Greater Good Gathering - Panel III

Consciousness: Effects on Thought and Thinking

Moderator
ESTER R. FUCHS

Panelists
JEANNETTE M. WING
ADRIAN WARD
CHRIS BAIL
JUDITH SHULEVITZ

The Greater Good Gathering held on February 6–7, 2019 explored the future of public policy and how best to advance the greater good in the 21st century in light of technological innovation, economic disruption, ideological polarization, and governance challenges. Event co-sponsors included Columbia University School of International and Public Affairs, Columbia Law School, Union Theological Seminary, The Academy of Political Science, and Public Works LLC/Brain Storm Consulting.

ESTER R. FUCHS: It's my pleasure to welcome you back to the greater good gathering 2.0. Before we start our panel this morning, I want to thank a couple of people for making this possible. First of all, Eric Schnurer who really is the originator of the whole idea of the greater good gathering. We are delighted to be co-sponsors here at Columbia, SIPA.

We've had a lot of help with this conference from the Academy of Political Science, Bob Shapiro playing two roles—president of the Academy of Political Science and professor at SIPA.

The Columbia Law School is also a sponsor, Public Works, Brainstorm Consulting and the Union Theological Seminary. And we want to thank them all for contributing to this really exciting conference. I also want to thank some of the people who work so hard behind the scenes to make this happen.

Kevin and Ashley who work with us at SIPA, without whom, this really wouldn't have happened at all. Loren, Caroline and Lacy who work at the Academy and with Public Works, all of whom really did an extraordinary effort to put these events together. Some of you who have done these kinds of things before know how difficult it is and I think they've done an extraordinary job. Thank you all.

It's my pleasure to be moderating our first panel of the morning. I'm Esther Fuchs. I'm a professor here at SIPA in the Department of Political Science. The title of our panel is Consciousness: Technologies Effects on Thought and Thinking. Not exactly a small topic but as you will see, we've got the right panel to have this discussion.

Now, just briefly, we all recognize the pervasiveness of technology in our personal relations, at our jobs and in our political life and whether it's family dinners where everyone immediately takes out their smartphones and puts them on the table, and it makes you wonder, "Do they think that little black thing is more interesting than I am?" And at work when we find

ourselves buried in email or start thinking about whether the new technology will actually someday replace us completely. Or in political life, as yesterday's panel explained, algorithms are being used to micro-target voters during elections, judges are using them to decide bail or prison for offenders who are awaiting trial. And of course, in my own field, smart cities are going to solve every urban problem from mass transit to poverty. I've yet to understand what a smart city really is because none of us really want to live in stupid cities, do we?

Most of us actually don't really understand how technology works especially how it impacts our cognition and thinking. When you start thinking about it, just you yourself as a simple human being, how can we really understand whether technology is ruining the quality of our life or producing a future for us all that we look forward to, if we don't understand how it affects us in the most fundamental ways?

It's fortunate for everyone here that our panel this morning can help us answer some of these critical questions. I'm really delighted to be doing this introduction with the perfect person, Jeanette M. Wing. Jeanette is the Avanessians Director of the Data Science Institute here at Columbia University and a professor of Computer Science. Before she came here, she was the corporate vice president for Microsoft Research and of course has had a distinguished academic career. We're very fortunate that she came to lead our Data Science Institute.

As a computer scientist and a data scientist, you've seen the evolution and explosion of the fields of artificial intelligence. So I'm going to ask you the simple question of the day, can you explain to us how AI impacts our human cognition and why maybe we should be worrying about this?

JEANETTE M. WING: Thank you. I am delighted to be here. I wanted to address this question of evolution and explosion of artificial intelligence and answer what can humans do that computers cannot? What can computers do that humans cannot and what can they do together that either cannot do alone? I think this will address some of the concerns and the capabilities of artificial intelligence as we know it.

Before I answer those questions, I wanted to give you a little history of how artificial intelligence has come about and, especially the recent few years, of why AI has exploded in terms of its prevalence and its applicability to everything. I am not going to go back to the dark ages, if you will. I'm just going to go back to 1956 which is when there was a conference at Dartmouth and basically. Those who we called the grandfathers of AI convened and where the term artificial intelligence was first coined by John McCarthy. In fact, the proposal that he and others wrote for holding this summer research institute was to study every aspect of learning, or any other feature of intelligence, that can be in principle so precisely described for a machine to simulate. In other words, the goal was to build a machine that could essentially simulate intelligence. Let's just say human intelligence. But early on, even at this conference, everyone realized that this was too big a problem to solve all at once.

The field very early on divided the tasks that humans are good at, like vision, speech, language, mobility and so forth, into separate buckets. Then the scientific community tackled each of these tasks to the point where each of these subtasks that humans are very good at became their own subfields of computer science. You had computer vision, computer speech, natural language processing, and machine translation. Each became their own complete little communities for decades until around the late 1990s, early 2000s, when a technique called machine learning really came to the fore. The reason that machine learning came to the fore then is not because it was a new technique, it's because the compute power was sufficient to actually make those algorithms efficient and practical.

And so what happened is that you saw the computer vision community using machine learning to do better computer vision. You saw the computer speech recognition community

using machine learning to do better speech recognition. You saw the machine translation people using machine learning to do better machine translation. You saw the robotics community using machine learning to do better robotics. And then everyone thought that maybe we can take this common method called machine learning to address the original 1956 vision of the grandfathers of AI and actually build a machine that puts it all together. This caused a huge resurgence of interest in AI.

I think the major breakthrough recently was around 2011 when the technique called neural networks showed how for the first time lots and lots of data could actually do tasks that were as good and eventually better than humans. Technology that for instance recognizes faces is now as good or better than humans. The computer program that plays Go is better than humans. And so everyone is scared thinking AI is going to better us in terms of what we can do.

On the one hand, there are certain tasks that are very well defined. The rules of the game are understood and everyone plays by those rules. That's like a game. And yes, you can make this computer program do better than a human. On the other hand, there are still a lot of things that computers are not good at no matter how much AI you throw at it because the learning is not there yet.

The canonical example is common sense reasoning. Little kids know how to do common-sense reasoning. We don't know how to get a machine to do common-sense reasoning. All of us who are rational and have some common sense. Our thought processes are not going to be easily taken over by AI yet.

The other thing that we are very good at, but very hard to get machines to do, is fine motor manipulation. For example, picking up a cup and knowing just how much pressure to put. Determining what grip to use and how much pressure to put on a paper cup versus a heavier plastic bottle is easy for us. Even little kids can do that. It's very difficult to get those robots to do that. We're making progress. A lot of the progress is made due to feeding lots and lots of data to these algorithms so that the robots can figure out to grasp and not crush the bottle or paper cup.

That gives you an idea of where the capability is today—of what machines can do over humans and what humans are still better at. How about when you put the two together? And this is what's very exciting in the computer science and artificial intelligence and data science community. When you put the intelligence of machines together with the intelligence of humans, can we actually do better than either alone?

One example that really impressed the community is, believe it or not, playing chess. There's a competition in the chess world where computers and machines together form a team. They play against other teams of computers and humans. These teams also have played against just machines and just humans. They showed that the combination of humans and machines outperformed humans alone and machines alone. In the competition you have to make your moves by a certain time. It turns out that human intuition was critical for saying, "I don't want to wait for the machine to spit out the answer, we're just gonna make this move." It's nice for us humans to know that there's something to human intuition. From a scientist's point of view, it's a puzzle. We don't understand.

A more concrete example is with big data. A lot of the deep learning algorithms and models are best when they're fed with data that is labeled. And today, who does that labeling? Usually, it's still people. Whether humans, Mechanical Turk, crowdsourcing or whatnot, there's still a lot of human input to making those machine deep learning models effective in the end. That will change over time. We will automate a lot of this. But right now, there's still an important boring task the humans play in making machine learning models of today effective.

FUCHS: Thank you, Jeanette. See my build up was really right on the mark. I think I just learned more in this brief presentation that you made about cognition and computer science than I did from reading several academic articles which I personally did not really understand. This is an extraordinary way to start. We're going into some applications and thinking about this.

Our next panelist is Adrian Ward. Adrian is an assistant professor of marketing. He's at the University of Texas at Austin. He studies technology, cognition, and how we can be better to those around us, which I thought was a very interesting description of your research. So, most people don't realize that they have relationships with their smartphone. We know we have smartphones but this idea of having relationships with technology is something we don't really think about much. We think about relationships with other human beings. What I learned from looking at your work briefly is that this can affect our knowledge, our beliefs, and even the decisions we make about other people. Could you explain how these relationships develop? How does it impact us in our day to day life in ways of course that most of us really don't think about?

ADRIAN WARD: I'm a professor of marketing but I'm a trained research psychologist. I'm a social cognitive psychologist. I do experiments to try to figure out how technology is affecting us. A lot of my work is about memory and attention, but in this sort of social context.

I want to start by not talking about that, but by talking about animal behavior. A metaphor that I think is really helpful for thinking about this sort of thing is that of a super normal stimulus. Something that did not exist in the environment when an organism was evolving, but now exists and surpasses anything available at that time. My favorite example of this is of gulls. Gulls have eggs and you've got to incubate those eggs by sitting on them. They decide which egg to sit on, which one to preferentially pay attention to, by just picking the biggest egg. It is a pretty good rule because the bigger eggs tend to be more viable. However, if a scientist shows up and paints a football white and puts it next to the gull, the gull is going to sit on the football and ignore all of its own eggs.

This makes sense because there's a biological limit on how big of an egg a gull can produce. So they just need to know bigger is better. They don't need to know what the best thing is or what the right thing is. The selection pressure can be manipulated when somebody comes in and gives you something bigger than what could ever exist in your world. Gulls don't know how to deal with that because they never had to deal with that. They were evolving this rule of bigger is better.

Luckily there aren't a lot of trickster scientists running around painting footballs white, but we are a lot like gulls. We have the same easy solutions to complex problems and those solutions evolved in a context where we didn't have a lot of these technologies. So this isn't just about computers or smartphones.

People have made the same argument about sugary and fatty foods. The reason that we are very attracted to sugary and fatty foods is they used to be pretty rare in the environment and calorie dense. If you could get them, you should get as much as you possibly could. But now we can just create those foods and we have an obesity epidemic.

If we're thinking about this in the context of computers, technologies, and smartphones, one of the things that they're really good at, and that we're really bad at, has to do with just how much information they can hold. One of the most fundamental constraints we have as humans is that we have limited cognitive capacity. We have limited cognitive resources. We can't pay attention to everything in the world. We can't notice everything. We can't know everything.

Lots of times it would be good to know stuff. We've never been able to know everything, but we've figured out ways to solve these problems. A very easy way to do that is to store information externally—whether it's in a book or a notepad. One source we use a lot is other people.

So, I don't need to know how to fix a car if my mom's a car expert. I have her phone number. We've developed what are called transactive memory networks where we know who knows things and that's pretty much just as good as knowing it. But, it's only as good as knowing if they're going to give you that information.

There are natural constraints on that network as well. If you don't know anything, eventually people are going to kick you out of that memory system because you're a free rider. There are things that force us to know things as well. If we think about who we should allow to know things on our behalf, you want somebody who has access to that information or knows that information and is going to give that information to you. We distribute information among our social networks.

Now, introduce computers in the internet into this same transactive memory system and it turns out that they surpass any human on these same dimensions. They know more than anybody else. It's instantly accessible. They are not going to get mad at you, right. They're not going on vacation. You can always get this information. What we see more and more is that people will offload a responsibility for information and memory onto the internet, not just at the expense of other people, but at the expense of their own memories. We no longer have this decision of whether I should know this myself because we have this sense that everything is saved.

I've done research showing this in learning situations. If you tell people this is on YouTube, they just don't remember it as well—even if you tell them you're going take the video down. We have this sense that we can find it if it's online and nothing ever actually disappears. If something is live, people will actually remember it a little bit better. Even with factual information, we do this financial decision making. Once you know that it's being saved specifically by a computer, you don't remember it even if we pay you to remember as much as possible. That's because there's this very basic human constraint. We have limited cognitive resources. We're trying to do the best that we can with our limited resources in a very demanding world.

So that's how we can think of people. We are worried that we're not remembering things. That's a well-founded worry. The other slice of this, which I think is also really interesting, is that there's a difference in retrieving information from a computer or a smartphone than from a person. If I have to go find the right person, I've got to track them down. Maybe they don't know that information. It's a very physical thing. I might learn something, but I know that I'm not the expert. I might become an expert over time, but I can still know that the information came from them.

In a lot of my research we find that when people Google information, they think that they remembered it themselves and they think that they thought of it themselves. A lot of this is because the Internet's faster than our own internal search processes. If you're asking Google a question, you're asking it to yourself. Before you realize you don't have that answer, Google gives you the answer and you say, "Oh, okay. I actually always knew that."

Even beyond that, it's a seamless interface. I don't know if anybody remembers Lycos. For a while, Lycos and Google had almost identical search results, but nobody used Lycos. So, I did an experiment where I had people either use a physical card catalogue or Lycos and Google. Then I just asked them a bunch of questions. Afterwards, I didn't ask whether or not they thought they knew the answers already. I asked whether you answered this yourself or used an external source? People who used Google or Lycos didn't even remember using Google or Lycos—an external source. When they used the external card catalog, they remembered looking up the answer. The process of looking up the answer reinforced that this was an external information source.

More and more we're offloading information to the Internet, but thinking that we actually know it ourselves. We can think about this at a cognitive level. We can also think about it at a social level. It's not very romantic, but one of the reasons we need people is because they hold

information for us. They allow us to know things and allow us to navigate the world. But if we don't need people for those reasons, that's one less reason to have social connections.

So another really fundamental constraint we have is not just on knowledge but on attention. There's a lot of stuff going on in this room right now but you're not noticing all of it. Hopefully you're noticing what's happening right here, but attention is selective. You can pay attention to some things and not others. A common problem we have is going into a new environment, having to figure out what's important. How do we direct our limited attention? And it uses some cognitive resource to do that sort of selection process.

Things that always tend to be relevant, we automatically attend to. Things like the sound of your own name. There's this thing called the cocktail party effect. If you're in a room, there are a lot of conversations happening. You filter out the things that are not relevant to you pretty automatically. If someone on the other side of the room says your name, it pops through that clutter. At a very basic cognitive level—you're not intentionally looking out for it—it's finding you because that's relevant information. Similarly, research shows that new mothers, if a baby is crying, they hear that. Research on new fathers is more mixed but you know that's something important which you should pay attention to. You shouldn't have to decide whether or not you need to pay attention to that. What we're seeing more and more as we put basically everything important and rewarding and satisfying into one device—the smartphones which you carry around everywhere—we automatically pay attention to our phones. There's even neuroscience research showing that the sound of your ringtone activates the same area in your brain as the sound of your own name. Not somebody else's ringing tone, but your ringtone. This is a very self-relevant thing.

We did some research to look at how just having your own smartphone around would affect the amount of cognitive resources you have available for other tasks. So this isn't multitasking. By the way multitasking is a myth. It's not good for you, don't do it. This is just having it around in your presence and not using it for the task. What we found is that the closer people's phones were to them—on the desk in front of them versus in their pocket versus in another room—the less available cognitive capacity they had for other tasks. That's because your brain is saying pay attention to this thing. This is your email, this is a cat video, this is social media, this is your calendar, this is everything. Overriding that automatic urge, in this context of trying to navigate a complicated environment, uses up some of your cognitive resource—leaving less available for the task at hand.

Even when you're doing a really good job at not being distracted, you have to realize that you're doing a job to not be distracted. By carrying around these sort of distracting items with us everywhere we go, we're constantly putting ourselves in this space of attentional tension where we have to override this automatic urge. So a lot of the stuff I do is thinking about how technology is interacting with our own limits and how we're producing unexpected outcomes, but from the same sort of cognitive systems that we've been using all along. Yeah, gulls sitting on white footballs.

FUCHS: Thank you. There's so many questions that I could now ask both you and Jeanette but we don't have a lot of time, so we're going to wait till all of our panelists get a chance to speak. I want to introduce our next panelist, Chris Bail. Chris is the Professor of Sociology and Public Policy and the Director of the Polarization Lab at Duke University where he helped create interdisciplinary data science program. I initially tried to ask Chris a really hard question that nobody really has answered yet, and fortunately for him, I sent it to him first.

I've changed it up to be more reasonable but I might come back to that really hard question, fair warning. Many people who are concerned about social media echo chambers and we heard a little bit about that last night in the panel moderated by Jim Fallows. Politics has a big

problem when it comes to echo chambers—the tendency for people to create social media networks that only expose them to views or opinions that they already agree with. This is not just an issue in politics, it's an issue in every aspect of our life.

I know your research is around this area and I'm hoping that this morning you can explain to us how these echo chambers shape people's political attitudes and their ability to empathize with those who have opposing views. It's a critically important question for us in the United States, in local communities, and globally.

CHRIS BAIL: Thank you. I might begin with something that Jeanette touched on, that humans and machines working together have come to improve the way we can classify things. That's one of the positive developments in AI. The negative developments have to do with embedding human bias in algorithms. Sometimes this is called human-in-the-loop dynamics. If we ourselves have some kind of bias in categorizing something, then a machine will reproduce that. And we see this all over the place from decisions about parole which increasingly are being made by algorithms, to how we build our social media networks which, as Esther mentioned, is my area of research.

It's been well-documented now that on social media sites like Facebook or Twitter, people tend to follow others who already share their views. This has the effect of us becoming repeatedly exposed to information that already conforms to our pre-existing views. The upshot of this—something that's implicit in a lot of what policymakers, academics and public intellectuals say—is if we could somehow break our social media echo chambers, if we could step outside them and listen to the other side, there would be more moderation.

We would be able to empathize with each other. We would be able to understand each other more effectively. On the other hand, if you spend any time on Twitter, you'll know that it's not exactly the paragon of the civil sphere that Rousseau might have been imagining. There's this other stream of research which has documented so-called backfire effects. So, when I'm exposed to something that conflicts with my view, I'm reminded of all the reasons I don't like that view, and in the end, I'm left with more reasons to disagree with that counter attitudinal message than I might have had had I not seen it in the first place.

The challenge for many years for social scientists has been to try to figure out how to do research on this issue. We can do it in the lab and plenty of good work can be done, but lab research has its limits. On the other hand, this brilliant new field called computational social science has identified a way of embedding experiments within social media platforms themselves. Some of these are more controversial than others. Everybody here probably knows about Cambridge Analytica and the so-called emotional contagion experiments on Facebook. But when done ethically and with careful human consent, there's a real opportunity to see what happens in the real world when we disrupt social echo chambers.

In the Polarization Lab that I lead at Duke, we recruited about 1600 Democrats and Republicans who use Twitter regularly. We paid half of them to follow bots our team built that retweeted messages from opinion leaders from the other side. So if you're a Republican, for one month you saw 24 messages a day from Democrats—elected officials, journalists, media organizations, advocacy groups—and vice versa for Democrats. They were exposed to Republicans.

During this time, about each day, someone retweeted cute cat pictures. Then at the end of each week we asked participants whether they could identify the cute cat pictures. We not only effectively disrupted people's social media echo chambers, but had some sense of who is paying attention, by testing who was able to recognize the animal pictures. So, what did we find?

Going in, we had hoped as a polarization lab that we might decrease polarization. We saw in no case a decrease in political polarization. Instead, we saw the Democrats become

slightly more liberal after following Republican thought for a month, and Republicans becoming substantially more conservative after following Democratic thought for one month. This was on the order of the change in public opinion over the last ten years, just through one month of exposure.

So what does this mean? Getting back to the practical upshots of this and how we're interfacing with technology, *The Washington Post* recently reported that Twitter has been experimenting with exposing its users to opposing views, and Facebook has done this as well. It's a very logical solution to try to counteract the tendency of people to reinforce their view. Our research suggests that is not going to be a panacea. We're not going to be able to simply shift the algorithm and move the needle.

Over time? Have we interviewed them? It hasn't been a year yet, but now you're giving me a great next project, right. We haven't looked at the persistence of it. When I describe the study, some people say: "You've confirmed what I already knew. You know Twitter is bad for democracy. Facebook is bad for democracy." I don't necessarily disagree in some respects, but I also think it's naive for us to say, "Let's all get off social media." At least for some of us older folks, maybe that's reasonable. But is anybody under 30, as Adrian has so eloquently documented, going to give up their phone? So what can we do? We have to tweak and probably not replace. What types of things might work?

We're currently pursuing follow-up studies. One that exposes people to moderates from the other side instead of the whole range of views. It could be that there's a curvilinear effect. If I'm a strong Republican and I'm exposed to a strong Democrat, that has a backfire effect. But if I'm exposed to a moderate Democrat, maybe that has a moderating effect.

Second, we can work to correct misperceptions about the other side. If you ask Democrats who Republicans are, the answer you typically get is that they are older, evangelical Christians who are wealthy. And if you ask Republicans who Democrats are, most will tell you that Democrats are younger, minorities and very progressive. It turns out that the average Democrat is a middle-aged Christian from the Midwest. The average Republican is also a middle-aged Christian from the Midwest. So some research shows that just simply correcting these myths and misperceptions can initiate change.

Lastly, to get back to something that Adrian touched on earlier. Some of the really interesting work on correcting these backfire effects suggests that if you can make someone think that the opinion was their own, then the more likely they are to embrace it. The people who had agreed with the misinformation, told the researchers one week later that they had actually disagreed with the misinformation. If people are led to believe that they themselves judged the information as true or false, then they're much more likely to endorse that kind of rejection.

There's something about being force-fed information. That's really the environment we inhabit on social media. It's just a constant fire hose of information. What if we can give some autonomy to social media users? I've been working with Twitter to develop tools to do this—where we can allow people to monitor their exposure to the other side. Allow them to monitor when there is disagreement about something they're seeing and make their own choice about whether to change their mind.

FUCHS: Thank you very much. I have so many questions and I'm holding off. The research is really fascinating. We are extremely fortunate today to have with us Judith Shulevitz, a *New York Times* contributing Op-Ed writer who covers feminism, culture, and science among those small issues we're all dealing with today. She formerly worked at Slate in its early years, which is when I got to meet her. I have followed her work ever since.

My favorite book that she wrote has a lot to do with this conversation, called *The Sabbath World: Glimpses of a Different Order of Time*. I have a complex question for you Judith, because

you're someone who thinks and writes about some of the really most important transformations that we're experiencing in the 21st century. You understand the upheavals that technology has created in a lot of different spheres. Some of the research that you've done involves affective computing. She explains it as a way of undertaking to program artificial, emotional intelligence into machines. You've written on smart appliances that we're anthropomorphizing. It's an extraordinary thing and I love how you talk about that. Things like empathetic refrigerators. Could you please explain to us what this really means for us personally in our social relations? If you want to jump into the politics—please go right ahead.

JUDITH SHULEVITZ: Well, thank you Esther. I want to just note something that just happened on this panel and then move into my cultural critic mode by taking as my objective research—the Super Bowl. One theme that both Jeanette and Adrian were talking about is of course, human obsolescence. It was fascinating to me how very slender the hooks Jeanette hung: human pertinence on fine motor skills and common sense reasoning. You're really pulling it back here and you were talking about the obsolescence of the human as a form of expertise—other people as a form of expertise.

How many people here watch the Super Bowl? I don't know if there's a big Super Bowl crowd. Okay, bigger than I expected. I personally don't understand football. I watch the ads. What I noticed was a common theme in the ads. The common theme is tremendous anxiety about computer robots taking our jobs. But not just taking our jobs, although there were several that were about robots taking our jobs. There were two that really struck me. One was an ad from Michelob beer. It started with people who were jogging. Then the robot came along and jogged faster than they did. Same with bicycling. Same with golfing. Then all the humans headed down to the bar. The robot did too, but stood outside looking through the window. The tag line was, "But it can't enjoy hanging out with the gang at a bar with a Michelob, so don't worry, drink up."

The second one that struck me was for Turbo Tax. It featured a sort of uncanny looking child-like robot—a very robotic thing with a human face—who said to his creator or father that he wanted to grow up to be an accountant like his dad. And a woman who was there, who wasn't his mother but was being shown this neat new device, said patronizingly: "All TurboTax live CPAs are human beings with real emotions. I'm sorry but you're never going to be emotionally complex enough for that job." And the robot child says, "Is that true papa?" And the father responds, "Yeah." And the robot child says, "I am sad." Then the woman said, "See what I mean?" And the father said, "He's still perfecting emotions."

What is affective computing? Affective computing is perfecting emotions for the robot child. I just have a funny story to tell about affective computing that I can't help telling. It was invented in the late 1990s by an MIT engineer named Rosalind Picard, who at the time was one of the few female engineers at MIT. She became interested in the question of how machines make decisions. She began reading Antonio Damasio and other psychologists of emotion, and it started soaking in at that point, a consensus that emotions are intricately and intimately linked to the decision-making process.

She decided she would try to program emotions into machine learning and machine decision-making or robot computer decision-making. But being a woman, she didn't want to call it emotional computing because you know—women and emotions and MIT. That was going to hurt her, so she called it affective computing which nobody understands. It is not actually in its infancy like we think it is. There is a part of it that is really well developed, has commercial applications, and is coming soon to a high-end car near you. That is the part that involves detecting emotions. You're all familiar with facial recognition, but you're possibly somewhat less familiar with the analysis of vocal intonation to identify people.

All these things can also be in body language, of course. All these things can be used to extract from this information a conclusion about an individual's emotional state. This is really sophisticated and they're really good at it. It's already being used in a lot of domains. A good example of a commercial application is in cars. Coming soon to a high-end car near you are cameras and recorders that can observe and hear you, and decide whether or not you're tired or distracted or overwrought. It can recommend that you pull off the road and calm down. Pull off the road and get a cup of coffee, and so on and so forth.

The next iteration is going to be what are called semi-autonomous or semi self-driving cars. A great concern in the automobile industry is what is called "the handoff." How does the car decide when it is okay to yield the wheel to a human? When does the car decide it better take the wheel back from the human? They will be using this emotion detection technology.

There are lots of applications. As you can imagine, Google and Amazon and Apple are all trying to imbue their digital assistants with this kind of analytical ability—mostly focused on the voice. That's the emotion detection side and the other side is emotion synthesis. That is, using this information to program robots or computers or personal assistants to simulate or reproduce appropriate emotions, usually in response to the emotions they detect. That is in it's early days, but it's advancing rapidly.

There's a lot of work being done on that at the MIT Media Lab. Google has something called an Empathy Lab. This to me is interesting in the sense that soon Alexa will be able to figure out what mood you're in, will be able to respond appropriately, and will be able to solidify bonds between you and your device.

We see the device. We hear the device. It says something vaguely approximate to an emotionally appropriate statement that a human would make, and we respond to it. There's a lot of information about people who confess their emotions to their Alexa, even though Alexa will then completely misinterpret by saying, "Oh, you'd like to watch a video of . . ."

People say: "I'm lonely." "I'm depressed." "I want to commit suicide." A lot of that. "Alexa, will you marry me?" This is to a device that is just really deeply stupid. Once they start to become less stupid—that's going to happen more and more—they'll be able to solidify the human-computer bond. That will give them a great deal of power persuasion and influence over our behavior.

It's also not just your Alexa or your Google home, it's a little assistant. It's in the internet of things like your smart refrigerator, your smart toilet. These are things that are going to be talking to you and are going to be programmed to detect your emotions and one day be able to respond appropriately.

Now, what does this mean for human obsolescence? I was fascinated to discover, at an affective computing conference held only a few weeks ago, a session on upskilling the human workforce. The argument being made was that manufacturing is such a small sector of our economy now, and service is such a growing sector of our economy. We can use affectively sophisticated robots and computers to train humans to be better at their service jobs. To train emotions to overcome their inability to be empathetic or place themselves in the minds of their customers. It was never mentioned in this discussion the fact that if robots and computers are good enough to train humans, what are the humans there for? Why bother?

There's already therapy being done with these devices. Mostly where there are avatars for humans. They're being controlled behind the curtain. But soon you will be having therapy with a robot and possibly an artificially intelligent entity on a screen. I think it does go back to the point of human obsolescence. The Super Bowl ads wanted to tell you, we have emotions and they don't. Don't worry. And I'm here to tell you, worry.

FUCHS: Thank you so much. I want to go one more round with everybody and give them a chance to comment on each other's presentations, and then we'll open it up briefly for questions.

I was hoping for a little more optimism here about the relationships that are being developed between technology and humans and the future it might present for us. Should we be thinking about turning this clock back in some way? And if you think we can't, clearly there is a role somewhere in this morass for humans.

Even the example of the courts that you gave, somebody's programming and deciding what the algorithm is. If we dig deep back enough into the process at least with AI, there is a human component to that. To what extent are we not talking about that and not recognizing that. To what extent are we not bringing that directly into our values discussion about what we want to see our society and our policy look like in this twenty-first century technological world that we are creating?

BAIL: I think the answer to the question really hinges on how much bad are we willing to accept alongside the good. In the case of parole decisions for example, there's wonderful work by the computer scientist John Kleinberg and the economist Sendhil Mullainatha which shows that when we use algorithms to make parole decisions, we actually let people out earlier who are going to do well outside of jail. Everybody gets out earlier. The problem comes in when we look at the differential racial outcomes. We're actually reducing the number of African-American men who are incarcerated, but we're doing that at a rate that's lower than their Caucasian counterparts. This is one of these utilitarian puzzles that I think is going to come with it. A computer can optimize for that. But as long as race is part of that calculation, that's always going to be a problem.

Another example, equally vexing, is health care. If we look at the delivery of health care—if we use machine learning to optimize decision-making or say testing for cancers—some of the latest research shows that we can actually make huge gains for lower income people. But those come at the expense of upper income people who will be monitored less than they are right now. These types of trade-offs I think are in our future. The question is going to be—are the people who stand to gain willing to give up their privilege?

FUCHS: But are we also going to pretend to delegate those decisions to computers and not take responsibility as a society for the values that are represented in these decisions.

WARD: I really think that it comes down to what is the point of being alive and being a human? If we are disposed to offload all of these tasks that we have to do to navigate the world—but that we don't have to do anymore—the research shows that we're not going to do it. So you could ask, what's the point of knowing anything? Is there a point? I don't know.

What is the point of having control over your attention? The reason we pay attention to things is because at some level, they're rewarding. Although people talk a lot about dopamine and say it's a reward thing, it's really motivation. But if you think of the business model of the major tech companies, it's all about capturing your attention. That is how they prove their value. Some are not financially valuable yet. They prove that they could be one day because they can sell you ads.

Is that really what we want to do with our most limited resource, which is attention and it's time. You can say is that valuable? But then you get to the question of what is valuable? And you could make an argument that if people are perfectly happy to look at screens for the rest of their lives—and they're not dying and it's not hurting anybody else—then maybe that's fine. I don't know that I want to make that argument, but I think it really does come to the question of

what is the point of our time and our attention. Do we want to have control over it and how do we want to spend it?

FUCHS: I'm going to move it to Judith.

SHULEVITZ: I have two answers to your question. One is I want to talk a little bit about the good that affective computing, or at least emotion detection, can do in the world. Then I want to talk about my theory of how we can bring values into it. Something I didn't talk about is that I just wrote an essay for *The Atlantic* on voice and how voice affects us. I said the robot revolution is here, it's just happening in voices. The embodied robots are a decade away or so, but we are surrounded by voices and voices affect us differently.

One of the things I learned in the course of this research is that there are amazing medical applications to voice analysis. There is a guy named Yuval Mor in Israel who has perfected techniques of analyzing voice to detect incipient Parkinson's disease, incipient cardiovascular disease, or diabetes. I don't understand how this works because I am not a computer scientist, but it's starting to have applications.

Another one is in a doctor's visit. There is a company that is focusing on applications for doctors and hospitals. A program exists in the cloud that can analyze the interaction between doctor and patient. Or recognize some emotional content—possibly things like hyper anxiety and depression—and signal the doctor during the course of the visit that maybe some questions should be asked about depression or something like that. That's a little invasive and creepy, but it is also Rosalind Picard at MIT doing all kinds of work on using this kind of facial and voice analysis to help autistic people respond in a socially appropriate way, and expand the social circles in which they can move and function.

All of this does raise the specter of the obsolescence of medical professionals, but that's a whole other thing. But how do we bring values back into this equation—regulation. All of this has to be regulated. We have a completely stunted regulatory apparatus, if it exists at all for these things. And we're going to need a regulation revolution before we can live with artificial intelligence in a way that maintains human dignity. Imagine how good China is at facial recognition. We've all read about that. Now imagine that they can analyze the faces they recognize and detect emotions. They can detect intentions or put emotions that they can algorithmically link to possible intentions.

That is a terrifying invasion of privacy. That is something we could, however, regulate theoretically. I haven't seen it going on much, but you're the political scientists, maybe you can tell me how far we could get with that.

FUCHS: I'll make a quick comment before we get to Jeanette because I think this is a critical piece for another panel, which is the role of government and the public sector. To a certain extent, in dealing with these transformations that are happening in science and that are being monetized through business ventures, there is little accountability to a broader public body, which historically we've called government. The more government weakens, the more it declines. The more it has no capacity to understand the conversation we've had today or to interject the values that society agrees upon, the less this is going to be a force for the good.

I agree with what you're saying and you're right about the need to think about purpose and meaning in the context of all this. We can do it personally or we can do it as a society, and those could produce very different outcomes for us. I'm going to let Jeannette have the last word on this set of questions.

WING: When Adrian was talking, I think he said what is the meaning? I was thinking even deeper or more philosophically, what is the purpose of humankind? What is our role in the universe? Judith is absolutely right, the progress, if you will, we've made in affective computing is tremendous and these kinds of emotional bots are already here. If you think about chatbots that have been proliferating around, we're going to see more of them.

First of all, I want to say that as a technologist and a computer/data scientist, I always am optimistic about using technology for good and making sure that we take responsibility for what it is we invent and try to ensure that the purpose of the technology is for good. I think we have tremendous capability with AI and data science, and so on, for healthcare, energy, climate, social justice, all these societal grand challenges.

There is potential for us with this technology to actually do better. We just have to make the right decisions. I think what's happening right now is because this technology is in our face as just lay people. There's a lot of media attention to this technology, the good and the bad. Now is the time to have this conversation. At the same time, we have the rise of China and our own political situation right now. Finally, I do believe IT companies are waking up and saying: "You know, we actually should take a step back and think about perhaps regulating facial recognition technology." Well, what would that look like?

I think it's a great thing that the conversation is now in the public. The people have to be in the room for this conversation. It's happening and I don't know what the answer is going to look like. But it's not going to be solved with technology alone or policy alone or government alone or industry alone or academia alone. It's got to be the mix in the conversation. I'm being an optimist—I just think that society, individually and societally, will make progress. I don't think there's an answer or a solution, but I think we all want to survive. So for the purpose of survival, we'll figure out something that will make progress in this complicated situation.

I think Esther you asked a really good question, which is how do we bring our value system into the technology and use of the technology? I think that the challenge there is that many cultures and societies have different value systems. But technology itself is—everyone has this kind of smartphone that has this capability. So do we have smartphones that project different value systems? Is that where we're going? I don't know.

FUCHS: Quick question regarding how to not feel judged by our robot servants. For example, there is a person driving in a car. The car says, "Hey sir, it looks like you're getting really angry." If I tell a person that who's driving—the person engages in an anger response. Calm down, is the worst thing you can say to someone. They're going to blame me—that I'm too sensitive—knowing that my robot car is not sensitive. How are we as a species able to deal with being told that we're angry without having this backlash? Or do you think there's going to be a revolution and coming to terms with our own emotions?

SHULEVITZ: How do we prevent the angry driver from getting even angrier once the intervention has been made? I got a lot of this from a company called Affectiva, which actually is a spinoff from the affective computing lab at MIT. They're working with automobile industries. They are very concerned about the exact question that you asked.

If the car detects emotional distress, maybe some calming music will come on. There's a lot of work being done on making sure that tonally the car is not projecting judgment. If you saw Ex Machina, you've seen this and in every science fiction movie that involves a subservient robot. They're highly skilled at managing the emotions of their human interlocutor.

There's a lot of work being done now using artificially intelligent interviewers to do triage in psychiatric situations. One of the findings is that people will actually tell an avatar more. People will yield more information about their emotional state to an avatar than to a human because

they feel less judged. Now, that is just about programming because you could certainly imagine programming an extremely judgmental computer. I know a person at Carnegie Mellon who does that. But, I would say commercially, the interest is in creating non-threatening computers. Then at the next level up—when we're talking about semi-autonomous and autonomous cars, and the handoff—the computer can just shut the car down. That may anger the driver but there's nothing they can do about it. There's definitely this question of managing the emotional response of the human that's being manipulated by the robot or the computer. This is a very big topic.

WARD: That was mostly about AI stuff, but I also have a lot of research on mind perception, which I like to call mind attribution. I think it was Descartes—the only mind that I actually know exists is mine. So when I'm looking at all of you, and I assume you have minds, I'm creating that out of something. It turns out people really are aggressive when it comes to mind attribution. I think my dog has a mind that is very clever and probably that's just me putting it on there. Thinking in the same ways about how technology is influencing our cognitive abilities, we can also think about this with social relationships—how we relate to others and how a lot of people are hungry for social connections.

I want to throw it back to this point, which was meant to be provocative, about what we value. There are social values, but there are also internal needs that we might value. For example, in Japan, there's a large aging population that has nobody to spend time with them. They are a lot more accepting of robots in Japan. You have caretaker robots being developed, which provide some sort of a social relationship for somebody that would have nothing otherwise. I think that it is interesting to see that maybe at the individual level this is providing something of value. There's a very real human value. But then if you scale that out—do we want our relationships to be with robots because they're easier than relationships with people? Then maybe we get afraid of it. Think about how these values will scale out at the individual level, at the societal level, and then beyond human society, and where we fit into this larger sphere.

WING: It could very well be that once we learn that the car is going to tell us to calm down in some way, whether it's music or talking to us, we all know it's just a car saying that. It is my guess that—just like we learn how to interact with individuals on an individual basis, that's I guess emotional intelligence—we will probably similarly learn how to interact with our self-driving cars or else our semi self-driving cars, as we will with our robots in our homes.

Then you were talking about values like in Japan. People are more accepting of robots and there are more robots helping to care for the elderly and the chronically ill. That's a cultural difference. So we wouldn't want to take that away from the Japanese, who are perhaps more accepting than we are right now in America. We have to be sensitive to those different value systems.

FUCHS: I just want to thank this extraordinary panel for this discussion kicking us off this morning at a very high level. I am optimistic about the future. I have an empirical reason for being optimistic, which is the interest in this day today. In the discussion about these issues—and the extraordinary intelligence of this panel in taking on very complex issues that we all worry about—that is something to be optimistic about. Thank you all very much.

ABOUT THE PANELISTS

ESTER R. FUCHS is Professor of International and Public Affairs and Political Science and Director of the Urban and Social Policy Program at Columbia University's School of International and Public Affairs. She previously chaired Urban Studies at Barnard and Columbia Colleges. Fuchs serves as Director of Whosontheballot.org, an online voter engagement initiative. Fuchs is also a member of Columbia's Just Societies Task Force, the Earth Institute, the Data Science Institute and its Smart Cities Center. She currently serves on the Committee on Economic Inclusion of the U.S. Federal Deposit Insurance Corporation (FDIC) and is a member of the boards of the Fund for the City of New York and the Citizens Union. Fuchs' academic research is in urban politics and policy, American parties and elections, smart cities and urban environmental sustainability policy. Fuchs served as Special Advisor to the Mayor for Governance and Strategic Planning under New York City Mayor Michael R. Bloomberg from 2001 to 2005. And was the first woman to chair a NYC Charter Revision Commission in 2005. Fuchs received a B.A. from Queens College, CUNY; an M.A. from Brown University and Ph.D. in Political Science from the University of Chicago.

CHRIS BAIL is the Director of the Polarization Lab at Duke University, where he is also the Lowey Associate Professor of Sociology and Public Policy. Chris's work combines social science and machine learning to identify new ways to reduce political polarization in the United States. Chris has been a Guggenheim Fellow, published work in the New York Times, and leads a variety of initiatives to teach people about the new field of computational social science.

JUDITH SHULEVITZ is a critic and journalist and author of *The Sabbath World: Glimpses of a Different Order of Time* (2010). Shulevitz was the editor of *Lingua Franca*, deputy editor of *New York* Magazine, and science editor of *The New Republic*. She has been a columnist at *The New York Times Book Review*, *Slate*, *The New Republic*, and *New York* Magazine, and contributed articles and essays to many other publications, including *The New Yorker* and *The Atlantic Monthly*. Her November cover essay for *The Atlantic* examined the social and psychological effects of voice-activated technology in the present and in the future.

ADRIAN WARD is a psychologist and assistant professor of marketing at the University of Texas at Austin. Professor Ward's work focuses on understanding how our relationships with technology and other people influence attention, knowledge, and decision-making. His research in this area has examined how the mere presence of one's own smartphone affects cognitive capacity, how Googling for information affects people's beliefs about their own knowledge, and how sharing responsibility with a relationship partner causes expertise to develop (or not) on a "need-to-know basis." His research has been published in top academic journals including *Psychological Science*, the *Journal of Experimental Psychology*, and the *Journal of Consumer Research*, and covered by popular press outlets including *The Wall Street Journal*, *The New York Times*, *The Atlantic*, and *Scientific American*.

JEANNETTE M. WING is Avanessians Director of the Data Science Institute and Professor of Computer Science at Columbia University. From 2013 to 2017, she was a Corporate Vice President of Microsoft Research. She is Adjunct Professor of Computer Science at Carnegie Mellon where she twice served as the Head of the Computer Science Department and had been on the faculty since 1985. From 2007 to 2010 she was the Assistant Director of the Computer and Information Science and Engineering Directorate at the National Science Foundation. Professor Wing's general research interests are in the areas of trustworthy computing, specification and

verification, concurrent and distributed systems, programming languages, and software engineering. Her current interests are in the foundations of security and privacy, with a new focus on trustworthy AI. She was or is on the editorial board of twelve journals, including the *Journal of the ACM* and *Communications of the ACM*.