AI and You

Transcript

Guest: Missy Cummings

Episode 147

First Aired: Monday, April 10, 2023

Hello, and welcome to episode 147! My guest today is the extraordinary Missy Cummings, whose resume reads like an action movie screenplay. In 1988 she became one of the US Navy's first female fighter pilots – yes, we're talking *Top Gun* here, the real thing – and she was so methodical about pursuing that goal that in high school she interned for her congressman, because she knew that she needed a congressional nomination to enter the Naval Academy later. After flying fighters, she got a PhD in space systems engineering. She later opened the Human Autonomy Laboratory at Duke University – yes, that spells HAL, that was intentional – and recently spent a year as a safety advisor for the National Highway Traffic Safety Administration, which is when her comments about Tesla received a great amount of attention, and you'll hear more about that in the interview. She just moved to George Mason University in the Department of Computer Science, running the Mason Autonomy and Robotics Center.

There are a lot of terms that get mentioned here, and one I'll just explain in advance for anyone who's not heard of it is LIDAR, which is a version of radar that uses light instead, in or near the visible spectrum, to map out the 3-D environment around the vehicle by bouncing beams off objects. Using light instead of radio makes the mapping much higher resolution than you can get with radar.

So let's get into the interview.

Missy Cummings, welcome to the show.

Thank you for having me.

And so your personal history, it reads like an overachieving science fiction novel that you were one of the first female fighter pilots in the Navy and irresistibly evoking images of *Top Gun*. And you were actually there two years after the movie came out. Is that right?

That's right. Indeed, *Top Gun* was a big motivator for me to make me want to become a fighter pilot.

And were they talking about that there? Did they do quotes from the movie around the school?

Oh yeah. It was all anyone could talk about and indeed, those few years right after *Top Gun*, they were extremely competitive to become fighter pilots because everybody wanted to do it and you got selected basically based on how well you had done in school.

So that trope in the movie of pilots running on excessive testosterone, how accurate is that?

Oh, I think it's pretty accurate, including women.

Okay. Now, it's not all as glamorous as it looks, of course. And you talked once about an incident that happened when one of your fellow pilots was killed and that was something that you were saying would happen on average once a month. And so the one that I read about, and maybe you can fill in the details, was where the aircraft flipped upside down and killed the pilot. Can you tell us that story? Because I think that there's something instructive in that.

Right. Particularly when we think about autonomy today, in those days, fly-by-wire was still relatively new and there was a flight control reset button. Literally, there was one button that if you had any problems with your flight controls, since they were fly-by-wire, it's almost like your blue screen of death. You would hit the reset button which would reset within a couple of seconds. But one of the problems that they had at the time that they didn't know was there was a failure mode: when you hit the reset button, the software would zero out, but the surfaces, the flight control surfaces of the aircraft would stay. They would remain in the place that you froze them. And so when the plane's system rebooted it thought it was in a zero position, and then it would move the flight control systems thinking they were in the zero position. But indeed, they were already severely, in some cases, at their limits. And so that is what happened multiple times. During my time flying fighters, I knew of two people that died. One person who was in training with me, and then another person, I didn't know him personally, but he was obviously flying F-18s off the front of a carrier. And the failure mode would happen. The flight control computer would be reset by the pilot, and then the plane would sense some kind of need to change motion. And then the plane would do a very serious yaw, flip upside down immediately and there's just zero chance. A human just doesn't have the reaction time to be able to pull that out. And for my compatriot, the day that happened for him was the day that all the wives and girlfriends had been invited to watch their significant others do this, these little touch-and-goes. And so she witnessed his death. So pretty, pretty compelling.

Yes. And that must have had a huge effect on you, but what was the outcome of the inquiry?

Well, then they figured it out. Of course, in any kind of accident and investigation, whether commercial or military, instantly the pilot is blamed. And then they started doing some digging and then realized that this flight control defect would happen. And this was kind of an early canary in the coal mine of how software is a lot harder to gauge its correctness, in whether or not you've fixed a problem or that you even have a problem. And so we did eventually uncover this problem. Several people died because of it. And it was early days and - I'm sorry he had to die, but that's what the military often does, right? They are often the proving ground for experimental technologies before we actually get them into the civilian world. Fly-by-wire was perfected in the military before it kind of made its way over to commercial airlines. And so I think that that's the way it should be. I think if we actually fast forward to today, and that's 30 years later, I think what's different is, it's this weird inversion where really advanced technologies are completely bypassing the research and development that they would get in the military and going straight to the commercial market. And indeed, this is so true of self-driving cars. We have not had the ability to work out the kinks in the higher-risk environments where people are highly trained and in theory, are there to prevent such accidents. Now we're just putting them into the hands of your everyday user and the outcomes are not good.

Right. A lot to unpack there. I've been through enough National Transportation Safety Board incident reports to have a sense of their methodology for approaching corrective actions and seen that they only conclude pilot error in a case where there is a corrective action that can be applied to have it not happen in the future. And so that would usually look like training or maybe selection. But that if that wasn't something that could be concluded from there, then the decision should not be that it was pilot error; instead, it should be something like a design error, maybe human interfaces. In the case that you were describing, superficially, it sounded to me more like a design error. What was the eventual outcome? Was it something where they said we needed to train the pilots better, or we need to redesign the system?

Oh, no, it was a software coding error and so they fixed it. But I do agree that accident analyses today are much better than they were back in the '90s. And the military then, and probably still to some degree today, just tends to hold the human accountable. So that's not all that surprising back then, but eventually, they were able to figure out that that problem was indeed a coding error. But it's layered; this is the Swiss cheese layers of accidents that yes, there was a software coding error that needed to be fixed, but there was a test and evaluation error that the company did not figure out a good enough test practice. The government did not do good acceptance testing to make sure. And then that was not the first accident that that had happened, then we didn't do good follow-up after other accidents with similar backgrounds. But I'll tell you what's so interesting about that accident. So that happened in 1995, and then most recently there was a whistleblower at TuSimple who released a video to the internet to let them see an autonomous trucking accident. Almost the exact same failure mode, 30 years later. And I think what concerns me is that the surface transportation community has not had the time (probably), or the wherewithal, meaning assigning people, to actually learn these lessons. So the surface transportation community and the autonomous world, they're learning a lot of the same lessons that aviation learned 30 years ago. And I find that very concerning, because people don't need to die. You just need to make sure that you're doing due diligence and have a safety division of whatever company that you've got to make sure you're not making these same mistakes.

There's a lot to unpack with that as well. And one of the things that I'm wondering is the NTSB seems to have a fairly rigorous methodology for approaching improving aircraft reliability. Does the National Highway Transportation Safety Administration have a similar ethos with respect to this? You were there until recently, am I right?

Right. Well, I think it's important to recognize that these agencies do two different things. So NTSB is an accident investigation group, and their job is to, in a very detailed manner, investigate accidents and make recommendations to various groups - policy groups, design, what have you. The National Highway Traffic Safety Administration is a regulatory agency. So its job is to make sure that companies are adhering to the letter of the law for putting whatever technologies out on the road and maybe making new laws and recalling vehicles when problems emerge. So they do have some accident investigation capability, but indeed it's not their only purpose. It's not their sole purpose. And I would actually say both NTSB and NHTSA, they both have a gap in their capability. They have the gap of having the right people investigate these technologies. So I don't always agree with everything that NTSB says. I think they're great, but I

think that they, like many other government agencies, including the Department of Defense, Department of Transportation, Department of Homeland Security, fill in all the three-letter identifiers after that; they don't have enough people who really fundamentally understand the core nature of artificial intelligence. And if you don't have people on staff who understand these technologies, you don't know how to test for them, you don't know how to find their weaknesses, you don't know how to keep an eye on. Are the systems that we're acquiring, for example, am I being sold a bad bill of goods? Have I asked for the right testing? Am I not repeating that 1995 problem of agreeing to take on a software system that is fundamentally fatal?

And again, with the parallels, the FAA has enough bureaucracy to ensure that no plane takes to the sky until it's accumulated its weight in paperwork. Does the Department of Transportation, I mean, if you look at the parallels with vehicles, are we behind in that respect? And is it possible that the development of autonomy and other technology in vehicles is going to exceed the rate at which regulation can keep up?

Oh, I think it already has. You just have to look at the Boeing 737 Max. That wasn't advanced AI, that was just baby AI. That was very basic and the FAA missed that. And so I definitely think the FAA is struggling to figure out how to think about air taxis, for example. So the Jobys of the world, the Liliums, they're aircraft companies proposing to have - effectively drones - fly humans around with no pilot in them. Well, that's the wild, wild west.

What could possibly go wrong?

Exactly. So, no, I really fundamentally believe that the US government - and they're not the only government that's really behind the eight ball in having capable people who really understand the fundamentals of these technologies - but I think that just because of our capitalistic market and innovation, and how much effort we put on that, that we are a country that's kind of at a crossroads, and we are basically the world's testing ground for how we're going to deploy autonomous technologies. Indeed, all of these problems are one of the reasons I moved from Duke University to George Mason in the Washington DC area, it's because I'm starting new degree programs that are focused on AI design and evaluation. And indeed, you can get a certificate, a master's degree, one day a PhD. So we're really looking at training a new workforce that doesn't necessarily need to be a computer scientist, a dedicated computer scientist, as much as they need to be technically competent people who understand how the technology is designed, built, so that they can evaluate it and also understand the public policy ramifications.

And while we're talking about your career now here, I want to link that to where you came from because you got a degree in mathematics at the same time you were flying fighter jets for the Navy. Is that right?

Oh, well, I'm glad you think I'm so capable all at once. No, I got my undergrad at the Naval Academy in mathematics, and then I graduated from Annapolis and then went to flight school.

Okay. They both showed up as 1988 in my research.

Well, that's when I finished one and started the other one.

Out of curiosity, how many of your fellow fighter pilots went into professorships and robotics?

There's just a handful of us and I could argue, not even that. It's very rare for a military person to leave and then become a full-fledged, I would say, an R1 university professor at the big research institutes. It's not a common skill set. I wish it were more common, and I wish the military would encourage more of that engagement because I think one of the reasons I've been successful is because I have a very clear understanding of how this technology affects people when it actually gets deployed. And I spend a significant part of my time working on teams of research proposals trying to reel people back in for like, you know, you can't propose that technology because it would have this kind of effect downstream. Or do you really want to do that as opposed to trying to tailor it to support people in another way? And this goes back to the DARPA Urban Grand Challenge. There's just this big push to replace people. And I get potentially the bottom line from a how much money are we spending on pilots? For example, if we could get rid of all pilots, then we would save a lot of money on training, insurance, et cetera. I get that. But I also think that people's thinking is very naive and superficial because even if you get rid of, for example, a truck driver, you're going to have to have someone else in a dispatch center somewhere else monitoring this. And indeed, those people who are the monitors of automation, whether it's in aviation or in surface transportation, they're a higher skill set. They get paid more because they're going to do more. And so there may be an ultimate cost savings, and I'm a big fan of automating, for example, rail. We need to do more of that in the United States. But it's just naive and you're making false promises if you think you're just going to completely wipe out an entire class of people and get rid of their jobs. And we see this right now with ChatGPT. ChatGPT is my third rail. I'm so sick of hearing all the hype over ChatGPT. ChatGPT is not going to replace anyone's job. Any company that thinks that you're going to replace someone with ChatGPT, I want you to declare that so I can deinvest in you because it means to me that you have no idea what you're doing with ChatGPT, you don't understand the problems that ChatGPT introduces, and now you're going to have to have even smarter, more clever people to figure out when ChatGPT is doing something wrong.

Correct. In addition to the cost savings, the benefit that's touted for automation of vehicles is safety, like the goal of making a car 10 times safer than a human driver. How far away are we from that? What's in the way?

So I'm going to call those claims performative safety claims. I think that in the past 10 to 15 years, there's been this huge push into Silicon Valley with the collapse of Silicon Valley Bank. We're starting to see exactly how much of that was built on a house of cards. So this big push of innovation, all this venture capital money out there, people looking for cool applications to invest money in, I think that a lot of people like the idea of a car driving itself. I personally would love to not have to drive a car around the Washington DC area. I'd like to be able to get in the back and get some work done. But the reality is that we could not be further away from having these cars be safer than humans. I don't care what any manufacturer wants to claim. I have yet to see any evidence that autonomy in cars that does lateral and longitudinal control is providing any safety benefits. I personally see the opposite. And I make a distinction there because there is

automation in cars that's excellent. Automatic emergency braking is very good. Blind spot warning detection, forward collision warning. I mean, there are some good, automated technologies, but I'm drawing the line between automated and autonomous. And for me, for the purposes of this conversation, autonomous is when neural networks show up.

Trying to find the truth about this has been so difficult because of the amount of money being thrown around, that the phrase that came to my mind was "fog of war," that I felt like I was trying to pull apart or peer through something that was deliberate smoke screening. Now, there are companies in China deploying what appear to be level 4, if their videos are accurate, autonomy in certain environments. And in the US, we've got some of the same companies and others in very limited environments, like Phoenix, again, deploying automated taxi services where there's no one in the driver's seat. There's a lot of qualification around that. But also, a lot of people - principally, executives of those companies - would like to say that this is the camel's nose in the tent and it's not so long before we get the whole thing. What do you think about what they're doing? Because they have invested a huge amount of money in this. Are they about to hit a reckoning?

So there's two parts of that question that I want to address. First is the China red herring. Oh, please, come on now. If China were so far along, especially in the climate that we've got now, if China could do level 4 or level 5 autonomous vehicle operations for real, meaning not a faked video, why don't we see that in operation here? Why haven't they applied for the operational permits in California? What we do see is Pony.ai, a Chinese company. They have been shut down by the state of California for unsafe business practices for well over a year now. And TuSimple had the crash that we spoke about earlier which was really kind of a shocking level of incompetence in coding and TuSimple is hanging by a thread in terms of viability as a company. So there's just no proof in the pudding. Anybody can fake a video. If China were really that good, they would be more than happy to show off their technical capabilities here in America, and they just haven't. So I don't think you can believe anything that you see in terms of a video for self-driving out of China. But then the question is, all right, but in America, we do see Waymo and Cruise making some inroads. And I've ridden in Waymo's cars and they've been very gracious to me, they've been very open, I've been inside their dispatch centers. So I think Waymo is genuinely trying very hard to do the right thing. I don't think that Waymo's trying to do any sleight of hand. That being said, Waymo is in Arizona, and they've just started operations in California but I don't see the breadth of operations that would make this scalable. And I think that this is actually what's really important is these technologies can work well in relatively narrow settings. And I'm not saying that we shouldn't have them. Indeed, I think the winner of this race will be those people who can figure out the operational domains that make the most sense, that are the most safe and provide the most value. I'm a big believer in last mile grocery delivery. If we can have the Nuro model, I think that's actually a pretty good model. You could imagine there's a big draw for that. The vehicles move very slowly in limited areas that are very well mapped. Good. But I think if you really want to know how scalable this is, so the California DMV has their companies put out annual reports about miles driven and disengagement and collisions and whatnot. And you can go see that Cruise, their cars are generating miles. I think

they have over maybe 250 plus cars that could be driverless. But I think the last time I looked at the data, there were only around 20 of the cars that were actually generating more than 10,000 miles a year. So if you have 20 cars generating 10,000 miles each, that's 200,000 miles. Even if you said a dollar a mile, which I think would be a good rough estimate, okay, so you've made $200,000. Even if you charge $2, you know, how much more would you need to charge to make that scalable? So I'm starting to see that technically, in some areas, this is doable but does it actually make sense, especially when you have to, as I said before, layer in the dispatch services? You have to have highly trained people. The state of California, or at least the city of San Francisco is up in arms over all the vehicle retrieval events and the blocking of emergency services that those vehicles have caused. So I think we're starting to move into the territory of, okay, it's technically possible in some limited areas, but is it feasible?

Do you think that this hinges to any degree on more mature definitions of operational design domains? It seems that even if you said that the only certified areas for autonomous operation were interstates, that there might still be a business case for trucking companies to run 18-wheelers in automated platoons across the country from on-ramp to off-ramp.

So I think that's the big question mark right now. Cruise and Waymo are having limited success in the urban environments where the cars are moving relatively slow. And then when there are collisions, they don't generally lead to serious injuries or fatalities. All you have to do is start looking at cars with driving assist, i.e. Teslas, to start seeing what happens when these cars crash at high speed. Then we start to see fatalities click in. And so, there's a big difference between a car moving at 65 miles an hour and at 30 miles an hour. And then if you put the weight of a truck behind that, we've got a serious potential for not just the deaths of one or two people, but an entire highway, lots of people. So I think that the reason that we haven't seen more progress there is not only is the risk higher when you're moving at higher speeds, but the computation and projection of correct action is also significantly higher. And I think that the typical sensor suite of LiDAR, radar, and computer vision; there's still major gaps in that. LiDAR, for example, does terrible when there's moisture in the air. Anywhere from rain to just mist in the air, that can destroy a LiDAR image. And I think one of the hidden gotchas that we're going to start seeing emerging in the next one to two years as the real measure of how companies are doing are the hard breaking events. These are the phantom braking events. We saw the Tesla pile up in the Bay Bridge Tunnel in November of 2022 on Thanksgiving Day in America. And fortunately, no one died but that wasn't just a Tesla problem. All the car companies are having these phantom braking events, even companies with LiDAR attached to them. And I think that this has become really the technical albatross behind people's necks, this phantom braking issue. And so I've been encouraging companies to quit being so hyper-competitive and start admitting this and maybe coming together. I think it could be solvable, but I think we would get to that solution faster and more comprehensively if we could have companies work together on the phantom braking problem instead of pretending that it (a) doesn't happen and (b) that it's not their problem.

That's the end of the first half of the interview; the rest will drop next week. It's a privilege to have someone on the show with such direct experience of the critical interactions between machines and the

physical world in such a challenging and dangerous environment as jet fighter dogfighting, and in Missy we have someone who not only has moved into an academic discipline that gives her the chops to reinterpret her fighter pilot experience in a whole new light, but who is also so uninhibited about giving us the straight dope on what she's learned and how it affects the buildout of systems in transportation that – well, that we're looking forward to so much, and if some of what she said runs counter to prevailing narratives from pundits about where we are in the state of the art in vehicle autonomy, I think it's important to be aware of whether we're projecting our desires onto our predictions. Professional poker players and investors know about this very well. Now, it's good to be aspirational. It's good to have ambition and lofty goals; and grand visions are what has driven human progress. But we shouldn't conflate our idea of what we *want* the world to be with our perception of what it *is,* and Missy definitely points out for me the importance of not cluttering perception with hope. Maybe it's because it's not just poker players and investors that can't afford to do that, but also fighter pilots who have to deal with the world the way it really is.

One of the companies Missy mentioned was TuSimple, which is a company developing autonomous trucks, and if you want to look it up it's spelled TuSImple. We also talked about Chinese AV operators, and if you've listened to the show much you probably remember me talking about a video from the company AutoX, which appears to be doing level 4 autonomy, the definition of which is that no involvement from a driver is required in certain environments, and they have no one in the driving seat, and their environment is parts of the city of Shenzhen. They are certainly very easy parts for a car to navigate. You might want to look at the video and decide what you think; just search YouTube for "AutoX." It's really striking how fraught with conflicting and inadequate information the space of vehicle autonomy is.

In today's news ripped from the headlines about AI, the White House's Office of Science and Technology Policy released a blueprint for "An AI Bill of Rights." The OSTP has identified five principles that should guide the design, use, and deployment of automated systems to protect the American public in the age of artificial intelligence. Those principles are Safe and Effective Systems, Algorithmic Discrimination Protections (which is what policymakers call preventing bias), Data Privacy, Notice and Explanation, and Human Alternatives, Consideration, and Fallback. Those last two call for a bit of elaboration. "Notice and Explanation" is headlined, "You should know that an automated system is being used and understand how and why it contributes to outcomes that impact you." "Human Alternatives, Consideration, and Fallback" is introduced with "You should be able to opt out, where appropriate, and have access to a person who can quickly consider and remedy problems you encounter." Fairly basic categories there, nothing contentious or likely to be as restrictive as the European Union's AI Act, which we discussed in episodes 139 and 140.

Next week, we will conclude the interview with Missy Cummings, when you'll hear what Missy thinks about Tesla, ChatGPT, and Boston Dynamics; the truth behind the dogfighting AI you may have heard of, the possibility of complete automation of air travel, how AI would handle air emergencies, and more. 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