Transcript

11 March 2025,

Interviewer 0:41
So thank you. First of all, I appreciate you participating in this and I just want to remind you the purpose of this interview is to gather in depth information on the subject of explanation from the perspective of human beings. OK, so this isn't a test for you. Like the I'm only gathering inputs, OK?

Stakeholder11_PolicyMaker 0:43
No worries.
OK.
OK.

Interviewer 1:05
No right or wrong answers, just your views, your perspectives, your thoughts and questions about a real life scenario. Hey, this is it. Anything you do is the best, OK, so.

Stakeholder11_PolicyMaker 1:09
OK.
Yeah, I'll do my best.

Interviewer 1:21
Let's just … I'm gonna' start with the description itself of the of the scenario that we're gonna' talk about. So it involves a real life case within the AI application domain of automated vehicles, I'm going to use the term AVs in short, so it involves the occurrences of actual car crashes involving one particular AV brand, Tesla, and its advanced Driver Assistance system, ADAS, called Autopilot. OK, so Tesla's Autopilot controls the steering, the braking, acceleration functions of the AV without any assistance from the human driver.

Stakeholder11_PolicyMaker 1:38
OK.
OK.

Interviewer 2:07
So it can just do it without the human in the loop. OK. Furthermore, note that Autopilot could at any time disengage and hand over the controls to the human driver now. According to USA NHTSA, National Highway Traffic Safety Administration's Office of Defects Investigation between January 2018 and January 2022, so a four year period, Tesla A VS with Autopilot engaged, were involved in 16 crashes where they struck highly visible stationary in road or roadside first responder vehicles that were attending to pre-existing collision scenes. So these are police, ambulance, fire trucks, maintenance, road maintenance vehicles, lights flashing, people with vests on all that now. Furthermore, on average in these crashes Autopilot aborted vehicle control less than one second prior to first impact. OK, less than a second average … average in these 16. So if you want, I can copy and paste the text in the chat or you're OK with this?

Stakeholder11_PolicyMaker 3:08
One second prior, well less than a second. Wow.
Yeah, I'm fine with this. Yeah.

Interviewer 3:22
OK, there is also a little website where I put.. reports and actual photographs of these crash sites, if you want to see it in Tesla video footage, but it's up to you. OK?

Stakeholder11_PolicyMaker 3:33
No, it's OK. No, I don't see it. That's good.
Yeah.

Interviewer 3:58
OK. So context wise, my research, as I explained in the background material involves the topic of explanation. And as you know, we humans are providing explanations to everyone, each other, every day, right about, why? We're asking Why questions. Why did our friends get divorced? Why is the economy tanking? Why did the plane crash? et cetera. And depending on the situation and the person, the answer and questions are all going to be different. All of us are, even if we're the same situation, different people with their different perspectives are going to ask for different explanatory

information. So what I'm doing is gathering inputs from different walks of life, different people, different lived experiences about this one incident.

Stakeholder11_PolicyMaker 3:59
Yeah.

Interviewer 4:23
To help AI developers develop more human centric and more tailored explainable AI. These are all black boxes that we're talking about, where even the engineers who design them don't know what's going on inside of these opaque systems that are making these decisions and operating in the world doing human tasks. OK, so, an initiative about 10 years ago from DARPA was saying, hey, this is a big deal. Let's start. It's called explainable AI, or xAI. So this is where my research is situated at the moment. So let's talk about this now. The first what I'm gonna do is ask one main question and then some secondary questions about the stakeholder group you belong to. So based on the scenario I've just described, you are seeking explanatory information when you're thinking of the why questions of why this car crash happened. You're seeking specific information about these car crashes from autopilot, the system that was doing all the motion control. So what kind... What specific questions would you ask of Autopilot?
Stakeholder11_PolicyMaker 5:30
Right. Sorry I lost you. I lost the connection. I didn't hear the question.

{Connection glitches}

....
Interviewer 7:59
So based on the description of the case study, you're looking for explanatory information. You know, when you ask why did this car crash happen, you have certain things in mind. What kind of information are you seeking from Autopilot itself? The thing that was doing the driving?

Stakeholder11_PolicyMaker 8:03
Yep. Yeah. What had happened? Yeah, yeah, yeah. Yeah. So, like, why didn't they see the .. the vehicles that were parked in front? Why didn't they see the humans?

Trying to think what else I would ask besides being… What speed were we going at the time? Were we within the speed limit? And were you following too close to the you know the car in front of you. So you didn't detect what was in front of you in terms of you know that there was maybe a car in front of you, I don't know.
I don't think is that….

Interviewer 9:02
Yep, you're headed in the right direction. Keep going. Yeah, all valid questions. So keep going. This is what this is very important. Yeah. Yeah.

Stakeholder11_PolicyMaker 9:04
OK. So… so we're talking about that there was an emergency vehicle there already with people. So why didn't they see the flashing lights? Why didn't they detect that? Were there warning signs that you know, even before you got? Because if … if emergency vehicles were already there, this crash must have happened way before the what… What did you say… The one second that they need or is it 16 seconds they need in order to engage the OR disengage the autopilot?

Interviewer 9:44
They can disengage at any point and

Stakeholder11_PolicyMaker 9:46
Right.

Interviewer 9:51
are you asking about human's ability to take over? Or are you?

Stakeholder11_PolicyMaker 10:12
Yeah. Well, that's the other thing, if there's a human being in the driver's, why wasn't the human being going, oh my gosh, flashing lights, let me take over, let me push the brake. So, what happened? Was the human being reading the newspaper? Why did the human being not disengage the autopilot? Was it not allowed? Was there, like, a there? Was there a glitch in terms of the transference from Autopilot to the human being like it's …it…?

Interviewer 10:27
No. OK, so let's, let's talk about the human to Autopilot takeover. And first of all, a human can take over at any instant. OK. Autopilot never overrides the human. The human overrides the …

Stakeholder11_PolicyMaker 10:31
Yeah. Any time, OK. … And what's the …what's? And it's within a second you said.

Interviewer 10:46
Yeah. And what happened, and OK, so what happened is we don't have information about what the driver was doing. So distracted in the realm of reading, sleeping. I don't know, you know, watching TV, whatever.

Stakeholder11_PolicyMaker 10:59
Yeah. OK.

Interviewer 11:03
So obviously the driver wasn't alert. Let's you know….

Stakeholder11_PolicyMaker 11:08
Assumption, yeah.

Interviewer 11:18
So from this scenario, Despite that the car just kept driving, driving didn't alter speed, didn't alter direction, just kept going and then less than a second to go so….
Stakeholder11_PolicyMaker 11:20
I'm done, ma'am.

Interviewer 11:22
Didn't… didn't say it, you know. Just release. That's it. Assuming you know, so the question here, you're right like what was the driver doing but … for …for the sake of understanding the decisions made or not.

Stakeholder11_PolicyMaker 11:37
Yeah.

Interviewer 11:39
….Actions taken or not, by this entity. What questions do you have about the driving functions? You know you've talked about the vision functions and I guess recognition functions of the objects around it, but what about the steering, braking and acceleration functions? Do you have questions about those and…?

Stakeholder11_PolicyMaker 11:52
Right, yeah. Well, I yeah. And that's the… where I was going at going is that so were they going the speed limit, did they attempt to like, did the autopilot attempt to break? I guess what I don't understand is how so the, the autopilot…

Interviewer 12:13
It did not. It did not break.

Stakeholder11_PolicyMaker 12:16
Did not break, so it just. Interesting. So I… I… I guess I know nothing about this stuff. So I guess if oh, I'm not sure, but I… I think if… OK. So I'm putting myself in that situation in terms I'm the human being responsible for this car.

Interviewer 12:26
Which is perfect, which is perfect.

Stakeholder11_PolicyMaker 12:37
So yeah, OK, given that that the driver might have been distracted doing whatever the car was, the main brain of the operation at this point. So why?
Why did it wait till the? So why did it disengage in the 1st place? The human didn't ask to take control over the autopilot. They …there must be some kind of command required. Is there command required to say OK I'm taking over so the human being that so who makes that decision?

Interviewer 13:04
They just… they can grab the steering wheel and it happens.
Stakeholder11_PolicyMaker 13:07
OK so.

Interviewer 13:09
They can hit the brake and it happens. It just automatically happens.

Stakeholder11_PolicyMaker 13:11
So, so, but in this case, in this case, the car made the decision to release control. So why did they? Why did it do that? What was what was in the… its,… its AV brain. Yeah. AV brain. What was it …was it was it ….like, why did it release?

Interviewer 13:16
Yes, yes. Processing. Yes, yes.

Stakeholder11_PolicyMaker 13:31
Because you said it released like less than a second before the impact. So why did it release? What? What triggered it to release?

Interviewer 13:36
Yes.

Stakeholder11_PolicyMaker 13:40
So was it so I can't even begin to conjecture. So you know it's, it's and how big does an object have to be? So is it so… Did they see the when they were approaching the the ….the emergency vehicles where they …that… they see a tiny little thing, you know how your side mirrors you say the objects appear larger than they are or smaller than they are. I can't even remember but it's. … It's …is it like that …that it seems like really far away and it's like 'cause it can't see obviously. So it has to sense everything I suppose. So what was the object? So did it think it was still safe? And so when did it not think it was safe anymore for its human being? So they released. But what? How did they come to that decision to release Autopilot?

Interviewer 14:40

Right. So you're asking for what kind of information it used to make the decision to release?

Stakeholder11_PolicyMaker 14:43
Yeah, yeah, yeah.
I think.

Interviewer 14:50
Yeah, that's good. OK, so the thing controlling the steering, braking acceleration, you're asking about the decision, what did it see or didn't see, recognise or didn't recognise you're asking?

Stakeholder11_PolicyMaker 14:59
Yeah.
Interviewer 15:03
Why? What made it release when it did? Yeah.

Stakeholder11_PolicyMaker 15:06
Hmm. What? Yeah, because the… the, the human didn't engage it. The human didn't put the hands on the steering wheel. I don't… didn't do voice control. Say, hey, I'm taking over. So what? What made it released? So it made some kind of a decision. Based on all this information that they had. So what was that information?

Interviewer 15:10
Yeah.
And you want to know which information? Yeah, OK. It made the decision.

Stakeholder11_PolicyMaker 15:27
Yeah, which information and yeah and….

Interviewer 15:30
Any other questions about the steering, braking or acceleration functions?

Stakeholder11_PolicyMaker 15:36
Well, I think the acceleration is… is really about like where they were they going the

speed limit where and like did they continue going the speed that they were going at the time before the crash?

Interviewer 15:51
The answer is yes to both, yes.

Stakeholder11_PolicyMaker 15:53
So they were going the …the same speed, but were they going above the speed limit?

Interviewer 16:00
No, Tesla doesn't do that. Yeah, these cars don't do that. They follow.

Stakeholder11_PolicyMaker 16:09
They don't allow it. Yeah, that's right. That's that's true. I'm not sure what kind of question to ask about the steering and in in terms of and the acceleration you said that they were going the same speed the whole time. So if it's programmed to go the same speed, I'm assuming they didn't accelerate. Did they accelerate at all?

Interviewer
No, no.

Stakeholder11_PolicyMaker
So it was that they didn't somehow sense the object.

Stakeholder11_PolicyMaker 16:33
And so like I said, maybe they thought that they were like several 100 feet or metres away from it and. So was that the case? What information that they have to determine that they were still a safe distance. You know. And if that helps.

Interviewer 16:52
OK, so yeah, definitely everything helps. So I'm gonna' give you a little bit of break and go to the secondary questions and then we'll double back. If you think of something else. OK, so.

Stakeholder11_PolicyMaker 16:59

OK.

Sure. OK.

End Transcription for analysis .... secondary questions and general discussion continued until 56:16 when Interviewer stopped recording and transcription