



Homo Deus: A Brief History of Tomorrow

February 22, 2017



Detail from book cover

Yuval Noah Harari

Joanne J. Myers

Introduction

JOANNE MYERS: Good morning, everyone. I'm Joanne Myers, director of Public Affairs programs, and on behalf of the Carnegie Council I'd like to thank you all for beginning your day with us.

Global Ethics Forum: Homo Deus with Yuval Noah Harari



It is a sincere pleasure to welcome the celebrated international sensation Yuval Noah Harari to this podium. He is the author of the critically acclaimed New York Times best-selling book entitled [Sapiens: A Brief History of Humankind](#). This book has been translated into 40 languages –quite an accomplishment. In this electrifying narrative, he told us about the human race, how it came to rule the planet, and how our values have been continually shifting since our earliest beginnings. *Sapiens* provided a very creative way to think about our collective past. Fans of his earlier work—and there are many—have been waiting anxiously for the sequel. Their patience has been rewarded. [Homo Deus: A Brief History of Tomorrow](#) is the title of his newest work, and it is just as riveting as it is provocative. Professor Harari does not disappoint.

Yet this time around it is not our past that he writes about, but humanity's future is what occupies our gifted speaker's mind. In *Homo Deus* Professor Harari writes about a data-centric world where computers, robots, and artificial intelligence (AI) reign supreme. In expanding our horizons he encourages us to think imaginatively about what is to come next and suggests that it is possible as we enter this new world to be in a world where a new super-caste could be created—a place where we are no longer in control, where *Homo sapiens* morph into *Homo deus*, humankind is no longer relevant, and we become obsolete. Science fiction? Not necessarily. As implied in *Homo Deus*, the journey has begun.

In expanding this next phase of our evolution, the project's dreams and nightmares that will shape the 21st century, the fundamental ethical questions are whether we have the right, or should we have the right, to alter humanity; how do we stay in control of a complex intelligence system; and how do we protect against unintended consequences?

To the question "Where do we go from here?" I thought about a quote from one of [Dr. Seuss's](#) iconic books of wisdom, which I would like to share with you. In [Oh, the Places You'll Go](#), he

wrote: "You have brains in your head. You have feet in your shoes. You can steer yourself in any direction you choose. You're on your own, and you know what you know. And you are the one who'll decide where to go."

For more on this, please join me in welcoming the quintessential thought leader of tomorrow, Professor Harari.

Thank you for joining us.

Remarks

YUVAL NOAH HARARI: Thank you. Hello, everybody. Welcome to this talk about the future of humankind, which we will try to cover in about 30 minutes and leave some time for your questions.

What I really want to discuss is the next big revolution in history. Of course there have been many different revolutions in the last thousands of years. We had revolutions in technology, in economics, in society, and in politics. But one thing remained constant for thousands, even tens of thousands, of years, and this is humanity itself.

Humanity has not changed really since the Stone Age. We are still the same people that we were in the Roman Empire, in Biblical times, or in the Stone Age; we still have the same bodies, the same brains, and the same minds as the people who painted the cave art at [Lascaux](#) and [Altamira](#).

The next revolution will change that. The really big revolution of the 21st century will not be in our tools, in our vehicles, in our society, or in our economy. The really big revolution will be in ourselves, in humanity. The main products of the 21st-century economy will be bodies, brains, and minds. We are now learning how to hack not just computers but how to hack organisms, and in particular how to hack humans; we are learning how to engineer them and how to manufacture them.

So it is very likely that within a century or two Homo sapiens as we have known it for thousands of years will disappear—not because, like in some Hollywood science fiction movie, the robots will come and kill us, but rather because we will use technology to upgrade ourselves—or at least some of us—into something different, something which is far more different from us than we are different from Neanderthals. Medicine is shifting, or will shift, from healing the sick to upgrading the healthy. You can say that the really big project of humankind will be to start manipulating or gaining control of the world inside us.

For thousands of years, humanity has been gaining control of the world outside us—gaining control of the animals, the forests, and the rivers—but with very little control of the world inside

us. We knew how to stop a river from flowing; we did not know how to stop a body from aging. We knew how to kill mosquitos if they buzzed in our ears and interrupted our sleep; we did not know how to stop buzzing thoughts in our own minds that interrupt our sleep—you go to sleep, you want to fall asleep, and suddenly a thought comes up. What to do? You don't know. It's not a mosquito that you can kill.

This will change in the 21st century. We will try to gain the same control that we had over the world outside. We will try to gain the same control over the world inside us. If this really succeeds, it will be not just the greatest revolution in history. It will actually be the greatest revolution in biology since the very beginning of life.

Life first appeared, as far as we know, on planet Earth around 4 billion years ago, and since its appearance nothing much changed in the fundamental laws of life. Yes, you had all the dinosaurs, all the mammals, and all these things, but the basic rules of the game of life did not change at all for 4 billion years. For 4 billion years you had natural selection. Everything—dinosaurs, amoebae, coconuts, and humans—evolved by natural selection, and for 4 billion years all of life was confined to the organic realm. Again, it doesn't matter if you're a Tyrannosaurus rex or a tomato; you were made of organic compounds.

In the coming century humankind may have the ability to, first of all, replace natural selection with intelligent design as the fundamental principle of the evolution of life. Not the intelligent design of some god above the clouds, but our intelligent design will be the main driving force of the evolution of life.

Second, we may gain the ability to create the first inorganic lifeforms after 4 billion years of organic evolution. If this succeeds, then this is really the greatest revolution since the very beginning of life.

Of course there are many dangers involved in this. One danger is that we'll do to the world inside us what we have done to the world outside us, which is not very nice things. Yes, over thousands of years humans have gained control of the world outside, of the animals, of the forests, and of the rivers; but they didn't really understand the complexity of the ecological system and they weren't very responsible in how they—us—behaved, which is why now the ecological system is on the brink of collapse. We used our control to completely destabilize the ecological system, to completely unbalance it, largely due to ignorance.

We are very ignorant of the world inside us also. We know very little about the complexity not just of the body and brain, but above all about the complexity of the human mind. We don't really understand how it functions and what keeps it in balance. The danger is that we will start manipulating the world inside us in such a way that will completely unbalance our internal ecological system, and we may face a kind of internal ecological disaster similar to the external ecological disaster that we face today.

Another danger is on the social level. Due to these new technologies we may end up with the most unequal society in human history, because for the first time in history it will be possible to translate economic inequality into biological inequality. For thousands of years there were differences between rich and poor and nobility and commoners, but they were just economic, legal, and political. The kings may have imagined that they are superior to everybody else, they are more capable, they are more creative, they are more courageous, whatever, and this is why they are kings and nobles. But this wasn't true. There was no real difference in ability between the king and the peasant.

In the 21st century this may change. Using the new technologies, it will be possible basically to split humankind into biological castes. Once you open such a gap it becomes almost impossible to close it.

Another related danger is that—even without bioengineering and things like that—we will see an extremely unequal society as elites and states lose their interest and lose their incentive to invest in the health, education, and welfare of the masses. The 19th and 20th centuries were the ages of the masses. The masses were the central force in politics and in society.

Almost all advanced countries, regardless of political regime, invested heavily in the health, education, and welfare of the masses. Even dictatorships like Nazi Germany or like the Soviet Union built massive systems of education, welfare, and health for the masses—hospitals, schools, paying teachers and nurses, vaccinations, sewage systems, and all that. Why did they do it? Not because [Stalin](#) and [Hitler](#) were very nice people, but because they knew that they needed the masses. Hitler and the Nazis knew perfectly well that if they wanted Germany to be a strong nation with a strong army and a strong economy, they needed millions of poor Germans to serve as soldiers in the army and as workers in the factories and offices, which is why they had a very good incentive to invest in their education and health.

But we may be leaving the age of the masses. We may be entering a new era in which the masses are just not useful for anything. They will be transformed from the working class into the useless class.

In the military it has already happened. Very often in history armies march a few decades ahead of the civilian economy. If you look at armies today, you see that the transition has already happened. In the 20th century the best armies in the world relied on recruiting millions of common people to serve as common soldiers in the army.

But today the best armies in the world rely on fewer and fewer humans, and these humans are not your ordinary common soldiers. They tend to be highly professional soldiers, all the elite special forces and super-warriors and the armies rely increasingly on sophisticated and autonomous technologies like drones, cyber warfare, and things like that. So in the military field most humans already in 2017 are useless. It is nothing to do with them; they are not needed to

build a strong army.

The same thing may happen in the civilian economy. We hear more and more talk about the danger of artificial intelligence and machine learning pushing hundreds of millions, maybe even billions, of people out of the job market. Self-driving cars that 10 years ago sounded like complete science fiction, today the only argument is whether it will take five years, 10 years, or 15 years until we see more and more self-driving vehicles on the road, and they will push all the taxi drivers, truck drivers, and bus drivers out of the job market. You won't need these jobs.

Same things may happen in other fields like in medicine. Computers and artificial intelligence are becoming better and better in competing, and even outperforming, humans in diagnosing diseases and in recommending treatment, which is what most doctors do. There will always probably be work for some doctors but maybe not for the vast majority of them.

Like in the army, you no longer need millions of GIs; you need small numbers of special forces. So maybe also in medicine you won't need millions of GPs; you will just need some elite special forces that research the latest treatments for cancer and the vast majority of the work of diagnosing people and recommending treatment is done by nonhuman doctors, by AI doctors.

Of course new jobs might appear. People say, "Okay, so you don't need truck drivers and you don't need your ordinary family physician, but there will be many new jobs, let's say in software engineering. Who will program all these new AI programs? And there will be lots of jobs designing virtual worlds and things like that." This is a possible scenario.

One problem with this scenario is that as AI becomes better and better, we have really no guarantee that even programming software is something that humans will do better than computers. The problem is not in having new jobs in 2050; the problem is having new jobs that humans do better. Just having new jobs that computers do better won't help in terms of the job market.

Another problem is that people will have to reinvent themselves again and again and again in order to stay relevant, in order to stay in the job market. This may not be easy. Think about an unemployed taxi driver or an unemployed cashier from Walmart at age 50 who loses his or her job to a new machine, new artificial intelligence. At age 50 to reinvent yourself as a software engineer is going to be very difficult.

When you look back in history, people constantly compare the threat of automation and job loss in the 21st century to what happened in the 20th century. In the 20th century you saw automation in agriculture, so lots of unemployed farmworkers moved to working in industry, and then when automation reached the industries they moved to working as cashiers at Walmart. But in those cases what happened was that people lost low-skill jobs and transferred to other low-skill jobs. Moving from being an agricultural worker to working in some car factory in Detroit you

moved from one low-skill job to another low-skill job. When you lost your job at the Detroit car factory and got a new job as a cashier at Walmart, again you moved from a low-skill job to a low-skill job.

But if the next stage means I am losing my job at 45 as a cashier at Walmart and now there is an opening as a software engineer at Google designing virtual worlds, this is going to be much more difficult than moving from the car factory to Walmart. It is very likely that, even if there are new jobs, most of the unemployed masses will not be able to make the transition.

It is also a big question about young people. Nobody really knows what the job market will be like in 20 or 30 years. It is really the first time in history that nobody has any idea what kinds of jobs and what kinds of skills people will need in 30 years, which means that we have absolutely no idea what to teach children at school. Most of what they learn is going to be irrelevant to the requirements of the job market and of society in 2050. What to teach them instead we just don't know.

The worst problem, of course, is not in the developed countries but in the developing countries. If you think about a country like, I don't know, Sweden—[which now gets a lot of attention in the United States](#)—I am not so worried about the Swedes. Even if millions of jobs are lost in Sweden, I think that, because of the tradition of the welfare state and so forth, the Swedish government will raise taxes on the big companies, and universal basic income or something like that. The Swedes will be okay, I think.

The really big question is what will happen to the Nigerians, to the Bangladeshis, to the Brazilians. If millions of textile workers in Bangladesh lose their jobs because of automation, what will they do? We are not teaching the children of Bangladesh today to be software engineers. What will they do in 20 or 30 years, and do you really think that the U.S. government will raise taxes on Google and Amazon in California and use that to pay basic income to the unemployed Bangladeshis? If you believe that, you could just as well believe that Santa Claus and the Easter Bunny will come and take care of the Bangladeshis. I don't think this is a realistic solution. Nobody knows what the solution is.

So we may be facing in the 21st century a completely new kind of inequality which we have never seen before in human history—on the one hand, the emergence of a new upgraded elite of superhumans enhanced by bioengineering and brain-computer interfaces and things like that, and on the other hand a new massive useless class, a class that has no military or economic usefulness and, therefore, also no political power.

Finally, there is the political question of authority. What we may see in the 21st century, alongside the processes I just discussed is a fundamental shift in authority from humans to algorithms.

There have been a few previous shifts in authority in history. Hundreds of years ago, say in the European Middle Ages, authority came down from the clouds, from God. You wanted to know who should rule the country or what to do, whether in terms of national policy or in terms of your personal life, authority to answer these questions came from God. So you asked God, and if you had no direct connection to God you read the [Bible](#) or you asked a priest or a rabbi who knew what God wanted, and this was the source of authority.

Then in the last two or three centuries rose a new worldview, a new ideology, called "humanism," and humanism said: "No, the source of authority is not above the clouds. The source of authority is inside you. Human feelings and human free choices, these are the ultimate source of authority. You want to know who should rule the country, you don't read the Bible, you don't ask the pope, you don't ask the chief rabbi; you go to every person, every human being, and ask, 'What do you feel? What do you think?' And based on that, we know who should rule the country."

Similarly in the economy, what's the highest authority? In the economy it's the customer—"the customer is always right." You want to know whether a car is good or not, who do you ask? You ask the customer. If the customers like it, if customers buy it, it means this is a good car. There is no higher authority than the customer in a modern humanistic economy.

It is the same thing in ethics—what's good and what's bad? So in the Middle Ages, it is what God says; it is what the Bible says. For example, if you think about homosexuality, why was it considered a terrible sin? Because God said so, because the Bible said so, and these were the highest authorities in the ethical field.

Then came humanism, which said, "No, the highest authority is human feelings, whether it makes humans feel good or bad. If two men are in love and they don't harm anybody else, both of them feel very good about their relationship, what could possibly be wrong with it? We don't care what's written in the Bible or what the pope says. We care only about human feelings."

So this was the ethical revolution of humanism, placing human feelings at the top of the ethical pyramid. This is also why the main ambition of humanist education was very different from education in the Middle Ages. In the Middle Ages the main aim of education was to teach people what God wants, what the Bible says, or what the great, wise people of the past have written.

The main aim of a humanist education is to teach people to think for themselves. You go to a humanist educational establishment—whether it's kindergarten or university—and you ask the teacher, the professor, "What do you try to teach the kids, the students?" The professor would say, "Oh, I try to teach history or economics or physics, but above all I try to teach them to think for themselves." Because this is the highest authority: What do you think? What do you feel? This is humanism, the big revolution in authority of the last two or three centuries.

And now we are on the verge of a new revolution. Authority is shifting back to the clouds, to the

Microsoft Cloud, to the Google Cloud, to the Amazon Cloud. Data and data processing is the new source of authority. "Don't listen to the Bible and don't listen to your feelings. Listen to Amazon, listen to Google; they know how you feel, they know you better than you know yourself, and they can make better decisions on your behalf than you can."

The central idea of this new worldview—which you can call "dataism" because it invests authority in data—is that given enough data, especially biometric data about a person, and given enough computing power, Google or Facebook can create an algorithm that knows you better than you know yourself.

How does it work in practice? Let's give an example so it doesn't stay too abstract. Let's say you want to buy a book—I'm in the book business, so it's very close to my heart—how do you choose which books to buy, which books to read? In the Middle Ages you go to the priest, you go to the rabbi, and they tell you, "Read the Bible. It's the best book in the world. All the answers are there. You don't need anything else."

And then comes humanism, which says, "Yes, the Bible, there are some nice chapters there, but there are many other good books in the world. Don't let anybody tell you what books to buy. You just go"—and "the customer is always right" and all that—"to a bookstore, you wander between the aisles, you take this book, you take that book, you flip, you look inside, you have some gut instinct that 'Ah, this is an interesting book,' take it and read it." You follow your own instinct and feelings.

Now you go online to the Amazon virtual bookshop, and the moment you enter an algorithm pops up: "Ah, I know you. I've been following you and following millions of people like you, and based on everything I know about your previous likes and dislikes, I recommend that you read this book. You'll find it very interesting." But this is really just the first baby step.

The next step, if there are people here who read books on Kindle, then you probably know—you should know—that as you read the book the book is reading you. For the first time in history books are reading people rather than vice versa. As you read a book on Kindle, Kindle is following you, and Kindle—which means Amazon—knows which pages you read slow, which pages you read fast, and on which page you stopped reading the book. And based on that, Amazon has quite a good idea of what you like or dislike. But it is still very primitive.

The next stage, which is technically feasible today, is to connect Kindle to face-recognition software, which already exists, and then Kindle knows when you laugh, when you cry, when you're bored, and when you're angry.

The final step, which probably will be possible in five to ten years, is to connect Kindle to biometric sensors on or inside your body which constantly monitor your blood pressure, your heart rate, your sugar level, and your brain activity. And then Kindle—which means Amazon—knows the exact emotional impact of every sentence you read in the book; you read a sentence,

what happened to your blood pressure.

This is the kind of information that Amazon could have. By the time you finish the book—let's say you read [Tolstoy's War and Peace](#)—you've forgotten most of it. But Amazon will never forget anything. By the time you finish War and Peace Amazon knows exactly who you are, what is your personality type, and how to press your emotional buttons.

And based on such information, it can not only recommend books to you; it can do far more spectacular and frightening things, like recommend to you what to study, whom to date, or whom to vote for in elections.

In order for authority to shift from you to the Amazon algorithm, the Amazon algorithm will not have to be perfect. It will just have to be better than the average human, which is not so very difficult because people make terrible mistakes in the most important decisions of their lives. You make a decision of what to study, whom to marry, or whatever, and after 10 years, "Oh no, this was such a stupid decision." So Amazon will just have to be better than that in order for people to trust it more and more, and for authority to shift from the human feelings and choices to these external algorithms that not only understand how we feel but even understand why we feel the way that we feel.

You can see all around you this process beginning to happen. For example, you want to go from here to the train station, more and more people rely on Google Maps to tell them where to go. Why? Because of empirical experience. They gave it a try. You reach a junction and your gut instincts say "turn right," but Google Maps says, "No, no, no, turn left. There is a traffic jam on the right." You follow your gut instincts and you're stuck in the traffic jam. You follow Google Maps and you reach there on time. So you learn "Better listen to Google and not to my gut instincts." By now there are many people who have absolutely no idea where they are if something happens to their smartphone and how to get anywhere because they no longer follow their gut instincts; they are just following whatever Google Maps is telling them to do.

One last comment before we open the floor for a few questions. It is very important to emphasize that nothing is really deterministic about all that. What I've outlined in this talk are not forecasts. Nobody really knows what the future will be like. They are more possibilities that we need to take into account, and we can still do something about these possibilities.

Technology is never deterministic; it gives us options. If you again look back to the 20th century, the technologies of the 20th century—trains, electricity, radio, television, and all that—could be used to create a communist dictatorship, a fascist regime, or a liberal democracy. The trains did not tell you what to do with them; electricity did not come with a political manual of what to do with it.

You have here a very famous [picture](#) taken from outer space of East Asia at night. What you see at

the bottom right corner is South Korea; what you see at the upper left corner is China; and in the middle, it's not the sea, it's North Korea. This black hole here is North Korea. Now why is North Korea dark while South Korea is so full of light? Not because the North Koreans have not encountered electricity before—they've heard of it; they have some use for it—but they chose to do with electricity very different things than the South Koreans chose to do with it. So the same people, same geography, same climate, same history, same technology, but different choices lead to such a big difference that you can actually see it from outer space.

And it will be the same with the 21st century. Bioengineering and artificial intelligence are definitely going to change our world, but we still have some options. And if you don't like some of the future possibilities that I've outlined in this talk, you can still do something about it.

Thank you.

QUESTION: Bill Raiford.

[Cambridge Analytica](#), a Texas firm, already has been engaged in the skills that you described. As it was [reported](#), they were contracted by [Jared Kushner](#) last summer, and he employed those skills, and it is reported that they provided the critical difference in the [2016 election](#). Are you familiar with that story?

YUVAL NOAH HARARI: I heard that, yes.

QUESTIONER [Bill Raiford]: Do you believe it's credible?

YUVAL NOAH HARARI: Yes, definitely there are political implications for all that.

To give another hypothetical example—I emphasize this is hypothetical; it's not a conspiracy theory; it's just a hypothesis—in principle, say in the next elections, Facebook could decide the election if it wanted to. I just came last week from Facebook. They say they have no interest in doing that and that they are not doing it, but—I am emphasizing it again—if they wanted to, they could.

As everybody knows, in order to decide the elections in the United States you basically need a little data. You need data to know which states are swinging; you need to know in the swinging states who are the swing voters; and finally, you need to know what should I say to swing the swing voters in my direction. So basically to decide the U.S. election you need to know who are the 100,000 people in Pennsylvania who still haven't made up their minds and what do I need to say to each of them—maybe different things—in order to sway them in my direction.

Who has this data? Facebook. Facebook knows—at least in the case of the people who have Facebook accounts, which is a very large proportion of the population. Facebook has exactly this data.

How does it have the data, the most valuable data in the world probably? Because we gave it to it for free in exchange for funny cat videos. We gave the most important information in the world to Facebook, which is maybe not necessarily that bad. It depends on what they are doing with the data.

It is a bit like they say—I haven't checked the facts; this may be a myth, but it does represent some actual historical events—that the entire island of Manhattan was sold by the local Indian tribe to the European merchants for a few pieces of textiles and colorful beads. It is happening again. We are selling the most valuable thing in the world—data—to a few merchants in exchange for funny cat videos.

QUESTION: My name is Larry Bridwell, and I teach international business at Pace University.

I read your book *Sapiens* a month ago, and then a couple of weeks ago I read a book about Israel, [My Promised Land: The Triumph and Tragedy of Israel](#). [Editor's note: For more on this book, check out author [Ari Shavit's](#) 2013 Carnegie Council [talk](#).]

YUVAL NOAH HARARI: That's a very good book.

QUESTIONER [Larry Bridwell]: I kept thinking, what does Professor Harari think about the kind of bleak forecast that the author has about the future of Israel with all of the historical forces which are facing Israel? So I'd be interested in your assessment of his forecast.

YUVAL NOAH HARARI: I would say that the entire Middle East is currently spiraling down, and Israel—like it or not—is part of the Middle East and is part of this spiraling down. You see many of the same phenomena that are occurring in neighboring countries of tribalism, religious extremism, and so forth also happening in Israel.

I don't know what the future may be. My skills as a forecaster are more in the field of long-term trends. It's funny in a way, but it's easier to make predictions about the really big things than about the smaller things. I can predict, I think with a lot of assurance, that artificial intelligence is going to change the job market in a fundamental way in the 21st century. This is a prediction I'm willing to stand behind. But what will be the political events in the Middle East in the next 10 years, nobody has any idea.

I would just say one thing about the Middle East in general and also about Israel, that the entire region is in danger of missing the train again. In the 19th century the [Industrial Revolution](#) was led by a few countries like Britain, Germany, and later the United States. And the Middle East and many other parts of the world missed the train, which is why they were oppressed, occupied, and exploited for more than a hundred years.

They cry about it and they cry about it, and now they are doing it again. The train is again about to leave the station. This time it's not steam engines and electricity. This time it's genetics,

artificial intelligence, and machine learning. You don't see Syria, Iraq, Libya, or Egypt getting on this train. If they miss this train, they will probably suffer even more than when they missed the Industrial Revolution of the 19th century.

QUESTION: Hi. I am Mahdi. I work in IT (information technology). I am a manager.

Professor Harari, you mentioned in your presentation that we don't have a future outside. You mentioned that maybe we should focus internally. The next step will be products like psychological products. Do you think this will lead to psychological jobs that will rise more in the future?

YUVAL NOAH HARARI: Again, there will be new jobs. The crucial question is whether humans will be able to do them better than computers. Even when it comes to things like emotions and emotional intelligence, AI is catching up with us. For example, if I want to diagnose your emotional situation, I mainly rely on external signals that you're giving me. I rely above all on vocal and visual cues. I listen to what you say—not just to the content, but even more to the tone of your voice. I look at your facial expression and at your body language, and based on that I assess what is your current emotional mood and what kind of personality you have.

It is basically recognizing patterns. You recognize a pattern of an angry person or of a happy person, and computers are getting very good at recognizing these patterns of voice and facial expression. But potentially they can do something that no human being is capable of doing, which is to recognize patterns inside the body as well.

I can't see what's happening in your brain or what your blood pressure is at present. Maybe if you blush, ah, okay, so I know something about your blood pressure. But this is very crude.

Using a biometric sensor, an AI can know exactly what happens to your blood pressure and what parts of your brain are active now. Based on that, an AI can assess your emotional condition much better, not only than the average human, but much better than the average psychologist or therapist.

There will be new jobs in the emotional market, but maybe these jobs too will go to the machines and not to the humans.

QUESTION: My name is George Below. I'm an entrepreneur, and I have thought about a lot of the same topics for quite a long time. So thank you for this.

My concern is that—let's use the military history side of it. Armies have to occupy land to ultimately stay and subject the peoples that are underneath them. That still requires sentries and all the rest of the lower-level duties for all the high tech that one may be able to achieve to have control, when it doesn't remain in control.

The issue then is: Are we going just to an overly large gated set of communities instead of having nations, instead of having anything that gives us a sense of having any responsibility for those who are not fortunate enough to have the money and have the connections and perhaps have the education – or at least the insight – to see the world in the larger sense that you've described it? Is this going to have to automatically lead to repression?

You say "no." But I look at this and I say to myself, "If you're going to keep anarchy at bay or the elements of traditionalism, religion, tribalism, and all the rest of this, in order for what you're describing to be the world that happens, there has to be a force in place to deal with those who react against." And I'm wondering how one goes about addressing that kind of question.

YUVAL NOAH HARARI: First of all, this depends on the political and ethical choices we make. We don't have to end up with a world of castes in which the upper class lives inside heavily guarded gated communities and you have oppressed masses and things. It is just a possibility.

The bad news is that it might become easier than ever to hold the restive masses in place with kind of an algorithmic occupation or an AI occupation. You don't need many humans to do that.

If you want to see it in action, you just need to go to Israel, which is now leading the world in algorithmic occupation. I think it was the [commander of the central command of Israel](#) a year ago who said – I am also Israeli – "[We are the world champions in occupation.](#)"

It is not a very well-known fact, but it is very interesting that the Israeli occupation has been transformed more and more into an algorithmic occupation. It is based on total surveillance, and it is based on more and more sophisticated algorithms that follow basically everything. The drones take footage of everything. If I pick the phone up in Ramallah and call Hebron, somebody is picking up that call and analyzing it – not a human – looking for patterns, patterns, patterns.

Every day I leave my house and go to the olive grove to water it, day after day after day after day. Then one day I do something different. [Makes sound] The eye in the sky opens. Something has broken the pattern. Something is happening here. And this is all done algorithmically.

It is very difficult to resist an occupation if you can't organize. And to organize you need to communicate, and all communication is under surveillance. It is very difficult to organize.

It is not a coincidence that the [last wave of unrest](#) was called – how to translate it into English – the "individual's intifada," or something like that, because you couldn't organize anything in a larger group than an individual. So the only resistance was like somebody wakes up in the morning, "I'm fed up with it," takes up scissors, and goes to stab a soldier. Because if you try to organize, [makes sound] they immediately know. With this, you don't need so many soldiers.

QUESTIONER [George Below]: But you're still oppressed.

YUVAL NOAH HARARI: Definitely. I said that the bad news is that with all these new tools you can occupy and oppress much more easily and cheaply than ever before.

If you think about what I discussed earlier, just think about the implications for North Korea, like everybody has to wear a biometric bracelet on the arm. You enter a room, you see the picture of [Kim Il-sung](#), and the biometric sensor monitors what happens to your blood pressure. If there is indication of anger [claps hands], that is the end of you.

QUESTION: Hi. I'm Edward Kabak. I'm in-house counsel for the Association of National Advertisers, a poet, and a public intellectual, perhaps.

Starting with the [Elon Musk comment](#) a couple of days ago about people being house cats for artificial intelligence unless they change, going beyond that, the question is: Who controls the algorithm? I'm thinking of books like [Olof Johannesson's](#) book [The Tale of the Big Computer](#), where computers 30-40 years ago wind up running mankind, and books like [Greg Egan's](#) [Permutation City](#), where people live forever in software. Who is going to wind up controlling the algorithms? Will computers wind up writing the algorithms and the humans effectively just executing them and not necessarily controlling them over a really long-term kind of view?

YUVAL NOAH HARARI: In the short term it's very likely that humans will stay in control, but in the longer terms the algorithms may become so sophisticated that humans just can't understand them, and authority will really shift, not from the masses to a small human elite, but even from the elite to the algorithms.

The whole idea of things like [DeepMind](#) and machine learning and neural network, which are now like the hottest buzzwords in the software world, is that the algorithms can evolve independently, they can learn from their experience, and they can learn, for example, to recognize patterns that make absolutely no sense to human beings.

In a way it is beginning to happen also in real life. You apply to the bank for a loan or a mortgage. In the United States, chances are your application does not go to a human; it goes to an algorithm. The algorithm analyzes all the data that it has on you, and it says "yes" or "no."

Let's say the algorithm said, "No, don't give him a loan," and you go to the bank and you ask, "Why did you refuse me a loan?" And the bank says, "We don't know. The algorithm said 'no.' And we don't understand why the algorithm said 'no' because the algorithm recognizes patterns—let's say: Who are the people who default on their loans?—that humans cannot recognize. This is the whole reason we employ the algorithm and not the humans, because they can do things we can't. But it also means that we just don't understand why the algorithm didn't give you a loan, but from our experience the algorithm tends to be right. So we just trust the algorithm without knowing why it decided the way it decided."

QUESTIONER [Edward Kabak]: Sort of like a democratic rule—the minority gets screwed, basically. That's on a primitive basis. But on a long-term basis no one will understand it except for the algorithms themselves, and that's the most frightening point; you're really run by a machine in terms of government.

QUESTION: Thank you very much. My name is Hardeep Puri. I'm a diplomat and now an author. I was fascinated by your presentation, and I'm looking forward to reading your book.

There's something that worries me, which is that at the end of the day what you say about how technology will determine our very existence is absolutely right. But I'm looking at the short term and medium term. You've got 1.3 billion people in China, 1.2 billion people in India, two countries I'm familiar with. There is a lot of technological development taking place there also. You have extreme developmental poverty issues to deal with.

This view in your book and from what you say is that you are flying at 30,000 feet and you're looking down, and you're looking at these highly successful technology-driven groups down there at ground level that are doing these things. But the interaction between that technology and the people and their lives is going to produce turmoil.

YUVAL NOAH HARARI: Yes.

QUESTIONER [Hardeep Puri]: As it is, when the [man who made Sweden famous spoke about revising the H-1B visa requirements](#), said, "We're going to raise the fees from \$65,000 to \$130,000"—I am somebody who has represented India here—I was one person who wasn't actually worried. I said, "If he does that, he's hitting the competitiveness of American industry straightaway."

Technology—unlike goods which cross borders and which are taxed at borders—can be produced. If you lift 100,000 people from Silicon Valley and put them next to Lac Léman in Geneva, they could do the same thing from there.

So I think the choices—and I am really impressed when you say it's a question of choices in the end—are not just political. The choices will be economic; the choices will be ethical; the choices will be whether democracies end up—the word you used—"disadvantaging" minorities, etc. It's going to create turmoil.

YUVAL NOAH HARARI: Yes.

QUESTIONER [Hardeep Puri]: I think at some stage we have to start looking at what those short- and medium-term things are. You talked about reinvention. Fascinating. We are reinventing ourselves, but we don't necessarily know where we are going, and that reinvention could be so costly to our very basics of existence.

YUVAL NOAH HARARI: I completely agree. We do need to think—not at some stage, but today—about the implications. I gave the example of the Bangladeshi textile workers. We have nationalist politics, but the economy is global as you say, and they depend for their living on exporting textiles to Europe or to the United States. If you have a new wave of inventions and automation and you can produce these textiles more cheaply in the United States with 3-D printers, robots, or whatever, then this means the people in Bangladesh lose their livelihood. What will they do?

We need to think about it today, not in 20 years, because if the answer is, "Oh, everybody will go and be a software engineer," we are not teaching the children in Bangladesh to be software engineers. So how can they be software engineers in 20 or 30 years?

I don't have a solution. But I think that yes, we need to think about these things now, and especially about the social and political implications, not about the technical issues. Most people who are leading this revolution are engineers, technicians, and scientists who specialize in things like computer science or biology, but they have no background, and sometimes no interest, in things like sociology, anthropology, and philosophy.

The problem is that the people in the political system, and the people also in the humanities and in the universities, often have no understanding, and also sometimes no interest, in what's happening in the computer science department. That is extremely dangerous.

I was following the [election campaign in the United States in 2016](#), and I saw Donald Trump scaring people that "The Mexicans will take your jobs, the Chinese will take your jobs." And he never said, "The robots will take your jobs," which potentially is just as scary. Why didn't he say it? It's part of the debate in Silicon Valley, it's part of the debate in the university, but it's still not part of the public debate.

So even if it's not true that the robots will take your jobs, it's still a very scary thing to say. So why not say it? Maybe because then you need to build a wall on the border with California, not on the border with Mexico. But it scared me that this is not yet part of the mainstream political debate because we don't have much time.

JOANNE MYERS: Yuval, it was just a fascinating morning. I have to thank you very much, and I encourage you all to buy his book. It's available for you at the back of the room. Thank you so much.

Copyright © 2017 Carnegie Council for Ethics in International Affairs. All Rights Reserved.