

## Technology, Culture, and Dignity — Reading #5

---

### Why Technology Favors Tyranny

Yuval Noah Harari

*The Atlantic* | October 2018 [adapted from Harari's recent book, *21 Lessons for the 21st Century*]

#### I. The Growing Fear of Irrelevance

[1] There is nothing inevitable about democracy. For all their success, democracies have had over the past century or more, 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.

[2] 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. . . might become obsolete.

[3] 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.

[4] 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.

[5] 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?

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

**[14]** 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.

**[15]** 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.

**[16]** 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.

**[17]** 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**

**[18]** 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.

**[19]** 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.

**[20]** 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. Adjacent to the image he wrote, “Good morning!” A Facebook translation algorithm made a small error when transliterating the Arabic letters. Instead of *Ysabechhum* (which means “Good morning”), the algorithm identified the letters as *Ydbachhum* (which means “Hurt them”). Suspecting that the man might be a terrorist intending to use a bulldozer to run people over, Israeli security forces swiftly arrested him. They released him after they realized that the algorithm had made a mistake. Even so, the offending Facebook post was taken down— you can never be too careful. What Palestinians are experiencing today in the West Bank may be just a primitive preview of what billions of people will eventually experience all over the planet.

**[21]** Imagine, for instance, that the current regime in North Korea gained a more advanced version of this sort of technology in the future. North Koreans might be required to wear a biometric bracelet that monitors everything they do and say, as well as their blood pressure and brain activity. Using the growing understanding of the human brain and drawing on the immense powers of machine learning, the North Korean government might eventually be able to gauge what each and every citizen is thinking at each and every moment. If a North Korean looked at a picture of Kim Jong Un and the biometric sensors picked up telltale signs of anger (higher blood pressure, increased activity in the amygdala), that person could be in the gulag the next day.

**[22]** And yet such hard-edged tactics may not prove necessary, at least much of the time. A facade of free choice and free voting may remain in place in some countries, even as the public exerts less and less actual control. To be sure, attempts to manipulate voters’ feelings are not new. But once somebody (whether in San Francisco or Beijing or Moscow) gains the technological ability to manipulate the human heart— reliably, cheaply, and at scale— democratic politics will mutate into an emotional puppet show.

**[23]** We are unlikely to face a rebellion of sentient machines in the coming decades, but we might have to deal with hordes of bots that know how to press our emotional buttons better than our mother does and that use this uncanny ability, at the behest of a human elite, to try to sell us something— be it a car, a politician, or an entire ideology. The bots might identify our deepest fears, hatreds, and cravings and use them against us. We have already been given a foretaste of this in recent elections and referendums across the world, when hackers learned how to manipulate individual voters by analyzing data about them and exploiting their prejudices. While science-fiction thrillers are drawn to dramatic apocalypses of fire and smoke, in reality we may be facing a banal apocalypse by clicking.

**[24]** The biggest and most frightening impact of the AI revolution might be on the relative efficiency of democracies and dictatorships. Historically, autocracies have faced crippling handicaps in regard to innovation and economic growth. In the late 20th century, democracies usually outperformed dictatorships, because they were far better at processing information. We tend to think about the conflict between democracy and dictatorship as a conflict between two different ethical systems, but it is actually a conflict between two different data-processing systems. Democracy distributes the power to process information and make decisions among many people and institutions, whereas dictatorship concentrates information and power in one place. Given 20th-century technology, it was inefficient to concentrate too much information and power in one place. Nobody had the ability to process all available information fast enough and make the right decisions. This is one reason the Soviet Union made far worse decisions than the United States, and why the Soviet economy lagged far behind the American economy.

**[25]** However, artificial intelligence may soon swing the pendulum in the opposite direction. AI makes it possible to process enormous amounts of information centrally. In fact, it might make centralized systems far more efficient than diffuse systems, because machine learning works better when the machine has more information to analyze. If you disregard all privacy concerns and concentrate all the information relating to a billion people in one database, you’ll wind up with much better algorithms than if you respect individual privacy and have in your database only partial information on a million people. An authoritarian government that orders all its citizens to have their DNA sequenced and to share their medical data with some central authority would gain an immense advantage in genetics and medical research over societies in which medical data are strictly private. The main

handicap of authoritarian regimes in the 20th century—the desire to concentrate all information and power in one place— may become their decisive advantage in the 21st century.

**[26]** New technologies will continue to emerge, of course, and some of them may encourage the distribution rather than the concentration of information and power. Blockchain technology, and the use of cryptocurrencies enabled by it, is currently touted as a possible counterweight to centralized power. But blockchain technology is still in the embryonic stage, and we don't yet know whether it will indeed counterbalance the centralizing tendencies of AI. Remember that the Internet, too, was hyped in its early days as a libertarian panacea that would free people from all centralized systems—but is now poised to make centralized authority more powerful than ever.

#### **IV. The Transfer of Authority to Machines**

**[27]** Even if some societies remain ostensibly democratic, the increasing efficiency of algorithms will still shift more and more authority from individual humans to networked machines. We might willingly give up more and more authority over our lives because we will learn from experience to trust the algorithms more than our own feelings, eventually losing our ability to make many decisions for ourselves. Just think of the way that, within a mere two decades, billions of people have come to entrust Google's search algorithm with one of the most important tasks of all: finding relevant and trustworthy information. As we rely more on Google for answers, our ability to locate information independently diminishes. Already today, "truth" is defined by the top results of a Google search. This process has likewise affected our physical abilities, such as navigating space. People ask Google not just to find information but also to guide them around. Self-driving cars and AI physicians would represent further erosion: While these innovations would put truckers and human doctors out of work, their larger import lies in the continuing transfer of authority and responsibility to machines.

**[28]** Humans are used to thinking about life as a drama of decision making. Liberal democracy and free-market capitalism see the individual as an autonomous agent constantly making choices about the world. Works of art— be they Shakespeare plays, Jane Austen novels, or cheesy Hollywood comedies— usually revolve around the hero having to make some crucial decision. To be or not to be? To listen to my wife and kill King Duncan, or listen to my conscience and spare him? To marry Mr. Collins or Mr. Darcy? Christian and Muslim theology similarly focus on the drama of decision making, arguing that everlasting salvation depends on making the right choice.

**[29]** What will happen to this view of life as we rely on AI to make ever more decisions for us? Even now we trust Netflix to recommend movies and Spotify to pick music we'll like. But why should AI's helpfulness stop there? Every year millions of college students need to decide what to study. This is a very important and difficult decision, made under pressure from parents, friends, and professors who have varying interests and opinions. It is also influenced by students' own individual fears and fantasies, which are themselves shaped by movies, novels, and advertising campaigns. Complicating matters, a given student does not really know what it takes to succeed in a given profession, and doesn't necessarily have a realistic sense of his or her own strengths and weaknesses.

**[30]** It's not so hard to see how AI could one day make better decisions than we do about careers, and perhaps even about relationships. But once we begin to count on AI to decide what to study, where to work, and whom to date or even marry, human life will cease to be a drama of decision making, and our conception of life will need to change. Democratic elections and free markets might cease to make sense. So might most religions and works of art. Imagine Anna Karenina taking out her smartphone and asking Siri whether she should stay married to Karenin or elope with the dashing Count Vronsky. Or imagine your favorite Shakespeare play with all the crucial decisions made by a Google algorithm. Hamlet and Macbeth would have much more comfortable lives, but what kind of lives would those be? Do we have models for making sense of such lives?

**[31]** Can parliaments and political parties overcome these challenges and forestall the darker scenarios? At the current moment this does not seem likely. Technological disruption is not even a leading item on the political agenda. During the 2016 U.S.

presidential race, the main reference to disruptive technology concerned Hillary Clinton's email debacle, and despite all the talk about job loss, neither candidate directly addressed the potential impact of automation. Donald Trump warned voters that Mexicans would take their jobs, and that the U.S. should therefore build a wall on its southern border. He never warned voters that algorithms would take their jobs, nor did he suggest building a firewall around California. So what should we do?

**[32]** For starters, we need to place a much higher priority on understanding how the human mind works—particularly how our own wisdom and compassion can be cultivated. If we invest too much in AI and too little in developing the human mind, the very sophisticated artificial intelligence of computers might serve only to empower the natural stupidity of humans, and to nurture our worst (but also, perhaps, most powerful) impulses, among them greed and hatred. To avoid such an outcome, for every dollar and every minute we invest in improving AI, we would be wise to invest a dollar and a minute in exploring and developing human consciousness.

**[33]** More practically, and more immediately, if we want to prevent the concentration of all wealth and power in the hands of a small elite, we must regulate the ownership of data. In ancient times, land was the most important asset, so politics was a struggle to control land. In the modern era, machines and factories became more important than land, so political struggles focused on controlling these vital means of production. In the 21st century, data will eclipse both land and machinery as the most important asset, so politics will be a struggle to control data's flow.

**[34]** Unfortunately, we don't have much experience in regulating the ownership of data, which is inherently a far more difficult task than regulating land or machines. Data are everywhere and nowhere at the same time, they can move at the speed of light, and you can create as many copies of them as you want. Do the data collected about my DNA, my brain, and my life belong to me, or to the government, or to a corporation, or to the human collective?

**[35]** The race to accumulate data is already on, and is currently headed by giants such as Google and Facebook and, in China, Baidu and Tencent. So far, many of these companies have acted as "attention merchants"—they capture our attention by providing us with free information, services, and entertainment, and then they resell our attention to advertisers. Yet their true business isn't merely selling ads. Rather, by capturing our attention they manage to accumulate immense amounts of data about us, which are worth more than any advertising revenue. We aren't their customers—we are their product.

**[36]** Ordinary people will find it very difficult to resist this process. At present, many of us are happy to give away our most valuable asset—our personal data—in exchange for free email services and funny cat videos. But if, later on, ordinary people decide to try to block the flow of data, they are likely to have trouble doing so, especially as they may have come to rely on the network to help them make decisions, and even for their health and physical survival. Nationalization of data by governments could offer one solution; it would certainly curb the power of big corporations. But history suggests that we are not necessarily better off in the hands of overmighty governments. So we had better call upon our scientists, our philosophers, our lawyers, and even our poets to turn their attention to this big question: How do you regulate the ownership of data?

**[37]** Currently, humans risk becoming similar to domesticated animals. We have bred docile cows that produce enormous amounts of milk but are otherwise far inferior to their wild ancestors. They are less agile, less curious, and less resourceful. We are now creating tame humans who produce enormous amounts of data and function as efficient chips in a huge data-processing mechanism, but they hardly maximize their human potential. If we are not careful, we will end up with downgraded humans misusing upgraded computers to wreak havoc on themselves and on the world.

**[38]** If you find these prospects alarming—if you dislike the idea of living in a digital dictatorship or some similarly degraded form of society—then the most important contribution you can make is to find ways to prevent too much data from being concentrated in too few hands, and also find ways to keep distributed data processing more efficient than centralized data processing. These will not be easy tasks. But achieving them may be the best safeguard of democracy.