Artificial intelligence could erase many practical advantages of democracy, and erode the ideals of liberty and equality. It will further concentrate power among a small elite if we don’t take steps to stop it.

I. The Growing Fear of Irrelevance

There is nothing inevitable about democracy. For all the success that democracies have had over the past century or more, they are blips in history. Monarchies, oligarchies, and other forms of authoritarian rule have been far more common modes of human governance.
The emergence of liberal democracies is associated with ideals of liberty and equality that may seem self-evident and irreversible. But these ideals are far more fragile than we believe. Their success in the 20th century depended on unique technological conditions that may prove ephemeral.

In the second decade of the 21st century, liberalism has begun to lose credibility. Questions about the ability of liberal democracy to provide for the middle class have grown louder; politics have grown more tribal; and in more and more countries, leaders are showing a penchant for demagoguery and autocracy. The causes of this political shift are complex, but they appear to be intertwined with current technological developments. The technology that favored democracy is changing, and as artificial intelligence develops, it might change further.

Information technology is continuing to leap forward; biotechnology is beginning to provide a window into our inner lives—our emotions, thoughts, and choices. Together, infotech and biotech will create unprecedented upheavals in human society, eroding human agency and, possibly, subverting human desires. Under such conditions, liberal democracy and free-market economics might become obsolete.

Ordinary people may not understand artificial intelligence and biotechnology in any detail, but they can sense that the future is passing them by. In 1938 the common man’s condition in the Soviet Union, Germany, or the United States may have been grim, but he was constantly told that he was the most important thing in the world, and that he was the future (provided, of course, that he was an “ordinary man,” rather than, say, a Jew or a woman). He looked at the propaganda posters—which typically depicted coal miners and steelworkers in heroic poses—and saw himself there: “I am in that poster! I am the hero of the future!”

In 2018 the common person feels increasingly irrelevant. Lots of mysterious terms are bandied about excitedly in TED Talks, at government think tanks, and at high-tech conferences—globalization, blockchain, genetic engineering, AI, machine learning—and common people, both men and women, may well suspect that none of these terms is about them.
In the 20th century, the masses revolted against exploitation and sought to translate their vital role in the economy into political power. Now the masses fear irrelevance, and they are frantic to use their remaining political power before it is too late. Brexit and the rise of Donald Trump may therefore demonstrate a trajectory opposite to that of traditional socialist revolutions. The Russian, Chinese, and Cuban revolutions were made by people who were vital to the economy but lacked political power; in 2016, Trump and Brexit were supported by many people who still enjoyed political power but feared they were losing their economic worth. Perhaps in the 21st century, populist revolts will be staged not against an economic elite that exploits people but against an economic elite that does not need them anymore. This may well be a losing battle. It is much harder to struggle against irrelevance than against exploitation.

The revolutions in information technology and biotechnology are still in their infancy, and the extent to which they are responsible for the current crisis of liberalism is debatable. Most people in Birmingham, Istanbul, St. Petersburg, and Mumbai are only dimly aware, if they are aware at all, of the rise of AI and its potential impact on their lives. It is undoubtable, however, that the technological revolutions now gathering momentum will in the next few decades confront humankind with the hardest trials it has yet encountered.

II. A New Useless Class?

Let’s start with jobs and incomes, because whatever liberal democracy’s philosophical appeal, it has gained strength in no small part thanks to a practical advantage: The decentralized approach to decision making that is characteristic of liberalism—in both politics and economics—has allowed liberal democracies to outcompete other states, and to deliver rising affluence to their people.

Liberalism reconciled the proletariat with the bourgeoisie, the faithful with atheists, natives with immigrants, and Europeans with Asians by promising everybody a larger slice of the pie. With a constantly growing pie, that was possible.
And the pie may well keep growing. However, economic growth may not solve social problems that are now being created by technological disruption, because such growth is increasingly predicated on the invention of more and more disruptive technologies.

Fears of machines pushing people out of the job market are, of course, nothing new, and in the past such fears proved to be unfounded. But artificial intelligence is different from the old machines. In the past, machines competed with humans mainly in manual skills. Now they are beginning to compete with us in cognitive skills. And we don’t know of any third kind of skill—beyond the manual and the cognitive—in which humans will always have an edge.

At least for a few more decades, human intelligence is likely to far exceed computer intelligence in numerous fields. Hence as computers take over more routine cognitive jobs, new creative jobs for humans will continue to appear. Many of these new jobs will probably depend on cooperation rather than competition between humans and AI. Human-AI teams will likely prove superior not just to humans, but also to computers working on their own.

However, most of the new jobs will presumably demand high levels of expertise and ingenuity, and therefore may not provide an answer to the problem of unemployed unskilled laborers, or workers employable only at extremely low wages. Moreover, as AI continues to improve, even jobs that demand high intelligence and creativity might gradually disappear. The world of chess serves as an example of where things might be heading. For several years after IBM’s computer Deep Blue defeated Garry Kasparov in 1997, human chess players still flourished; AI was used to train human prodigies, and teams composed of humans plus computers proved superior to computers playing alone.

Yet in recent years, computers have become so good at playing chess that their human collaborators have lost their value and might soon become entirely irrelevant. On December 6, 2017, another crucial milestone was reached when Google’s AlphaZero program defeated the Stockfish 8 program. Stockfish 8 had won a world computer chess championship in 2016. It had access to centuries of
accumulated human experience in chess, as well as decades of computer experience. By contrast, AlphaZero had not been taught any chess strategies by its human creators—not even standard openings. Rather, it used the latest machine-learning principles to teach itself chess by playing against itself. Nevertheless, out of 100 games that the novice AlphaZero played against Stockfish 8, AlphaZero won 28 and tied 72—it didn’t lose once. Since AlphaZero had learned nothing from any human, many of its winning moves and strategies seemed unconventional to the human eye. They could be described as creative, if not downright genius.

Can you guess how long AlphaZero spent learning chess from scratch, preparing for the match against Stockfish 8, and developing its genius instincts? Four hours. For centuries, chess was considered one of the crowning glories of human intelligence. AlphaZero went from utter ignorance to creative mastery in four hours, without the help of any human guide.

AlphaZero is not the only imaginative software out there. One of the ways to catch cheaters in chess tournaments today is to monitor the level of originality that players exhibit. If they play an exceptionally creative move, the judges will often suspect that it could not possibly be a human move—it must be a computer move. At least in chess, creativity is already considered to be the trademark of computers rather than humans! So if chess is our canary in the coal mine, we have been duly warned that the canary is dying. What is happening today to human-AI teams in chess might happen down the road to human-AI teams in policing, medicine, banking, and many other fields.

What’s more, AI enjoys uniquely nonhuman abilities, which makes the difference between AI and a human worker one of kind rather than merely of degree. Two particularly important nonhuman abilities that AI possesses are connectivity and updatability.

For example, many drivers are unfamiliar with all the changing traffic regulations on the roads they drive, and they often violate them. In addition, since every driver is a singular entity, when two vehicles approach the same intersection, the drivers sometimes miscommunicate their intentions and collide. Self-driving cars, by
contrast, will know all the traffic regulations and never disobey them on purpose, and they could all be connected to one another. When two such vehicles approach the same junction, they won’t really be two separate entities, but part of a single algorithm. The chances that they might miscommunicate and collide will therefore be far smaller.

Similarly, if the World Health Organization identifies a new disease, or if a laboratory produces a new medicine, it can’t immediately update all the human doctors in the world. Yet even if you had billions of AI doctors in the world—each monitoring the health of a single human being—you could still update all of them within a split second, and they could all communicate to one another their assessments of the new disease or medicine. These potential advantages of connectivity and updatability are so huge that at least in some lines of work, it might make sense to replace all humans with computers, even if individually some humans still do a better job than the machines.

The same technologies that might make billions of people economically irrelevant might also make them easier to monitor and control.

All of this leads to one very important conclusion: The automation revolution will not consist of a single watershed event, after which the job market will settle into some new equilibrium. Rather, it will be a cascade of ever bigger disruptions. Old jobs will disappear and new jobs will emerge, but the new jobs will also rapidly change and vanish. People will need to retrain and reinvent themselves not just once, but many times.

Just as in the 20th century governments established massive education systems for young people, in the 21st century they will need to establish massive reeducation systems for adults. But will that be enough? Change is always stressful, and the hectic world of the early 21st century has produced a global epidemic of stress. As job volatility increases, will people be able to cope? By 2050, a useless class might
emerge, the result not only of a shortage of jobs or a lack of relevant education but also of insufficient mental stamina to continue learning new skills.

III. The Rise of Digital Dictatorships

As many people lose their economic value, they might also come to lose their political power. The same technologies that might make billions of people economically irrelevant might also make them easier to monitor and control.

AI frightens many people because they don’t trust it to remain obedient. Science fiction makes much of the possibility that computers or robots will develop consciousness—and shortly thereafter will try to kill all humans. But there is no particular reason to believe that AI will develop consciousness as it becomes more intelligent. We should instead fear AI because it will probably always obey its human masters, and never rebel. AI is a tool and a weapon unlike any other that human beings have developed; it will almost certainly allow the already powerful to consolidate their power further.

Consider surveillance. Numerous countries around the world, including several democracies, are busy building unprecedented systems of surveillance. For example, Israel is a leader in the field of surveillance technology, and has created in the occupied West Bank a working prototype for a total-surveillance regime. Already today whenever Palestinians make a phone call, post something on Facebook, or travel from one city to another, they are likely to be monitored by Israeli microphones, cameras, drones, or spy software. Algorithms analyze the gathered data, helping the Israeli security forces pinpoint and neutralize what they consider to be potential threats. The Palestinians may administer some towns and villages in the West Bank, but the Israelis command the sky, the airwaves, and cyberspace. It therefore takes surprisingly few Israeli soldiers to effectively control the roughly 2.5 million Palestinians who live in the West Bank.

In one incident in October 2017, a Palestinian laborer posted to his private Facebook account a picture of himself in his workplace, alongside a bulldozer.