HUMANS NEED NOT APPLY

JERRY KAPLAN

A Guide to Wealth and Work in
the Age of Artificial Intelligence

n a nutshell, after fifty years of effort and billions spent on research, we're cracking the code on artificial intelligence. It turns out that it's not the same as human intelligence, or at least it looks that way right now. But that doesn't matter. In the words of computer scientist Edsger Dijkstra, "The question of whether machines can think is about as relevant as the question of whether submarines can swim." Whether the website that finds you a date or the robot that cuts your grass will do it the same way you do doesn't matter. It will get the job done more quickly, accurately, and at a lower cost than you possibly can.

Recent advances in robotics, perception, and machine learning, propelled by accelerating improvements in computer technology, are enabling a new generation of systems that rival or exceed human capabilities. These developments are likely to usher in a new age of unprecedented prosperity and leisure, but the transition may be protracted and brutal. Without adjustments to our economic system and regulatory policies, we may be in for an extended period of social turmoil.

The warning signs are everywhere. The two great scourges of the modern developed world—persistent unemployment and increasing income inequality—plague our society even as our economy continues to grow. If these are left unchecked, we may witness the spectacle of widespread poverty against a backdrop of escalating comfort and wealth. My goal is to give you a personal tour of the breakthroughs fueling this transition and the challenges it poses for

society: I will also suggest some free-market solutions that promote progress while reducing government intrusion in our lives.

The work in artificial intelligence is advancing on two fronts. New systems of the first class, many of which are already deployed, learn from experience. But unlike humans, who are limited in the scope and scale of experiences they can absorb, these systems can scrutinize mountains of instructive examples at blinding speeds. They are capable of comprehending not only the visual, auditory, and written information familiar to us but also the more exotic forms of data that stream through computers and networks. Imagine how smart you would be if you could see through thousands of eyes, hear distant sounds, and read every word as it is published. Then slow the world down to a pace where you can sample and ponder all of this at your leisure, and you'll get an idea of how these systems experience their environment.

As we amass data from an expanding array of sensors that monitor aspects of the physical world—air quality, traffic flow, ocean wave heights—as well as our own electronic footprints such as ticket sales, online searches, blog posts, and credit card transactions, these systems can glean patterns and grasp insights inaccessible to the human mind. You might reasonably describe them as exhibiting superhuman intelligence, but that's misleading—at least for the foreseeable future—because these machines aren't conscious, self-reflective, and don't exhibit any hint of independent aspirations or personal desires. In other words, they don't have minds, as we commonly understand the word. They are incredibly good at specific tasks, but we don't fully understand how they do what they do. In most cases, that's because there is literally no explanation that can be comprehended by simple creatures like us.

This area of research doesn't have a universally accepted name. Depending on the focus and approach, researchers call it machine learning, neural networks, big data, cognitive systems, or genetic algorithms, among others. I will simply refer generically to the product of their efforts as *synthetic intellects*.

Synthetic intellects are not programmed in the conventional sense. You cobble them together from a growing collection of tools and modules, establish a goal, point them to a trove of examples, and set them loose. Where they wind up is unpredictable and not under their creator's control. Synthetic intellects will soon know more about you than your mother does, be able to predict your behavior better than you can, and warn you of dangers you can't even perceive. I will describe in some detail how synthetic intellects work and why they transcend our common preconceptions of what computers can do.

The second class of new systems arises from the marriage of sensors and actuators. They can see, hear, feel, and interact with their surroundings. When they're bundled together, you can recognize these systems as "robots," but putting them into a single physical package is not essential. In fact, in most cases it's undesirable. The sensors may be sprinkled throughout an environment, on the tops of streetlights or in other people's smartphones, with their observations harvested and siloed in some distant server farm, which then uses this information to formulate a plan. The plan may be executed directly, by controlling remote devices, or indirectly, for example, by coaxing you to take some desired action. Often, the results of these actions are immediately sensed, leading to continuous revision of the plan, just as you do when you guide your hand to pick up an object.

You are part of such a system when you follow automated driving directions. The program, monitoring your location and speed (usually by GPS), directs you, often pooling your information with that of other drivers to detect traffic conditions, which it uses in turn to route you (and them) more efficiently.

Perhaps the most remarkable of these systems will appear deceptively simple, because they accomplish physical tasks that people consider routine. While they lack common sense and general intelligence, they can tirelessly perform an astonishing range of chores in chaotic, dynamic environments.

To date, automation has mostly meant special-purpose machines relegated to performing repetitive, single tasks on factory floors, where the environment is designed around them. In contrast, these new systems will be out and about, tending fields, painting houses, cleaning sidewalks, washing and folding laundry. They may be working in concert with human workers to lay pipes, harvest crops, and build houses, or they may be deployed independently in dangerous or inaccessible places to fight fires, inspect bridges, mine the seabed, and fight wars. I will refer to these embodied systems as *forged laborers*.

Of course, these two types of systems—synthetic intellects and forged laborers—can work in unison to perform physical tasks that require a high level of knowledge and skill, such as fixing cars, performing surgery, and cooking gourmet meals.

In principle, all these developments will not only free you from drudgery but make you more efficient and effective, if you're lucky enough to be able to afford them. Bespoke electronic agents may promote your personal interests, represent you in negotiations, and teach you calculus—but not all such systems will be working on your behalf.

Humans are suckers for the quick win. What Jaron Lanier presciently calls "siren servers" will custom-tailor short-term incentives to your desires, persuading you to do things that may not be in your long-term interests.[1] The irresistible lure of temporary bargains and faster delivery may obscure the gradual destruction of the lifestyle that you hold near and dear. You can order a new rice cooker online tonight and get it delivered tomorrow, but the cost doesn't include the gradual closing of retail stores near your home and the neighbors it puts out of work.

It's one thing for these systems to recommend what, music you should listen to or what toothbrush you should buy. It's quite another when we permit them to take action on their own—or, to use today's buzzword, make them autonomous. Because they operate on timescales we can barely perceive, with access to volumes of data we can't comprehend, they can wreak havoc on an unimaginable scale in the blink of an eye—shutting down electrical grids, placing all airplane takeoffs on hold, canceling millions of credit cards.

You might wonder why someone would build a system that could do such things. It's simple prudence to design in safeguards that protect against rare events, such as simultaneous short circuits in two or more critical power transmission lines. These catastrophic, once-a-century events somehow seem to happen with alarming regularity. When they do, there isn't time for a so-called human in the loop to review the decision in context because the damage is done literally at the speed of light. As scary as it sounds, the launch of a Russian nuclear missile would at least afford us a few minutes to consider an appropriate course of action, but a cyber attack on a nuclear power plant could disable its control systems in an instant. So we have little choice but to trust the machines to protect us.

In the untamed wilds of cyberspace, you never know when two or more autonomous systems whose goals are in conflict may encounter each other. The scale and speed of the resultant electronic brawl can take on the characteristics of a natural disaster. This isn't hypothetical—it's already happened with horrific effects.

On May 6, 2010, the stock market inexplicably plunged 9 percent (one thousand points on the Dow Jones Industrial Average), most of the drop taking place in a matter of minutes. Over $1 trillion in asset value temporarily evaporated, money that represented the retirement savings of millions of workers, among other things. Specialists on the floors of the stock exchanges were left scratching their heads in disbelief.

It took the U.S. Securities and Exchange Commission nearly six months to figure out what had happened, and the answer is hardly comforting: Competing computer programs, buying and selling stocks on behalf of their owners, had gotten out of control. In the murky, secretive world known as high-frequency trading, these systems not only reap small profit opportunities that appear and disappear in an instant but also detect and exploit each other's trading strategies.[2]

What the creators of these electronic card sharks couldn't anticipate was the effect of their programs on each other. Designers develop and test their sophisticated models using historical data, and so cannot predict the presence and behavior of equally capable opposing forces. The seemingly random clash of these titans shook the very foundation of our financial system, which is our faith in its fairness and stability. Economists give this strange new phenomenon the unassuming name of "systemic risk," which makes it sound like something that can be fixed with a shot of regulatory penicillin and a good night's sleep.

But the root cause is much more sinister—the emergence of invisible electronic agents empowered to take actions on behalf of the narrow self-interests of their owners, without regard to the consequences for the rest of the world. Because these agents are stealthy and incorporeal, we can't perceive their presence or comprehend their capabilities. We'd be better off with robotic muggers —at least we could see them coming and run away.

The "Flash Crash" of 2010 may have caught regulators' attention, but it did nothing to slow the application of similar techniques to a wide variety of other domains. Any time you buy something, visit a website, or post a comment online, a hidden army of electronic agents, working for someone else, is watching you. Whole industries have sprung up that do nothing but sell weapons in the form of programs and data to companies brave enough to wade into these never-ending melees. Later in this book I will describe one such arena in detail: the monumental cluster fight that takes place behind the scenes for the right to show you an ad every time you load a web page.

The emergence of powerful autonomous agents raises serious ethical questions. Much of the way we allocate shared resources among people follows unstated social conventions. The town regulations near my house permit me to park in a spot for up to two hours, on the assumption that it is inconvenient for me to move my car that often. But what if my car can change spots by itself? Will my personal robot be permitted to stand in line at the movie theater on my behalf?

Autonomous cars, which are only a few years from broad deployment, raise much more serious issues. The split-second decisions these contraptions will have to make pose ethical questions that have bedeviled deep thinkers for millennia. Imagine that my car is crossing a narrow bridge and a school bus full of children suddenly

enters from the other side. The bridge can't accommodate both vehicles, so to avoid destroying both it's clear that one of them will have to go over the edge. Would I buy a car that is willing to sacrifice my life to save the children? Will the aggressiveness of a self-driving car become a selling point like gas mileage? Moral quandaries like this, no longer confined to the musings of philosophers, will urgently arrive on our courthouse steps.

The emergence of synthetic intellects and forged laborers that act as our individual agents will raise a raft of practical conundrums. What should "one per customer" mean when a robot is the customer, and I own a whole fleet of them? Can my personal electronic assistant lie on my behalf? Should it be required to report me if I direct it to serve my twenty-year-old daughter wine at Thanksgiving dinner?

Society crafts laws and regulations on the assumption that people can occasionally exercise a certain amount of individual discretion. How will you feel about a dog-walking robot that fails to save your child from being mauled because it is obeying a "Keep off the grass" sign? Or an autonomous car that refuses to speed you to the hospital to save you from a heart attack? Our institutions will soon have to grapple with balancing the needs of individuals against the wider interests of society in a whole new way.

But all of this pales in comparison to the economic dangers these systems pose. A broad cross section of today's blue-collar and white-collar jobs will soon come under threat from forged laborers. and synthetic intellects respectively. An astonishing range of productive activities, both physical and mental, will become vulnerable to replacement by these new devices and programs. Why should someone hire you instead of buying one of them?

We're about to discover that Karl Marx was right: the inev-

itable struggle between capital (whose interests are promoted by management) and labor is a losing proposition for workers. What he didn't fully appreciate is that we're all workers, even managers, doctors, and college professors. As an economist, Marx understood that industrial automation substitutes capital for your labor, even if he didn't quite have forged laborers in mind. But what he couldn't foresee is that synthetic intellects can also substitute capital for your mind. So the conflict he characterizes between poorly paid workers and highly compensated managers—people against people,—cuts the wrong way. The real problem is that the wealthy will need few, if any, people to work for them at all.

As bizarre as this sounds, the future will be a struggle of assets against people, as the resources accumulated by our creations serve no constructive purpose or are put to no productive use. As I will explain, the so-called 1 percent may be the beneficiaries of these trends today, but without some careful precautions as to who—or what—may own assets, there's a real possibility that the 1 percent will shrink to the 0 percent, just as the pyramids of ancient Egypt drained the resources of an entire society to serve the personal whims of a single ruler. The economy we know today, as difficult to manage as it may be, is in danger of motoring on without us, throwing ever more of us overboard. Will the last human dismissed please turn off the lights? Actually, no need—they can turn themselves off.

But there are even greater risks. When we think of AI, we tend to picture a future full of robots as docile servants or malevolent overlords (take your pick), or giant computer brains buried in fortified bunkers. There's nothing quite like the image of a grimacing cyborg to raise your hackles. But this is just our anthropomorphic bias and countless Hollywood treatments leading us astray. The real dan-

ger comes from distributed armies of tiny forged laborers organized like swarms of insects, and disembodied synthetic intellects residing on remote servers in the cloud. It's hard to worry about threats that you can't see or perceive. Things can just seem to inexplicably get worse. The Luddites of the early Industrial Revolution could smash the looms that took their jobs, but how can you fight back against a smartphone app?

Modern policy makers are wringing their hands over the root causes of persistent unemployment and economic inequality, but certainly one underappreciated driver is accelerating technological progress. As I will demonstrate, advances in information technology are already gutting industries and jobs at a furious clip, far faster than the labor markets can possibly adapt, and there's much worse to come. They are also substituting capital for labor in completely new ways, delivering a disproportionate share of new wealth to the already rich.

The usual rejoinder to this is that the improved productivity will increase wealth, floating all our boats, and that new jobs will emerge to cater to our expanding desires and needs. True enough—in aggregate and on average. But when you dig deeper, this doesn't necessarily mean that we are better off. With labor markets, as with global warming, it's the pace that matters, not the fact. Current workers may have neither the time nor the opportunity to acquire the skills required by these new jobs. And average income doesn't matter if a small cadre of superwealthy oligarchs takes the lion's share while everyone else lives in relative poverty. Increasing wealth may float all yachts while sinking all rowboats.

In the first chapters of this book I will lay out some foundational concepts and ideas required to reframe the ongoing policy debates. I will attempt to demystify the magic under the hood by explaining why most of what you believe about computers is wrong. Unless you understand what's really happening, you can't appreciate what's likely to happen in the future.

Then I will suggest practical solutions to some of the most serious issues, such as how we can extend our legal system to regulate and hold autonomous systems responsible for their actions. But the economic consequences are by far the most serious problems we will have to address. The obvious simple solution, to redistribute the wealth from the rich to the poor, is a nonstarter in our current political environment. It also doesn't address the underlying cause of the problems; it just stirs up the pot in the hope of preventing it from coming to a boil. Instead, I will present a framework that applies free-market solutions to address the underlying structural problems we are creating.

Unemployment is going to be a serious problem—but not, surprisingly, because of a lack of jobs. Rather, the skills required to do the available jobs are likely to evolve more quickly than workers can adapt without significant changes to how we train our workforce. Our current sequential system of education and work—first you go to school, then you get a job—was fine when you could expect to do more or less the same thing for a living throughout your working life. But looking forward, it simply isn't going to work. The nature of the jobs available will shift so rapidly that you may find your skills obsolete just when you thought you were starting to get ahead. Our current system of vocational training, largely a holdover from medieval apprenticeships and indentured servitude, is in need of significant modernization.

I will propose an approach to this problem in the form of a new type of financial instrument, the "job mortgage," secured exclusively by your future labor (earned income) similar to the way

your home mortgage is secured exclusively by your property. Out of work? Payments are suspended for some reasonable grace period, until you find another job.

In the proposed system, employers and schools will have in-centives to collaborate in a new way. Employers will issue nonbind-ing letters of intent to hire you if you acquire specified skills, and they will get certain payroll tax breaks if they ultimately follow through. These letters of intent will serve the same purpose for job mortgage lenders as an appraisal serves for a home mortgage lender. Training institutions will have to craft their curricula around the specific skills required by sponsoring employers in order to meet the requirements of the loans, or else students won't enroll. You won't be committed in advance to accepting a particular position if someone else makes you a better offer, but at least you have the comfort of knowing that you are acquiring the skills valued by the marketplace. In effect, this scheme introduces a new form of feedback and liquidity into labor markets, enforced through the discipline of the free market.

But our greatest societal challenge will be to rein in growing income inequality. I will propose an objective, government-certified measure of corporate ownership, which I will call the public bene-fit index, or PBI, which can serve as the foundation for a variety of programs to keep society on a more even keel. By scaling corporate taxes based on how many stockholders benefit from a company's success, we can tilt the scales in favor of broad public participation in an asset-based economy. But how can the average Joe and Jane afford to buy assets? For starters, they already own more than you might expect, in the form of pension funds and Social Security—they just don't know it because an opaque system of fiduciaries manage their wealth instead of them. We need to give people more visibility

and control over their nest eggs, with incentives to direct the assets toward high-PBI companies. A side benefit is social stability. The temptation to riot and loot the local department store is greatly di-minished if you know you are a stockholder.

We don't need to take from the wealthy and give to the less fortunate because our economy is not standing still; it's continually expanding, and this growth is likely to quicken. So all we need to, do is distribute the benefits of future growth more widely, and the problem will slowly melt away. A carefully crafted program of tax incentives, portfolio transparency, and increased individual control over asset allocation based on the PBI offers us a way to keep from capsizing in the rising tide of concentrating prosperity.

So why can't our chosen leaders better assess the situation and take corrective actions? Because you can't steer when you can't see, and you can't discuss what you can't articulate. At the moment, our public discourse lacks the concepts and exemplars to properly describe what's likely to happen as technological progress acceler-ates, much less to guide us to reasonable solutions.

Letting nature take its course—as we did during the Industrial Revolution of the late eighteenth and early nineteenth centuries—is a dangerous gamble. Per capita income rose dramatically, but the changes entailed untold human suffering during an extended period of economic transformation. We can ignore the coming storm and eventually everything will work out fine, but "eventually" is a long time. Without some foresight and action now, we may condemn our descendants to half a century or more of poverty and inequality, ex-cept for a lucky chosen few. Everyone likes to play the lottery—until the losers are identified. We can't wait to see who's winning before we take action.

The holy grail of Silicon Valley entrepreneurs is the disruption of entire industries—because that's where the big money is to be made. Amazon dominates book retailing; Uber decimates taxi services; Pandora displaces radio. Little attention is paid to the resulting destruction of livelihoods and assets because there's no incentive to do so. And what's cooking in the research labs is quickening the hearts of investors everywhere.

My goal in this book is to equip you with the intellectual tools, ethical foundation, and psychological framework required to successfully navigate these challenges. Whether we wind up as desperate paupers, willing to gamble our last dime for a chance to join the haves on Easy Street or as freethinking artists, athletes, and academics tenderly cared for by our own creations will largely depend on the public policies we put in place over the next decade or two.

Of course, many talented and thoughtful writers have already rung the alarm about the risks of recent technological advances. Some have expressed this in the form of engaging stories;[3] others have brought to bear the analytical skills of economists.[4] My goal here is to add a different voice to the growing chorus of concern, mine from the perspective of a technology entrepreneur.

Despite this litany of plagues, I remain an optimist. I'm confident we can craft a future of eternal peace and unbounded prosperity. I truly believe the world will be *Star Trek*, not *Terminator*. In the end, the tsunami of new technology will sweep in an extraordinary era of freedom, convenience, and happiness, but it's going to be a rough ride if we don't keep our hands firmly on the wheel of progress.

Welcome to the future, which begins in the past.

1. Teaching Computers to Fish