AI and You

Transcript

Welcome to episode 41! In this episode we'll conclude the interview with Peter Asaro. He is a professor at the New School in New York, and is a philosopher of science, technology and media. His current research focuses on the social, cultural, political, legal and ethical dimensions of automation and autonomous technologies, from a perspective that combines media theory with science and technology studies. He has written widely-cited papers on autonomous weapons from the perspective of just war theory and human rights, and the legal and moral issues raised by law enforcement robots and predictive policing. He also examines agency and autonomy, liability and punishment, and privacy and surveillance as it applies to consumer robots, industrial automation, smart buildings, UAVs and drones, and autonomous vehicles. That's exactly what we talked about in the last episode and we'll get into more of it now. Let's get back to talking with Peter Asaro.

You mentioned a couple of organizations earlier and I'd like to understand the relationship between them and their achievements. The International Committee for Robot Arms Control and their Campaign to Stop Killer Robots. I'd like to know more about what has happened with that, what it achieved. I believe Amnesty International was involved. Can you tell us something about its activities and aims?

Sure. So I'll start with my own NGO, which is the International Committee for Robot Arms Control, which I co-founded with some other scientists and researchers in 2009. And we're really a group of academics from robotics, AI, philosophy, law, international relations. I think we got some anthropologists in there as well. And so we're primarily working as academics but we're concerned about this issue, and sort of organize ourselves through this as an expert group. And we're part of a coalition of non-governmental organizations called the Campaign to Stop Killer Robots, and that was founded with nine groups back in 2012. You put me on the spot to remember them all, but it was a Human Rights Watch, the Nobel Women's Initiative, The International Committee for Robot Arms Control, this group called Article 36 in the UK, Canadian Mines Action Canada group, the Pugwash International group, AAR Japan which is a large Japanese NGO. And I think that's most of them, I'm sure I'm forgetting… oh, the WILPF - Women's International League of Peace and Freedom. And now we've grown; we've added Amnesty International to the steering committee. And we've been joined by, I think we're about 170 NGOs internationally, representing, I think, 70 different countries now around the world. And we've had a few big international meetings to coordinate all these national-level campaigns around the world. And we started out at the UN, there was a report that was presented to the Human Rights Council at the UN by the Special Rapporteur for extrajudicial summary executions, Christof Heyns back in 2012. And the Human Rights Council doesn't really like to deal with arms control issues. So they sort of said, "Oh, that's a nice report, but we're not going

to do anything about it." And a different body within the United Nations called the Convention on Certain Conventional Weapons, which is a treaty body that's designed to govern conventional weapons systems, was created in the 1980s, took it up as one of their issues. And they started with a series of what they call informal expert meetings and then elevated that to formal discussions called a Group of Governmental Experts or GGE. And there are lots of acronyms in the UN, so CCW, GGE. And they went with lethal autonomous weapon systems, which is LAWS, which is kind of odd because we're talking about laws to govern LAWS, but I usually just drop the "L" and say autonomous weapons systems. And so since 2013, there's been a series of meetings within this body of the CCW every year to discuss whether it's necessary to have regulation of these things, and if so, what are they, what would be regulated? And we're now sort of several years into the GGE and we're really hoping that it would escalate to the next level, which would be the negotiation of a treaty element, which would either be a new protocol to the existing CCW treaty. They have done in the past certain protocols for landmines and cluster munitions and permanently blinding lasers on things like that. Unfortunately, the ones they have done in the past for landmines and cluster munitions were ineffective, essentially. And so even though they did something about it, it was necessary to have a stronger international treaty. And in those cases, the countries went out on their own and started what's called an outside process. And so they have the Oslo process that Norway initiated. And you have the Ottawa process that Canada initiated, which was for land mines, and Oslo was for cluster munitions. And they basically got an international treaty signed by enough countries around the world that it becomes international law. And then once that happens, it kind of comes back into the UN. The UN is kind of an umbrella organization that oversees all of these treaties that meet annually and things like that. So that's one option. The other option would be like the Nuclear Ban Treaty, which went through the General Assembly of the United Nations. And before that, the Arms Trade Treaty also went through the General Assembly. The problem with the CCW and why their previous protocols were weak, and why we have had trouble getting any kind of protocol over the past eight years now, is that it's a body that is ruled by consensus, which means effectively, every state has a veto. So any one state whose party can just say no, and nothing happens. And so what winds up getting passed is usually pretty watered down. What's better about the sort of the outside process or the General Assembly works by majority. So you can get just a majority vote to pass some of these things. And then outside, you can just get like-minded states to it. And if a significant number of states - 50 or 70 states - are willing to sign on, then even if your states don't sign it, it's recognized as international law. There's been a lot of controversy around the Nuclear Ban Treaty because the states that have nuclear weapons haven't signed on to it. But you can imagine, there's a handful of states that have nuclear weapons, and if all the other states in the world are saying, "But nuclear weapons should be illegal. They are illegal. We've all signed a treaty that says they're illegal. You guys are the outliers in this, and if you use nuclear weapons, they are illegal, and you're sort of violating this treaty." And that treaty just went into effect in January.

There's quite a difference, though, in that you can spot a uranium refining facility from space but there's no way that you could detect artificial intelligence, or even [a] manufacturer of killer drones, which actually makes me think about the title Campaign to Stop Killer Robots seems

somewhat polarizing in that there are a lot of AI professionals who are like, "Oh, no, stop talking about killer robots." And then Elon Musk talks about killer robots, which makes them all the more energetic about "Stop talking about this. We're nowhere close to Terminators walking through the streets." And Stuart Russell at Berkeley who is an artificial intelligence expert and teacher and author and says that he will not do any interviews if they're going to run a picture of the Terminator next to him and they do anyway. And then he produces the "Slaughterbots" video, which is literally about killer robots. Does that title get in the way? Because it's quite sensationalized. I can't think "killer robot" without a picture of the Terminator coming into my head. I can't separate that.

Yeah, I mean, it was an issue for us as far as the press running that Terminator photo every time they talked about us for many years. I think they've mostly gotten over that. And now there are enough other things that can show pictures of these new drones and so forth. And I agree that it doesn't capture exactly what I think the core concern is. But I also teach in the media studies department that I recognize that the fact that it grasps public attention, that people want to talk about it, the media wants to cover it, and the UN has been willing to discuss it even though they are going to substitute the phrase lethal autonomous weapon systems shows that it's a real concern. In one sense, it should be kind of a no-brainer to say, "Well, of course we don't need killer robots. That sounds like a terrible idea." And yet we see militaries developing the key components and elements and deploying them and testing them. And you're like, "Well, they're building killer robots."

Exactly.

And, of course, I mean, there are all these other concerns that we might have around technology and AI and we started to see a lot of that. And I've written also about that, in terms of bias and predictive policing, and all sorts of things in which automated algorithmic decision making is going to impact human life. And these questions about human responsibility, about the ability to appeal decisions that have been made to find out the transparency and the accountability behind those decisions, I think are crucial across society. And this is a case in which the decisions that are being made are irrevocable and fundamental. I mean, it's about taking human life. And they get to these deep questions about what is human dignity? What are our rights with respect to automated decision processes, with respect to holding to account the humans who designed and built and implement those systems? And so I think it sort of transcends a whole lot of questions. It's easy to sensationalize to be sure, and it's been in science fiction for a while. But in fact, the "Slaughterbots" video itself, I think, has that sort of mixed. Well, it's captivating, and everybody wants to talk about it. People bring it up all the time. We didn't produce it as part of the campaign because it doesn't make the argument that we need to make to convince states to sign treaties about this, but it captures the public's attention. And most of those technologies are required to implement the system that's displayed on that kind of *Black Mirror*-esque video are readily available. We could build systems like that if we really wanted to. The big hurdle from what you see in the video, I would say is the battery life of drones. That's the constraining factor from implementing--

Or sourcing a plastic explosive.

Yeah, well.

Well, I wonder if it illustrates perhaps a bigger issue is not the use of these things in state warfare, or the accidental consequences of its use in a democracy or even by non-state actors, but in states that are oppressive regimes. I'm thinking particularly about China and what they could do against the Uyghurs in particular, and they've shown no hesitancy in applying AI for all the kinds of applications that give us palpitations over here. And even at the end of the "Slaughterbots" video, those drones are deployed against the students by actors that are not apparently working for the government, except that they achieve the aims of suppressing dissent against the government by the students that were protesting the use of that very technology. So one could say maybe that was being done by some under-the-table arrangement there. Is that perhaps something that we should look more at? Is there a way of addressing that concern? Because a nation could say, "Look, it's our business what we do with our own people. Stay out."

Yeah, well, I think there is a serious concern about the use of these kinds of systems in policing, and especially in say, crowd control, and disrupting peaceful protests, and killing protesters. And we've seen a range of different events in the last few years starting in the Arab Spring, where, in Egypt in particular, the military was ordered to turn on the public that was in Tahir Square - a million people peacefully protesting - and refused. The military actually refused to do that and that led to the fall of the Mubarak regime. Would robots refuse the orders? No, of course not. They're going to just do what they're told to do. And so the idea that you could kind of win a popular uprising by appealing to the sentiments or the politics or the humanity of the military is sort of precluded by these systems. And that would be one of many bad outcomes. And even then, I think, in China's use of military in Tiananmen Square in the 1980s, they had to pull in troops from far-flung regions of China because local troops would not have attacked the protesters at Tiananmen. But I think a lot of really bad things have been done by human beings that convinced other human beings to do them, and that's not going to stop. But you add a whole other dimension to it with machines. And I think what the "Slaughterbots" video also captures that you were getting out a little bit in your question was the lack of attributability. Like these "slaughterbots" are coming in and killing everybody, but nobody really knows who's controlling them or why. They kind of figure out that people who watched some video have been targeted. And that's a human rights video. But is it the government? Is it some terrorist group? And I don't think the video resolves this question on purpose because I think we see this already with cyber attack that we often don't know, or don't know for certain who the source of an attack is. We may have an initial idea of who is behind an attack, and then we may develop a more clear picture of that over time. And now we can say, look back on hacks that happened two years ago, and say, "Oh, yeah, we're pretty certain now that it was this group." But at the time, of course, you have to respond. You have to respond in terms of mitigating the attack, but also whether or not you should retaliate and who should you retaliate against. And that level of uncertainty makes that all the more difficult, which means it's much easier to sort of get away with an attack, if you will. And this is, again, like this idea that I think we could see very likely a lot more

assassinations, which again, are potentially not attributable. We saw an attempted assassination of the president of Venezuela by drone a few years ago, some small, explosive drones. And now, just strapping plastic explosives to DJI, hobbyist drones is incredibly common through Syria, and Iraq, where ISIS is still operating. They are using this a lot to try to attack US forces or other forces that are in the area. And that sort of thing is obviously very difficult to try to regulate, and international treaties already banned improvised explosive devices and things like that. So these groups aren't going to follow that. And I think this kind of comes back to a question you asked before, which is what is the treaty going to do to actually ensure that you're not coding something in the basement or designing a bunch of drones in secret or something. And I think that's where it really comes to states, and what states, one, feel is necessary for their own security or for international stability to enforce. And that's questions about verification. If I'm really afraid that you're going to build a robot army and invade me, then I need something that's going to reassure me that you're not. And what is that? What does that look like? If what I'm really afraid of is simply that you're going to use these things, and that use is going to be atrocious, then I can just wait until you use it, and then use the condemnation of the international community, and sanctions and things like that, after you've used the technology. And we've seen a mixture of these across the treaties around biological weapons, chemical weapons, nuclear weapons in the past, but chemical weapons don't have really strict regimes where we get to go in and inspect. It does require states to track the production of chemicals that have industrial uses but could also be turned into chemical weapons. But they're not prohibited from having those or using them for their normal uses. But if they use a chemical weapon, which even states like Syria that were using them on their own people, and were not party to that treaty, there was a huge international movement to condemn that and to hold them accountable and to get them to commit to destroying those weapons and not using them again. So I think there's a range of options.

We could spend so much more time on this but unfortunately, we don't have it. In the time we have remaining, I would like you to just tell us about a documentary you made a few years ago called *Love Machine*. Where can we find it? What does it tell us?

Right, *Love Machine*. So yeah, that was my transition I talked to you about earlier, where I kind of got into thinking about outside of intelligence as a human property, like what other essential human properties are we thinking about trying to replicate in machines? Of course, creativity people talk about a lot, but I got really interested in the embodied forms of humanity. So love, emotion, sexuality. And of course, there are all kinds of machines for sexual pleasure and things like that, but also people trying to develop robots and AI as companions, potentially replacing humans in terms of not just certain kinds of social relations or work relations, but in terms of intimate human relations, and friendship and love and companionship, and things like that. So it's a video that explores that. I talked to a lot of philosophers and roboticists. And that was shot in 1999, 2000, so 20 years ago. And I was thinking the 20th anniversary of the release is coming up and we were in discussions before the pandemic to do a re-reissue. And that was, of course, all before streaming was around, so thinking about getting it streaming. Right now, it's on DVD. You can email me if you want a copy, and I'll hook you up. But yeah, I think it needs to get out on the web, and more people should see it. And it's funny because a lot of the people in that

video that I interviewed then are still talking about those things today. And back then they were like, "Oh, yeah, it will be three to five years and we're going to have robots that do this, and that talk and do that." And 20 years later, we don't. The technology turns out to be much more difficult than people anticipate in certain respects, and in others, much easier. The idea that we have Alexa and Siri and these conversational agents that are pretty sophisticated, even if they are really just satisfying command requests, for the most part, they're not actually really having conversations. But that they're so ubiquitous, and that they're able to collect so much data from use to improve over time makes them much more effective. It's much harder when you get to the machinery - things that move. We don't have a robot that can stand at a sink and do dishes, right? But we can also just make dishwashers.

Right. But it still takes a lot of time to put things in and take them out. I speak from some experience, not as much as my wife's. So just in conclusion here, you've been doing a tremendous amount of research, thinking, teaching, writing, and filmmaking about the social, psychological impacts, and military aspects of artificial intelligence here. Where can people follow you, find more about what you're doing, learn about the things that you care about, and get involved?

Yes, so my website is peterasaro.org. And you'll find a link to my Twitter handle there. You'll also find all of my published papers are on my website. And actually, all the syllabi for all of my courses. And to the extent that the readings are publicly available, they are linked right there in the syllabi. So if you're interested in any of those topics, I probably taught a class about it at some point and you can find it there. And yeah, I'm on Twitter to some extent and that's about it for social media. I'm trying to limit that as much as possible. But yeah, I'm trying to figure out how to get that movie online eventually.

I'd love to see it myself. It's been fantastic having you on the show. The time has flown by. Maybe we could do this again sometime because there are so many more things we could talk about.

I'd be happy to come back.

Thank you, but for right now, Peter Asaro, thank you for coming on the show.

Thanks for having me.

That's the end of the interview. I know from the courses I've taught that people can get down from hearing about these sorts of concerns and risks, and I don't want that to happen to you; after all, we have enough to deal with with the pandemic at the moment, and our collective mental health is – well, another thing to get concerned about. This is starting to get very meta. No point in getting down about being down, though. And what people like Peter and I have learned to do, and what I want our listeners and my readers to do, is to treat this objectively, to look at it – well, kind of like an AI, if you will – logically, to take the way you feel about it and use it to motivate you and not hold you back. If you want to make a difference, like the title of a book that came out 20 years ago, you can't afford the luxury of a negative thought. The stakes are too important to let you be held back by being down. Peter focuses on the issues of these weapons every day and recognizes the importance of that. Similarly, Thomas Homer-

Dixon who was on episodes 25 and 26, wrote a whole book on the topic of staying effective in the face of these kinds of existential issues, a book called *Commanding Hope*. So my challenge to you is to find that way you can engage with the important issues that we face and make a difference to them without, so to speak, 'getting it on you.'

In today's news ripped from the headlines about AI, the battle between facial recognition and privacy is heating up as an AI tool has been released to anonymize your face. If your face is online it has probably already been imported into the AI deployed by the Clearview company, which made the results of a massive sweep of that data available to law enforcement agencies for searching, and attracted a great deal of attention and not a little outrage and lawsuits.

Now comes a startup called Generated Media, who can give you an image to use online that is based on your face but won't trigger facial recognition algorithms. It uses AI to take your image and make a face that looks kinda sorta like you, enough for people to recognize you, close enough so that someone meeting you on a blind date won't feel cheated, but to Clearview, will look like someone else. And you can get as many of these images as you like and they'll all look kinda like you but different from each other to facial recognition. Your move, Clearview.

In next week's episode, I will be talking with science fiction author and screenwriter David Gerrold. In s-f fandom David is a household word for his work especially on Star Trek, for which he wrote the landmark hilarious episode *The Trouble With Tribbles*, plus stories for the Animated Series, the Next Generation, and appeared as an extra in Star Trek: The Motion Picture. He has written award-winning science fiction novels such as The Man Who Folded Himself, which explores numerous permutations of time-travel paradoxes, and series including the *War Against the Chtorr*, about an alien invasion, and the *Dingilliad*, a Young Adult series about the adventures of the Dingillian family in space. It's mostly for his 1969 novel *When H.A.R.L.I.E. Was One*, and its substantial rewrite in 1988, though, that I'll be talking with him, because that novel examines the interaction between a newly-awakened artificial consciousness and its human creator, and how the conversations change them both. That's next week on *AI and You*.

Until then, remember: no matter how much computers learn how to do, it's how we come together as *humans* that matters.

http://aiandyou.net