Transcript

6 March 2025,

Interviewer 1:03
So this case study scenario. Let me just. We go back and read this, OK, so this case study scenario involves a real life case study within the AI application of automated vehicles, which I'll call AVs.
OK, so it involves the occurrences of actual car crashes involving one particular AV brand, which is Tesla and it's advanced driver assistance systems (ADAS) called Autopilot. So Tesla's Autopilot system controls the steering, braking and acceleration functions of the AV without any assistance from the human driver. Furthermore, note that Autopilot could at any time disengage and hand over the controls to the human driver.
So according to USA NHTSA, which is National Highway transport…er… Traffic Safety Administration's Office of Defects Investigation. Between January 2018 and January 2022, Tesla AVs, with Autopilot engaged were, involved in 16 crashes where they struck highly visible stationary in road or roadside first responder vehicles. So these are police, ambulance, fire, trucks, road maintenance vehicles, you know with lights flashing, people with vests that were attending to pre-existing collision scenes so they.

Stakeholder8_AV Driver 2:31
Yeah.
Yep.

Interviewer 2:38
The car just kept driving, and then on average, in these crashes, Autopilot aborted vehicle control less than one second prior to the first impact. OK, so this is the scenario across America, 16 different crashes and the reason NHTSA opened this particular engineering analysis investigation was because of the pattern of.

Stakeholder8_AV Driver 2:48
Yep.

Interviewer 3:05
… You know, striking, highly visible stationary vehicles, right and….

Stakeholder8_AV Driver 3:11
Yeah.

Interviewer 3:12
Basically like you know, like Tesla's, with or without AV control, have you know, like other car manufacturers have crashes, etcetera, but this is an official investigation that was opened up and you know, looking into this. So it provides a nice case study scenario to probe people on, you know, the concept of explanation from the human perspective.

Stakeholder8_AV Driver 3:39
Yeah.

Interviewer 3:40
If you want I can put this text into the chat so you can reference it if you want, or you're OK with the verbal description.

Stakeholder8_AV Driver 3:50
I'm OK with the verbal.

Interviewer 3:51
OK, I also have a website where I've basically attached photographs of the traffic's scenes, if you will. The crash site if you want that also. Yeah. OK, so let me just.

Stakeholder8_AV Driver 4:03
Sure, sure.

Interviewer 4:08
I compiled it about a year ago based on… there's a crash… There's a bunch of basically from newspaper reports and TV reports, so there is also there is also. Video of. … of a Tesla driving up to a scene like this that someone got from, you

know, Wall Street Journal or Washington Post or something like that. But anyway, it's, you know, it's a … it's just a visual reminder. Some people want to see this, some don't. OK.

Stakeholder8_AV Driver 4:28
Yeah. Yep. Yeah, yeah, that's good. It's helpful. See, what's the environment?

Interviewer 4:42
So based on this scenario, OK, that I've just described to you, "What kind of explanatory information?" OK, the main question is: "What kind of explanatory information are you seeking about these crashes from Autopilot, the system that was doing the motional control functions of this vehicle?" OK, so when you ask a why question, why did this accident happen? What kinds of information are … do… do you have in your mind?

Stakeholder8_AV Driver 5:10
Yeah. Yeah, sure.
There's a couple of things that come to mind. First one is… is I've…. I've observed I use Autopilot every day, but I every once in a while I observe it doing things that make me say it's not ready for real time … full time. Maybe two weeks ago, there I… I ran into a scenario where just on an evening I'm driving on a clear day and there's a car in front of me and I just observed that it wasn't slowing down like it normally does the majority of the time. And I had to intervene and brake. Why? Why? I don't know. And so the questions I ask myself is it training? So the analogy I like to use is. You know when… when… AV a year ago, or Autopilot a year ago versus Autopilot today, Autopilot today is in a … in a much better state. Is it because a year ago the analogy is it's you're it's a 16-year-old and now it's a 25-year-old? It's got more time, it's got more experience, it's got more data. Is that what it's about? So I observed it, for example a week ago, not stop at a railroad crossing, even though the … the arms are down and the lights are flashing. Is it because it's missing data to stop? Is it because the cameras were blinded by flashing lights? Is there some sort of camera obstruction? I know in the past it's asked me to take over when it's really sunny outside and you can, you know, the sun is blinding your eyes, so it must be blinding the front camera as well. And … and it'll know. It'll know it's being blinded and asks me to take over. So the … the why's for… for that crash are along those lines. Is there

... is it... is it... is it camera obstruction? It... it ... was it... a really sunny day? Is it a lack of data on how to handle vehicles parked? I've also observed it not stop at a ... so you've got a bus, a bus with a the... the stop extended out flights lights are flashing, ... it's slowing down. I've had to take over. Is it because it's lacking the data too? Is it lacking the training? I don't know how it works to stop for bus stops.
And I guess what makes it opaque is I've observed it do something wrong one day and then the next day, the same intersection, the same manoeuvre was done right.

Interviewer 7:59
OK, so let's go back you you've talked about, let's say, the vision functions, OK, of the vehicle. So let's assume that all of the hardware is working accurately the....

Stakeholder8_AV Driver 8:15
Yeah.

Interviewer 8:15
You know the cameras, et cetera. So if we assume all the hardware, all the other systems on the Tesla are working fine, what questions do you have about the perception functions or vision functions? The decisions it's making about acceleration, steering and braking? So can you probe into those aspects, those particular functions?

Stakeholder8_AV Driver 8:44
As it relates to these crashes?

Interviewer 8:46
Yes, and even your experience using Autopilot or FSD.

Stakeholder8_AV Driver 8:52
Yeah. So I... I again, I... I observe it with pretty good consistency. If I'm driving on a one way on a two way, so one way each way. So on a two-way road, which I do almost every day and let's say there's a car parked on the side stopped so it could be like they're doing road work. And so it's half of the car is on the street and the other half is on the side. And you can have somebody standing there with that slow stop sign.

Interviewer 9:31
Right.

Stakeholder8_AV Driver 9:31
I've observed it recognise that, cross the median even though it's a solid yellow line. I've observed it manoeuvre, to go around that obstruction and …and drive. So I … I observe it to doing ….doing that with consistency. But I don't observe it if that…. let's say there's a person standing there with a stop sign or a slow sign. I don't observe it follow the direction of slowing down. I'll continue to… I'm going 60 kilometres an hour, someone's there standing with a slow sign, it'll go around the person, it'll continue to drive at 60 rather than slowing down at 30. So the question, if you're telling me that … that it's a … there's no visual issues, hardware issues. The question that always comes to my mind is … is it a lack of training? And the reason I always go to that is because I've observed it getting better and I assume that's because of training data. But I just don't know where the training comes from. Is it coming from … the… the camera recording all the data. Is it coming from?

If I take over there's an it gives us the ability to press a button and give it feedback. I don't know where the feedback goes. Does it go to AI? Does it go to someone who's actually a human observer who's taking the data and doing something with it. That's always been that black box for me. Where's that data going?

Interviewer 11:00
So, assuming it's getting the data somehow into that, whether it's a human taking the data and making alterations to the algorithm, the software itself or the software's doing it itself. What questions do you have about the decisions it's making and the actions it's taking between the input and output? What questions would you ask about that aspect?

Stakeholder8_AV Driver 11:28
And … and just make it clear for me, is it, is it we're if we're using the these case studies and as an as an example?

Interviewer 11:33
Yes, yes. For these particular case studies, what questions do you have for these 16 case studies? For these 16 crashes, this pattern of 16 crashes, in particular, what questions do you have about the decisions made? And actions taken or not? Between input and output. If it helps you.

Stakeholder8_AV Driver 11:59
Well, yeah, yeah.

Interviewer 12:02
Imagine, a human did this, went across America, right? Crash 16 times. Right? What would you ask that human? What would you ask?

Stakeholder8_AV Driver 12:04
All right, right. Yeah, yeah. Yeah. So I mean, like, you know, it's I think it's why didn't you? I'm kind of putting myself in the shoes of like, parent teaching my daughter how to drive. Well, first it would be: “Why didn't you slow down?” And again that's that's... that's literally the text message I sent to my friend. I sent him the video of that again from a week or two ago where it just did it. So I was like, why didn't it slow down and then and then it ...it ...was it weather related?

Interviewer 12:38
Mm hmm.

Stakeholder8_AV Driver 12:39
Was it the visual real? Was it camera related?

Interviewer 12:43
Let's assume all of those answers were no. So keep focusing on the inside the thinking part. You know the decision part, the action part. Deciding to take action, deciding not to take action. Ask from that narrow perspective.

Stakeholder8_AV Driver 12:46
So that's. Assuming everything, all the variables are perfect.

Interviewer 13:04
Yes. Yeah. The motion controlled questions about the motion control functions.

Stakeholder8_AV Driver 13:05
And that's... that's the tough part for me because then it's just really why, if you've been trained.

Interviewer 13:17
So assume the eyes are functioning properly, but it's just what the thing inside is doing with the data it receives, you know?

Stakeholder8_AV Driver 13:20
Yeah. I mean, so again, I'll try to put myself in the shoes of ...as a pilot. And I tend to listen to a lot of, like NTSB investigations. And, you know, you got the TV show Mayday. A lot of you know the end of the day or they say something like 50% of plane crashes are pilot error and so you have everything in the plane working properly. But it could be as simple as fatigue. Pilot had a got into a fight with their wife the night before, didn't sleep properly, got into the plane exhausted and as a result of pilot.. pilot fatigue, all of these issues happened. That's not the first thing that comes to mind when I think of AI like.

Interviewer 13:45
Mm hmm mm.
Right.

Stakeholder8_AV Driver 14:07
Is ... is AI ... is the algorithm fatigue? But I guess that's a interesting thought. Can it... can it...Can it get fatigued? I've never thought about it. I've always just thought, hey, it's a programmable robot that follows a set of instructions. And so if everything is working properly, why did you not stop so?

Interviewer 14:26
Yeah, but it's making decisions, right. Remember. OK, you can take the human aspects of fatigue out of this. And what you have is an entity. That is ... It's a driver, OK, it and so.

Stakeholder8_AV Driver 14:33
Yes. Yeah.

Interviewer 14:42
It is performing the human tasks of controlling the vehicle, so it's making decisions. It's making decisions we're not privy to why it made the decision. So what kind of information are you looking for from this thing?

Stakeholder8_AV Driver 14:46
Yes..

Interviewer 15:00
Like, it's not fatigued. It's a machine. It's not fatigued, but it's doing human functions. Without a human in the loop. So what questions do you have for it?

Stakeholder8_AV Driver 15:01
Sure, I'll get, I'll, I'll get, I'll sure.
Sure. Yeah.
So I'll start off, I guess I'll go with just what came off to the top of my head that I don't really think about, but it's one of the reasons why I always keep my hand on the steering wheel is are you are you trying to commit suicide or you know if we're gonna' get dark here.

Interviewer 15:21
Yes.

Stakeholder8_AV Driver 15:25
Did you did this algorithm, all of a sudden, just have a because it's trained, it has all the data, it knows that people kill themselves. It knows of... I'm assuming it knows of the concept of suicide. ... did the did the thought of suicide just come into your head AI? And you've ... you've rammed into the car. It's not something I think about as a daily driver, but it's something that's always in the back of my head. As in what ... that's why I can't relax behind Autopilot in its current state. So. So that I guess that would be one of the why's.

Interviewer 15:37
Mm hmm. OK, what else? Let's assume it's not suicidal.

Stakeholder8_AV Driver 16:01
And again. We're making this hard now, 'cause it's not whether it's not camera, it's not suicide.

Interviewer 16:06
Yes.

Stakeholder8_AV Driver 16:11
You know, it's. Why didn't you slow down? Don't know why. Why didn't you go around?

Interviewer 16:16
Mm hmm mm hmm.

Stakeholder8_AV Driver 16:19
We don't. Why?

Interviewer 16:20
Mm hmm.

Stakeholder8_AV Driver 16:21
So you've got all the perfect variables. And…. What else would I ask it if I was talking to it? If it's, if…. If it's telling me, hey has nothing to do with the weather, has nothing to do with camera obstruction, has nothing to do with weird thoughts in the algorithm.

Interviewer 16:42
Right. So it let's… let's recall the scenario, so 16 crashes pattern of 16 crashes into highly visible vehicles. So you could see it from a distance.

Stakeholder8_AV Driver 16:50

Yes.

Interviewer 16:59
It kept driving, kept driving, kept driving, didn't slow down, didn't brake, didn't steer differently, just kept driving towards the object and then less than a second before the crash it released control.

Stakeholder8_AV Driver 17:15
How have you been trained to deal with seeing those objects on the road?

Interviewer 17:21
OK.

Stakeholder8_AV Driver 17:22
Have you been trained to ... is there, ... is there training for how to deal with stopped police cars and stop ambulances and stop fire trucks? Yes or no?

Interviewer 17:32
OK,

Stakeholder8_AV Driver 17:34
If the answer is no, I think that answers it. If the answer's yes, which I presume the answer is yes, there's been training. Why have you? What is the training?

Interviewer 17:43
Mm hmm.

Stakeholder8_AV Driver 17:43
Is the training to be hey, is there a set of rules? What are those rules? Are the rules slow down, pass slowly.

Interviewer 17:48
Right.

Stakeholder8_AV Driver 17:57
If, … if… if again if there if there is a set of rules, why didn't you follow the rules?

Interviewer 18:03
Mm hmm. Mm hmm mm hmm.
Good. Thank you. OK, I know I'm making this difficult, so I'm gonna give you a little bit of a break. We're gonna go to the secondary questions and then we'll come back to this. If you think of anything else, OK, so.

{Secondary Questions and General Discussion}

Stakeholder8_AV Driver 35:37
Going back to the first question, I wrote something down.

Interviewer 35:39
Yeah, yeah.

Stakeholder8_AV Driver 35:41
You know … the why something that I noticed happens and I don't understand is have you regressed? So, there's … there's a concept called regression where there's a software update and AI the Autopilot behaves better than it did before, but then it completely regresses in other areas that it was it used to do well in. And I had again, zero transparency, zero understanding why, zero understanding of why, how that even happens.

Interviewer 35:50
Yes.

Stakeholder8_AV Driver 36:07
And so you so….My question is so going back to 16 crashes, I have to imagine that it's been in a scenario way more than 16 times where it's successfully done the right thing. Is… is it a matter of regression in those sixteen cases and why? Why does the regression happen? I think that's important for me to understand. I'm that's, yeah. So,

on the road every day there's got it. ...it's... it's got to have success. It's, you know, if I had to guess the success rate I got to imagine it's north of 90%, I'm guessing. So why did it fail in this very ... very few times and you know it's... it's got to be zero.

Interviewer 36:57
You know, I ... I love that last question. Truly. I'll tell you every single one of you guys that I interviewed. You're giving me new perspectives. So absolutely. Thank you. OK. Are there any other explanatory information or questions you have for Autopilot or Tesla's engineers, you know the ... the guys who came up with these algorithms, like any, any questions about the features or functions or methodologies that you have questions about?

Stakeholder8_AV Driver 37:35
It just for the engineers. Again, it just goes back to how did, like, how's the feedback loop work? That's always my number one question, if I'm disengaging and I'm giving feedback, where is that feedback going? Is... is it? Is it just going straight into the neural net for training? Is it just listening to my words? So that's ... that's something that I've never heard an engineer really explain. I ... I think in the past they, from my understanding it was it when it was programming an engineer would take the feedback, watch the computer data. And then reprogram.

Interviewer 38:11
Yeah, that's the old fashioned algorithms these are yes, there is feedback and it does get fed back into the input and it's told the outputs are accurate or more accurate or less accurate. I'm... I'm giving you a very top level abstract description of how machine learning happens and reinforcement learning and recurrent neural nets and so on, and how these algorithms are trained but engineers aren't getting into the nitty gritty code. They're ... It's just not possible because it's all stochastic, so it's.

Stakeholder8_AV Driver 38:51
Can't, can't, can't. Can, are. Are there rules? Can it make its own?

Interviewer 38:52
Yes, it ...it... it is.

Stakeholder8_AV Driver 38:57
And it all of a sudden. So if if it's trained to not crash into fire trucks, can it all of a sudden on a whim, change its mind and train itself that it's OK to do that like one of the things that comes up is let's you're... you're in a dangerous situation and It's ... it's ... It's deciding between your life and having to step on the gas and run someone over. Save your life. It could be a criminal. Somebody with a gun, they've got a gun pointed at you. They're about to start firing in your car. Will the AI say it's OK for you to step to gas and hit that person? Because that's what you have to do. Or will it stop it because it recognises there's a person? That's a question. Me and my friends ask ourselves, like, what is it thinking?

Interviewer 39:23
Mm hmm.

Stakeholder8_AV Driver 39:46
How would it behave in a scenario like that? We're not committing murder. It's defence self defence. But what is it in the eyes of AI? You're running somebody over and it's trained not to run somebody over, right?

Interviewer 39:58
It's not trained for those kinds of decisions. The short answer, the short answer, it's not trained for those kinds of decisions.

Stakeholder8_AV Driver 40:02
Right.

Interviewer 40:09
It's not making those kinds. It's not making those kinds of, ...it's not making those kinds of decisions. It's making decisions to accelerate, decelerate, steer based on all of the data it's received from other human drivers. So the multitude of all that data goes into these algorithms and they just fine tune.

Stakeholder8_AV Driver 40:09
But will it stop you as a human driver? Will it stop?

Interviewer 40:33
It's trained on basic traffic signals, and you know, rules of the road, et cetera, but it's not trained on the ethics of individual decisions. Humans are. We think in those terms. But these algorithms do not think in those terms.

Stakeholder8_AV Driver 40:51
Now or for now, for now.

Interviewer 40:55
AI is not one monolithic thing, so the AI that does facial recognition cannot do. Driving functions, the AI that does the driving functions cannot do facial recognition. The AI that plays the game of go against the world's champion cannot drive a car and cannot do facial recognition. None of those can do the recommender systems for your Netflix saying, watch this movie next versus that one. None of these. Like and, these AI cannot do those. You know those robots, those dancing robots that you see on YouTube? OK, different functions, different input data, different training sets, different types of algorithms. A human being can walk and talk and chew gum and drive a car and play Go, and dance and, you know, recommend movies to their friends and have these ethical conversations. There is no one AI that can do all of that.

Stakeholder8_AV Driver 41:45
Yeah.
Got it.

Interviewer 42:07
OK. So everyone that's talking about the whole Skynet Terminator scenario.

Stakeholder8_AV Driver 42:12
Yeah, that's what comes to my mind.

Interviewer 42:14
No, we're not even close. Remember these AI you have first hand experience with automated vehicles. They're not even close. They're not even at level 2. OK, a 16-

year-old is performing at level 3-4 and five OK. A 16-year-old driver with a driver's permit. These AI are just at level 2 right now.

Stakeholder8_AV Driver 42:40
Right. But I've observed it. And so if you were to ask me today, Sameir, do you see a future of driverless taxis? And if I had to pick an answer, there's only yes or no, there's no in betweens. I would lean towards yes and the only reason I would lean towards yes is because I've just... just observing the changes over time. You go back to how many years ago to two years ago to a year ago and to how it's performing today and you can see there's an evolution.

Interviewer 42:41
Yeah. Agree.

Stakeholder8_AV Driver 43:05
And you just multiply that down the road and you're like, yeah, I can see. I can see that.

Interviewer 43:11
I... I agree with you. I agree with you. I do believe the technology will get there. It's just not as fast as people want, right. And... and also just so you know about those driverless taxis you hear about in San Francisco, the Waymo cars, very small subset of San Francisco, a few blocks only.

Stakeholder8_AV Driver 43:35
Yes, that's right.

Interviewer 43:37
It's memorised all of the static terrain, the weather is great.

Stakeholder8_AV Driver 43:41
Right, right.

Interviewer 43:45

Not snow slush, et cetera. Visibility is good and the only variables are the pedestrians and you know that … that it's ….so it's a very….

Stakeholder8_AV Driver 43:47
Right, exactly.

Interviewer 43:55
Constrained environment for test purposes. Remember that those driverless.

Stakeholder8_AV Driver 44:00
Right.
Interviewer 44:02
Level 4, three or four that it's at so.

Stakeholder8_AV Driver 44:06
Yes.

Interviewer 44:08
The world is very big, much … more bigger than a few blocks in San Francisco, so.

Stakeholder8_AV Driver 44:11
Yes.
OK.

Interviewer 44:15
So when you're driverless and people go, but San Francisco and I'm going, OK, this square block.

Stakeholder8_AV Driver 44:16
Yeah.

Interviewer 44:22
And then the whole world, right and.

Stakeholder8_AV Driver 44:23
I think that's kinda like the blue cruise as well, where it's just dedicated same we can only it's not go anywhere it wants it's it can only work on specific and and and another thing that just came to mind with it with those 16 crashes, it's like what about the driver like why?

Interviewer 44:30
Right.

Stakeholder8_AV Driver 44:40
Again, was it was it trust is it. Was it just? Hey, I trusted the the system to do its thing again, I....I ...I'll tell you.

Interviewer 44:46
You know the root cause of why they were distracted, that's not in the output report. The output report simply says the expectation and understanding of its capabilities was here, and the reality was over here. OK, big gap.

Stakeholder8_AV Driver 44:51
Got it. I'll tell you something interesting.

Interviewer 45:03
So yeah. So they've made Tesla put more alerts. All these software updates here automatically getting is because they've forced Tesla to put more alerts and the driver is tested more and they're taking away.

Stakeholder8_AV Driver 45:10
Yes. Yeah, yeah, yeah.

Interviewer 45:20
The FSD capabilities automatically if the driver's too distracted, and they're not. You know they let it go, so test right, so Tesla is.

Stakeholder8_AV Driver 45:24
OK. Yes, that's been a big change. I can't even look at my screen to change music anymore. Just looking for a second. It'll tell me. Pay attention. And that it wasn't like that a year ago.

Interviewer 45:39
Right, so this is in response to all of the crashes and all the concerns they're having. So that's the gap. And So what Tesla's responded is not about focusing on the inside of these non transparent boxes. They're focusing on the driver responsibility and forcing the driver to pay more attention until the technology gets to where you know, you're saying it needs to get to.

Stakeholder8_AV Driver 45:45
Yeah.
Right.

End Transcription for analysis .... secondary questions and general discussion continued until 64:33 when Interviewer stopped recording and transcription