w5

Sun, Oct 30, 2022 10:07AM 1:57:13

SUMMARY KEYWORDS

ai, people, stakeholders, technology, values, system, privacy, ethics, users, data, human, important, software, accountability, design, lecture, online, deployed, vsd, recording

---

00:04
And

00:07
I'm sending you the link

00:24
anything that should work

00:25
and if you join I will make your co host right away did my email with the link come through this marvelous a half day thank you both landings silent procedure so then can you try recording it on your hard drive

01:42
no

01:45
yes

01:51
myself hear me just

01:57

so I mute here because then otherwise you won't be recorded

02:04

okay hello people online sorry that was the biggest lecture just left so we just had time to set it up

02:21

on my last

02:25

Day Good evening this is test newbies find us to all of this ancient consumables I am not enough obviously I'm so

04:11

you may someone else

04:13

maybe well couche is getting shot slides um hang on How's everyone doing YAY or may no response

04:47

so, the assignment specs released

04:53

right There's no audio soon

05:19

So has anyone not seen the

05:21

05:21
group assignments thing? Right

05:26
zero responses

05:27
you have all signed up okay ask people

05:38
can you hear us online sure mute NO

05:45
NO

05:46
NO NO NO ME here right Yep. Cool No you can't no echo Oh great.

05:55
Yeah I think that's because the last time I was like

06:00
all right oh great perfect very nice fancy setup

06:05
okay not a fancy setup here he cannot live with the student Okay you ready

07:20
I'll hit recording now

07:26
so yeah screen us

07:39
yes

07:40
you can now and I will see my screen

07:54
just checking again you can see the screen

07:57
yeah too. Right perfect. You see the slides

08:04
let me see if I can hit record

08:08
Sorry to keep you waiting

08:11
this is pretty quick by all standards

08:13
there's always a delay in getting

08:16
between the two lectures the previous one and this one so

08:21

watch this what's the record button is recording started

08:28

on not yet so recorded how to do here

08:31

idle

08:35

on every track record the cloud but

08:38

trying to do cloud and I'm recording to my hard drive with my Zoom stopped sort of started again record again straightaway. What is that okay,

08:52

so hello again everyone.

08:57

Welcome to

08:59

the second lecture my my second lecture so it's week by and before I start, I like to acknowledge her video people or traditional custodians of this land. Also like to pay my respects to the elders both past present and extended respect to other Aboriginal Torres Strait Islanders for present here

09:19

today.

09:21

So today, we're going to talk about value sensitive design VSD AR AR ethics. Also, also we will have a deeper dive on liability and accountability as some of the components of AI framework and governance. Very quick is on payments and there will be a dedicated

09:56

in May 7, and also mercy In the

10:01

end, so now

10:08

if you've seen your asana specification, one of the jargon in that spec is PSD. So I have orange actually,

10:20

copper.

10:22

BSD. Hang on. There's like an echo they say

10:28

for you. So you done okay. So what

10:36

is BSD? So BSD is basically a theory that has been studied by and proposed by Dr. Friedman for almost three decades now. With me. So the, I mean, for people online, I was interrupted by chocolate. So if your sound droplets come to the lecture,

11:05

it'd be nice to be here

11:08

every night. So yeah, I'm going back to value sensitive design. Beta Freeman actually proposed this, and this was a paper from 1996. But if anyone remember, two weeks ago, during my first lecture, I was actually start I started the lecture about discussing about values remember what values were that we discussed? You Gianna? These are the things that you think are important. And this actually constitutes loneliness. In in athletics framework. Right. So that's great, thank you. So whenever you see an ethics framework, there are values that are principles, right? Their purpose. And they're basically certain activities that you want, you want to set the practices day on abide to those values, but values are basically the one that I'm saying that values are basically the heart of an AI ethics. So you whenever you see a code in let's say, you are in a company or you are going to, you know, larger startup, you have your own, you have something very innovative and you want to launch a startup and you want to have your own company code of practice in terms of ethics, you need to know what your values

12:39

are. What do you want to you want to

12:43

uphold? What do you think is important? Now? So value Friedman, she's she's a professor at Euro, University of Washington, and this was what she wrote in, in this article values instead of design in interactions. He had interactions magazine in 1986. She say value values emerge from the tools that we built and how we choose to use

13:12

a Yeah. Most of the current

13:17

practice in designing computer technology and related are in front of infrastructure of cyberspace leaders is saying about values. So when we actually discuss about values also last we asked you whether technology is neutral. I didn't give you a direct answer because I'd like you to consider and have a discussion about it during the tutorial. I don't know if you you have that discussion in your tutorial whether technology is neutral, neutral.

13:51

What was the

13:55

what came out of that discussion? What is your view now? Whether technology is neutral or not?

14:01

Anyone online or here in the room?

14:12

What's your view is technology neutral? Raise your hand if you agree

14:22

somebody's just doing this.

14:26

So no, so no one seems to think that technology is neutral. And I will say technology is not neutral. So I agree with you. Technology is not neutral. So you are nine was had my sort of realization yourself about this. Because if we think technology neutral, we can always design and that's in our innovation advancement of our computing infrastructures. Technology as much as we like. And they say, because it's neutral, it's the users fault, or it's the one adopted at fault. But in fact, bad equipment in line, you know, sick brokers, that we need to start embedding values, right as far as the design of technology. And so VSD is an interactional theory and method that accounts this human values in a very personable and structured manner throughout the design process. So that VSD, so human values. This is a snapshot I took from her paper in 2013. So, so I mentioned earlier from that paper in the United States, and this is a like almost two decades after she was still working on on this topic, even she should still be doing keynote on BSD in like 2019, in even in 2020. So the principle and the theory on BSD is still as valid as ever, even almost three days from the time she actually proposed it.

16:15

So

16:18

a wider range of human values that we might want to consider. So there's a long list here, the Texas South Node, sorry, but let's see if I can zoom. Apparently can't but but you can read the paper. So the point I put put them here is not for you to read them. But I hope there'll be some interest in the room for you to actually read the paper. So some of them include human welfare, privacy, trust, freedom from bias, courtesy, identity, and commerce, caring, autonomy, environmental sustainability. So there's a wide array, these values actually a lot broader than ethics in AI. So these are basically all the human values that we think are important, and how

we as a creative technology, developer technology, how do we embed these values in the things that we create? So then how do we actually adopt VSD, we think this is a good approach. So I've used these an approach to rigorously account for human values and technical design engineering process. So I really show you this slide before. So in judgment call, this is one of the game right? In BSC, it begins with identification of stakeholders. And then you start to look at the values. The thing is, if you mentioned, one is less than one values important, say, for example, autonomy is important. So for example, there's an ability to be able to plan and achieve a goal in in an independent way. So I can have information at my fingertips, I can be more independent in making a decision. That might mean if this is very important, autonomous encoded, the information system need to be able to call me to enable that and if they have to have access to my data, and the value of autonomy might have a friction with privacy, for example. So some values that you think are important, maybe in the friction of another value. So this is why when you're thinking about these different parameter values, you may want to look at who are the stakeholders of the system? Who are the potential users? And that's why we discuss how do we actually identify users in our lecture last time in our first lecture, in my first lecture, not so much, and how do we identify the stakeholders? And how do they serve as their value so what what is the what's an important value for these stakeholders? I'm just giving you a framework for VSD This is from the point of view of

19:31

better treatment. So,

19:34

when when you actually do a VSC from a solid design process, you will do these three things. First is a conceptual investigation. Second, is an empirical investigation and third, technical investigation, but these three actually will inform one another in Alou. So come Essentially, you start with the stakeholders and use a and you can see the what the values are, you do research on this, you might, you might have to do a study of previous technology, you might have to start asking the potential users. So these are conceptual investigation. So even if the technology hasn't been created yet, you can have you can create a story a storyboard, for example, and ask people if let's say this technology is to be deployed, um, what is your opinion? How would you accept it? How will these be valued to you how this views you can ask you can you can you survey interview and all those things, focus group. So a lot of companies love focus groups, there's a way to validate the kind of technology with just a small number of potential users. There's a good thing about focus group, of course, there's also negative witnesses our focus group, because the sample is very small. Then there's also empirical investigation, which means you actually definitely technology, you prototype it, and you monitor the use and see how an impact on the life of the users and the stakeholders before an actual before the actual system is deployed. And the technical investigation is you need to look at each of the feature, the property of the technology. And whether each of these units in your system sends out the test, right? It's like you're doing your unit testing is like you're actually doing a comparability. Real reliability checking.

22:00

For example, if you

22:03

before every flight, if you want to know, if you're taking a flight, they normally have a random video, they normally have a very thorough safety check for every single thing. And this is what actually needs to be done an audit on our own software system or app or whatever system that we develop with it, what not just technical functionality, but whether it's to do whether it's still a down to our value. So this is something I think, quite

22:36

different approach.

22:40

Now, someone asked, Will these

22:43

slides be in slip CMS? Yep, we will share it after the lecture.

22:47

So the question here, What is interactional theory from that was on the earliest slide about value

22:53

system design. Right, okay.

22:58

So instructional theory comes from interaction design. So, interaction design interactional theory, there are so many of them, they look at how human can interact with computers, or computer system or information system around them. There's a lot of different theory, but one of them is basically VSD that basically focus on the actual of it between interaction of a user or human and the computer or computer technology. The focus is on the human

23:35

bathroom itself. Right now,

23:43

so with VSD, since better, Friedman proposed VSD in 1996, there were tons of VSD tokens that have been proposed by researchers, I even some companies, and but but DEP herself, and the group will have actually come up with this VSD, topical, envisioning pods. So envisioning cards, think about when you try to envision of envision a technology. These are the different criteria you need to consider. So you can go to envisioning cards.com, it's all there. You can even download it, you can even print it. One is stakeholders. One most egregious stakeholders are who are the stakeholders of this supposedly design or basically it could be less even existing design that you want to retrofit or improve? What are the stakeholders who are the current stakeholders who could be the potential stakeholders who are the current users? Where Where do we where do we miss in the market share is up So you ask all these question questions. And you're thinking about the timeframe. perspective, is it the short term, the International short term? perspective, this is what it's going to be in the long term is how the technology guy can be. So this is basically the timeframe. And then you actually look at the value. So basically, across time, these are the stakeholders. So what are the values? And you consider each of the value according to the each of the stakeholder

25:36

separately.

25:38

And then think also for pervasiveness of the system? Whether how pervasive is this system? Is this going to be all of this is going to be designed in a very generic way that can be deployed in all industry, for example? Or is it just going to be a very specific technology for one thing? Right, so obviously, the more pervasive it is, the more we have to account and be responsible for the value that we embed on the technology. If it's only a system that, you know, will only be used for, let's say, it's a system only for high school students. Will it be used for university students? No, maybe not only for high school students, and only for that, so then you only have to care about the stakeholders of that school. But it's a system that, you know, can have a major impact. And that is where all these other values have stakeholder thinking about the sticker from the different views stakeholders as part of the water, and multiple lifespan thinks about, you know,

26:51

the potential

26:53

the potential ripple effects on on the evolution of technology. Of course, you can't you cannot we all can have a crystal ball, right? So for example, when Twitter was first launched, it was only used to launch as a micro

27:13

bot echo. Who still does blogging here.

27:21

I mean, it used to be great, well done. But, um, Twitter was actually at heightened of blogging. You know, the trend of blogging, so Twitter was like, basically, you know, you can blog with very short spurts of text messages. But no one knew right? That how Peter took the world by storm event and was use in a heightened revolution of app springs and all those things, right. So even if now could be used to propagate misinformation disinformation. So this is a multiple lifespan of Twitter. We all know, the kind of technology that you and I are currently that we'll be researching, we potentially won't know what the impact will be. But if we start embedding values as desired from the beginning,

28:19

it will have at least

28:24

you know, those safety net right in the beginnings being sort of embedded. Now, talking about safety net, this is a hypothetical scenario, again, taken from Friedman's article, another article, this is now in Chi. So 2007. So I mean, I love Friedman's work. She's got a lot of amazing papers, safety net, a hypothetical scenario that she used to run or use a sudden.

28:59

So let's read this

29:01

safety net as a hypothetical commercial software platform, which leverages publicly available demography and criminal data mapping technology. satellite tracking capability to degree maps for display on various mobile technologies, such as cell phones, blackberries, contacts, alternativen lifeberries was still very hot in the car navigational systems. These maps are used to alert urban travelers that venture into potentially unpleasant or dangerous areas using their travel now let's think out loud here what do you think other stakeholders sorry, generally people

29:54

general but can you?

29:57

Everything can be general people, but let's be more specific. Who are these general people? Yes, urban travelers urban regions

30:16

so, what could be urban travelers here? Who might be interested? If last time you tried to sell? Let's say you're trying to sell this software platform, who do you think will be your market share? What kind of people walk out personas? Tourist? One is tourists and very good tourists. Yeah. Well, police police interesting. Yes. Um,

30:48

I can tell you one,

30:49

and that this, this came out of our study as well which, which was in very interesting out of her study. So she used it and then she went out and asked people,

30:56

moms with kids. So

31:04

So basically,

31:06

for kids,

31:07

you know, they need to be walking to school. And alone with a, if this app can be installed on the phone, at least the parents can see and track where the kids off. And it could be it could be used to alert the kids it Hey, don't go there, because that's kind of dangerous.

31:27

Take this route instead.

31:31

SAP says real estate agents. Right? Any other idea? Well, in fact, you're right, it can be all them can't be all of them. Now what are the values for each of those people

31:56

safety net

31:57

one value that is so clear there is safety. So that's the value they want to end that in the system, I want you my user to feel safe when you go around the city and basically, for parents who install this on on their kids phones or blackberries when you're still hot and in so they found it to be really useful. But then real estate agents and police as you mentioned, start seeing interesting and weird trends in the map that some some areas seem to be full of, let's say Somali games or African gangs or some certain and this was in the in the hours just outside I didn't make it up it was in the paper and basically in fact, he was used to he was used this was actually used to become a game this was made become a game by these people. So I didn't know how to turn into a game but what happened was then the data that went into the the data that went into the system was basically became became false and misleading and very bias on the system how difficult Has this happened in real life? In fact Yes

33:45

How can I go

33:49

with Pokemon Go was first launch Oh want to put these to share now you can be proud of who helped me I'm used to playing Pokemon

34:00

Go here. Athletes I can see one hand right anyone online though. Well, I still have friends that play poker my door and not not me on my family. But um What is interesting here you can see from Ingress. Ingress shared is on their own map. The port portals. Let me just move this this so the poorest sister because it's really in majority white census tracts.

34:37

So, in fact,

34:40

biggest impact in my yard. It creates a positive impact in the urban area. It activates places that people don't tend to go, which means it became another stakeholder that then became eventually a psycho They're a pokeymon go was actually shot owners, they never automatically get beginning but then they realize, hey, is to be a revenue at revenue as well for, for, you know, our supporters to be, you know, put near our shop, so people can come to our shops. So there's a lot of different stakeholders potentially, you know when when you know, this is just going to be again an augmented reality game. But apparently the portals are considering a majority what census tracts because this just

35:33

highlights in this part of the city

35:39

the bias of the users the representation of the users, they only go to areas that are majority majority white. And in fact, they actually lead to more segregation in cities. So we saw that they are publicly knowledge that there was no website and they say, We are changing this, we're not, we don't want to see inequality that become worse and because of our game. So this is how you can I mean, even once a system is lunch, right, okay, my goal was lunch, if you know what is your value, you could rectify the system. It is important therefore to have that value sensitive design when you develop something. So I want to replay this I played in a lecture two weeks ago, I mean, this just want to call the game. And rather than I mean, I've given you a specification based on judgment call game card, focusing on

36:42

AI ethics.

36:45

But if for some reason you prefer the envisioning cards, you have

36:51

the choice,

36:52

you can do so small, I just think that a judgment call a game is it's actually much easier to explore envisioning card has lots of different things to explore with judgment call you in your in your work, let's say your group of three you have to consider all the kinds of stakeholders that might be possible in the hypothetical system or maybe a real system that you know is about to be deployed or was is already deployed or was deployed with a commission and then you need to do a roleplay or act like a stakeholder. So, pick the top three stakeholders that will have the highest friction between them potentially and then you need to argue from the point of view of ethics holder.

37:41

So, for example, this facial

37:45

recognition technology for for next generation TB

37:55

so

37:57

Thisara wants to do face facial expression analysis that can be sold to the likes of Samsung or Panasonic all the TV manufacturers saying that you know we can monitor viewer engagement level levels basically, and this could be used for recommender system recommend movies, so it could be used for or you know, this can be the use as a service for streaming providers. So you can then share the data add additional revenue for you and save time you know the end user they will enjoy you know, they will be there they actually have more option to share the emotion their reaction online with people they watch it they you know, they do maybe they might be watching the TV together the same show together but from different places. So they could be share emotions, share, share joy, or share anguish or share horror. Now, let's ask around the room. You're thinking about a stakeholders who are day

39:24

one yes.

39:26

Yeah. So the end consumer, the end consumer who buy the CD?

39:35

Well, forecasters,

39:37

the broadcaster's yet I want to know the actual reaction because they this is one thing they never learned. They only know how many viewers they don't know how much they enjoy which part of the

39:52

movie yet

39:53

we're going conspiracy maybe the government All right, very good. Who else? Marketing teams, marketing teams from any company that are advertising?

40:07

Advertising? Very good.

40:11

All right. So you become a nice that's very straightforward. Yes.

40:19

We production companies. Yeah. movie production

40:23

companies. You know, one important stakeholder that is always there is the investor investor, we started this startup, right. So this should make money for investors.

40:38

Anyone else?

40:40

Out idea online? A normal man to mention viewers, TV manufacturers, content producers, advertisers Very good. You just summarizing what's happening and what's being discussed in the room? Social media platform. Yeah, interesting. Could be. Now, then what you need to

41:02

do is

41:06

be before think about the value, you need to think about the goal for each of these stakeholder. So if someone mentioned before advertiser, how is

41:15

this gonna be valuable? I think you mentioned advertiser, right? I need to see if people are engaged with the product. Yeah.

41:30

And, you know, let's say in the James Bond movie, right, there's always Avitus it's full of elitism of actually all kinds of branded stuff there with, with the cars and certain gadgets with a watch. So, you know, if there is such a TV that can monitor facial expression, then people seem to lie in the car, advertisers will be gleam on it, right. So a lot of these different things then, but then let's think about the value that is how you what you need to do them as a as in your assignment will be the value for the end user value for each of the stakeholder and see whether that friction is there and then you decide whether to deploy

42:15

if you want to deploy it, how would you deploy it in a more responsible way?

42:23

If let's say you still decide to deploy right, maybe it is useful maybe how to ensure that this is done in a responsible manner. Now in this why we need to have a good knowledge of these different principles, ACI principles for algorithmic transparency and accountability. Um, so, this is even different from the ACM code of ethics. So, this has to do with algorithms. So, include some awareness of how this algorithm our algorithm came to be. And how is access How can it can it be is it open access, can it be actually re improve over time accountability, explanation of algorithm data provenance auditability and validation testing.

43:22

So,

43:24

so I think this goes into aspects of owners, designers, builders, aware of the full biases as well as regulators. So this is also applicable to regulators that can actually then question whether an algorithm can provide a standard proof of audits if something goes wrong. Now, moving further from algorithms, then I like to talk about AI ethics. So why AI is such a big thing because AI has been has been adopted by industry widely because it's actually improved productivity improves. Basically, economic value out of a lot of activities every day. Ai actually speed up the work that the value of AI comes when the equal human labor labor required to do the same task. It will be done through multiple hundreds of hours more than what AI can do. So if AI can do it in a much faster way, then it's it's a no brainer for a company

44:45

will adopt AI.

44:48

Now, and this is the global survey by McKinsey and the different companies sorry, again, the texts are very small. They If you look at the report, but it basically shows you the proliferation of AI and adoption in ministry, the highest one is high tech. And then the second one is automatic. So automatic production, you can see there's a lot of AI and robotics in manufacturing, you can see as well, it's too obvious. And also, down the line there. There's also things like consumer goods, and also banking, financial transactions, AI has been widely adopted, which means everyone use banking and financial transactions. If we need to actually look more at how this adoption, still are actually ethical. Again, I mentioned about this, that AI adoption is to increase revenue and decrease costs.

45:55

So this is where

45:57

the, there's a huge cost decrease, then the highest cost decreases in manufacturing, because it's actually then removing the human labor, and the highest revenue increases in marketing and sales.

46:18

So

46:21

and this is why Australian government has started to look at the need for an AI ethics principle. And these are the eight blocks of the ultra AI ethics principle, versus human social and environmental wellbeing. Human Centered values, fairness, privacy, protection of security, reliability, and safety, transparency, and explainability, compressibility and accountability. So what are these I need these are the things that is in your reading for this week. So I'm not going to try to read here, but I assume you all have read it. Otherwise, you will access a slide later. And Altria is not the only one trying to come up with an ad explain work. So these are the global effort taken from CSR report. You can see in the global back. You UK in 2018. November, the center of data ethics and innovation announced that they will advise government on this governance and regulation to go AI, France in India June 2017, earlier that UK and France, Canada and 2017. Eu. So Japan, China. So you could see this a global effort on looking at ethical AI framework in order to regulate adoption of AI because it goes so fast. And as we know from our discussion last week, there's ethics there's also law right? So what is legal and may not be ethical and what is

48:17

you know,

48:19

what is actually unethical, maybe illegal, so, the other way around. So this one is never even tried to first have the ethical framework course before we even talk about illegal because without adequate framework, it's very hard to

48:32

legalize things. Now,

48:37

let's talk about accountability, promising governance. So, what those two are one of those AI was human values already covered quite a bit of human centered values, accountability, promising governance and one of the aspects in in this slide, he is reliability and safety, solid reliability and safety one of the most interesting stats case study on reliability, safety has actually predated before AI. This is a this was a computer system in the 60s, Herrick 25. So, what is what did the system do? The system was known to be the cancer Zappa. So it is used to automate the amount of radiation that to be used on cancer cancer pack patients. So once they got into the bed on a big to the system that will regulate how much radiation is needed to kill

the cancer under the to tar humor of the patient. But unfortunately, during its operational time, it has killed several people and many, several. And there was there were dozens though there were actually burned

50:15

badly, so major injuries. And he was deemed to be one of the worst software bugs in history.

50:28

So

50:30

the problem with this is the, the Def Leppard didn't have a proper safety checks. There was sort of, I think, a gap of communication between the developer and the domain expert to download the radiation, the radiation, technology, and the medical doctor. I think at one time, there was a function that actually was not properly checked that it could also I think it was 10 times, it could give a radiation at the scale of 10 times

51:09

more than it's needed.

51:11

So it was the worst or the worst history of software bugs. So I think there's a lot of syndromes was a long time ago, many decades ago, based on these in software engineering practice, you know, there's an importance on unit testing for each of the unit and each of the function, and then only when you deploy their real world, you need to know what's the actual implication, because sometimes what happened in the lab is not the same as when you deploy in the real world. So you need to have multiple checks before something is released in the wild. Now, let's, let's come to the very contemporary example. That's just in the news, again, Robo debt. So this is a famous case,

52:02

from here,

52:04

and this is not AI Bay. So just like power 25, that was not AI base, it was just an if and else function is that it was just a heuristics rule of thumbs. Robert, that was not AI based leader, but

it's an ADM, it's an automated decision making system. So carrot 25 was an ATM. So if replaced, so with care, 25 could replace the operator, the human operator of the radiation machine, with Robo debt, this was deployed to replace a manual check, by people essentially, to see if someone is overpaid with a Centrelink payment. Now, what normally happen is, it will, it will look at history of your salary. And then if someone's overpay, with human operator, they can actually see the nuances of all the different data at hand. But she encountered that. And this was the unless it's properly properly properly designed. And this one was essentially it was just like a database search and filtering and matching system based on heuristics. It was not AI based, as I told you, and, and on yesterday, it was announced that 200,000 probate cases will be wiped. So why was the implication proper debt and a story of this lady for example. She one day I got a message, you owe the government

53:59

$1,000 $8,000.

54:05

And I'm not even employed, how can I pay $1,000 And so she is there. She doesn't have she's got no jobs because she's been having severe seizures and epilepsy. So she's been out of work previously, she was a research assistant at University of Newcastle. She's got

54:24

a PhD. And

54:27

apparently, after that, she went to the lawyer and they went to check robot that look at hers scholarship stipends that was given to her and was supposed to be tax free. But of course, is if someone is behind rather than a machine, you'll be able to know this. This is supposed to be tax free, but they don't know the machine doesn't have that clause. So the the robot that says those three years you didn't pay any tax, so we have to play have $1,000 from you. I know many cases like hers satisfy 10s of 1000s, not just 8000. And they were and the scheme claim $2 billion from more than 400,000 Australians, there was a lot of money because that was including 750 million, all depths and the rest are from the lawsuits. Last year, the federal court approved $1.87 billion settlement. So we're wiping all the remaining debts and the legal costs of $8.1 million, because there was a lawsuit from all the

55:50

people affected by Robo debt. Now in this

55:57

55:57

scenario, this goes back not beyond reliability, it goes to

56:01

accountability.

56:05

The question is who was accountable moreover, that if and until now, no one's happy the right answer? Because number one is the problem of many heads. This was, again, this is actually an overview of our panel, this involves the problem many heads. So the problem but we know that there was 35. But who is on the database? How is the heuristic set in place? Is the heuristic valid all the time because people's salaries get changing? You know, the rules today may not apply tomorrow. There's so much thing so who deploys who made the decision for the system to be used on the go

56:53

problem for many years.

56:56

So this is still a very much open problem and accountability of software systems

57:03

and AI.

57:07

Anyway, I think we are at one o'clock and I have low battery to run from office to get my charger. So let's stop here and come back at 110 minutes part one

57:21

running with my dog

57:25

probably just leave that running. Recording

57:29
later but he wants to charge up all right, thank you.

57:35
Thank you.

57:38
But yeah, please pick up

57:42
other slots lots of chocolate No no no

57:53
this is not a problem.

57:54
I can't love this. It doesn't

58:00
I had the same it's one of the

58:07
Is it true I might have a separate society.

58:11
Let me see second one I don't know if this one we do the same our output file. It's a little longer than

58:30
computer science I do come back and I

computer science I do come back and I

58:34

this one here not everyone

1:10:49

Hello, hello. Okay, so thank you, Tom who say that, hey, I have to actually walk back in office share my screen again

1:11:16

okay, um, accountability. So

1:11:22

one of the important thing to consider also is accountability isn't so much fast foot. And, for example, the problem reading hands these days is so true, especially of anything that you make autonomous like for example, autonomous vehicle. In fact, car manufacturers like, you know, follow or Mercedes and all these companies, they basically just an integrator off software hardware system.

1:12:02

For example,

1:12:05

there is one company kind of remember, oh, it's called. It's called Mobileye. Right, so mobile AI was a spinoff from Hebrew University of Jerusalem. It was acquired by Intel, I think for $18 billion a couple years ago, because of the pervasiveness of mobile II system, in all the ad companies, they own every single company, almost every single company that I know actually adopts Mobileye system for the situational awareness of the car, because it's actually has a 90 scenario of vision to actually detect objects and things that move around the car. And yet, there's only one that sensor system just one of the many, in fact all mobilizes sensor system and actually multiple has cameras, infrared and all those things, but be an autonomous vehicle there are at least 600 sensor systems. So, the problem enhance becomes very true, because if if a lot a navy here a colorable pedestrian the amount of traces led to have to be done it was it will be a quite a tedious work to actually understand the root cause, because it could be one of the many systems that can cause a failure. So, the problem enhanced is something that is still an open problem. He has to be assured and deployment of systems

1:13:45

of systems especially

1:13:48

now with data being the core of AI, and this is taken from CSR radio 61 ai ethics framework discussion paper. This is the fact that AI becomes so powerful like people will because of postural data. So how the power of AI comes from your data and more data. And less that means this could be related to privacy issues. And this is a principle that we we all need to be aware of. There is a trade off between privacy and utility. So if you want to be more private, then that means they'll you won't be able to access high end features. So the software that you're using may not be as useful as those who actually sharing the data with the software. So the more access to your data that you allow You then do more, you will be able to access all the functionality of that software. You might notice now, because of the requirement of the consent process that is quite specific, this was just released a couple years ago, whenever you install an app on your phone, or you use a new set of software, it will tell you, for example, this app will use your microphone or this app will use your camera or will have access to your photo library. And then you have to take whether you allow it or not. And it will tell you why. Why it needs your microphone or for example, if you don't allow it, or even as simple as a location serves, if it allowed location sharing, then you need to manually enter the address every time you go right in and on the road, it automatically captured on where you are to enable a faster search. So there's a trade off between privacy and utility.

1:16:13

Just a quick hands on here, because I I think the demography in the room is pretty much within a similar age range. But I found that whenever asked this question, the difference would be a very different answer. Are you if you're if you if you're have to choose which one do you focus on? Which which one? Would you actually put more priority on privacy or utility? If you're more private person? Privacy focus person, right hand or utility focus

1:16:48

left? Quite a mix.

1:16:55

See, I can see some of you do right size 5050 in the room,

1:17:00

online. Where's the chat?

1:17:07

Privacy normal messes? Well, you can, you know, your reaction anomaly will tell you whether you're more privacy utility, let's say you install something in it says you know, you accesses data, or FISA cookies, accept all cookies, you just don't care about antiquing are all these things. If you actually a more your utility person, you will not you will not care, sometimes you just get except Well, I don't care. Trust me, you know, I'm just going to use a software. So what does it mean measuring privacy? There are a couple of things. So this is still taken from CN Tower? No, privacy is not really my expertise. So privacy is one is to do with secrecy. So it's about how can people learn about us? How can people infer information about us from data that about us that, you know, there's been access and secrecy. Secrecy is important because sometimes you want to hide information from others. Because if people know that information about you, they might change their view about you. So this is secrecy. And it could be secrecy to do it, for example, whether you have kids or not, you know, because you don't want any software company laws, because that means you might be presented without advertising that you don't want to. You don't have any implication on your kids, for example, right? So this is secrecy. Some people want to have separation between their private life and you know, their public life. And I don't want maybe you're not so much a secret person. You know, you can see you can find any information you want. But as long as I'm anonymous, so it's about basically, if I share this information, I can share this, as long as I'm not name on it. So let's anomaly mentee. So in terms of public gaze, do people know who we are?

1:19:23

And then the last

1:19:25

is, solitude is basically the degree to which others have physical access to us. So how remote Are they from US soldiers has to do with being able to access you. So the more contact information we have online, the more people less audited you are, the more reachable you are. So this is basically the different ways of measuring privacy. Now, this seems to strike seems to be straightforward. Let's like look at one case study.

1:19:57

Excellent recognition technology, my Amazon Alexa

1:20:12

so in October 2018. So I just saw Jessamine shout utility. So we do have 5050 as well on the chat. So in October 2018 Amazon issue Pater allow Aleksa to decipher user physical characteristics and emotional state based on their voice.

1:20:36

So using the voice features

1:20:38

like like accent, ethnic origin, emotion, gender, age, background noise will immediately be extracted from your engagement with Alexa. And then you that feature that will be tagged to your user data. And this will help more targeted advertising. So this is why you see advertising is like one of the ones that has the highest increase revenue because there's so much more data being learned about us even from Amazon. Now the algorithm will also consider customers physical location based on their IP address shipping address, browser settings to help desk mundane Acts

1:21:21

Now,

1:21:23

what is the problem here to deal with a problem if you look at this measure of privacy? does it violate secrecy on a mini solitude? What do you think Alright, secrecy Yeah. And you want to expand on that oh, boy, I'm really

1:21:57

quiet. So today Yeah. So the technology is trying to

1:22:03

reveal something that we have not revealed to the technology right. So revealing the ethnic origin for example. Not not maybe not all of us are comfortable with that with that fact that you know, I think origin should be revealed to the technology and then use for advertising so the questions I have for you so can can you just do it in pairs? I'm just for three minutes to discuss these questions. And for you online I'm sorry I can put you in groups of pairs but feel free to chat on using the chat try to answer these questions go and have a chat. So in relation to the again when asked like a privacy versus utility

1:23:17

understands There. There Guys I don't

1:25:19

know think for the same discussion if any of you are still bearing for for the utility camera

obviously you can also say why well the sitter right

1:26:26

if you want to know

1:26:28

the payment document you can just look it up it's available on Amazon excellent I love a discussion online Jasmine said life is too short to consider privacy

1:26:59

he wants to have them

1:27:00

all all the features available in performances. Don't all of us use data online All right, alright, let's read the discussion. So anyone in the group is more on a utility camp who wants to have this feature? And don't mind that Amazon is actually doing this on you.

1:27:47

Jasmine, Jasmine, do you want to activate the camera because you say life is too short to consider privacy or just men online? You want to speak

1:27:57

about that?

1:28:01

My dog ate my camera. No. Camera. Just audio just microphone or microphone as well. Okay, so she she wants to have them all but she wants to be solitude so she doesn't want people to reach her. So

1:28:18

no worries. Okay. Um, anyone want to share something? Yep.

1:28:24

So it's so concerning the accent recognition thing? Yep. If someone comes up to me in face to face, and he can see my face, he can hear to my voices and my accents. And he can judge the love from Asia. And then he can like talk about things from Asia to try and engage with me into a conversation and maybe get something from him friendship or money or something. That's all right, but AI doesn't dissolve itself. Okay, I don't

1:28:54

I don't see the like, it's maybe the AI is stealing things from me that I don't want to show

1:28:59

back. All right, yes. I think with with AI, like it's gonna hang on