Artificial Intelligence and the Changing Realities of the Virtual and Physical Worlds

KAREN WEATHERMON: Now to introduce our speaker. Dr. Taylor is the Allred Distinguished Professor in Artificial Intelligence in WSU's School of Electrical Engineering and Computer Science. He's also the director of the Intelligent Robot Learning Laboratory. As those titles might suggest, his research focuses on how agents, or robots, interact with their environments to learn new tasks, to perform multiple tasks, and to coordinate with other agents.

Dr. Taylor graduated magna cum laude with a double major in computer science and physics from Amherst College. After working as a software developer, he then earned his PhD in computer science from the University of Texas at Austin. Before arriving at WSU in 2013, Dr. Taylor also held research and teaching positions at UCLA and Lafayette College.

He's been the recipient of several honors, including a Google Faculty Research Award, and National Science Foundation Career Award, and a WSU Award for Excellence in Advising. Please help me welcome him this evening.

[APPLAUSE]

MATT TAYLOR: Great. Thank you for that introduction. We'll have to talk later about how many points I can get for giving a lecture during the comment. So I was really excited to be invited here. I'm a big nerd. And I love playing games. We're actually doing some research on gamification and how to use video games to get students to be more engaged. But what I want to talk about is artificial intelligence.

So it's been increasingly in the news the last few years, because people are worried that-- well, first of all, that the robots are all going to kill us. Other people are worried about that the robots are all going to take our jobs. So I wanted to spend some time talking a little bit about artificial intelligence and some of the things that I think might be interesting.

But at a high level, the idea of artificial intelligence can kind of be classified in four different ways. We want to get machines to do something. So what is that something? We could say, well, we want the machines to either think or act like humans-- so whether they're just doing something kind of virtually, or whether they're acting in the world and interacting with things around them.

Or do we want them to be like humans? Or, humans are slow, squishy, and suboptimal. Maybe we want them to be better than us. Maybe we want them to act or think rationally.

So one example I'd like to show you is actually from the 1960s. This came out of MIT. And this is the Eliza chat program. So let me see if I can make this a little bit bigger. There we go.

So, "Hello, I'm Eliza." "Hi, I'm Matt." And it thinks. "Did you come to me because you are Matt." "Sure." So Eliza is simulating a Rogerian psychotherapist. "Well, it all started back in the day."

So Eliza is going and basically coming up with these very generic prompts. "Can you elaborate on that?" And this was basically an exercise in natural language processing and people did this to have fun. And when they looked at the computer logs, they found out that when people would sometimes log on to the system and spend hours talking to this-- so this program is not intelligent at all. All it's doing, it's basically a script that will look at what you say and try to come up with things to get you to keep talking. But some people actually seem to find it useful, or at least, amusing.

So another example of something a little more recent is a chat bot so some of you might have interacted with these. There's one recent one called Cleverbot. And the goal of this chat bot, in the general sense, is to try to have a good conversation with you. But you could see how this might be really useful for businesses. So if you call your bank, if the bank, instead of going through those god-awful menus-- press 1 for such and such-- if you could actually have a conversation with the agent and get your things solved without having to talk to a human, that could save them tons of money.

So this one Cleverbot chat bot did pretty well. It was able to make reasonable-- have reasonable conversations with people. And then someone did a simple experiment.

So you can tell the state of the art is not exactly convincing. But this is one of the really interesting things about AI. I'm going to give you some examples where there are kind of spectacular failures. But there are also things that we can do really well right now.

And one of the difficult things is, when you're not an expert, understanding what's going to be hard and what's not going to be hard. So for instance, geotagging pictures and building up a map of where you went and took it took a bunch of photos with is pretty easy. But going and looking at a photo and saying, is there a bird, we can't do that. At least, we can't do it very accurately.

So a lot of the things we're interested in in AI are hard in some matter. Because if the problem was easy, then it wouldn't be AI. At least, that's how it often happens. So a lot of these problems, if you wanted to solve the problem perfectly, you'd need a ridiculous amount of compute, or a ridiculous amount of memory.

So a common thing would be, oh yeah, we'd need the same number of operations as particles in the universe. So there's no-- some of these problems we're tackling, there's no way to solve optimally, but we'd like to solve as well as possible. AI has been gaining a lot of traction in the last few years, in part because of advances in things like GPUs, Graphical Processing Units, and memory and computational advances. But still, we need better algorithms to get to more difficult problems.

And there's one class of problems that's informally called AI-complete, basically saying, if you could solve this type of problem, you could solve everything else. So here's an example of Deep Blue, the computer that beat the world's best chess champion. And then we saw, recently, IBM Watson beating some of the best people at Jeopardy! And even more recently, we saw AlphaGo program beat some of the world's best Go players. And these are all problems which are hard, but not nearly as hard as driving a car.

So throughout today, I want to mostly talk about agents. So we could think of Watson, which is just a question answering program. So all you're doing is throwing a question at it and seeing if it gets back a good answer. I'm more interested in things that interact with the environment. So you've got, for instance, a Roomba, which will go and clean for you. Or maybe you've got a program that can go and help you bid on eBay, or help you beat your favorite first-person shooter.

So all of these agents somehow have to sense the world around them and act. And one thing we'll be focusing on today is thinking about how they can reason. What does it mean for an agent to reason? How can they do that? What do we want them to do? And you can think of kind of a spectrum. So something very simple like a very simple thermostat or the Nest thermostat, it really just senses the temperature and acts to change the temperature, all the way up to something which, you could argue, is very similar to human.

So in artificial intelligence, we talk about rationality. We really would like-- often we say, well, we could have a agent mimic a human. But maybe we want them to be rational. So we say they've got some predefined goals, and we want them to try to maximize those goals, or maximize the likelihood of achieving them.

We don't typically care about the thought process behind them. So if our machine goes off and thinks for half an hour and comes back, if it comes back with the right answer, we're usually happy with that. But you think about these goals in terms of their utility.

So an easy mapping is often money. Think about, OK, here's how much you get paid. Let's think about what the best action is so I can get the most money. And you want to maximize your expected utility.

All right, so for this next step, I need a volunteer. Yeah. Perfect. OK, so in front of me, I have two envelopes. Which envelope would you like?

AUDIENCE: That one.

MATT TAYLOR: OK. So take it. Open it. I sealed it really well. OK, so what's inside?

AUDIENCE: $4.00.

MATT TAYLOR: Congratulations, you have $4.00. Now, you can either take that $4.00, or you can take this envelope, take whatever's inside. And I'll take that back. Now, the one thing I'll tell you is the envelopes have x and 2x dollars. So if that's $4.00, this either has $2.00 or $8.00 in it.

So how many people think he should keep the same envelope? How many people think he should switch? How many people think it makes no difference? OK, now, what would you like to do?

AUDIENCE: Um, I think I'll keep it so I don't have to rip through another envelope.

MATT TAYLOR: OK, thank you.

[CLAPPING]

OK, so that was good. So he also had-- his utility function was not just monetary. There was also some-- possibly some motivation to sit down and get back to the lecture instead of standing up here awkwardly with this weird guy. OK, so let's open this envelope. And we see, oh, this was actually $8.00.

Now, and here's an interesting thing. You could say, oh man, he just lost out on $4.00. Or you might be saying, he just got $4.00 for doing almost nothing. So that's kind of cool.

So it turns out the correct thing to do is actually switch. So this is kind of weird. So if you think about it this way. You've got some amount of money in your envelope Y. And you want to figure out what's in the other envelope.

OK, well there's a 50/50 chance it's either half of what I have or twice of what I have. So it could be half of $4.00, or it could be twice $4.00. And if I end up doing the math, it ends up being better to switch. So this is kind of counterintuitive. This is kind of like the Monty Hall problem. You pick a door, then you see a goat. And then you have to decide whether you change the door or pick a door. And before looking in that envelope, they were equally good. But getting this extra information tells you something.

The other thing is we're assuming that all values were equally likely. So it could have been that I put-- if he chose the envelope with $4.00, it could be I put $2.00 in the other one, or I put $8.00 in the other one. How likely is that? And part of that's going to depend on, how much do I want to make you guys enjoy this talk by buying your enjoyment. And how much money do I have to spend on making you want to enjoy this talk?

Anyway, so this is an example of, the rational thing was to switch envelopes. And it's not obvious. Hopefully, that was non-obvious to most people. So this is the kind of reasoning that we ultimately would really like our agents to be able to do.

OK, so I want my agents to do cool stuff. And then, there's this guy who has a big-- well, he's done a few things. You've probably heard of him. So Elon Musk says AI is the "biggest risk we face as a civilization." "With artificial intelligence, we are summoning the demon. In all these stories where there's a guy with the pentagram and the holy water, and it's like, yeah, he's sure he can control the demon. Doesn't work out." So that's cool. I mean, everyone likes being compared to a demon summoner, right?

So a lot of people in the AI community-- well, first of all, think he's wrong. He's not an AI expert. But it is a legitimate concern, right? Anytime you're introducing new technology, it can be used for good or bad. And it can have unintended consequences. So when people invented cars, they didn't immediately think of global pollution. But that's one of the results of that invention. And if we can at least understand, or think about, what might be coming, maybe we can work to make sure we don't-- that we achieve the outcomes we want as a society.

So one of the things that people often think about is, well, machine learning is scary. Once the-- so an example is the shoe-making robot. This robot goes, and it wants to make as many shoes as possible. So it builds this factory and starts producing all these shoes. And then these pesky humans come and take some of those shoes away. So the robot learns that it needs to kill the humans in order to produce more shoes.

So that's the kind of situation where, if you have something that the robot wants to maximize, really, we want to make sure that's the right thing. Because the machine is going to learn whatever you want it to. So one of the examples where learning is really useful is for these unanticipated situations.

So if I build a robot, there's only so many things I can think of that someone would want it to do. Whereas once you bring that robot home, if you can learn what you want it to do, then it can learn to be more useful. It can learn about your home instead of somebody else's home.

Learning can also be useful, because as much as I love everyone learning to program, programming can be slow. If a robot can learn how to do something rather than someone sitting there and coding away, that could be a lot faster. The other thing is, in machine learning-- and I'll talk about some examples of this in a minute-- you can actually learn to have superhuman performance. So you can sometimes come up with unexpected solutions that are actually better than what the human programmer envisioned.

Here's a simple example from my PhD advisor's lab. And what these robots are doing are they're moving back and forth across the field. And what they're doing is they're timing themselves. And their reward-- they're trying to learn to minimize their time. So what they can do is learn things like, how far should I step, how high should I step.

So you change these parameters. And over time, you can learn to walk faster. And the robots are just timing themselves as they're going back and forth. So they can fully autonomously learn how to go faster. So that's pretty cool.

And this type of learning, and in general, these agents, can work in at least three distinct settings. So I want to talk a little bit about the physical world, because you know, that's what we're stuck with living in most of the time. There's also augmented reality. Now, there's not a whole lot. How many people played Pokemon Go? OK, so a few people.

So one of the things, in Pokemon Go, you could do is, when you had it turned on, you could see the image of the grass or whatever and see the Pokemon superimposed on it. So that was-- and then, the other thing you could do is go to certain locations and get more Poke Balls or battle at a gym. So here it was combining the virtual and the physical. And there'll probably be more and more cases where we're combining the virtual and physical. And finally, there's the fully virtual reality, like OASIS in the book series.

So thinking first about the physical, it turns out robotics interacting with the physical world is a lot harder than a lot of people first think. So a few years ago, DARPA, a big defense agency, spent millions of dollars and got a bunch of teams to compete in this grand challenge where they would have these multimillion-dollar robots come down to California and do some complicated tasks. So this was going to take the state of the art robotics and push it a little bit further. And by and large, it was pretty successful. However, there were a few things that didn't go quite as planned.

Multiple schools that spent millions of dollars of research over multiple years, they come down with their really expensive robots. And there were some good things too. But when things go wrong, they go really wrong. And it turns out even something as simple as picking up this tape dispenser is actually relatively hard for a robot.

Because first of all, I had to look under here and see it. So there's not a whole lot of light here. There's a bunch of shadows. You need to figure out where to pick it up-- probably somewhere in the middle. Because if you pick it up on the side, it's going to fall down. And there's the question of, how hard do you grip it. If it's too weak, it slips. If it's too hard, you crush it.

So even simple things like picking up objects can be difficult for robots. And so when people say, you know, I'm worried about the robots coming and taking our jobs, those are not very complicated jobs at this point, let's put it that way.

[VIDEO PLAYBACK]

[ENGINE RUNNING]

Here's an example of something my lab's doing at WSU. And here, what we have is, this is a fully autonomous robot that goes around apple orchards. So If you've been to an apple orchard in the Pacific Northwest, you know they're often in very straight, regular rows. What this robot does is it can go through the orchard, and then pick up bins full of apples that weigh about 800 pounds, and then also bring empty bins for the human pickers.

So when you think of automating apple picking, the first thing you think of is getting the apples off the tree. But it turns out this is actually also a really useful thing to do. And it's a lot easier. Just for your reference, this machine costs about $75,000.

So at this point, again, it's still research. If we actually show it works, then we could go, and work, and try to make that price-- the price of it lower. And the reason we're looking at this specifically is, right now, in Washington State, we can't hire enough workers for apple picking.

If you're not using a robot, there's a little forklift than someone drives around the orchard that picks up these bins full of apples. And you pay somewhere around $20 an hour for driving a forklift around the orchard. And they have trouble getting people to fill that job.

So this is a case where robots--

[END PLAYBACK]

We're trying to get robots to take over the jobs that people don't want. And that's something we'll come back to. Now, so that that was successful in part because we were wheeled. And once you have wheeled robots, things are a lot easier.

One of the most impressive companies right now is Boston Dynamics outside of Boston, like you might expect. And here they've got a range of robots, from a couple of bipedal robots and also some really freaky dog-looking robots. But here is a recent video from CNN. Because I wanted to show you the robots actually working well too.

Complex robots that we can just bring into someone's house and work-- the only thing we really have is the Roomba, which can only vacuum. But hopefully, as we get more successes with these super-expensive robots, hopefully we'll see that filter down into more commercial and residential settings.

Another example of-- oh, the other thing I wanted to mention about robotics is it's computer science, it's mechanical engineering, it's electrical engineering, and it's also things like psychology and sociology. Because these robots need to be able to interact with humans. So robotics is one of those really interdisciplinary things if this is at all interesting to you.

Another way we're using robots at WSU is trying to get them to help with a smart home. So we've got increasing-- we have what's called the Silver Tsunami in the US. The population distribution is getting older and older. You have more adults that need help at home. Instead of just hiring a person, maybe you can have a robot that could help. And if we can keep people in their homes a little bit longer, they'll have better health outcomes, they'll be happier, and they should be able to save money.

But the prototype we're using is only about $4,000 and is a wheeled platform. So we ignore a lot of the mechanical issues with things like stairs.

So that was talking a little bit about the physical. Then there's the kind of in-between, the augmented reality. And there's actually some really interesting stuff happened. So did anyone else mess around with the Google Glass when it came out a few years ago? No? OK.

There was a whole lot of articles when it came out about Glass holes, and how people really didn't like the idea that someone could be filming them all the time. But it turns out Google Glass has kept going on, but it's mostly been for industrial settings. So things like, if you were repair-- I was talking to people at Xerox. If you want to repair a copier, it's very difficult. You need a lot of training to know how to do it. But if you could have a little video that's playing, reminding you, oh, check this, then check that, if you see this, then do this thing, then you can have much more effective repair people.

We're also talking with people in construction management about trying to detect trip hazards and point those out. Or if there is a live wire and there shouldn't be, if your device can sense that and then throw up a little alert, right, and highlight where you should not touch, that could be useful. What I'd really like is to be able to go to a party and have my computer, first of all, tell me their names, second of all, remind me if I know them. Because I've gone up to the same person three times over a few-month period and reintroduced myself, because I'm horrible with names.

But you could also think of why virtual reality would be useful, for instance, if you're on a bike, or if you're doing comparison shopping. So I think we're going to see a lot more of this type of augmented reality, both in video games and in the real world, which finally brings me to the virtual-- so when we're not dealing with the physical world. And one of the benefits here-- so as in the book, the architects for OASIS were able to control everything. So for instance, they had different physics in different regions.

Now, of course, one of the drawbacks is you have to control everything. So if you've ever written a program before, you know that you need to think about all of the things that might happen. And if you have a virtual world, there's a whole lot of things that can happen. And you need to figure out how to handle those situations.

So one of the things we've been thinking about is how you'd like to try to teach agents. So one of the things we were think-- saying, OK, if I have an agent-- and in this case, it's a virtual agent-- and I want to teach it to do something, maybe I could train it so that it could learn, kind of the same way I could train a dog to do something. So let me give you one example of this.

[VIDEO PLAYBACK]

- In this demo, the agent moves at constant fast speed at first. The first task the user needs to train the agent to complete is moving the blue chair to the purple room. At first--

[END PLAYBACK]

OK, so we're going to train this virtual agent, our dog, to move the blue chair to the purple room. And we're going to do that by providing rewards and punishments, just like you would train a dog.

[VIDEO PLAYBACK]

- We punish the agent since it moved to the green room without anything. Then it moves to the blue room without chair, so we punish it again. Then it starts to move the blue chair, so a reward is given. However, it stops at the green room. So now the punishment's given.

Finally, the agent moves to the purple room. Then we check whether the agent has successfully completed the command without any feedback based on last training. It shows that agent reaches the goal state.

Now we will experience another agent that moves at an adaptive speed.

[END PLAYBACK]

So you could also think, with dogs, there is some two-way communication. It's not just the dog doing its thing. You could also think, well, if the dog is confused, it might look at me more often, or it might go more slowly. So we also looked at saying, when the agent knows what it's doing, it should go faster. And when it doesn't know what it's doing, it should go slower, giving some implicit feedback to the trainer.

[VIDEO PLAYBACK]

- It adjusts its speed based on its confidence level of its current action choice. Recall, the other agent will still be trained in sensory environments presented in random order. The first task is moving the ramp back to the blue room.

The agent first move slowly since it is unsure about its current action choice at beginning. But after receiving several punishments for its incorrect behaviors, it moved faster. Then it becomes really sure about its current action choice and starts to move very quickly to the ramp back, so some rewards are given.

[INAUDIBLE] it successfully moves the ramp back to the blue room. Then, when the agent executes the same command the second time, it moves faster at beginning since it is more sure about its current action choice after last training.

[END PLAYBACK]

So hopefully, someday soon, you'll be able to train virtual agents to do things for you, maybe in video games, or physical agents to do things for you in the real world, which brings me to video games. So this has been of interest to me, in part, because I like video games. Some of the work that's been popular recently has been looking at, how can an agent learn to play StarCraft, or Pac-Man, or Mario. And it's just a hard problem. You think, OK, the agent can take all these actions, and it's trying to maximize this reward, this score that it gets. And there's been a lot of work recently on getting agents to be able to accomplish these difficult tasks.

I mentioned this recently-- or earlier-- where there's this program called AlphaGo from a Google company, which was able to be one of the world's Go champions-- actually, a few of the world's Go champions-- recently. So this was something that a lot of people in the AI field did think would happen in this decade, because Go is go is a very complex game. So the idea is you place stones on this large grid and try to capture areas. And it's very hard, because there's so many possible places to put these stones. But it turns out a new technology called deep learning allowed DeepMind, this Google company, to be able to beat world-class players.

Oh, if you're looking to make lots of money, you should learn about deep learning. So deep learning is a recent machine learning technique. And there's lots of people who get PhDs in it and make tons of money. There's also-- I know a number of-- I know a few cases where students just learn about it on their own and go and start companies with just an undergraduate degree. So good phrase to remember, "deep learning."

Other things about video games-- there's also the non-player characters. So I was thinking about, in the Ready Player One when-- what's his name-- Wade goes and battles the lich, right? And every time he goes up to him, it's exactly the same. And so that's an example of this non-player character, which is just really scripted and very simple.

And really, AI is starting to be used to make much more interesting non-player characters. So if you want to have a more-- a better experience, part of this is going to be having the environment learn from you. And part of that environment is going to be the storekeeper, right? It's very unlikely you're going to get someone who wants to play your game and play being the storekeeper. That's got to be the most boring game ever. But you want to have some kind of intelligence there, or it's not as much fun.

Another place where artificial intelligence is starting to be used is in automated development-- so figuring out, OK, it takes a long time to program a game. How can I use AI to help me? So one way of thinking about that is procedural generation. So in some games, you go, and the map is randomized for you.

So like, in Civilization, in some of the levels-- or some settings-- you can go and play on a random map. And this increases replayability. And it also-- in some sense, you could think it's saving the developers' time. Because they don't have to go and draw out a map by hand. They can come up with rules that can generate a fun map. And then you can play an infinite number of those.

This is really big right now. EA Is spending a lot of time and money on trying to detect cheaters. So when you have people-- for those of you who haven't heard of this, like, in first-person shooting games, you're running around, and you're trying to shoot people. And some players will try to download something which will help them aim, an aim bot. And that makes it more fun for that one person, but less fun for everyone else.

So in order to try to keep the game fun for everyone, they try to detect cheaters. But there's a very high cost. Because if you accuse someone of cheating who's not, they're going to get really upset with you. Even when you accuse someone of cheating who is, they're going to get upset with you.

So there's this idea of a Turing test, that, can a robot or agent act like a human so that you can't tell the difference. There's one interesting version of this in video games where you have an agent trying to play like a person. So think about this, if the agent is always shooting perfectly, then you're going to be able to tell it's an agent. So it's got to figure out how often to miss.

But this could be used in figuring out, how can I make automated enemies. So instead of playing against other people, can I play against agents that are interesting, that act similar to people, and who I can automatically turn the dial up? As I get better, my opponents get better. And if I'm having a really off day, they'll get worse. So these are some of the ways that AI is and will help video games.

So we were talking about rationality. And we want our robots to be rational. Do they want to maximize the reward, or their-- achieve their goals. Now, one thing is, rational is not necessarily the same as ethical. So when we think about, what should an agent do, what research should we do, maybe there's parts of AI that we shouldn't spend time on, like making really freaky-looking robots that pick people up out of beds.

So self-driving cars-- super popular right now. There's lots of pros and cons. There's lots of possible benefits to society, but there's also some costs. So when we think of, how is this going to help us, well, if you're drunk, be great to have your car bring you so you don't have to bother with an Uber. If you're blind, if you're a child, now you can have the car bring you places.

If your car can go park itself, maybe instead of parking itself, it can go help somebody else. So you could have multiple people share the same car. So you reduce traffic.

On the other hand, with these self-driving cars, you're going to be putting some people out of business-- so long-haul truckers, taxi drivers. So there is some societal cost to this new technology. Another thing to think about is, when these cars get in an accident, what happens. So there's definitely a cost. The legal system is going to have to figure out how to deal with that kind of cost.

Another example is autonomous weapons. So here, something not a lot of people know about, a gun by Samsung. First deployed in 2006. Autonomous gun sounds like a horrible idea, because you know, it just shoots whatever's moving. They deployed it at the South Korea-North Korea demilitarized zone. Because that's one of the few places on Earth where no human should ever be. And if there's a human there, you're allowed to shoot them.

But when you think about it, autonomous weapons, well, they could be good, right? If you deploy autonomous weapons, then our soldiers aren't getting shot at. A nice thing about them, robots, they can shoot second, right? If you see something which might be an enemy, wait until it shoots at you. Then, once it shoots at you, you know you can return fire.

However, a number-- I'm not sure whether it's most-- people in artificial intelligence think this is not a research direction we should pursue. Because it will develop into an arms race. So now we're going to be spending lots of money on something which-- of questionable ethics. There's a pretty low barrier to starting conflict. So like with nuclear weapons, you need a lot of money, a lot of resources in order to build a nuclear bomb, whereas getting an automated weapon could be relatively cheap.

Oh, and if people aren't getting shot at, politicians may be more eager or more willing to start a conflict. So again, technology could be good, could be bad. It's probably worth thinking about before we decide to go down this path.

So in general, robotics, but AI in general, is good for automating things that we don't want people to do or people don't want to do-- so things that are dirty, dangerous, dull. A lot of things, particularly in agriculture, there aren't enough people to do the jobs we currently have. So let's see if we can automate some of these jobs. And hopefully there will be different types of jobs created, like robot engineer, and will lead to a higher quality life because of these robots.

This is by iRobot, the company that makes the Roomba vacuum. But this iRobot robot was deployed in the Fukushima nuclear disaster. So that's a good example of a place where no human should go, because the radiation will kill them. Let's get the robots to help us out.

The last thing I'll mention is one thing, I think, that is a problem. So I don't think robots are going to take over the world and take all our jobs or kill us anytime soon. One thing in AI that we could be worried about-- a couple of things that I want to just touch on are, right now, data is becoming more and more synonymous with power. So the more data someone has, the more things they can do. The more data Facebook has about you, the better they can manipulate you.

So when you think of Google, Facebook, Amazon, these huge companies, and we're not necessarily the consumers, they want our data to do different things so that they can sell it to marketers. So there's risk there. I don't know of anything that's egregious that's happening, but we need to recognize that it could be a problem.

There's been ongoing rumors, for the past year or two, that the Facebook app listens to you when it's not on and then shows you ads of what you've been talking about. And this was just an article three days ago, that Facebook still denies it. And from what I've read, it's probably completely false. Facebook is probably not listening to you. But the fact that they could be, and the fact that it's really hard for us to show they aren't, is kind of scary.

The other thing I'll mention is, these data-driven algorithms can be racist. So for instance, recently, Microsoft had a chat bot. Remember Cleverbot. This was their version. And their version was Tay. Tay was a teenage girl, and she was awesome. And she learned through talking with people on Twitter.

So "Can I just say that I'm stoked to meet you? Humans are super cool." "Chill, I'm a nice person. I just hate everybody." "I effin' ate-- I effin' hate feminists, and they should all die and burn in Hell." "Hitler was right. I hate the Jews."

So Microsoft's chat bot went full-on Nazi within about a day, because people on Twitter are horrible, right? This is the internet. They should have realized that people are awful. But this is an example. Tay was learning from the data that was fed her.

Another example is this recent article six months ago from ProPublica looking at-- there's this one company that sells an algorithm which will go and give judges a score-- how likely is this person to commit a crime again? And OK, so we can make better sentencing decisions. And ProPublica did this expose basically showing that the algorithm was more-- for very similar crimes, the algorithm was more likely to say a black guy was going to commit the crime again than a white guy.

And when they went and looked at the data, it turns out that when-- it wasn't true. Go figure. So the algorithm learned from the data. But if the data is flawed, even if you're not training, even if their intent was not to be racist, depending on what your data is, you can come up with bad outcomes. So that's something to be aware of.

So where are we going in the future? We'd like to have better transparency and trust. If I get a low score from that algorithm, I want it to be able to tell me why it thinks I'm going to commit a crime. And in fact, in the EU, there's this law coming up where you get a right to an explanation.

So this would be great. If you have a low credit score, the credit bureau should be able to tell you why your score is low. Or in 2001, if Hal says, I'm sorry, Dave, I'm afraid I can't do that, and I have to kill you, it would be nice if it at least gave you a justification.

[VIDEO PLAYBACK]

All right, so--

[MOTOR RUNNING]

--with that, I'll just finish up and say, if this was at all interesting, I encourage you to take computer science classes. There is a robotics club on campus. They meet regularly-- roboticseecs.wsu.edu. And one other club I'll highlight is there is a robo sub club. So again, like the robotics club, they take all majors.

[VIDEO PLAYBACK]

And they go and build an autonomous submarine that goes and competes once a year down in San Diego. So here's a video of it-- of this sub autonomously going through the start gate and then performing a few tasks. So if you think AI or robotics is interesting, I encourage you to get involved.

[END PLAYBACK]

And both the robotics club and robo sub have lots of tutorials. So if you don't know anything, that's totally fine. Hopefully you'd find this interesting though.

So with that, I'll wrap up and say thanks for listening. I'll put these slides up at irll.eecs.wsu.edu if you want. So thank you. And I'd be happy to take questions now. Yeah?

AUDIENCE: What got you first interested in [INAUDIBLE]?

MATT TAYLOR: Ooh, what got me first interested? Probably because I kind of see life as an optimization problem, right? So like, if you think, I've got to get to work-- so of course, you should be going and figuring out what's the fastest route to get there. Or I've got to do this thing on my computer. Oh, I don't want to have to type this whole thing out again. How can I get the computer to do it for me-- so thinking about how I was motivated to get computers to help people to do things faster and better. Yeah?

AUDIENCE: Do you read the webcomic xkcd?

MATT TAYLOR: I do read xkcd. It's a good comic.

AUDIENCE: Yes, it is.

MATT TAYLOR: How do I tell if people online had questions on the--

AUDIENCE: [INAUDIBLE] .

MATT TAYLOR: Oh, it's you. OK, awesome.

AUDIENCE: So the live stream had a question-- they were wondering about if you could re-explain deep learning.

MATT TAYLOR: Yeah, so I completely ignored deep learning. So neural networks are roughly based on the human brain, and they started in the '70s. And now we have better neural networks which use modern computation to do things like learn direct-- how to play video games directly from pixels. So if people are interested in deep learning, there's actually a Coursera class on it and a few other online classes. Or they could take the graduate course in computer science.

All right, well, thanks for coming. And hopefully you found this interesting. And if you are interested in AI, hopefully you find ways to either go through these clubs, or online, or through classes. So thanks again.

[APPLAUSE]