Transcript

13 February 2025, 11:53pm

Interviewer 0:03
And it's wonderful. So, um.
OK, so first of all, thank you. I received your um... consent form, your signed consent form so you consent to be interviewed and recorded for this study.

Stakeholder1_Engineer 0:16
Of course.
Absolutely.
Interviewer 0:17
OK. Thank you. Yeah. You emailed it to me on February 9th. I received it. It's filed officially.

Interviewer 0:26
OK, so this just to OK. So have you gone over that background information paper that I sent to you?

Stakeholder1_Engineer 0:27
uh .... OK.

Interviewer 0:40
I sent you 2 attachments, the consent form and the yeah. OK, so the background kinda situates you in what explainable AI is a little bit about my research and how I'm gonna use your data to support my final dissertation. Pete, for the PhD. OK, So what I'm gonna do for this interview is I'm going to give you.

Stakeholder1_Engineer 0:41
Yeah. Yeah, absolutely. Yeah. Yeah, yeah.

Interviewer 0:42
OK, so the background kinda situates you in what explainable AI is a little bit about my research and how I'm gonna use your data to support my final dissertation. Pete, for the PhD. OK, So what I'm gonna do for this interview is I'm going to give you a scenario of a real life example of incidents with automated vehicles.

.

Stakeholder1_Engineer 1:13
Mm hmm.

Interviewer 1:14
And so these are where the dr... uh... the car drives. You know, the AI algorithm drives the car, essentially , and what happened. And I'd like your inputs on what your thoughts are, your views, your opinions, what kind of questions you have about the decisions made and the actions taken by the AI itself. So what I'm going to do is I'm going to put in the chat. Um...I'm gonna put in the actual scenario so you can reference it if you need to, OK.

Stakeholder1_Engineer 1:52
Yeah.

Interviewer 1:53
In in the chat as well, and then we'll go through it. Here we go copy.
Chat. Here we go.
Paste. Here you go. OK, so.
Let me just go back to, I hope we're transcribing.
OK, here we go. Yes, we are transcribing. OK, so. Um...
So here's the scenario. We'll go through that first and then we'll go through the questionnaire. So this case study uh...scenario involves a real life case within the AI application of automated vehicles or AVs. It involves the occurrences of actual car crashes involving one particular AV brand which is Tesla and its Advanced Driver Assistance System, ADAS, it's called Autopilot, so Tesla's Autopilot system controls the steering, the braking and acceleration functions of the AV without any assistance from the human driver. So the car can drive without the human doing anything. Furthermore, note that Autopilot could at any time disengage and hand over the controls to the human driver.

Stakeholder1_Engineer 2:49
Uh huh. Yeah.

Interviewer 3:14
So according to USA NHTSA's Office of Defects Investigation, so NHTSA is the National Highway Transportation Safety um...

Stakeholder1_Engineer 3:26
Agency, yeah.

Interviewer 3:26
Agency. Yeah. And they opened engineering analysis, 22-002 about um...incidents involving Tesla, AVs between January 2018 and January 2022.
So this EA 22002 involves Tesla AVs with Autopilot engaged. They were involved in 16 crashes where they struck highly visible stationary in road or roadside first responder vehicles. So, police, ambulance, fire, trucks, road maintenance vehicles, they were that were attending to pre-existing collision scenes so you know , bright lights flashing and all of that.

Stakeholder1_Engineer 4:14
Ok .... Yeah, yeah.

Interviewer 4:16
Furthermore, on average in these crashes, Autopilot aborted. Vehicle control less than one second prior to first impact, so the driver had to take over within you know, one second.
Stakeholder1_Engineer 4:26
Wow

Interviewer 4:27
So this is the scenario if you wish ... there are reports of these incidents. There are photographs and videos if you want. I can also in the chat put in the link to that if you wish to see those before we begin the interview. It's up to you.

Stakeholder1_Engineer 4:44
I think it's OK. I understand you. Mm hmm.

Interviewer 4:45
You're OK. Alright, OK, alright. So this is the scenario and ... um...
Let's get through the questionnaire. So remember, this isn't a test for you. It's gathering information about your views, your perspectives, your thoughts, questions, opinions about the case study scenario. So there are no right or wrong answers. OK. The purpose of this interview is to gather in depth information on the subject of explanation from the perspective of humans. OK.

Stakeholder1_Engineer 4:56
Mm....Hmm.

Interviewer 5:19
So in order to identify the varied explanatory information that multiple stakeholders would have regarding these crashes, here are the questions. So, based on the scenario described above, you're seeking explanatory information about these car crashes from Autopilot , the AI system that controls steering, braking, acceleration, function of the AV so the motion control functions. So what questions? What are specific questions would you ask of and or what types of information would you seek from this AI system about the decisions made and actions taken? OK, so what did Autopilot, which in essence what is an agential entity that is performing the human driving tasks of motion control? What question do you have about this scenario? For the thing that was doing the driving?

Stakeholder1_Engineer 6:21
I think the first obvious question is how come ... uh.... the AV couldn't see this stationary uh... let's say obstacle in general. In this case it's vehicles. But in general because the whole system I do understand is based on that I mean.

Interviewer 6:36
Mm hmm. OK.

Stakeholder1_Engineer 6:45
It's identifying obstacles and of course avoid them. So how come they didn't do, I mean, they didn't avoid these obstacles. It doesn't matter that doesn't matter if there are vehicles, first responder, vehicles, whatever. It's a volume on the road.
Which is the primary function and ADAS should have.

Interviewer 7:15
Right.

Stakeholder1_Engineer 7:16
Now my second question is why first responder vehicles? Because I understood that the whole cases are first responders, responder vehicles, no?

Interviewer 7:30
Yeah, there to be fair, to be fair, yeah.

Stakeholder1_Engineer 7:32
Or they they are this. These are the 16 out of many others. OK. OK.

Interviewer 7:37
Yes, yes, to be fair. A lot of Tesla crashes. Hundreds have been reported to NHTSA OK.

Stakeholder1_Engineer 7:43
OK, OK, OK. OK, I understand. So I get back to my first question. I mean, why, how come?

Interviewer 7:46
These are, yeah. Why did I choose these? Sixteen. Yeah. Or yeah. Yeah.

Stakeholder1_Engineer 7:51
No, that's fine. It's OK. But how come a system that is mainly built to avoid obstacles, go crashing into obstacles?

Interviewer 8:02
OK, so your question for the AI system is why didn't you see the obstacles? OK, OK.

Stakeholder1_Engineer 8:08
Yeah, yeah, yeah, these obstacles because ... If... if it doesn't see obstacles, means that it will crash the first minute. I mean because the . . . the road is full of obstacles.

Interviewer 8:23
All right, so.

Stakeholder1_Engineer 8:24
So why not these ones? I mean which is?

Interviewer 8:28
Well, the reason I chose these 16 is well, first of all compares you other things on the road. They're highly visible, right? Because light's flashing and there's a lot of them, right. There's an a .... If you're driving like a human being, you can see an accident. OK, in a distance. So. OK, so why didn't it see it highly visible? So, beyond that, like.

Stakeholder1_Engineer 8:35

Yeah, yes, absolutely.
Yeah, sure.
OK.

Interviewer 8:53
OK, so the task of motion control in a car involves ,you know, seeing the surroundings, seeing what's on the road, but it also involves steering, braking, acceleration. So, you're assuming that by not seeing it, it didn't change its steering, braking or acceleration, right? You're you're assuming that's what happened.

Stakeholder1_Engineer 9:01
Absolutely, yeah, yeah. Yeah.I can. I can. I can imagine that. Yeah, I cannot imagine that with in a, let's say, a truck.

Interviewer 9:22
Mm hmm.

Stakeholder1_Engineer 9:24
The system, necessarily detected obstacles and avoid the . . . them, but not these ones. I mean, it's not a question of the efficiency of the . . . the system to avoid obstacles. I mean to steer to, to brake. But the problem is why it doesn't detect it, why it didn't detect these ... these obstacles. And it went crashing then.

Interviewer 9:55
So you're assuming, you're ... are you assuming that if it had detected them, it would have steered, braked or accelerated accordingly?

Stakeholder1_Engineer 10:03
Because I say if ... if ... it if it didn't.

Interviewer 10:07
Mm hmm.

Stakeholder1_Engineer 10:09
If it doesn't see obstacles, it wouldn't arrive until this point. I mean, it could crash into the first one.

Interviewer 10:16
Yeah, yeah. Tight. You do realise that at some point it did release control?

Stakeholder1_Engineer 10:24
Yes, in one . . one of the cases I think, yeah.

Interviewer 10:25
No, no. On average in these 16, on average.

Stakeholder1_Engineer 10:28
Ah yes, on average. OK. Yes, on average, yes, absolutely. Yes, I missed that one second before.

Interviewer 10:32
For these 16 less than a second, it just released it. It just released it right?

Stakeholder1_Engineer 10:36
Yeah, that's weird.
Interviewer 10:37
So what questions do you have about that?

Stakeholder1_Engineer 10:51
Hmmm....uh ... Well for that, why? I don't know how it works, but does it notify the driver before ... before aborting the aborting the autopilot?

Interviewer 11:02
Sometimes yes, and sometimes no. In these scenarios in these scenarios.

Stakeholder1_Engineer 11:04
OK. Nice.
I mean, yeah, it means that they . . . they notify but on the spot, I mean, not . . . not enough before for the driver to take control.

Interviewer 11:10
Yeah, yeah. Right.

Stakeholder1_Engineer 11:19

Because I try to understand. What could happen? Because if I imagine that the system went to a situation where it couldn't make a decision, I do imagine that in its programme or its structure, whatever ... it could anticipate, the fact that it's get into a situation of difficult decision and then it should stop. But before that it should notify the driver that OK, now you take control.
So why this .. that didn't happen also I mean because I believe that as especially as these obstacles are stationary, so it's not something going through the .. the pathway, let's say of of the of the of the vehicle that created a ... a problematic situation out of nothing. But these they these obstacles are there. So I do imagine that if the system uh...probably detected it, but didn't understand what it is. It could make a decision. I mean normally it should be programmed to just uh...Just give control to the driver and notify the driver. But of course as it's stationary, the situation is... Well, I mean not controlled, but maybe in well described way before getting to this point. So probably it should have notified the driver early enough to take control of instead of the Autopilot. So question #2 why considering the fact that these obstacles are stationary and... and...supposing that the system saw them or detected them enough early before getting to this point, why it didn't notify the driver that yes, we're getting to a situation where probably the driver should take control.

Interviewer 13:48
OK, so your first question is why didn't it's? Yeah, why didn't it see it? And your second question is if you saw it, why didn't you notify the driver, OK.

Stakeholder1_Engineer 13:48
Second question, yeah. If it's. Yeah, yeah, yeah.

Interviewer 13:59
As I said to you, remember this AV can control the steering, braking and acceleration. Do you have any questions about the decisions relating to the steering, braking and acceleration in this scenario?

Stakeholder1_Engineer 14:06
Yeah. Yeah, I would. But first I need to to understand why this ...because if the system get to this situation means that it could not interpret what's going on, so I'm not questioning the consequences because the braking stuff is OK, there is a situation that I analyse and understand and I make a decision braking, accelerating whatever.

Interviewer 14:29
OK.OK. Yeah.

Stakeholder1_Engineer 14:43
I suppose in this situation that the system couldn't describe what's happening, and of course didn't know what ... I don't know why and I don't understand why and I don't understand how a second before it realises that there is a problem and try to give control to the driver by aborting the the control.

Interviewer 15:06
Yeah. OK, So what questions?

Stakeholder1_Engineer 15:09
It's. It's weird, it's weird.

Interviewer 15:11
Yeah. So what exactly .. So what questions do you have about the algorithm itself to say the designers or?

Stakeholder1_Engineer 15:20
Yeah.

Interviewer 15:24
How would you go about improving the algorithm? So what kind of features, functions, methodology would you like to ...information about in in order to improve this algorithm that was involved in this car crash. Like clearly you want to know why it didn't see it, and if it did see it, why it didn't recognise what it was seeing, right? So.

Stakeholder1_Engineer 15:42
Yeah Exactly. So there are there are multiple, there are multiple system component that are probably involved. First of all the sensors before getting to the algorithm.

Stakeholder1_Engineer 16:00
Is there a problem because the fact that it disconnected the second before means that it finally realises that there is something, so probably the sensor.

Interviewer 16:10
Mm hmm.

Stakeholder1_Engineer 16:13
Was not enough, or maybe was there was an obstacle or something on the sensor. I don't know what exactly happened because at the same sensor that didn't see the obstacle, so it finally a second before the collision. Let's say maybe. I mean just interpreting.

Interviewer 16:17
Yeah. OK. Yeah. That's. That's a yeah, that's a really excellent assumption. But for the purposes of this experiment itself, let's assume that all the systems, all of the systems in the car, are operating and functioning correctly. So everything hardware engine, steering, vehicle, pedals, cameras, all the sensors, all the software such as engine optimization, fuel efficiency, all of that's working.

Stakeholder1_Engineer 16:40
Yeah. OK. OK, so the algorithm, the brain, the brain. OK.

Interviewer 16:59
It's just, yeah, the algorithm that that controls the steering, the braking and acceleration. What question do you have about those . . . those decisions? Because remember it just kept driving same speed.

Stakeholder1_Engineer 17:05
Yeah, yeah. Yeah.

Interviewer 17:17
Same direction, same lane, you know? No braking.

Stakeholder1_Engineer 17:18
Yeah, yeah. Yeah, which is.

Interviewer 17:22
So what ... what questions do you?

Stakeholder1_Engineer 17:30

As a... I don't know. It's... it's. It's a weird situation. I mean, if you suppose the sensors are working, the cameras, etcetera and that the algorithm had the information that there's an obstacle and didn't react, that's weird. That's really weird because if because this is a ... anomaly that happened.

Interviewer 17:57
16 times.

Stakeholder1_Engineer 17:59
Yeah, but, but but....No, this is not the the the right stats. From the starting point of this ... uh... drive how many obstacles this system avoided before getting to this one that it ignored. This is this is the problem. This is my problem.

Interviewer 18:12
Mm hmm. Mm hmm mm hmm. Mm hmm.

Stakeholder1_Engineer 18:24
Because it's it was a problem of the algorithm. It could crash. The first obstacle if you suppose the sensors are working.

Interviewer 18:27
Mm hmm.

Stakeholder1_Engineer 18:33
And it gets all the informations about the obstacles and ignore them.

Interviewer 18:38
Mm hmm.

Stakeholder1_Engineer 18:39
I don't even ... That's why I'm laughing, because I don't. I don't imagine ...a code.

Interviewer 18:47
Mm hmm. Mm hmm mm hmm.

Stakeholder1_Engineer 18:48
Is that selective in ... in recognising or in reacting to an obstacle. Or hundreds, and then choosing one and not reacting to it. You see because.

Interviewer 19:08
Yeah.

Stakeholder1_Engineer 19:09
I can't imagine that these. I cannot imagine that for all these cases this stationary first responders, vehicles where the first obstacle that this car uh .... faced since the start of the driving.

Interviewer 19:26
Mm hmm.

Stakeholder1_Engineer 19:26
So it's not. It already avoided a lot before get into these and this and this and decided not ... not to not to avoid them. This is my problem and that's why I'm getting to the sensor because I cannot imagine that.

Interviewer 19:38
Mm hmm. Mm hmm. Mm hmm mm hmm.

Stakeholder1_Engineer 19:46
An algorithm that is built to avoid obstacles and avoid obstacles, avoid obstacles. Get to this one and stop working. There's something. There's something weird.

Interviewer 20:05
Well, let's assume that the. Yeah. Let's assume that the sensors are working OK.

Stakeholder1_Engineer 20:05
Now. Yeah, yeah, that's that's what I'm yeah.

Interviewer 20:11
And yeah, So what in the algorithm didn't recognise?

Stakeholder1_Engineer 20:14
And as you said, it's not only this kind of obstacles cause you said that these cases are only taken from from others where other different obstacles. So it's not selective, it's not linked to the kind of obstacles.

Interviewer 20:18
Yeah. OK. Yes. I don't know. We don't have that information. Remember, this is a black box, OK? This AI algorithm is a black box.

Stakeholder1_Engineer 20:33
Yeah. I mean, it's not, yeah. Yeah, yeah, yeah. And we cannot say that it's that. It's.

Interviewer 20:44
Remember, AI algorithms that are black boxes are not deterministic software programmes, right? And the programmers themselves do not know how it made the decisions it made. OK, they do not know and they cannot go line by line code to determine how it took the actions it took or didn't take. So within that constant uh … context, what information do you want about?

Stakeholder1_Engineer 20:47
Yeah, yeah, I know. Yeah. Yes. Yeah, yeah. Yes. I understand.

Interviewer 21:10
These algorithms about the decisions made and the actions taken. Are there any other questions beyond why didn't it see this obstacle?

Stakeholder1_Engineer 21:24
Well the the questions.

Interviewer 21:27
Mm hmm.

Stakeholder1_Engineer 21:28
That could be. I mean just like objectively, probably silly, but objectively is “OK. Why did you see? OK, I saw them. OK. Why didn't you avoid?” questions.

Interviewer 21:43
Mm hmm.

Stakeholder1_Engineer 21:45
uh …. How many, many, many possibilities, 2 main ones of course is changing direction or or braking , Why it doesn't happen? These are the questions.

Interviewer 21:57
Mm hmm mm hmm.

Stakeholder1_Engineer 21:59
Which I believe. Are too basic I mean but.

Interviewer 22:04
No, no, no. They're not basic. Assume it's a human being driving.

Stakeholder1_Engineer 22:05
But anyway, these are these are the normal. Yeah. I mean, these are the normal, yeah.

Interviewer 22:06
Assume it's a human being, driving. What questions would you ask?

Stakeholder1_Engineer 22:11
Exactly. Yeah. Yeah, that's that's exactly what I'm doing. Yeah, I'm in. I'm in the. Yeah. If I was me, if it was me, I would brake. I would change direction. I would do something, even at even at. Even in the last second. I mean, I would react do something.

Interviewer 22:28
Not go

Stakeholder1_Engineer 22:29
Not, not just.

Interviewer 22:32
Yeah, not put your hands up and go. OK, I I I'm gonna stop driving. Yeah. Yeah. I'm not gonna touch the steering wheel or the gas pedal or the brake. I'm just gonna let go, right. Yeah. OK. So your questions are why didn't you see the? Yeah.

Stakeholder1_Engineer 22:32
Absolutely. Yeah. Yeah. Just say, OK, maybe somebody else will come and save the situation. Exactly. Yeah. Yeah, so. Now OK. Now for me, and there's also another. Variable which is the learning process.

Interviewer 23:01
Mm hmm. Mm hmm. Mm hmm mm hmm.

Stakeholder1_Engineer 23:02
Because an algorithm is you know, processing and learning to create some decisions or some actions in the world.

Interviewer 23:07
Yes.

Stakeholder1_Engineer 23:14
And and ...this is also a maybe a question, an investigation, that we can go through, which is what could.

Interviewer 23:22
Mm hmm. Mm hmm mm hmm.

Stakeholder1_Engineer 23:27
Could happen in the learning process that um... but then I I'm asking questions to, but I I don't have information, I mean about all the cases because is it a problem with the stationary? Is it problem of obstacles? Could not because normally a tree is a stationary obstacle? A ... I mean I don't know anything coming in that's.

Interviewer 23:54
If you look, yeah, if you listen to if you listen to the Tesla. Uh... claims millions and millions of driver hours and all of that right, so it recognises, it recognises obstacles moving and stationary on roads all the time, right? So ... and despite that, there's been hundreds of crashes and these sixteen are very particular because these are not just any stationary objects.

Stakeholder1_Engineer 24:07
Yeah. Yeah. Exactly. Exactly.

Interviewer 24:26
They're very bright lights, very um...like the road is blocked by big things. Remember that right? Like I mean, you've got fire trucks and.

Stakeholder1_Engineer 24:37
No, no, of course, but I'm trying. I'm trying to understand.

Interviewer 24:40
Yeah.

Stakeholder1_Engineer 24:44
I'm trying to. I'm trying to understand that the perception of an algorithm is not the same as for us.

Interviewer 24:53
Yeah. Yes.

Stakeholder1_Engineer 24:56
And that what kind of learning process could?

Interviewer 25:05
OK so.

Stakeholder1_Engineer 25:06
It could drive the algorithm to make this decision of not doing anything in this case.

Interviewer 25:13
So you're questioning the software engineers and the data that they're using to teach this?

Stakeholder1_Engineer 25:19
I ... I would be questioning everything because this is the system is what is. The algorithm, the learnings, the sensors, these, these are the three things that that collaborate all the time. So I'm not sure it's.

Interviewer 25:25
Yes. Yeah. Right, right. So we've.

Stakeholder1_Engineer 25:40
Only the the algorithm, unless it I don't. I mean because nobody knows. As you said. I

mean it's a black box. We don't. We don't know exactly what's happening. I mean, we're interpreting what we are seeing from outside. But I would question all this system. Yeah.

Interviewer 26:02
Right. Well, remember the question is what would be an explanation? What kind of information are you seeking as an explanation for these crashes, for the decisions it made or didn't make for the actions it took or didn't take? What information would you seek from this algorithm or from the designers of this algorithm? What kind of information would satisfy you? From them, you know, like, how would they go about digging into this? And you know, but they need to know what you want to know.

Stakeholder1_Engineer 26:32
Yeah. Oh, I will. As I said, I mean obvious questions are. For me it would, it would be. First of all, are these obstacles that do these obstacles have something specific? Number one so, just to remove this eventually the I mean something probably.

Interviewer 27:08
Mm hmm. Mm hmm mm hmm.

Stakeholder1_Engineer 27:13
That made the the system mot to recognise these obstacles as what.

Interviewer 27:22
So you're questioning the learning methodology, you're questioning how.

Stakeholder1_Engineer 27:25
Yeah, yeah.

Interviewer 27:28
The evolution of the algorithm itself you want to question the designers also right about what they use to train the algorithm, right? Yeah. OK.

Stakeholder1_Engineer 27:34
Yeah, exactly. And see how what could? Yeah, exactly what could in the training make the algorithm think think.

Interviewer 27:46
Yeah.

Stakeholder1_Engineer 27:47
That these are not obstacles.

Interviewer 27:49
Right, right.

Stakeholder1_Engineer 27:50
Until, but until it get close to them because you know when it get close to them , tinally, it recognised that there's something wrong in this situation.

Interviewer 28:01
But maybe it. Yeah. Yeah, that's the thing. It didn't recognise what it was, right? So it just let go.

Stakeholder1_Engineer 28:05
Yeah. Yeah. Yeah, it's it's something. It's a,... and when it ...and then let go means that it got to a situation that it's not described.

Interviewer 28:18
Yeah.

Stakeholder1_Engineer 28:19
It's in its in in the algorithm I mean or in the learning.

Interviewer 28:22
Yeah.
Right. OK.

Stakeholder1_Engineer 28:26
There is a there is a mismatch between the reality. I mean the situation and the learning.

Interviewer 28:33
OK.

Stakeholder1_Engineer 28:34
It like like exactly a driver. I mean when when I'm a human, when, when, when he or she.

Interviewer 28:37
Mm hmm. But when you had a driver's permit, right when you had a learner's permit, right. So you're questioning, what did it learn? Is what your your. Yeah. OK, got it. Yeah. Yeah. OK.

Stakeholder1_Engineer 28:43
Yeah. Yeah, exactly, absolutely, absolutely. 'cause, when you get in a situation where we you don't know what to do means that in your learning because some other people would would find a solution. But you can means that in your learning or in in your processing of your learning.

Interviewer 29:00
Mm hmm mm hmm. Yeah. Right. Yes, yes. Understood. OK.

Stakeholder1_Engineer 29:11
You didn't cover this.

Interviewer 29:19
Before we go on to any other more questions. Do you have anything else to add beyond what you've just said?

Stakeholder1_Engineer 29:33
Well, if it was a small scale, I would. I would put these cars back into a test. But I know that there are billions of data. I mean about already so.

Interviewer 29:43
Yeah, yeah. Yeah, yeah, understood. And by the way, it's still being tested. We know this for a fact, right? It's still gathering information. It's still learning, right. So, OK. All right.

Stakeholder1_Engineer 29:48
Yeah. Yeah, yeah, yeah. Learning. Yeah. Yeah, exactly. Yeah. Yeah.
{general Discussion}

Stakeholder1_Engineer 41:46
I have I have a a broad opinion on on the on the AI system in general and especially. the learning process. And if I and I think this is, I think that probably people should

think of it because if I compare the learning process of AI systems to the one for humans. There is something I mean, it's like I take a child.

Interviewer 42:24
Mm hmm mm hmm.

Stakeholder1_Engineer 42:25
I....and you know that in the education system they the values and are the first layer that that we normally teach a child before start starting, giving them for her the knowledge. But for the AI, it's completely the opposite. Or at least it's a mix. So we just like taking a child of three years and start maths or driving or whatever. Before any framework, it's just. A volume of informations coming in and we are expecting the system to structure them to classify them, to interpret them and to use them.

Interviewer 43:13
Mm hmm mm hmm.

Stakeholder1_Engineer 43:14
To do something that we are expecting to be logical.

Interviewer 43:18
Hmm.

Stakeholder1_Engineer 43:19
I think that this is a big problem that we people from AI industry. Would probably consider as a subject.

Interviewer 43:32
Yeah.

Stakeholder1_Engineer 43:33
Of of the reflection. So that's why I'm asking about for these cases about the learning process. It's very important to understand. Of the algorithm is ...is interacting with all the informations that all the because most of it is behaviours of drivers.

Interviewer 43:56
Yes.

Stakeholder1_Engineer 43:57
In different situations that probably are not are not are not coherent, are not consistent to each other. Maybe one is driving to the snow, the other is driving along in the mountain. The other one is I don't know.

Interviewer 44:04
Yes. Yeah, yeah.

Stakeholder1_Engineer 44:11
And. How the system is combining all this?

Interviewer 44:16
Yes.

Stakeholder1_Engineer 44:17
To decide what to do in a situate in a single situation. I don't we don't know where and when and how, so this.

{General Discussion}

Stakeholder1_Engineer 45:00
Yeah, exactly. That's that's the learning curve. That's exactly what we do. I mean, as a human beings, we start with simple things and then we go complex. Absolutely not.

{General Discussion}

Stakeholder1_Engineer 47:15
Yes, of course. Yeah, yeah.
From my academic perspective, yes.
I need to understand how the I mean how as I told you I mean.
I would like to understand.
How the AI systems are getting knowledge?
'Cause, I think that the let's say the coding part is not the most important, most important is kind of information we are given to the system in the learning process. My opinion is very important because I make comparison with the human beings cuz you can take whatever human being. It's the same brain, same but then depends on how

what kind of values you give them, how how, what kind of knowledge you give them how.

Interviewer 48:04
Mm hmm.

Stakeholder1_Engineer 48:19
Kind of framework of behaviour. That's what makes a human being.

Interviewer 48:27
Oh, OK, so I understand. I understand you're going into the depth. As an engineer, you're viewing this as an engineer. But here's a question for you. And I like that you're giving me the analogy to a human brain. So if this was a human driver.

Stakeholder1_Engineer 48:28
Mm hmm mm hmm. Yeah, yeah. Yeah.

Interviewer 48:46
And had these 16 accidents. And you assume, OK, the eyesight is fine. You know, they're sitting on the side of the road after they've crashed into these vehicles, and you're coming out to them and you're going. Why did you crash? Right, a human being can give you reasons right? Can justify their actions. So what kind of questions do you have for that person?

Stakeholder1_Engineer 49:08
Yep. What happened to you? What's what's the problem? Or as a human being, I don't know what do you have again, do you have against this guy? Something? I'm something against this. I don't know. Many questions come to me, I mean.

Interviewer 49:35
OK, when you ask what happened to you, what kind of information are you seeking from that person? Of what? Be more specific, what?

Stakeholder1_Engineer 49:40
Yeah, his state. His sensors, have you seen? Have you seen these obstacles or not? Is there anything that that prevent? Did you from seeing them?

Interviewer 49:49
I did. No, I can see them. I could see them.

Stakeholder1_Engineer 49:56
OK, so well, that's weird. I mean, what happened to you, I mean?

Interviewer 49:58
But I still kept driving. I still. Well, I got confused, so I let go of the steering wheel and push back.

Stakeholder1_Engineer 50:05
What do you mean by confused? What exact?

Interviewer 50:11
Is that an adequate explanation for you?

Stakeholder1_Engineer 50:13
But what happened to you, I mean? Why did you do this? I mean what?

Interviewer 50:18
'Cause I was confused. I didn't understand.

Stakeholder1_Engineer 50:22
Yeah, but by by what? By what? What, what? What made you confuse?

Interviewer 50:25
I didn't. I didn't understand what I was seeing in front of me.

Stakeholder1_Engineer 50:35
Hmm... So you were seeing something but you, but it's something on the road, no?

Interviewer 50:37
Yes, yes. Yeah.

Stakeholder1_Engineer 50:42
There was something on the road. And you didn't brake?

Interviewer 50:43
No, I didn't know what to do.

Stakeholder1_Engineer 50:44
Or or change direction or whatever.

Interviewer 50:47
No.

Stakeholder1_Engineer 50:48
You couldn't. You didn't. And you didn't think of avoiding this obstacle.

Interviewer 50:53
No, I didn't think about that.

Stakeholder1_Engineer 50:55
Wow. Yeah, that's a mental problem.

Interviewer 51:06
See what I'm telling you? I I'm. I'm being a black box, right?

Stakeholder1_Engineer 51:07
Yeah, yeah, yeah. Yeah. Yeah, exactly. Yeah, I understand.

Interviewer 51:11
I'm pretending to be a human being that did this, and I'm not giving you any answers. But listen, to listen to what you're doing. You're actually probing me about the decisions I made and the actions I took or didn't take, right? That's what I'm asking you to ask the AI algorithm. What information do you want from them? Not about its training necessarily.

Stakeholder1_Engineer 51:15
Yeah, but. But you know the problem here.

Interviewer 51:35
Not about it. You know the same information or different information.

Stakeholder1_Engineer 51:41
Do we have a a log of what? Decisions this algorithm made or didn't make.

Interviewer 51:49
No. Remember, it's a black box.

Stakeholder1_Engineer 51:52
Yeah, but there's no log. I'm I don't wanna understand how it works, but is there a log?

Interviewer 51:58
No, but yeah, it's a log. It says it will release control less than one second.

Stakeholder1_Engineer 52:03
I need just a list of of what happened. I mean chronology of OK, #1 sensors 1 send this information #2.

Interviewer 52:07
Yeah. No, you don't get that information.

Stakeholder1_Engineer 52:14
OK.

Interviewer 52:15
That information is not provided. Black boxes don't provide that. You're talking like it's a step by step deterministic system, it's not.
A black box in an airplane gives you step by step. OK. You're thinking like an aeroplane when I'm saying black box, I mean it's opaque. OK, that's another metaphor that they've borrowed from. But a black box on an airplane provides tonnes of info. This is not no. OK, what I'm saying is this is an opaque system.

Stakeholder1_Engineer 52:34
Yeah. Exactly. OK.

Interviewer 52:50
It's not giving you this kind of information cause the algorithm.

Stakeholder1_Engineer 52:53
It's not a black box, it's it's a black hole.

Interviewer 52:57
Yeah, it's a black hole. Exactly. Exactly, exactly. Now, what questions do you have for this black hole?

Stakeholder1_Engineer 53:08
Does it answer? A black hole?

Interviewer 53:11
What? What questions would you want for the designers of the black hole?

Stakeholder1_Engineer 53:19
Well ... Internal. I would like to understand why there has been. If we compare with the human being, to this computer in this situation.

Stakeholder1_Engineer 53:29
What happened? When the algorithm is facing this kind of situation.
But still I need to understand how the algorithm perceives this situation. How? How did it perceive the situation? In a way that it couldn't enter or or maybe it could... I don't know. I don't know what happened, but.

Interviewer 53:57
Right.

Stakeholder1_Engineer 53:57
Number one, how the algorithm perceived the... the external environment around it.

Interviewer 54:04
Yes.

Stakeholder1_Engineer 54:04
At that moment. And why the conclusion of this interpretation was just to.. to give up?

End Transcription for analysis .... general discussion continued until 1:31:23 when Interviewer stopped recording and transcription