit tremendously from AI—use these examples from the past to dismiss claims of AI-induced unemployment in the future. They point to millions of inventions—the cotton gin, lightbulbs, cars, video cameras, and cell phones—none of which led to widespread unemployment. Artificial intelligence, they say, will be no different. It will greatly increase productivity and promote healthy growth in jobs and human welfare. So what is there to worry about?

The End of Blind Optimism

If we think of all inventions as data points and weight them equally, the techno-optimists have a compelling and data-driven argument. But not all inventions are created equal. Some of them change how we perform a single task (typewriters), some of them eliminate the need for one kind of labor (calculators), and some of them disrupt a whole industry (the cotton gin).

And then there are technological changes on an entirely different scale. The ramifications of these breakthroughs will cut across dozens of industries, with the potential to fundamentally alter economic processes and even social organization.

These are what economists call general purpose technologies, or GPTs. In their landmark book *The Second Machine Age*, MIT professors Erik Brynjolfsson and Andrew McAfee described GPTs as the technologies that “really matter,” the ones that “interrupt and accelerate the normal march of economic progress.”

Looking only at GPTs dramatically shrinks the number of data points available for evaluating technological change and job losses. Economic historians have many quibbles over exactly which innovations of the modern era should qualify (railroads? the internal combustion engine?), but surveys of the literature reveal three technologies that receive broad support: the steam engine, electricity, and information and communication technology (such as computers and the internet). These have been the game changers, the disruptive technologies that extended their reach into many corners of the economy and changed how we live and work.

These three GPTs have been rare enough to warrant evaluating on their own, not simply to be lumped in with millions of more narrow innovations like the ballpoint pen or automatic transmission. And while it’s true that the long-term historical trend has been toward more jobs and greater prosperity, when looking at GPTs alone, three data points is not enough to extract an iron-clad principle. Instead, we should look to the historical record to see how each of these groundbreaking innovations has affected jobs and wages.

The steam engine and electrification were crucial pieces of the first and second Industrial Revolutions (1760–1830 and 1870–1914, respectively). Both of these GPTs facilitated the creation of the modern factory system, bringing immense power and abundant light to the buildings that were upending traditional modes of production. Broadly speaking, this change in the mode of production was one of deskilling. These factories took tasks that once required a high-skilled worker (for example, handcrafting textiles) and broke the work down into far simpler tasks that could be done by low-skilled workers (operating a steam-driven power loom). In the process,

these technologies greatly increased the amount of these goods produced and drove down prices.

In terms of employment, early GPTs enabled process innovations like the assembly line that gave thousands—and eventually hundreds of millions—of former farmers a productive role in the new industrial economy. Yes, they displaced a relatively small number of skilled craftspeople (some of whom would become Luddites), but they empowered much larger numbers of low-skilled workers to take on repetitive, machine-enabled jobs that increased their productivity. Both the economic pie and overall standards of living grew.

But what about the most recent GPT, information and communication technologies (ICT)? So far, its impact on labor markets and wealth inequality have been far more ambiguous. As Brynjolfsson and McAfee point out in *The Second Machine Age*, over the past thirty years, the United States has seen steady growth in worker productivity but stagnant growth in median income and employment. Brynjolfsson and McAfee call this process “the great decoupling.” After decades of productivity, wages, and jobs rising in almost lock-step fashion, that once tightly woven thread has begun to fray. While productivity has continued to shoot upward, wages and jobs have flat-lined or fallen.

This is leading to increasing economic stratification in developed countries like the United States, with the economic gains of ICT increasingly accruing to the top 1 percent. That elite group in the United States has roughly doubled its share of national income between 1980 and 2016. By 2017, the top 1 percent of Americans owned almost twice as much wealth as the bottom 90 percent combined. While the most recent GPT proliferated across the economy, real wages for the median American have remained flat for over thirty years, and they’ve actually fallen for the poorest Americans.

One reason why information and communication technologies may differ from the steam engine and electrification is because of ICT's “skill-bias.” While the two other

GPTs ramped up productivity by deskilling the production of goods, ICT is instead often—though not always—skill-biased in favor of high-skilled workers. Digital communications tools allow top performers to efficiently manage much larger organizations and reach much larger audiences. By breaking down the barriers to disseminating information, ICT empowers the world’s top knowledge workers and undercuts the economic role of many in the middle.

Debates over how large a role ICT has played in job and wage stagnation in the United States are complex. Globalization, the decline of unions, and outsourcing are all factors here, providing economists with fodder for endless academic arguments. But one thing is increasingly clear: there is no guarantee that GPTs which increase our productivity will also lead to more jobs or higher wages for workers.

Techno-optimists can continue to dismiss these concerns as the same old “Luddite fallacy,” but they are now arguing against some of the brightest economic minds of our present day. Lawrence Summers has served as the chief economist of the World Bank, the treasury secretary under President Bill Clinton, and as the director of President Barack Obama’s National Economic Council. In recent years, he has begun warning against the no-questions-asked optimism around technological change and employment.

“The answer is surely not to try to stop technical change,” Summers told the New York Times in 2014, “but the answer is not to just suppose that everything’s going to be O.K. because the magic of the market will assure that’s true.”

Erik Brynjolfsson has issued similar warnings about the growing disconnect between the creation of wealth and jobs, calling it “the biggest challenge of our society for the next decade.”

AI: Putting the G in GPT

What does all this have to do with AI? I am confident that AI will very soon enter the elite club of universally recognized GPTs, spurring a revolution in economic production and even social organization. The AI revolution will be on the scale of the Industrial Revolution but probably larger and definitely faster. Consulting firm PwC predicts that AI will add \$15.7 trillion to the global economy by 2030. If that prediction holds up, it will be an amount larger than the entire GDP of China today and equal to approximately 80 percent of the GDP of the United States in 2017. Seventy percent of those gains are predicted to accrue in the United States and China.

These disruptions will be more broad-based than prior economic revolutions. Steam power fundamentally altered the nature of manual labor, and ICT did the same for certain kinds of cognitive labor. AI will cut across both. It will perform many kinds of physical and intellectual tasks with a speed and power that far outstrip any human, dramatically increasing productivity in everything from transportation to manufacturing to medicine.

The main thrust of AI's employment impact is not one of job creation through deskilling but of job replacement through increasingly intelligent machines

Unlike the GPTs of the first and second Industrial Revolutions, AI will not facilitate the deskilling of economic production. It won't take advanced tasks done by a small number of people and break them down further for a larger number of low-skill workers to do. Instead, it will simply take over the execution of tasks that meet two criteria: they can be optimized using data, and they do not require social interaction. (I go into greater detail about exactly which jobs AI can and cannot replace below.)

Yes, there will be some new jobs created in this process—robot repairing and AI data scientists, for example. But the main thrust of AI's employment impact is not one of job creation through deskilling but of job replacement through increasingly intelligent machines. Displaced workers can theoretically transition into other industries that are more difficult to automate, but this is itself a highly disruptive process that will take a long time.

Hardware, Better, Faster, Stronger

And time is one thing that the AI revolution is not inclined to grant us. The transition to an AI-driven economy will be far faster than any of the prior GPT-induced transformations, leaving workers and organizations in a mad scramble to adjust. Whereas the Industrial Revolution took place across several generations, the AI revolution will have a major impact within just one generation. That's because AI adoption will be accelerated by three catalysts that didn't exist during the introduction of steam power and electricity.

First, many productivity-increasing AI products are just digital algorithms: infinitely replicable and instantly distributable around the world. This makes for a stark contrast to the hardware-intensive revolutions of steam power, electricity, and even large parts of ICT. For these transitions to gain traction, physical products had to be invented, prototyped, built, sold, and shipped to end users. Each time a marginal improvement was made to one of these pieces of hardware, it required that the earlier process be repeated, with the attending costs and social frictions that slowed down adoption of each new tweak. All of these frictions slowed down development of new technologies and extended the time until a product was cost-effective for businesses to adopt.

Digital algorithms can be distributed at virtually no cost, and once distributed, they can be updated and improved for free.

In contrast, the AI revolution is largely free of these limitations. Digital algorithms can be distributed at virtually no cost, and once distributed, they can be updated and improved for free. These digital algorithms—not advanced robotics—will roll out quickly and take a large chunk out of white-collar jobs. Much of today’s white-collar workforce is paid to take in and process information, and then make a decision or recommendation based on that information—which is precisely what AI algorithms do best. In industries with a minimal social component, that human-for-machine replacement can be made rapidly and done en masse, without any need to deal with the messy details of manufacturing, shipping, installation, and onsite repairs. While the hardware of AI-powered robots or self-driving cars will bear some of these legacy costs, the underlying software does not, allowing for the sale of machines that actually get better over time. Lowering these barriers to distribution and improvement will rapidly accelerate AI adoption.

The second catalyst is one that many in the technology world today take for granted: the creation of the venture-capital industry. VC funding—early investments in high-risk, high-potential companies—barely existed before the 1970s. That meant the inventors and innovators during the first two Industrial Revolutions had to rely on a thin patchwork of financing mechanisms to get their products off the ground, usually via personal wealth, family members, rich patrons, or bank loans. None of these have incentive structures built for the high-risk, high-reward game of funding transformative innovation. That dearth of innovation financing meant many good ideas likely never got off the ground, and successful implementation of the GPTs scaled far more slowly.

Today, VC funding is a well-oiled machine dedicated to the creation and commercialization of new technology. In 2017, global venture funding set a new record with \$148 billion invested, egged on by the creation of Softbank's \$100 billion "vision fund," which will be disbursed in the coming years. That same year, global VC funding for AI startups leapt to \$15.2 billion, a 141 percent increase from 2016. That money relentlessly seeks out ways to wring every dollar of productivity out of a GPT like artificial intelligence, with a particular fondness for moonshot ideas that could disrupt and recreate an entire industry. Over the coming decade, voracious VCs will drive the rapid application of the technology and the iteration of business models, leaving no stone unturned in exploring everything that AI can do.

Artificial intelligence will be the first general purpose technology of the modern era in which China stands shoulder to shoulder with the West in both advancing and applying the technology.

Finally, the third catalyst is one that's equally obvious and yet often overlooked: China.

Artificial intelligence will be the first general purpose technology of the modern era in which China stands shoulder to shoulder with the West in both advancing and applying the technology. During the eras of industrialization, electrification, and computerization, China lagged so far behind that its people could contribute little, if anything, to the field. It's only in the past five years that China has caught up enough in internet technologies to feed ideas and talent back into the global ecosystem, a trend that has dramatically accelerated innovation in the mobile internet.

With artificial intelligence, China's progress allows for the research talent and creative capacity of nearly one-fifth of all humanity to contribute to the task of distributing and utilizing artificial intelligence. Combine this with the country's gladiatorial

entrepreneurs, unique internet ecosystem, and proactive government push, and China's entrance to the field of AI constitutes a major accelerant to AI that was absent for previous GPTs.

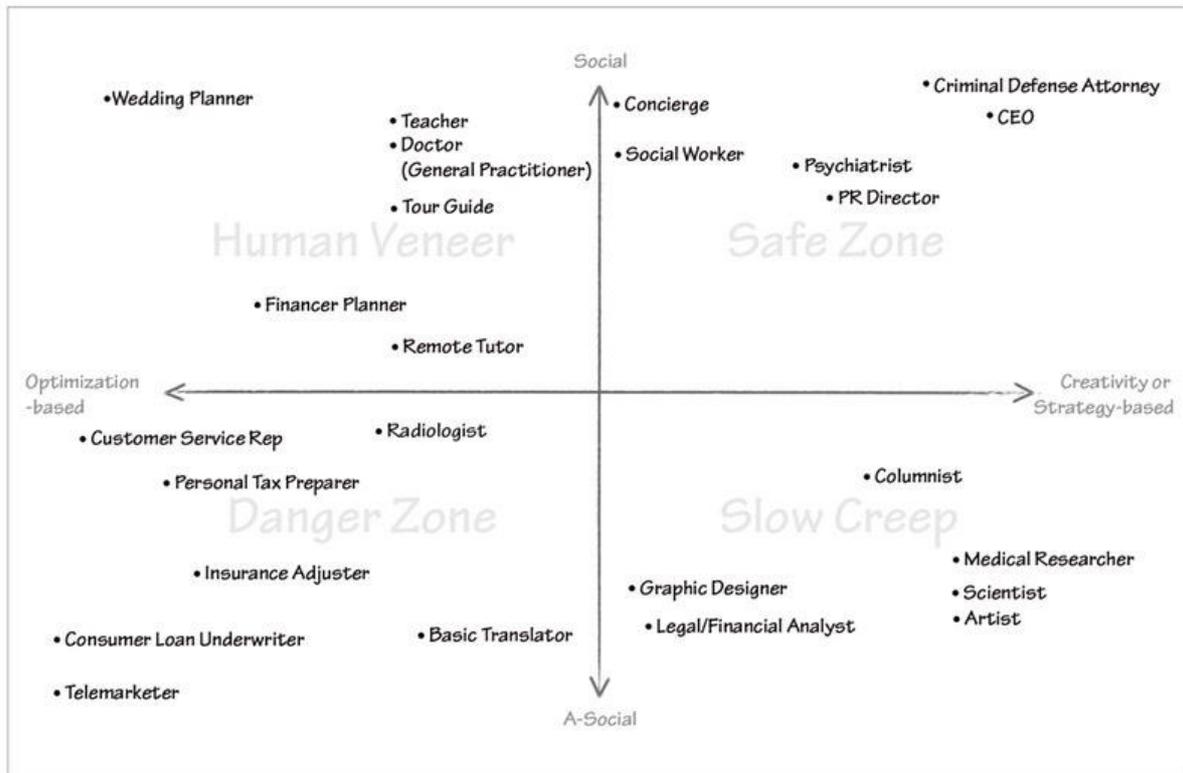
Reviewing the above arguments, I believe we can confidently state a few things. First, during the industrial era, new technology has been associated with long-term job creation and wage growth. Second, despite this general trend toward economic improvements, GPTs are rare and substantial enough that each one's impact on jobs should be evaluated independently. Third, of the three widely recognized GPTs of the modern era, the skill biases of steam power and electrification boosted both productivity and employment. ICT has lifted the former but not necessarily the latter, contributing to falling wages for many workers in the developed world and greater inequality. Finally, AI will be a general purpose technology, a GPT whose skill-biases and speed of adoption— catalyzed by digital dissemination, VC funding, and China— suggest it will lead to negative impacts on employment and income distribution.

If the above arguments hold true, the next questions are clear: What jobs are really at risk? And how bad will it be?

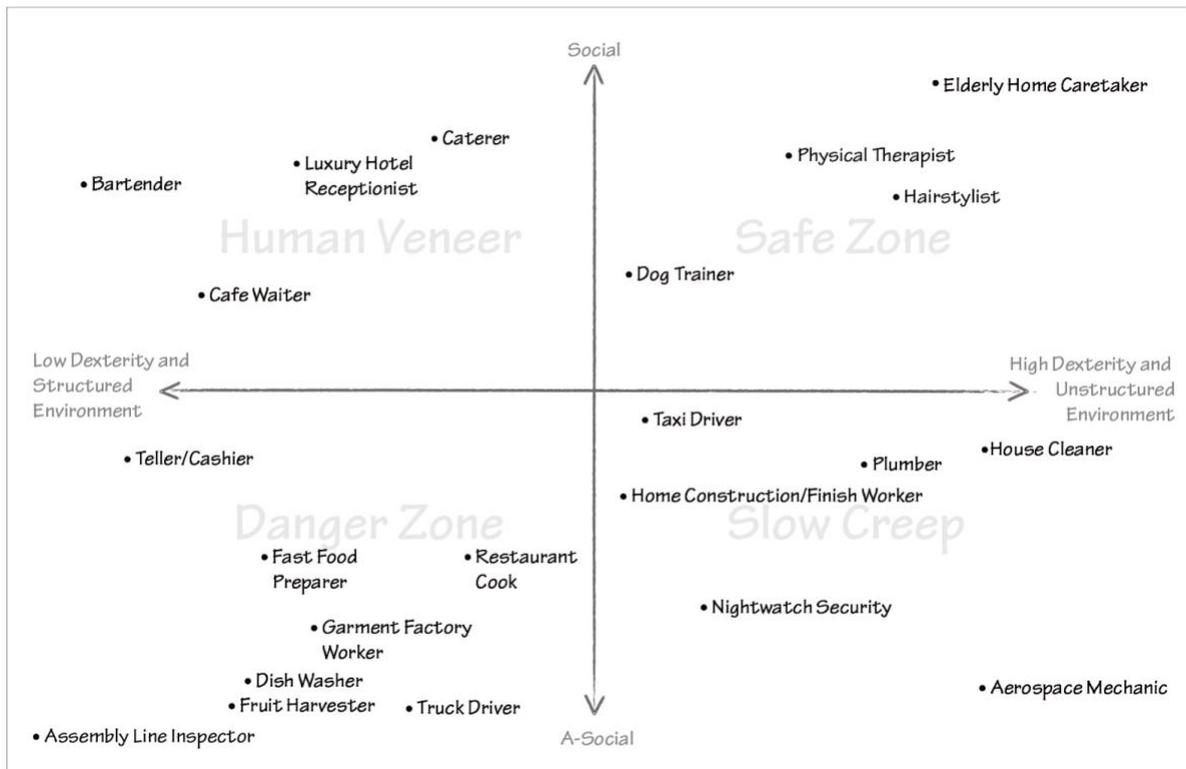
What AI Can and Can't Do: The Risk-of-Replacement Graphs

When it comes to job replacement, AI's biases don't fit the traditional one-dimensional metric of low-skill versus high-skill labor. Instead, AI creates a mixed bag of winners and losers depending on the exact content of job tasks performed. While AI has far surpassed humans at narrow tasks that can be optimized based on data, it remains stubbornly unable to interact naturally with humans or imitate the dexterity of our fingers and limbs. It also cannot engage in cross-domain thinking on creative tasks or ones requiring complex strategy, jobs whose inputs and outcomes aren't easily quantified. What this means for job replacement can be expressed simply through two X-Y graphs, one for physical labor and one for cognitive labor.

Risk-of-Replacement Graph: Cognitive Labor



Risk-of-Replacement Graph: Physical Labor



For physical labor, the X-axis extends from “low dexterity and structured environment” on the left side, to “high dexterity and unstructured environment” on the right side. The Y-axis moves from “asocial” at the bottom to “highly social” at the top. The cognitive labor chart shares the same Y-axis (asocial to highly social) but uses a different X-axis: “optimization-based” on the left, to “creativity- or strategy-based” on the right. Cognitive tasks are categorized as “optimization-based” if their core tasks involve maximizing quantifiable variables that can be captured in data (for example, setting an optimal insurance rate or maximizing a tax refund).

These axes divide both charts into four quadrants: the bottom-left quadrant is the “Danger Zone,” the top-right is the “Safe Zone,” the top-left is the “Human Veneer,” and the bottom right is the “Slow Creep.” Jobs whose tasks primarily fall in the “Danger Zone” (dishwasher, entry-level translators) are at a high risk of replacement in the coming years. Those in the “Safe Zone” (psychiatrist, home-care nurse, etc.) are likely out of reach of automation for the foreseeable future. The “Human Veneer” and “Slow Creep” quadrants are less clear-cut: while not fully replaceable right now,

reorganization of work tasks or steady advances in technology could lead to widespread job reductions in these quadrants. As we see below, occupations often involve many different activities outside of the “core tasks” that we have used to place them in a given quadrant. This task-diversity will complicate the automation of many professions, but for now we can use these axes and quadrants as general guidance for thinking about what occupations are at risk.

For the “Human Veneer” quadrant, much of the computational or physical work can already be done by machines, but the key social interactive element makes them difficult to automate en masse. The name of the quadrant derives from the most likely route to automation: while the behind-the-scenes optimization work is overtaken by machines, human workers will act as the social interface for customers, leading to a symbiotic relationship between human and machine. Jobs in this category could include bartender, schoolteacher, and even medical caregiver. How quickly and what percentage of these jobs disappear depends on how flexible companies are in restructuring the tasks done by their employees, and how open customers are to interacting with computers.

The “Slow Creep” category (plumber, construction worker, entry-level graphic designer) doesn’t rely on human beings’ social skills but instead on manual dexterity, creativity, or ability to adapt to unstructured environments. These remain substantial hurdles for AI, but ones that the technology will slowly chip away at in the coming years. The pace of job elimination in this quadrant depends less on process innovation at companies and more on the actual expansion in AI capabilities. But at the far right end of the “Slow Creep” are good opportunities for the creative professionals (such as scientists and aerospace engineers) to use AI tools to accelerate their progress.

These graphs give us a basic heuristic for understanding what kinds of jobs are at risk, but what does this mean for total employment on an economy-wide level? For that, we must look to the economists.

What the Studies Say

Predicting the scale of AI-induced job losses has become a cottage industry for economists and consulting firms the world over. Depending on which model one uses, estimates range from terrifying to totally not a problem. Here I give a brief overview of the literature and the methods, highlighting the studies that have shaped the debate. Few good studies have been done for the Chinese market, so I largely stick to studies estimating automation potential in the United States and then extrapolate those results to China.

A pair of researchers at Oxford University kicked things off in 2013 with a paper making a dire prediction: 47 percent of U.S. jobs could be automated within the next decade or two. The paper's authors, Carl Benedikt Frey and Michael A. Osborne, began by asking machine-learning experts to evaluate the likelihood that seventy occupations could be automated in the coming years. Combining that data with a list of the main "engineering bottlenecks" in machine learning (similar to the characteristics denoting the "Safe Zone" in the above graphs), Frey and Osborne used a probability model to project how susceptible an additional 632 occupations are to automation.

The result—that nearly half of U.S. jobs were at "high risk" in the coming decades—caused quite a stir. Frey and Osborn were careful to note the many caveats to their conclusion. Most importantly, it was an estimate of what jobs it would be technically possible to do with machines, not actual job losses or resulting unemployment levels. But the ensuing flurry of press coverage largely glossed over these important details, instead warning readers that half of all workers would soon be out of a job.

Other economists struck back. In 2016, a trio of researchers at the Organization for Economic Cooperation and Development (OECD) used an alternate model to produce an estimate that seemed to directly contradict the Oxford study: just 9 percent of jobs in the United States were at high risk of automation.

Why the huge gap? The OECD researchers took issue with Osborne and Frey's "occupation-based" approach. While the Oxford researchers asked machine-learning experts to judge the automatability of an occupation, the OECD team pointed out that it's not entire occupations that will be automated but rather specific tasks within those occupations. The OECD team argued that this focus on occupations overlooks the many different tasks an employee performs that an algorithm cannot: working with colleagues in groups, dealing with customers face-to-face, and so on.

The OECD team instead proposed a task-based approach, breaking down each job into its many component activities and looking at how many of those could be automated. In this model, a tax preparer is not merely categorized as one occupation but rather as a series of tasks that are automatable (reviewing income documents, calculating maximum deductions, reviewing forms for inconsistencies, etc.) and tasks that are not automatable (meeting with new clients, explaining decisions to those clients, etc.).

The OECD team then ran a probability model to find what percentage of jobs were at "high risk" (i.e., at least 70 percent of the tasks associated with the job could be automated). As noted, they found that in the United States only 9 percent of workers fell in the high-risk category. Applying that same model on twenty other OECD countries, the authors found that the percentage of high-risk jobs ranged from just 6 percent in Korea to 12 percent in Austria. Don't worry, the study seemed to say, reports of the death of work have been greatly exaggerated.

Unsurprisingly, that didn't settle the debate. The OECD's task-based approach came to hold sway among researchers, but not all of them agreed with the report's sanguine conclusions. In early 2017, researchers at PwC used the task-based approach to produce their own estimate, finding instead that 38 percent of jobs in the United States were at high risk of automation by the early 2030s. It was a striking divergence from the OECD's 9 percent, one that stemmed simply from using a slightly different algorithm in the calculations. Like the previous studies, the PwC

authors are quick to note that this is merely an estimate of what jobs could be done by machines, and that actual job losses will be mitigated by regulatory, legal, and social dynamics.

Around 50 percent of work tasks around the world may already be automatable.

After these wildly diverging estimates, researchers at the McKinsey Global Institute landed somewhere in the middle. I assisted the institute in its research related to China and coauthored a report with it on the Chinese digital landscape. Using the popular task-based approach, the McKinsey team estimated that around 50 percent of work tasks around the world are already automatable. For China, that number was pegged at 51.2 percent, with the United States coming in slightly lower at 45.8 percent. But when it came to actual job displacement, the McKinsey researchers were less pessimistic. If there is rapid adoption of automation techniques (a scenario most comparable to the above estimates), 30 percent of work activities around the world could be automated by 2030, but only 14 percent of workers would need to change occupations.

So where does this survey of the literature leave us? Experts continue to be all over the map, with estimates of automation potential in the United States ranging from just 9 percent to 47 percent. Even if we stick to just the task-based approach, we still have a spread of 9 to 38 percent, a divide that could mean the difference between broad-based prosperity and an outright jobs crisis. That spread of estimates shouldn't cause us to throw up our hands in confusion. Instead, it should spur us to think critically about what these studies can teach us—and what they may have missed.

What the Studies Missed

While I respect the expertise of the economists who pieced together the above estimates, I also respectfully disagree with the low-end estimates of the OECD. That difference is rooted in two disagreements: one in terms of the inputs to their equations, and one major difference in the way I envision AI disrupting labor markets. The quibble causes me to go with the higher-end estimates of PwC, and the difference in vision leads me to bring that number up higher still.

My disagreement on inputs stems from the way the studies estimated the technical capabilities of machines in the years ahead. The 2013 Oxford study asked a group of machine-learning experts to predict whether seventy occupations could likely be automated in the coming two decades, using those assessments to project automatability more broadly. And though the OECD and PwC studies differed in how they divided up occupations and tasks, they basically stuck with the 2013 estimates of future capabilities.

Those estimates likely constituted the best guess of experts at the time, but significant advances in the accuracy and power of machine learning over the last five years have already moved the goalposts. Experts back then may have been able to project some of the improvements that were on the horizon. But few, if any, experts predicted that deep learning was going to get this good, this fast. Those unexpected improvements are expanding the realm of the possible when it comes to real-world uses and thus job disruptions.

One of the clearest examples of these accelerating improvements is the ImageNet competition. In the competition, algorithms submitted by different teams are tasked with identifying thousands of different objects within millions of different images: birds, baseballs, screwdrivers, and mosques. It has quickly emerged as one of the most respected image-recognition contests and a clear benchmark for AI's progress in computer vision.

When the Oxford machine-learning experts made their estimates of technical capabilities in early 2013, the most recent ImageNet competition of 2012 had been the coming-out party for deep learning. Geoffrey Hinton's team used those techniques to achieve a record-setting error rates of around 16 percent, a major leap forward in a competition where no team had ever gotten below 25 percent.

That was enough to wake up much of the AI community to this thing called deep learning, but it was just a taste of what was to come. By 2017, almost every team had driven error rates below 5 percent—approximately the accuracy of humans performing the same task—with the average algorithm of that year making just one-third of the mistakes of the top algorithm of 2012. In the years since the Oxford experts made their predictions, computer vision has now surpassed human capabilities and dramatically expanded real-world use-cases for the technology.

Those amped-up capabilities extend far beyond computer vision. New algorithms constantly set and surpass records in fields like speech recognition, machine reading, and machine translation. While these strengthened capabilities don't constitute fundamental breakthroughs in AI, they do open the eyes and spark the imaginations of entrepreneurs. Taken together, these technical advances and emerging uses cause me to land on the higher end of task-based estimates, namely PwC's prediction that 38 percent of U.S. jobs will be at high risk of automatability by the early 2030s.

Two Kinds of Job Loss: One-to-One Replacements and Ground-Up Disruptions

But beyond that disagreement over methodology, I believe using only the task-based approach misses an entirely separate category of potential job losses: industry-wide disruptions due to new AI-empowered business models. Separate from the occupation- or task-based approach, I'll call this the industry-based approach.

Part of this difference in vision can be attributed to professional background. Many of the above studies are done by economists, whereas I am a technologist and early-stage investor. In predicting what jobs were at risk of automation, economists looked at what tasks a human completed while going about their job and asked whether a machine would be able to complete those same tasks. In other words, the task-based approach asked how possible it was to do a one- to-one replacement of a machine for a human worker.

My background trains me to approach the problem differently. Early in my career, I worked on turning cutting-edge AI technologies into useful products, and as a venture capitalist I fund and help build new startups. That work helps me see AI as forming two distinct threats to jobs: one-to-one replacements and ground-up disruptions.

Many of the AI companies I've invested in are looking to build a single AI-driven product that can replace a specific kind of worker, for instance, a robot that can do the lifting and carrying of a warehouse employee or an autonomous-vehicle algorithm that can complete the core tasks of a taxi driver. If successful, these companies will end up selling their products to companies, many of whom may lay off redundant workers as a result. These types of one-to-one replacements are exactly the job losses captured by economists using the task-based approach, and I take PwC's 38 percent estimate as a reasonable guess for this category.

An entirely different breed of AI startups reimagines an industry totally from the ground up.

But then there exists an entirely different breed of AI startups: those that reimagine an industry totally from the ground up. These companies don't look to replace one

human worker with one tailor-made robot that can handle the same tasks; rather, they look for new ways to satisfy the fundamental human need driving the industry. Startups like Smart Finance (the AI-driven lender that employs no human loan officers), the employee-free F5 Future Store (a Chinese startup creating a shopping experience comparable to the Amazon Go supermarket), or Toutiao (the algorithmic news app that employs no editors) are prime examples of these types of companies. Algorithms aren't displacing human workers at these companies, simply because the humans were never there to begin with. But as the lower costs and superior services of these companies drive gains to market share, they will apply pressure to their employee-heavy rivals. Those companies will either be forced to adapt from the ground up—restructuring their workflows to leverage AI and reduce employees—or risk going out of business. Either way, the end result is the same: there will be fewer human workers.

This type of AI-induced job loss is largely missing from the task-based estimates of the economists. If one applied the task-based approach to measuring the automatability of an editor at a news app, you would find dozens of tasks that can't be performed by machines. They can't read and understand news and feature articles, subjectively assess appropriateness for your app's audience, or communicate with reporters and other editors. But when Toutiao's founders built the app, they didn't look for an algorithm that could perform all of the above tasks. Instead, they reimagined how a news app could perform its core function—curate a feed of news stories that users want to read—and then did that employing an AI algorithm.

I estimate this kind of from-the-ground-up disruption will affect about 10 percent of the workforce in the United States. The hardest hit industries will be those that involve high volumes of routine optimization work paired with external marketing or customer service: fast food, financial services, security, even radiology. These changes will eat away at employment in the "Human Veneer" quadrant of the earlier chart, with companies consolidating customer interaction tasks into a handful of

employees, while algorithms do most of the grunt work behind the scenes. The result will be steep—though not total—reductions in jobs in these fields.

The Bottom Line

Putting together percentages for the two types of automatability—38 percent from one-to-one replacements and about 10 percent from ground-up disruption—we are faced with a monumental challenge. Within ten to twenty years, I estimate we will be technically capable of automating 40 to 50 percent of jobs in the United States. For employees who are not outright replaced, increasing automation of their workload will continue to cut into their value-add for the company, reducing their bargaining power on wages and potentially leading to layoffs in the long term. We'll see a larger pool of unemployed workers competing for an even smaller pool of jobs, driving down wages and forcing many into part-time or "gig economy" work that lacks benefits.

This—and I cannot stress this enough—does not mean the country will be facing a 40 to 50 percent unemployment rate. Social frictions, regulatory restrictions, and plain old inertia will greatly slow down the actual rate of job losses. Plus, there will also be new jobs created along the way, positions that can offset a portion of these AI-induced losses, something that I explore in coming chapters. These could cut actual AI-induced net unemployment in half to between 20 and 25 percent, or drive it even lower, down to just 10 to 20 percent.

These estimates are in line with those from the most recent study (as of this writing) that attempted to put a number on actual job losses, a February 2018 study by the consulting firm Bain and Company. Instead of wading into the minutia of tasks and occupations, the Bain study took a macro-level approach, seeking to understand the interplay of three major forces acting on the global economy: demographics, automation, and inequality. Bain's analysis produced a startling bottom-line

conclusion. By 2030, employers will need 20 to 25 percent fewer employees, a percentage that would equal 30 to 40 million displaced workers in the United States.

Bain acknowledged that some of these workers will be reabsorbed into new professions that barely exist today (such as robot repair technician), but predicted that this reabsorption would fail to make a meaningful dent in the massive and growing trend of displacement. And automation's impact will be felt far wider than even this 20 to 25 percent of displaced workers. The study calculated that if we include both displacement and wage suppression, a full 80 percent of all workers will be affected.

This would constitute a devastating blow to working families. Worse still, this would not be a temporary shock, like the fleeting brush with 10 percent unemployment that the United States experienced following the 2008 financial crisis. Instead, if left unchecked, it could constitute the new normal: an age of full employment for intelligent machines and enduring stagnation for the average worker.

U.S.-China Comparison: Moravec's Revenge

But what about China? How will its workers fare in this brave new economy? Few good studies have been conducted on the impacts of automation here, but the conventional wisdom holds that Chinese people will be hit much harder, with intelligent robots spelling the end of a golden era for workers in the "factory of the world." This prediction is based on the makeup of China's workforce, as well as a gut-level intuition about what kinds of jobs become automated.

Over one-quarter of Chinese workers are still on farms, with another quarter involved in industrial production. That compares to less than 2 percent of Americans in agriculture and around 18 percent in industrial jobs. Pundits such as *Rise of the Robots* author Martin Ford have argued that this large base of routine manual labor could make China "ground zero for the economic and social disruption brought on by the rise of the robots." Influential technology commentator Vivek Wadhwa has

similarly predicted that intelligent robotics will erode China's labor advantage and bring manufacturing back to the United States en masse, albeit without the accompanying jobs for humans. "American robots work as hard as Chinese robots," he wrote, "and they also don't complain or join labor unions."

These predictions are understandable given the recent history of automation. Looking back at the last hundred years of economic evolution, blue-collar workers and farmhands have faced the steepest job losses from physical automation. Industrial and agricultural tools (think forklifts and tractors) massively increased the productivity of each manual laborer, reducing demand for workers in these sectors. Projecting this same transition out into the age of AI, the conventional wisdom views China's farm and factory laborers as caught squarely in the cross-hairs of intelligent automation. In contrast, America's heavily service-oriented and white-collar economy has a greater buffer against potential job losses, protected by college degrees and six-figure incomes.

While China will face a wrenching labor-market transition due to automation, large segments of that transition may arrive later or move slower than the job losses wracking the American economy.

In my opinion, the conventional wisdom on this is backward. While China will face a wrenching labor-market transition due to automation, large segments of that transition may arrive later or move slower than the job losses wracking the American economy. While the simplest and most routine factory jobs—quality-control and simple assembly-line tasks—will likely be automated in the coming years, the remainder of these manual labor tasks will be tougher for robots to take over. This is because the intelligent automation of the twenty-first century operates differently

than the physical automation of the twentieth century. Put simply, it's far easier to build AI algorithms than to build intelligent robots.

Core to this logic is a tenet of artificial intelligence known as Moravec's Paradox. Hans Moravec was a professor of mine at Carnegie Mellon University, and his work on both artificial intelligence and robotics led him to a fundamental truth about combining the two: contrary to popular assumptions, it is relatively easy for AI to mimic the high-level intellectual or computational abilities of an adult, but it's far harder to give a robot the perception and sensorimotor skills of a toddler.

Algorithms can blow humans out of the water when it comes to making predictions based on data, but robots still can't perform the cleaning duties of a hotel maid. In essence, AI is great at thinking, but robots are bad at moving their fingers.

Moravec's Paradox was articulated in the 1980s, and some things have changed since then. The arrival of deep learning has provided machines with superhuman perception abilities when it comes to voice or visual recognition. Those same machine-learning breakthroughs have also turbocharged the intellectual abilities of machines, namely the power of spotting patterns in data and making decisions. But the fine motor skills of robots—the ability to grasp and manipulate objects—still lag far behind humans. While AI can beat the best humans at Go and diagnose cancer with extreme accuracy, it cannot yet appreciate a good joke.

The Ascent of the Algorithms and Rise of the Robots

This hard reality about algorithms and robots will have profound effects on the sequence of AI-induced job losses. The last century of physical automation largely hurt blue collar workers, but the coming decades of intelligent automation will hit white-collar workers first. The truth is that these workers have far more to fear from the algorithms that exist today than the robots that still need to be invented.

In short, AI algorithms will be to many white-collar workers what tractors were to farmhands: a tool that dramatically increases the productivity of each worker and

thus shrinks the total number of employees required. And unlike tractors, algorithms can be shipped instantly around the world at no additional cost to their creator. Once that software has been sent out to its millions of users—tax-preparation companies, climate-change labs, law firms—it can be constantly updated and improved with no need to create a new physical product.

Robotics requires a delicate interplay of mechanical engineering, perception AI, and fine-motor manipulation.

Robotics, however, is much more difficult. It requires a delicate interplay of mechanical engineering, perception AI, and fine-motor manipulation. These are all solvable problems, but not at nearly the speed at which pure software is being built to handle white-collar cognitive tasks. Once that robot is built, it must also be tested, sold, shipped, installed, and maintained onsite. Adjustments to the robot's underlying algorithms can sometimes be made remotely, but any mechanical hiccups require hands-on work with the machine. All these frictions will slow down the pace of robotic automation.

This is not to say that China's manual laborers are safe. Drones for deploying pesticides on farms, warehouse robots for unpacking trucks, and vision-enabled robots for factory quality control will all dramatically reduce the jobs in these sectors. And Chinese companies are indeed investing heavily in all of the above. The country is already the world's top market for robots, buying nearly as many as Europe and the Americas combined. Chinese CEOs and political leaders are united in pushing for the steady automation of many Chinese factories and farms.

Whatever gaps exist between China and the United States, those differences will pale in comparison between these two AI superpowers and the rest of the world.

But the resulting blue-collar job losses in China will be more gradual and piecemeal than the sweeping impact of algorithms on white-collar workers. While the right digital algorithm can hit like a missile strike for cognitive labor, robotics' assault on manual labor is closer to trench warfare. Over the long term, I believe the number of jobs at risk of automation will be similar between China and the United States. American education's greater emphasis on creativity and interpersonal skills may even give it an employment edge on a long enough time scale. However, when it comes to adapting to these changes, speed matters, and China's particular economic structure will buy it some time.

The AI Superpowers versus All the Rest

Whatever gaps exist between China and the United States, those differences will pale in comparison between these two AI superpowers and the rest of the world. Silicon Valley entrepreneurs love to describe their products as "democratizing access," "connecting people," and, of course, "making the world a better place." That vision of technology as a cure-all for global inequality has always been something of a wistful mirage, but in the age of AI it could turn into something far more dangerous. If left unchecked, AI will dramatically exacerbate inequality on both international and domestic levels. It will drive a wedge between the AI superpowers and all the rest of the world, and may divide society along class lines that mimic the dystopian science fiction of Hao Jingfang.

As a technology and an industry, AI naturally gravitates toward monopolies. Its reliance on data for improvement creates a self-perpetuating cycle: better products

lead to more users, those users lead to more data, and that data leads to even better products, and thus more users and data. Once a company has jumped out to an early lead, this kind of ongoing repeating cycle can turn that lead into an insurmountable barrier to entry for other firms.

Chinese and American companies have already kick-started this process, leaping out to massive leads over the rest of the world. Canada, the United Kingdom, France, and a few other countries all play host to top-notch talent and research labs, but they often lack the other ingredients needed to become true AI superpowers: a large base of users and a vibrant entrepreneurial and venture-capital ecosystem. Other than London's DeepMind, we have yet to see groundbreaking AI companies emerge from these countries. All of the seven AI giants and an overwhelming portion of top-notch AI engineers are already concentrated in the United States and China. They are building huge stores of data that are feeding into a variety of different product verticals, such as self-driving cars, language translation, autonomous drones, facial recognition, natural-language processing, and much more. The more data these companies accumulate, the harder it will be for companies in any other countries to ever compete.

The U.S. and China are set to capture a full 70 percent of the \$15.7 trillion that AI will add to the global economy by 2030, with China alone taking home \$7 trillion.

As AI spreads its tentacles into every aspect of economic life, the benefits will flow to these bastions of data and AI talent. PwC estimates that the United States and China are set to capture a full 70 percent of the \$15.7 trillion that AI will add to the global economy by 2030, with China alone taking home \$7 trillion. Other countries will be

left to pick up the scraps, while these AI superpowers will boost productivity at home and harvest profits from markets around the globe. American companies will likely lay claim to many developed markets, and China's AI juggernauts will have a better shot at winning over Southeast Asia, Africa, and the Middle East.

I fear this process will exacerbate and significantly grow the divide between the AI haves and have-nots. While AI-rich countries rake in astounding profits, countries that haven't crossed a certain technological and economic threshold will find themselves slipping backward and falling farther behind. With manufacturing and services increasingly done by intelligent machines located in the AI superpowers, developing countries will lose the one competitive edge that their predecessors used to kick-start development: low-wage factory labor.

Large populations of young people used to be these countries' greatest strengths. But in the age of AI, that group will be made up of displaced workers unable to find economically productive work. This sea change will transform them from an engine of growth to a liability on the public ledger—and a potentially explosive one if their governments prove unable to meet their demands for a better life.

Deprived of the chance to claw their way out of poverty, poor countries will stagnate while the AI superpowers take off. I fear this ever-growing economic divide will force poor countries into a state of near-total dependence and subservience. Their governments may try to negotiate with the superpower that supplies their AI technology, trading market and data access for guarantees of economic aid for their population. Whatever bargain is struck, it will not be one based on agency or equality between nations.

The AI Inequality Machine

The same push toward polarization playing out across the global economy will also exacerbate inequality within the AI superpowers. AI's natural affinity for monopolies will bring winner-take-all economics to dozens more industries, and the

technology's skill biases will generate a bifurcated job market that squeezes out the middle class. The "great decoupling" of productivity and wages has already created a tear between the 1 percent and the 99 percent. Left to its own devices, artificial intelligence, I worry, will take this tear and rip it wide open.

We already see this trend toward monopolization in the online world. The internet was supposed to be a place of freewheeling competition and a level playing field, but in a few short years many core online functions have turned into monopolistic empires. For much of the developed world, Google rules search engines, Facebook dominates social networks, and Amazon owns e-commerce. Chinese internet companies tend to worry less about "staying in their lane," so there are more skirmishes between these giants, but the vast majority of China's online activity is still funneled through just a handful of companies.

AI will bring that same monopolistic tendency to dozens of industries, eroding the competitive mechanisms of markets in the process. We could see the rapid emergence of a new corporate oligarchy, a class of AI-powered industry champions whose data edge over the competition feeds on itself until they are entirely untouchable. American antitrust laws are often difficult to enforce in this situation, because of the requirement in U.S. law that plaintiffs prove the monopoly is actually harming consumers. AI monopolists, by contrast, would likely be delivering better and better services at cheaper prices to consumers, a move made possible by the incredible productivity and efficiency gains of the technology.

AI will bring a monopolistic tendency to dozens of industries, eroding the competitive mechanisms of markets in the process.

But while these AI monopolies drive down prices, they will also drive up inequality. Corporate profits will explode, showering wealth on the elite executives and engineers lucky enough to get in on the action. Just imagine: How profitable would Uber be if it had no drivers? Or Apple if it didn't need factory workers to make iPhones? Or Walmart if it paid no cashiers, warehouse employees, or truck drivers? Driving income inequality will be the emergence of an increasingly bifurcated labor market. The jobs that do remain will tend to be either lucrative work for top performers or low-paying jobs in tough industries. The risk of replacement cited in the earlier figure reflects this. The most difficult to automate jobs—those in the top-right corner of the “Safe Zone”—include both ends of the income spectrum: CEOs and home-care aides, venture capitalists and masseuses.

Meanwhile, many of the professions that form the bedrock of the middle class—truck drivers, accountants, office managers—will be hollowed out. Sure, we could try to transition these workers into some of the highly social, highly dexterous occupations that will remain safe. Home healthcare aide, techno-optimists point out, is the fastest growing profession in America. But it's also one of the lowest paid, with an annual salary of around \$22,000. A rush of newly displaced workers trying to enter the industry will only exert more downward pressure on that number.

Pushing more people into these jobs while the rich leverage AI for huge gains doesn't just create a society that is dramatically unequal. I fear it will also prove unsustainable and frighteningly unstable.

A Grim Picture

Scanning the economic horizon, artificial intelligence promises to produce wealth on a scale never before seen in human history—something that should be a cause for celebration. But if left to its own devices, AI will also produce a global distribution of wealth that is not just more unequal but hopelessly so. AI-poor countries will find themselves unable to get a grip on the ladder of economic development, relegated

to permanent subservient status. AI-rich countries will amass great wealth but also witness the widespread monopolization of the economy and a bifurcated labor market divided into economic castes.

Make no mistake: this is not just the normal churn of capitalism's creative destruction, a process that has previously always helped lead to a new equilibrium of more jobs, higher wages, and a better quality of life for all. The free market is supposed to be self-correcting, but these self-correcting mechanisms break down in an economy driven by artificial intelligence. Low-cost labor provides no edge over machines, and data-driven monopolies are forever self-reinforcing.

These forces are combining to create a unique historical phenomenon, one that will shake the foundations of our labor markets, economies, and societies. Even if the most dire predictions of jobs losses don't fully materialize, the social impact of wrenching inequality could be just as traumatic. We may never build the folding cities of Hao Jingfang's science fiction, but AI risks creating a twenty-first-century caste system, one that divides the population into the AI elite and what historian Yuval N. Harari has crudely called the "useless class," people who can never generate enough economic value to support themselves. Even worse, recent history has shown us just how fragile our political institutions and social fabric can be in the face of intractable inequality. I fear that recent upheavals are just a dry run for the disruptions to come in the age of AI.

Taking It Personally: The Coming Crisis of Meaning

The resulting turmoil will take on political, economic, and social dimensions, but it will also be intensely personal. In the centuries since the Industrial Revolution, we have increasingly come to see our work not just as a means for survival but as a source of personal pride, identity, and real-life meaning. Asked to introduce ourselves or others in a social setting, a job is often the first thing we mention. It fills our days and provides a sense of routine and a source of human connections. A

regular paycheck has become a way not just of rewarding labor but also of signaling to people that one is a valued member of society, a contributor to a common project.

In the centuries since the Industrial Revolution, we have increasingly come to see our work not just as a means for survival but as a source of personal pride, identity, and real- life meaning.

Severing these ties—or forcing people into downwardly mobile careers—will damage so much more than just our financial lives. It will constitute a direct assault on our sense of identity and purpose. Speaking to the New York Times in 2014, laid-off electrician Frank Walsh described the psychological toll of intractable unemployment.

“I lost my sense of worth, you know what I mean?” Walsh observed. “Somebody asks you ‘What do you do?’ and I would say, ‘I’m an electrician.’ But now I say nothing. I’m not an electrician anymore.”

That loss of meaning and purpose has very real and serious consequences. Rates of depression triple among those unemployed for six months, and people looking for work are twice as likely to kill themselves as the gainfully employed. Alcohol use and opioid overdoses both rise alongside unemployment rates, with some scholars attributing rising mortality rates among uneducated white Americans to declining economic outcomes, a phenomenon they call “deaths of despair.”

The psychological damage of AI-induced unemployment will cut even deeper. People will face the prospect of not just being temporarily out of work but of being permanently excluded from the functioning of the economy. They will watch as algorithms and robots easily outperform them at tasks and skills they spent their

whole lives mastering. It will lead to a crushing feeling of futility, a sense of having become obsolete in one's own skin.

The winners of this AI economy will marvel at the awesome power of these machines.



But the rest of humankind will be left to grapple with a far deeper question: when machines can do everything that we can, what does it mean to be human?

That's a question that I found myself grappling with in the depths of my own personal crisis of mortality and meaning. That crisis brought me to a very dark place, one that pushed my body to the limit and challenged my deepest-held assumptions about what matters in life. But it was that process—and that pain—that opened my eyes to an alternate ending to the story of human beings and artificial intelligence.



About the Author

Dr. Kai-Fu Lee's Unique Perspective on AI and the Two Superpowers

For anyone who wants to understand the dynamics of AI today, and how AI is likely to affect how people will work and live – and what jobs are going to thrive or fall away – Dr. Lee is the ideal guide. Through his wide personal experience in the growth of artificial intelligence in the United States and in China, he knows better than anyone the social and political implications of AI and new technologies and how they will influence the future global economy.

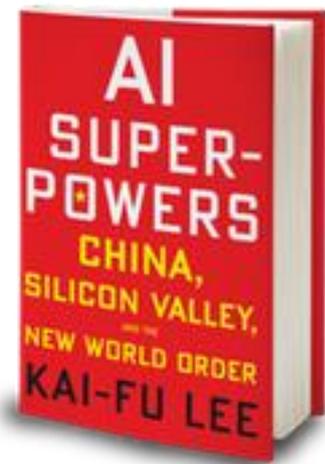


Dr.

Lee's Passionate Vision for the World in AI Superpowers

[AI Superpowers](#) has given Dr. Lee the chance to connect the dots on four decades of work, growth, and evolution. As he reflects on developments in AI through his own personal journey in different companies and cultures, from AI researcher and business executive to venture capitalist, author, and cancer survivor, Dr. Lee touches on issues both global and deeply personal: the rise of artificial intelligence, the intertwined fates of the places that he's called home, and his own evolution from a workaholic to a more loving father, husband, and human being.

Dr. Lee's academic and work background



Dr. Kai-Fu Lee moved to the U.S. from China when he was 11, attending middle school and high school in Oak Ridge, Tennessee. He studied computer science at Columbia University in New York and later got his PhD in computer science from Carnegie Mellon in Pittsburgh, where as his thesis he developed the world's first speaker-independent, continuous speech recognition system. He went on to become one of the world's early experts in artificial intelligence. Dr. Lee later worked as an executive, first at Apple, then SGI, Microsoft, and then served as founding president of Google China, where as one of the most prominent figures in the Chinese internet sector, he was the founding president of Google China. Dr. Lee helped establish Sinovation Ventures, a Chinese early-stage ventures firm, with a presence in Beijing, Shanghai, Shenzhen and Silicon Valley.

About the book

Dr. Kai-Fu Lee—one of the world’s most respected experts on AI and China —reveals that China has suddenly caught up to the US at an astonishingly rapid and unexpected pace.

In AI SUPERPOWERS, Kai-fu Lee argues powerfully that because of these unprecedented developments in AI, dramatic changes will occur much sooner than many of us expected.

As the US-Sino AI competition begins to heat up, Lee urges the US and China to both accept and to embrace the great responsibilities that come with significant technological power. Most experts already say that AI will have a devastating impact on blue-collar jobs. But Lee predicts that Chinese and American AI will have a strong impact on white-collar jobs as well.

Is universal basic income the solution? In Lee’s opinion, probably not. But he provides a clear description of [which jobs will be affected and how soon](#), which jobs can be enhanced with AI, and most importantly, how we can provide solutions to some of the most profound changes in human history that are coming soon.

[Buy the Book](#)

Will AI Take Over Your Jobs?

Yes, depending on your job. The one certain outcome is that AI will change the job market dramatically.

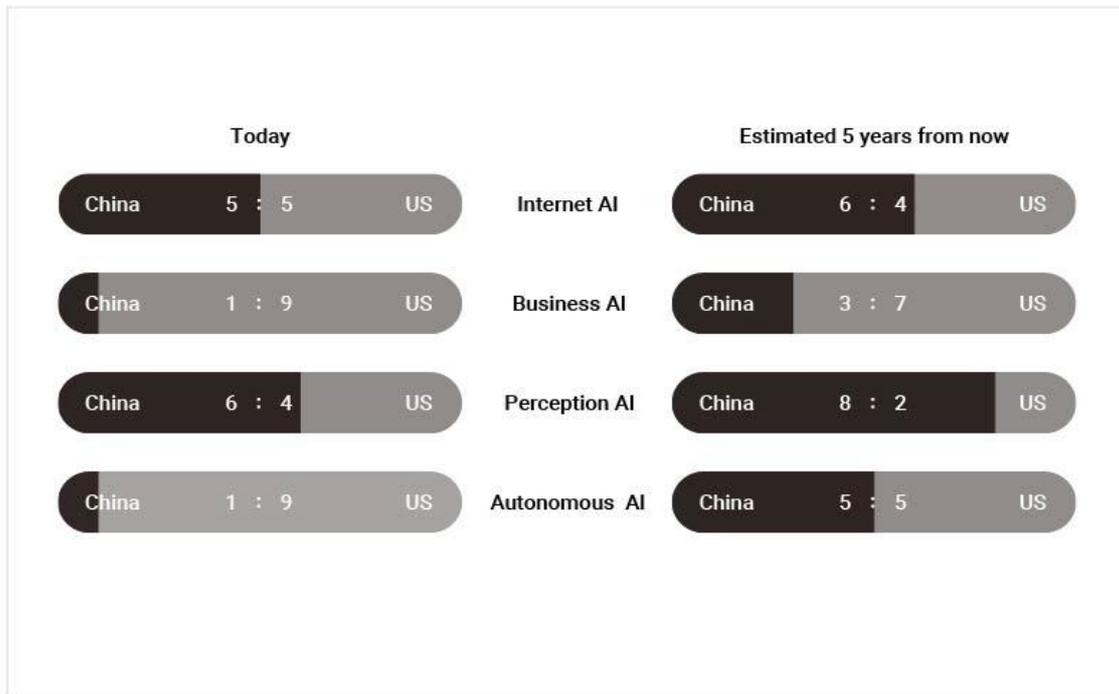
It *will* eliminate thousands of professions, and will impact hundreds of millions of people who work in those jobs, *but* it will also create new ones. But which jobs? What markets? Is your job at risk? Certain white-collar professions such as finance will see massive job loss. But others, such as robotics, will grow. Where are there opportunities? Nursing? The arts? Where are there risks? What will the new jobs be? AI Superpowers will give you insight into the



The Global Impact of AI - China Vs. the U.S.

The two giants of AI, China and the U.S., have two completely different attitudes towards AI in terms of research and development. Which one is better? China's do-anything, copy-anything approach? Silicon Valley's drive to monetize? What will lead to more powerful innovation? China's national support for technology? Or the U.S.'s drive to disrupt?

Dr. Kai-Fu Lee has lived and worked in both countries and knows the markets better than anyone. He steers you through the two very different worlds that both have one giant goal: world-changing AI breakthroughs.



Why Is AI Superpowers So Important and Timely?

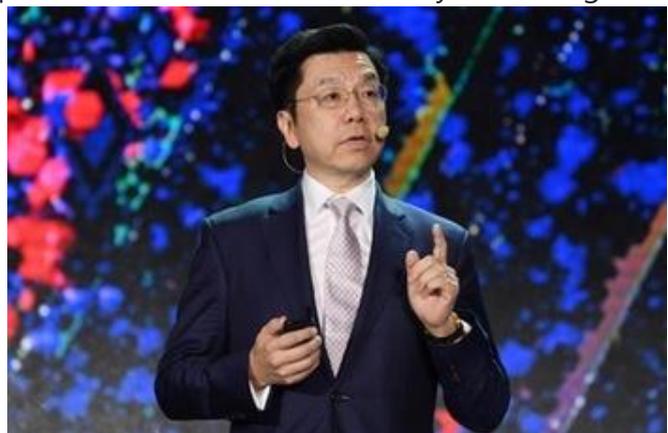
Technology is transforming our society at breakneck speed.

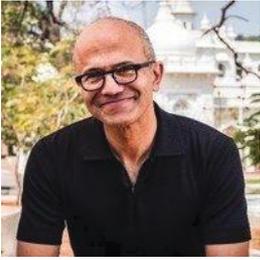
AI isn't just a story of robots and algorithms. It isn't a story of profits and losses. It tells a bigger story, a story about human beings. It's a story about how what makes us human, and why that matters to our future and our relationship with AI.

It explores whether there's space for our humanity in a world run by intelligent machines, and offers inspiration and hope to those that are troubled by the changes that new technology will bring.

[Buy the Book](#)

What people are saying about the book





Satya Nadella

CEO, Microsoft

"Kai-Fu Lee's smart analysis on human-AI coexistence is clear-eyed and a must-read. We must look deep within ourselves for the values and wisdom to guide AI's development."



Marc Benioff

Chairman & CEO Salesforce

"In his brilliant book, Kai-Fu Lee applies his superpowers to the future and proposes a revolutionary social contract that forges a new path forward."



Michelle Rhee

Founder, HuffPost, and founder & CEO, Thrive Global

"Michelle Rhee, a cancer survivor has led to this brilliant book about global tech - the future of work, the future of education, and the future of humanity. This is a must-read for anyone who cares about the future of our world."



Peter Diamandis

Executive Founder, Singularity University; author of NY Times bestseller *Abundance*

"Peter Diamandis is helping us transcend the limitations of thought, reach, and imagination. This book is a must-read for anyone who wants to see the future of technology and humanity."



Chris Anderson

CEO, TED

"Chris Anderson is a master storyteller and a visionary leader. His book is a must-read for anyone who wants to understand the future of technology and humanity."



Prof. Max Tegmark

MIT, author of bestseller Life 3.0: Being Human in the age of AI

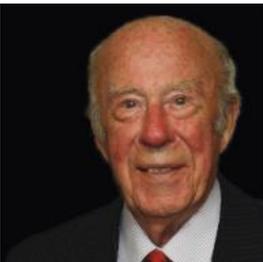
“In this riveting page-turner, one of the founding fathers of China’s AI industry tells the inside story of China’s rise as an AI superpower. This is a must-read for anyone interested in the future of AI and the world.”



Alan Murray

President of FORTUNE

“Kai-Fu Lee may have the most comprehensive view of the global technology scene of any living person.”



George Shultz

Former U.S. Secretary of State

“AI is already presenting new economic opportunities—and new problems of governance—for the United States and China. AI is being applied to transform our lives and our economies. AI Superpowers is a superb primer on this important driver of change.”



John Chen

Chairman & CEO, Blackberry

“Kai-Fu Lee has great insights on one of today’s most exciting and important technological trends. China’s rapid development and commitment to the future should bring an interesting and exciting challenge to the U.S.: the game is on.”



Daniela Rus

Director, MIT Computer Science and Artificial Intelligence Laboratory

“Kai-Fu Lee has been part of the AI revolution for decades. Now, in this fascinating and galvanizing book, he puts into perspective China’s role in the emergence of AI superpowers onto the global scene. He also gives us an optimistic view of a future where, working together with AI systems, people can augment and amplify many aspects of work and life.”

Are you at Risk of losing your job

Over and above the international economic and geo-political changes AI will bring, it will also change the job market dramatically. In the next 15 years up to 40% of jobs could be replaced by AI.

Is your job next?

Take this short [AI Superpowers quiz](#) and discover if your job might be at risk. In addition, you will learn more about your unique *human* superpowers so that you can leverage them in the AI future.

Take the Quiz

