

THE BROOKINGS INSTITUTION
BROOKINGS CAFETERIA PODCAST

“Will artificial intelligence lead to utopia or dystopia?”

Washington, D.C.

Friday, July 31, 2020

PARTICIPANTS:

Introduction:

FRED DEWS
Managing Editor, Podcasts and Digital Projects
The Brookings Institution

Host:

BENJAMIN WITTES
Senior Fellow, Governance Studies
Editor-in-Chief, Lawfare

Guests:

JOHN R. ALLEN
President
The Brookings Institution

DARRELL WEST
Vice President and Director, Governance Studies
The Brookings Institution

* * * * *

PROCEEDINGS

DEWS: Welcome to the Brookings Cafeteria. The podcast about ideas and the experts who have them. I'm Fred Dews. On this episode, I'm re-broadcasting a recent episode of the Lawfare Podcast in which Brookings Senior Fellow Benjamin Wittes interviews Brookings President John R. Allen and Brookings Vice President Darrell West about their new book, *Turning Point: Policymaking in the Era of Artificial Intelligence*.

In their book just published by the Brookings Institution Press, Allen and West discuss both the opportunities and risks posed by artificial intelligence. And how near term policy decisions could determine whether the technology leads to a utopia or dystopia.

My thanks to Benjamin Wittes for sharing his excellent interview with me to share with you all. The Lawfare Podcast is an audio production of Lawfare, a blog about hard national security choices that you can find online at lawfareblog.com

You can follow the Brookings Podcast Network on Twitter @policypodcasts to get information about and links to all our shows including Dollar and Sense: The Brookings trade podcast, The Current, and our events podcast. And now, here's Benjamin Wittes with John R. Allen and Darrell West.

WEST: Many of the data sets that are being used to train AI are either unrepresentative or incomplete and that becomes a big source of bias. I mean, we see this very clearly in the facial recognition software area where a lot of the pictures that are used to train a facial recognition system come from Caucasians. And so, it turns out the facial recognition software actually is pretty accurate for White people but not for non-White people.

So, there is an obvious recommendation here which is in order to reduce the bias in AI in areas like this, we need to improve the representativeness of the data that go into those

algorithms.

WITTES: I'm Benjamin Wittes and this is the Lawfare Podcast July 17, 2020. Darrell West is Vice President and Director of Governance Studies at the Brookings Institution. John Allen is the President of the Brookings and a retired U.S. Marine Corps 4-star General. They are together the authors of the book, *Turning Point: Policymaking in the Era of Artificial Intelligence*.

The book is a broad look at the impact that artificial intelligence systems are likely to have on everything from the military, to healthcare, to vehicles and transportation and to international great power competition. It's also a deep dive on the question of how we should govern AI systems. What makes for ethical uses of AI, what makes it scary, what are the anxieties that people have about artificial intelligence and to what extent are the fears legitimate? It's the Lawfare Podcast July 17th. John Allen and Darrell West on artificial intelligence.

I want to start with some of the big overarching themes that you guys deal with in this book. And the first of them is just the question of whether the concept of AI is discreet enough and definable enough to organize policymaking around since, you know, we've had computers doing big data analytics and making decisions for a long time. What distinguishes AI of the sort that you guys are talking about here from the stuff that's been going on ever since we've had, you know, kind of cloud computing and mass marketing and, you know, big platform companies that are making decisions regarding what ads to serve us every moment of every day?

WEST: Well AI is automated software that learns from data text or images and then makes intelligent decisions based on that. So, I think the key distinguishing features are intelligence learning and adaptability. I mean that's really what distinguishes AI from some of these past technological solutions.

In our Turning Point book, we argue that AI is the transformative technology of our time. We're seeing it deployed in many different sectors from healthcare and education to transportation and e-commerce and national defense. Some of the issues raised are cross cutting in the sense that like privacy, it cuts across every particular sector. Some of the other policy issues though are really sector specific. And the policy solutions will need to be tailored to that particular sector.

WITTES: And so, what do you think of, Darrell, as the crosscutting issues? When you say some of the issues are crosscutting, there are certain themes that come up again and again in these chapters. You know, for example, the possibility of algorithmic bias or, you know, sort of governance of the algorithms. But how would you describe the sort of crosscutting policy challenges of AI?

WEST: I mean, the things that people worry about with AI are issues such as fairness, bias, especially racial bias is a big concern these days. The transparency of AI. Like nobody really understands what's inside the black box and how decisions are being made. Human safety is an issue. We have a chapter on autonomous vehicles and transportation so the question is, you know, are autonomous vehicles going to be safe? That is a big concern.

And then the question of who decides. I think one thing that's distinctive about our book is we emphasize both the policy aspects but also the governance questions. The question of who should decide. In the past, we pretty much delegated a lot of the decisions to private companies. Now the public has a greater need for public engagement. Policymakers are coming in regulators are starting to think about this space. So, those are, I think, the big issues that animated our discussion of AI.

ALLEN: And Darrell is 100 percent correct. Many of the functions that you described in

the first question where you talk about big data analytics and supercomputing. Many of the, what would appear to have been decisions, were really automatic decisions. What we're facing in the future will be artificial intelligence backed by big data and supercomputing where the algorithm will take on the capacity of making autonomous decisions.

And how we develop that algorithm and how we ultimately train that algorithm with data sets will be how we can have a confidence that the autonomous decisions that are being made are decisions that we can live with. And that goes to the very policy issues that Darrell has raised.

WITTES: So John, one of the fascinating things about this project is and I was trying to think of another example of this. I suppose internet connectivity more generally as an example. But there aren't that many issues that you can at once go through discreet domestic policy areas like how is this going to affect education and how is it going to affect healthcare. But then also frame it in terms of great power competition. Is China going to be the world leader in this, or are we? And also, frame it in terms of pretty granular national defense policy questions.

And so, I guess my question is, is this fundamentally an issue of international competition and sort of great power rivalry. Is it fundamentally an issue of domestic governance, is it both, is it, you know, sort of what is the broad rubric through which we should think about AI both at a domestic and international level?

ALLEN: Well, it's all of those things that you described, Ben. Artificial intelligence will be one of the most transformative aspects of emerging technology in the 21st century. And it's about 50 years old. I mean, the concept of artificial intelligence with some pretty good clarity and definition was proposed years and years ago, decades ago.

But I don't think that in those days big data, data sets and computing power were such that many of the scientists who proposed the various qualities and potential benefits of artificial

intelligence, I don't think that they envisaged that we'd achieve many of those for some time. Truth of course is that with supercomputing, with cloud computing, giving us power that we could not have imagined just years ago.

And with access to data sets that we then used to train the algorithms, virtually anything that is of value to us in our society has the capacity of being enhanced through artificial intelligence. And Darrell just gave us a very good list. Everything from transportation to smart cities to education. I think education will be transformed in many ways through artificial intelligence and augmented in virtual reality.

So, we're going to see many transformational dimensions unfolding within our domestic world. But very importantly, when we look across the international scene, in most cases and we propose to do it in all cases but we're still learning, we seek to build ethical AI. We seek to create environments and applications of this technology that enhance the human experiment that operates within the rule of law. That in fact improves societies.

That's not the case with certain authoritarian states, that's not the case and we're seeing that today. China is the most sophisticated surveillance state on the planet. And in many respects, it's because of their capacity to employ AI algorithms backed by almost unimaginable sized data sets and the appropriate levels of supercomputing to keep tabs on the population.

And frankly, that's not who we are, that's not what we'll tolerate and the community of democracies, whether it's the transatlantic relationship or the community of democracies more broadly incorporating the East Asian democracies. A very powerful and sophisticated democracies there. That's not who we are.

And so, in many respects, the mastery of this emerging technology and the use to which it is put is going to define the struggle, I think, in terms of great power competition in the 21st

century. And we have to be attentive to this. And frankly, you know, I've always been one who has sought to define the relationship with China as one where we should seek cooperation. But we cannot permit ultimately the Chinese to dictate the technology of artificial intelligence in the 21st century for all the reasons that I've just attempted to explain.

WITTES: All right, so when people talk about AI and you mention this in your introduction, there is a lot of sort of images that come up whether it's the Terminator or how or some kind of self-aware malevolent network that is trying to seize control. You know, in fact AI is seldom general purpose, is never self-aware in the science fiction sense of the word, at least not yet. AI's tend to be a very specific purpose and they tend to be very good at one thing.

And so, I'm curious to what extent you guys think that the fears of AI that we have are realistic fears. And to what extent they're kind of more, I don't want to be disparaging of it but more luddite kind of disparagement of and fear of technological progress of precisely the sort that we always see when technology is evolving very quickly in fashions that make certain human jobs less valuable or even obsolete.

ALLEN: I mean I would say the Hollywood fears of the terminator or a malevolent robot that's going to enslave humanity is vastly overblown. I personally don't worry about that. I think it will be generations before humans actually have to worry about that because it is going to take a long time for the AI to really develop.

But that doesn't mean that the more specific fears of are the AI (audio skip) in the financial area biased. Like that is a very reasonable concern. We actually see evidence of that. I'll say for our autonomous vehicles going to be? How are we deploying technology in healthcare and in education? I mean, there are a lot of questions related to the equity of access to digital technology.

We see a lot of benefits in various areas but if people don't have equitable access it means they're going to be winners and losers. So, these fundamental questions of fairness and bias in human safety are very legitimate. We talk a lot about possible solutions to these problems. We present a policy blueprint on things that we think would be necessary.

I mean, we entitled our book *Turning Point* because we do feel that we're at major junction point where things could go in a positive direction or in a negative direction. But we think the crucial factor in determining the future is really the policy response. And there are very reasonable policy steps that could be undertaken to address a lot of people's concerns.

WITTES: And so, let's talk about the bias concern because it really does show up recurrently in the book. Why the fear of bias in AI and to what extent are AI's in general written with significant biases of one sort or another? And, you know, to what extent is that a function of intentional coding discrimination versus more subtle unintentional, I would say, subconscious biases on the part of the coders. Or even superficially or notionally unbiased algorithms that have highly biased affects. What's driving the bias concern here?

WEST: I mean, the AI is only as good as the data that goes into it. And the big problem we face right now is many of the data sets that are being used to train AI are either unrepresentative or incomplete. And that becomes a big source of bias. I mean, we see this very clearly in the facial recognition software area where a lot of the pictures that are used to train a facial recognition system come from Caucasians.

And so, it turns out the facial recognition software actually is pretty accurate for White people but not for non-White people. So, there's an obvious recommendation here which is in order to reduce the bias in AI in areas like this, we need to improve the representativeness of the data that go into those algorithms.

ALLEN: And as Darrell said again, the algorithm itself may have no inherent bias to it at all but it's no better than the data by which it's trained. And data that is extracted from segments of our society that have no access and have no presence in the data sets themselves or are vastly underrepresented in the data sets themselves. Or if the data frankly has corruption in it and I don't mean corruption in the context of malfeasance I mean corruption in the context of it's just not representative and doesn't accurately reflect the population that created the data set itself.

If that exists and that's the basis for the training of the algorithm then the algorithm will make decisions which are inherently biased and that's the problem that Darrell explain and that's one of the issues that we hit very hard in the book. Because it's important to understand that AI in and of itself is not going to be inherently biased. But if the information used for that algorithm is biased in the context of it's not representative or is corrupted in some form or another then we can't imagine that the decisions will be otherwise than reflective of those disparities coming out of the data sets.

WITTES: And yet some AI's themselves are biased in their judgements. You design an AI to A, take a whole lot of data and, you know, learn off of it but you're also teaching it to make judgements. And if you create an algorithm that makes judgements that are informed by your own prejudices or preconceptions, you could have bias inherent in the AI itself, right, even before you get to the data.

ALLEN: I mean, I think in the finance area there are some problems that need to be addressed. So, we know in the bricks and mortar world of banks and financial institutions that banks often would redline certain neighborhoods and simply not provide mortgages or loan to people who lived in low income areas and that was a real problem.

The risk today is that we may be creating digital redlining. That there is so much

information about people that even if you are explicitly excluded from incorporating issues of race or gender into the algorithm. There are lots of proxy measures that can be incorporated in an AI solution that end up biasing against racial minorities or women.

So, this is an area where we need to develop the policy and regulatory means to deal with bias and redlining in a digital space. It's often hard to do that and to make sure that the algorithm works fairly but that's a very important priority in terms of moving forward in order to avoid these types of problems.

WITTES: John, you referred earlier to ethical AI which is a term that will not mean a lot to a lot of people without explanation. Ethical AI is sometimes contrasted with black box AI and it has connotations both of, you know, ethics as we traditionally understand them but also of a certain degree of algorithmic transparency. So, talk us through what you mean when you say ethical AI and what examples of ethical versus unethical AI would look like.

ALLEN: The term ethical AI or open AI or explainable AI is an objective as we move forward in this technology. And what it means is we want the people who ultimately who are served by the particular algorithm, we want them to feel as though both the coding and the training, the coding of the algorithm and the training of the algorithm was done in a way that's explainable. That's done in a way that's transparent, that's the open AI, and was done in a way that gives us the capacity to have high levels of confidence in the ethical orientation of the algorithm and outcome.

And sometimes we use the acronym ELIi. The little i stands for implications. And when we think about any emerging technology but in this case artificial intelligence, we talk about the ethical component of the implementation and the application of this particular algorithm. We talk about the legal propriety of it. In fact, it isn't legal in the way in which we are proposing the use

of this particular algorithm.

And finally, what are the societal implications. And as coders, as elements within the private sector who are contemplating the development of algorithms or the perfection of algorithms, improvements over time. If we are governed by a bias, if you will, to ensuring that we understand the ethical implications. The legal implications and the societal implications of an algorithm or an outcome then it is explainable, it is open and the users, those who are affected by the algorithm can have a high level of confidence that their lives will be enhanced by it rather than their life being threatened by it.

WEST: Yeah, I think John is exactly right on that on distinguishing ethical from unethical AI. I would say the worst example that I have seen is really in the criminal justice area. You know, that obviously is a very important area because people go to prisons over bad decisions in that area or they end up on probation.

And in the criminal justice area, there are many jurisdictions that are engaging in what are called predictive analytics in which they essentially mine a bunch of existing data and make predictions in terms of who is at risk of committing another crime. And they use these data analytics to direct law enforcement resources and, you know, where the police should engage in various activities.

But when you look at the actual algorithms that they're developing, you know, in the criminal justice area there is a tremendous amount of racial bias. And so, the algorithms basically predict that some minorities are likely to commit another crime because they have been convicted of a crime in the past.

And given the fact that there is so much bias and discrimination in the criminal justice area, like I think that's an example of how AI is being used for unethical purposes and are developing the

algorithms in ways that almost reaffirm the existing bias. So, we have to be very attuned to those types of negative uses of AI and make sure that, you know, people are not being harmed and their future is destroyed because of the algorithm.

WITTES: So, I'm struck as I hear you talk about this that a huge percentage of the general anxiety that you guys regard as legitimate as opposed to the sort of science fiction computers are taking over anxiety actually coalesces around the areas where AI's are making judgements about people. As opposed to, for example, you know, autonomous vehicles where the real question comes down to can we figure out an appropriate liability regime and are autonomous vehicles likely to be safer or less safe than human driven vehicles.

But the crunch of a lot of what you're saying is a sort of anxiety about situations in which AI's are making judgements that affect actual people's lives and kind of evaluating their fitness for things in one way or another. And that that makes us uncomfortable, perhaps rightly so, particularly when we don't understand the bases for the judgements. Is that fair?

ALLEN: I think that is completely fair. I think that's a good summary of a lot of the problems in AI as well as people's anxieties about AI. And, you know, a I give a lot of talks about AI around the country and around the world and people worry about a loss of human agency. And they worry that humans are no longer really in control of the software. And I think that's something that we really need to address.

One of the things we develop in our concluding chapter is we talk about the need for annotated software and annotated AI. I mean, most AI solutions have millions of lines of code and even professional coders could not look at all the code and figure out what is going on. But we suggest that software designers start to annotate their code so that when they reach a major decision point, they're drawing on certain types of data, they annotate it with how they made that

decision and on what basis they made that decision.

So, that we can then go back if there are problems or if there are harms to humans and try and figure out what went wrong and how we can go about correcting it. So, I think you're absolutely right that a lot of the anxiety does come from people worrying about the software making judgements but we really have to put an emphasis on how we can improve the basis on which the algorithms operate.

WITTES: All right, I want to talk about some of the discrete policy areas that you guys address in the book which are actually in and of themselves huge subjects and not discrete areas. They're each subjects that you could you have written an entire book about themselves. But I want to focus because we are Lawfare on the ones that have obvious security implications.

So, let's start with the national defense arena. The debate over lethal autonomous weapons systems is a relatively old one and it's one actually that Lawfare was, you know, an early part of the sort of public arguments over the propriety of these systems. To what extent and this is a loaded question, to what extent is the AI debate with respect to the military and national defense applications limited to lethal autonomous systems and to what extent is it broader than that?

ALLEN: Well let me jump in on this. It's quite broad, actually. It's a very healthy debate in my mind. If you build a firewall bed, if you envisage a firewall between the application of combat power on the one side of the firewall and I'll come to that in a moment. And all other military functions on the other side of the firewall such as medical service, maintenance, supply discipline, traffic management. There are many, many functions that are performed in the military where artificial intelligence can both improve effectiveness, efficiency, and safety.

And we're seeing a robust conversation about that. Predictive and preventative

maintenance, for example, has really come a long way where there are smart systems that can be placed on many of our dynamic systems. Whether it's a generator on board a ship or an aircraft or a tank or any kind of a vehicle which can monitor the operating systems on board that particular platform. And detect at the earliest possible moment anomalies in the dynamic systems which can give sufficient warning to prevent catastrophic failure.

Now that enormously valuable to the military in the context of efficiency in terms of keeping systems up and running for the operational efficiency of the unit itself but also in the expenses associated with the maintenance. And that applies to aircraft, that applies to terrestrial vehicles, that applies to ship borne systems as well. So, just that alone is a tremendous opportunity.

WITTES: Let me pause you right there for a second because I think the implications of what you just said are pretty vast. Which is if you snap your fingers tomorrow and said there will never be lethal autonomous systems used by the U.S. military and, you know, so we can just end that debate, we can get rid of that debate. What you're saying is by doing that, you wouldn't begin to scratch the surface of the impact of the AI conversation on the U.S. military.

ALLEN: Yeah and there's really very little, that is correct, that's a fair statement, Ben. And I would not say that there's much debate, actually, about whether AI has a role in the U.S. military in the context.

You know, an aircraft, for example, that is not self-aware because it's AI but certainly has the capacity in flight to determine whether it's beginning to have certain -- once again, the artificial intelligence that is now inherent in the dynamic systems. Monitoring the dynamic systems, are able to detect at the earliest possible moment anomalies in that operation that permit that aircraft on its own to project down to the wherever it may be, the airfield or the maintenance

facility that it is coming in with a potential maintenance problem.

And the outcome of that particular signal is that when the ground crew meets that aircraft in the pit for maintenance, the artificial intelligence has already, because of natural language processing, selected the pages from the maintenance manuals that those mechanics are going to have to use to do the maintenance. The parts have already been selected, the parts are ready to go, the pages of the maintenance manual are already ready to go, the diagnostics have already occurred, they're already loaded onto the tablet. And when that aircraft pulls into the pit for the maintenance that will be necessary, it's because we're out ahead of the potential failure of a component rather than experiencing a catastrophic failure that could require hours and hours.

This means that we can keep an aircraft in the air longer. This means that the turnaround between sorties is shorter. This means that the pilot will be safer. It means we'll be much more efficient in the application of parts. That's just one example that could fly across the entire military.

Also, maintenance of the supply system knowing where everything is and pushing those supply elements in the direction that is necessary based on an overarching awareness of an operational environment. Recognizing that there may be elements within the force that are engaged in extensive combat operations and beforehand.

Having done the analysis that we're going to need medical support for burn victims. We're going to need to be pushing fuel resupplies forward to a particular area so they'll be ready and in position anticipating a need based on the training of certain algorithms on what a tank battle dose in terms the consumption of petroleum and tank ammunition, main gun ammunition, et cetera.

This is an area of both predictive maintenance, predictive supply that was done manually and required lots of people to do it. Now it could be inherent to the planning process to give

commanders a sense of what it will cost them to select a certain course of action and have that kind of decision support very, very quickly.

WITTES: You offer another example that I was very struck by which is hypersonic missiles. Where the implication of what you're saying is that the guidance of these systems is impossible to imagine without AI. And I was struck by that because, you know, even that implies that, you know, there is a generation of weapon systems that are adversaries will be developing even if we are not. That are pervasively dependent on AI not merely for their lethal decision making but for their functional operation. Is that reasonable to anticipate?

ALLEN: It is. And, you know, we talk about the concept of hyper war as one that we're going to have to face. Hyper war is conflict that is conducted at speeds so great that the presence of the human in the decision making loop slows down that particular side to the extent that it's fighting at a disadvantage. Sometimes it's called conflict or a war at the zero phase.

And without spending too much time on this because we just don't have the time in this particular podcast, there is something called the OODA loop, the observe, orient, decide and act sequence and it's in a loop, it's always circular. The leaders and decision makers are making these decisions in a loop constantly in combat.

It is the intent of the one side of the other to compress the OODA loop to a point where if the commander can make decisions cycling through that loop more quickly than the opponent. Warfare is a time competitive process, that's the bottom line. The side that can generally move more quickly can decide more quickly and act more quickly will generally be the side that can prevail.

And that often is completely apart from the size of armies or the size of military forces. It often is decided by the commander who can decide more quickly and move more quickly

through the OODA loop. If we can position artificial intelligence both in the gathering and the orientation on information which is the intelligence process, artificial intelligence in decision support and decision assistance for the commander to weed out extraneous information and provide a feasible and credible course of action to that commander.

If we have selected that commander early in that person's career, that woman's career, that man's career and educated and trained that individual to be comfortable in a hyper war environment. So, when given that decision or offered those courses of action at speeds that we could never have imagined in previous eras of warfare.

And that officer has been trained to be comfortable making these quick decisions and accepting these courses of action. We can compress that OODA loop and be faster than our opponent as a result of artificial intelligence being embedded in the mechanisms that provide us the ability to observe the enemy, orient on the enemy, ultimately decide what to do about the enemy and then act. And we're back to firewall, the acting piece is where artificial intelligence will play in the larger application of force and combat power.

WITTES: So, Ken Anderson likes to say that if you want to understand the military and robotics and you might as well say AI here as well. You really have to understand the autonomous vehicles industry because every aspect of the two has real parallels to each other, including the lethal potential of the technologies.

So, Darrell, you guys have a whole chapter on autonomous vehicles here. This does not raise the discrimination or bias questions but it does raise huge public safety questions as well as significant opportunities. So, how do you understand the potential and dangers of the incredible advancements that we've made in autonomous vehicles and their relationship? The roll of AI in those vehicles both now and in the future?

WEST: I mean, you're right that the whole controversy over autonomous vehicles is all about the AI. I mean, it's the software that integrates the lidar information, the camera information and tries to keep the car or the truck on the road in the right lane so as to avoid hitting people. But people do worry about the human safety of autonomous vehicles.

And the technology is coming along. It's actually getting better. The funny thing is if you look at, America has something like over 40,000 highway fatalities every year. 90 percent of the accidents involve some type of human error. Either drunk driving, distractedness or something like that.

Autonomous vehicles are going to be much safer than human begins in the sense that the AI is not going to get drunk and it's not going to get distracted. So, at a fundamental level, autonomous vehicles are definitely safer than human driven cars but people worry about being killed by a machine.

And so, the whole notion of an autonomous vehicle is more worrisome to people. We actually have done surveys on public attitudes towards riding an autonomous vehicle. And it's something like only about 21 percent of Americans say they, at this point in time, would be willing to ride in autonomous vehicle. So, this is going to be a tremendous problem for the auto industry as they start to rollout autonomous cars just overcoming those fears.

WITTES: One of the things that this area has and presents quite pointedly is what the liability regime is going to look like when you create an autonomous vehicle and something goes wrong. Who carries the risk of that? Is it, you know, in our normal highway situation, the driver does and to the extent there's a product defect, the manufacturer of that product does.

In an environment in which it's not entirely clear who the driver is and we have layers of liability protections for software and hardware manufacturers as well as a general quite

protective environment for internet carriers. You know, when you get in a Tesla and put it in autopilot mode and it fails, who's going to be responsible for that?

WEST: Well, that is the big question for autonomous vehicles. And you're right that the answer to that is almost completely unresolved right now. Because in the case of human drivers it's easy because most of the time it is a human error that leads to the accident. And so, the liability almost always is going to lie with one of the humans involved with that. And, of course, we spent a lot of time trying to figure out which human is actually liable.

When you move to autonomous vehicles, then it's a question of if there's a human operator in the car is that person responsible, is it the hardware meaning the car itself or the computers that are powering the car or is it the software. You know, did the AI integrate the information properly, did somehow the lidar misread road conditions which actually has been a problem in some of the cases that have hit the news.

And so, resolving those liability issues is one of those policy and legal types of questions that need to get resolved. One, to reassure people that autonomous vehicles are actually safe and then two, just to deal with that important question of who is going to be responsible.

But it all illustrates kind of the larger point we make in the book of humans are in control in the sense that the way we deal with technology harms, either in the case of AI or any other type of technology, is through laws, policies and regulations. It is within our power to work on resolving those liability questions so it's really not a problem of the technology it's a problem of us needing to clarify the liability associated with the technology.

WITTES: Yeah and so two other policy area specific questions and then I want to talk a little bit about the governance. One is, you know, people like Andrew Yang and many others have talked about robotics as a, you know, sort of major threat to American jobs. And I'm

wondering whether there's a bit of a conflation here, you know. A lot of what he's talking about in the industrial sectors regarding, you know, robots is first of all actually derivative of the AI issues that you guys are talking about.

But secondly, whether AI's potentially pose as well as a giant growth opportunity in certain areas of the economy a parallel threat to office jobs that robots pose in the manufacturing sector. Is this a, you know, another step down the road to the expendability of labor?

WEST: Robots and AI are going to take over the repetitive, boring and dangerous tasks. So, from that standpoint, it's actually a positive development. And on the jobs front, I mean, there's no question automation, AI and robots are going to take some jobs and certainly some of the entry level jobs are at risk. But as you point out, some professional jobs like radiologists. It turns out, AI can read CAT scans and x-rays pretty arcuately. So, radiologists represent an occupation that is going to be at risk.

But at the same time that some jobs are going to be lost, AI and emerging technologies are going to create new types of jobs. Certainly, anything related to data analytics will be in very hot demand. But also, we're going to need more ethicists just to figure out these ethical issues associated with AI.

I think the thing I worry the most about is not lost jobs but a mismatch between the new jobs that are being created and the skills that people have. We are going to move to a model of lifelong learning where people are going to have to upgrade their job skills at age 30, 40, 50 and 60.

WITTES: We're going to have to retrain all of those radiologists as ethicists.

WEST: Perhaps.

WITTES: You mentioned radiologists and I want to talk briefly about healthcare.

Because, you know, other than the autonomous vehicles industry and perhaps financial services, this is really the area where this is really likely to touch everybody's lives in a way that's very direct. And so, when you think about AI in the context of healthcare, are we talking mostly opportunity or are we talking mostly downside and risk?

ALLEN: I mean, I think there is tremendous opportunity for AI in the healthcare area. I mean in this COVID world, we're already seeing AI being used to help develop new drugs and new vaccines. Because the software can review scientific literatures much faster and more efficiently than humans and find promising new chemical compounds and things that might have been used in the past to deal with one issue that could be relevant for new types of issues.

We're also seeing the use of data analytics to track the spread of the pandemic and also show problems, racial disparities, geographic breakdowns and so on. A lot of the healthcare applications are basically drawing on wearable devices, remote monitoring devices that might record your health symptoms, automatically transmit them to health providers. And then that provider on a proactive basis can contact you to say hey, you may not realize it but your blood pressure is spiking and, you know, here's an issue that we need to address.

So, I do think there are many potential benefits in the healthcare area. Although we do have to still address the equity issues of whether everybody will share in the benefits of those new applications.

WITTES: All right, before we wrap up, I want to, you know, we've talked about such an amazingly diverse range of questions here. I want to come back to where we started which is, is this subject so broad that it defies any cohesive governance principles? Or is there some set of ideas about governance and about oversight and management that can be in some meaningful way cross cutting as we think about how to manage this transition.

WEST: It certainly is important to address the question of who decides. I mean, that is a basic governance question that really cuts across every one of these sectors. So, I think that's important, I think. In the book, we talk about bringing together experts from a variety of different areas, the coders, the ethicists, the social scientists who work on these types of issues. So, I think in many of the AI areas, we need a new type of governance that can address the problems that have developed.

WITTES: The book is *Turning Point: Policymaking in the Era of Artificial Intelligence*. The authors are Darrell West and John Allen, my two bosses. John, Darrell, pleasure speaking to you. Thanks for joining us.

ALLEN: Thanks, Ben.

WEST: Thank you, Ben.

WITTES: The Lawfare podcast is produced in cooperation with the Brookings Institution and when we have the President and Vice President of the Brookings Institution on for a show, that is a good illustration of what in cooperation with means.

The Lawfare podcast is produced and edited by Jen Potchia Howell. Our audio engineer this episode is Zachary Frank of Goat Rodeo. You need to do your part to promote the Lawfare podcast so tweet about us, share us on Facebook, pin us on Pinterest and upload us on Reddit. Make Tic Tok videos about it on your way to the lawfarestore.com where you can get all the Lawfare merch your little heart could desire.

But don't forget to stop at whatever podcast distribution service you got us from and rate and review the Lawfare podcast. You know how many stars we want from you. Our music is, of course, performed by the one, the only Sophia Yang and as always, thanks for listening.

* * * * *

CERTIFICATE OF NOTARY PUBLIC

I, Carleton J. Anderson, III do hereby certify that the forgoing electronic file when originally transmitted was reduced to text at my direction; that said transcript is a true record of the proceedings therein referenced; that I am neither counsel for, related to, nor employed by any of the parties to the action in which these proceedings were taken; and, furthermore, that I am neither a relative or employee of any attorney or counsel employed by the parties hereto, nor financially or otherwise interested in the outcome of this action.

Carleton J. Anderson, III

(Signature and Seal on File)

Notary Public in and for the Commonwealth of Virginia

Commission No. 351998

Expires: November 30, 2020