Transcript provided as a free resource by 

Data Mesh Radio Episode #175: Ethical Data Usage - Informing and Educating Consumers

Interview with Esther Tham

Listen (link)

Transcript provided as a free community resource by Starburst.
To check out more Starburst-compiled resources about Data Mesh, please check here: https://www.starburst.io/info/data-mesh-resource-center
To get their Data Mesh for Dummies book (info gated), please see here: https://starburst.io/info/data-mesh-for-dummies/

0:00:01 Scott Hirleman
The following is a message from George Trujillo, a data strategist at DataStax. As a reminder, DataStax is the only financial sponsor of Data Mesh Radio, in the Data Mesh Learning Community at this time. I work with George and I would highly recommend speaking with him, it's always a fun conversation.

0:00:18 George Trujillo
One of the key value propositions of a Data Mesh is empowering lines of business to innovate with data. So it's been really exciting for me personally, to see Data Mesh in practice and how it's maturing. This is a significant organizational transformation, so it must be well understood. Empowering developers, analysts, and data scientists with downstream data has been part of my personal data journey that reemphasized the importance of reducing complexity in real-time data ecosystems, and the criticality of picking the right real time data technology stack. I'm always open and welcome the opportunity to share experiences and ideas around executing a Data Mesh strategy. Feel free to email or connect with me on LinkedIn if you'd like to talk about real time data ecosystems, data management strategies, or Data Mesh. My contact information can be found in the notes below. Thank you.
LinkedIn: https://www.linkedin.com/in/georgetrujillo/
Email: george.trujillo@datastax.com.

0:01:11 Scott Hirleman
A written transcript of this episode is provided by Starburst. For more information, you can see the show notes.

0:01:18 Adrian Estala
Welcome to Data Mesh Radio with your host, Scott Hirleman, sponsored by Starburst. This is Adrian Estala, VP and Field CDO at Starburst and host of Data Mesh TV. Starburst is the leading contributor to Trino, the open source project and the Data Mesh For Dummies book that I cowrote with Colleen Tartow and Andy Mott. To claim your free book, head over to starburst.io.

0:01:49 Scott Hirleman
Data Mesh Radio, a part of the Data As A Product Podcast Network is a free community resource provided by DataStax. Data Mesh Radio is produced and hosted by Scott Hirleman, a cofounder of the Data Mesh Learning Community. This podcast is designed to help you get up to speed on a number of Data Mesh related topics. Hopefully you find it useful.

Bottom line up front, what are you going to hear about and learn about in this episode? I interviewed Esther Tham, who's an experienced designer at Thoughtworks. I reached out to her to talk about data ethics based on a post Esther had made on LinkedIn. Here's some key takeaways or thoughts from Esther's point of view. Number one, when designing your UX, your user experience, companies should aim for as little as friction as possible when signing up with somebody or transacting with them in general. For an ethical company, that means collecting as little information as possible to still maximize value of the service to the user. Number two, companies, if you don't need it, don't collect it. It isn't ethical, but also it increases your attack surface for a data leak and potentially lowers consumers trust. Number three, we don't have the proof points yet of many companies doing the right thing and disclosing potential issues of sharing information with them in an understandable way. It makes sense because that might be something that would scare people off or would prevent them from sharing some of their information, but that would probably increase consumer trust if they did do that. But is that increased trust worth more than the hassle to a company or worth more than the risk? We need companies willing to try being more ethical to really know about this. Again, we don't have the proof points, but being that upfront with their disclosure, it's a cost with a very uncertain upside. So it's hard to see many companies really going for that just yet. Hopefully we can get a few that will.

Number four, people need to learn that their personal data has value and risk associated with it. Don't give it over to companies without thinking about how it might be used or misused. But most people are nowhere near that thought process yet. Right now, most people are only worried at most about getting scammed, not should this company have my data and how might they misuse it. Number five, ethics isn't just about collection or even usage. Protection is also crucial. If you can't protect sensitive information, you shouldn't be collecting it. Number six, how can we encourage the general population to really care about ethical collection and use of their data? Is it just a better explanation of how it's going to be used? With greater understanding, will most people actually care? We still don't really know. Number seven, the question of who is responsible for ethical data collection is an interesting one. On the one hand, companies should be behaving ethically. On the other hand, they often don't. So how much responsibility to protect sensitive information is on

the consumer, not handing it over in the first place? You know, that kind of fool me once, shame on you, fool me twice, shame on me. Yet more companies fooling me every day about not doing things ethical with my data. When should the shame fall to me?

Number eight, a designer's role is to advocate and build for the user. But we still don't really know exactly what most users want when it comes to being ethical around data, right? Do they really care about ethics around their data? Or are they willing to trade their data, pretty much almost any amount of data, for certain services? Is it about more education, communication, or do users genuinely not care? We need brave companies willing to test this and provide us some of this information and insight. Number nine, how do we press companies to be more ethical in the data they collect, how they protect the data and how they use data? Have many companies suffered damage, reputational or otherwise, from ethics breaches? We clearly can't trust every organization. We talked a little bit about the Equifax thing. And Equifax's business went up after they had this massive, massive breach. And they had all these distractions, but it just meant that people saw that they had more and more data and their business didn't really suffer much from it. They didn't do a whole lot of fines or anything. So, we need to have a little bit more of a hammer around the unethical use, right?

Number 10, on the flip side, how do companies that are actually doing the right things ethically communicate with those ethics? Like, what are they doing? Is the cost of behaving ethically worth it? Does it result in a tangible benefit? We assume there is an additional cost to behaving ethically too. So, there needs to be an upside for companies to consider it. But I think one that you might look at as a case study around this would be DuckDuckGo, the browser that's really focused and the search engine that's really focused on consumer protection, or at least they're telling us that. What do we really know? We're not sure. Number 11, it's easy for consumers to have a false sense of security online relative to their data. While identity theft and similar issues are on the rise, companies are still asking for, and consumers are regularly freely giving, sensitive PII. Number 12, very few people really think about potential misuse of data we give to private companies, often with little explanation by those companies of what they will use our data for. Can we really expect companies to fully explain their projected use of data when that might simply confuse more people? We can press them to do it, but likely not expect them to do it willingly.

Number 13, however, when companies do try to explain their use of data, does anyone read it? Are EULAs actually useful? Do we need something that is in addition to a EULA to explain how data will be used and what will be collected? That kind of idiot's guide to, here's what we're going to do with your data, because again, that can scare people. Number 14, and finally, when most people seem to not really be all that

concerned with the data they share until it seems it was used improperly, especially if a company sold their data to a partner or some scammer got a hold of it. Esther talked about that genie is already out of the bottle. And most people don't expect a scam to happen to them specifically, so they only look to react after the data is already out there. Can we change their approach and view? Is there really any benefit to pushing on consumers to be more conscious? And is there any benefit to pushing on companies to do this? Does anybody really care? That's kind of what we talk about, and there's a lot of discussion around it. But I think this is the beginning of the conversation, not the end of. With that bottom line up front done, let's jump into the interview.

I'm very, very excited for today's episode. I've got Esther Tham here, who's an experienced designer at Thoughtworks. And I saw that she was giving a presentation talking about some stuff around ethics and how important ethics is when we start to think about data. So I reached out to her and we're going to be talking about how can we actually start to get better around data ethics? And so much of I've put out a Mesh Musings about ethics and ethics isn't only about bias. Like, how can we be ethical in how we're actually leveraging data and how can we still drive value and that it's not, oh, we can either be ethical or have value. And so we're going to be talking about a lot of things around data ethics. But before we get to that, Esther, if you don't mind, if you could give people a bit of an introduction to yourself and then we can jump into the conversation at hand.

0:11:20 Esther Tham
Sure, Scott. Hi, everybody. My name is Esther. I'm an experienced designer at Thoughtworks. I'm based in Singapore, even though I've worked in the States for about maybe five years. So I originally started out in media and advertising, but given that I used to like push like fast food, pharmaceuticals, like products to people, it kind of left me a little bit like a distaste, I guess, that industry. Like, why am I kind of pressuring people to buy products that they don't need? So hence, I kind of switched over to experience design and hope that I can be like practice more ethical design, like in my current career track.

0:12:04 Scott Hirleman
Yeah, I think that moral compass, it can be a little bit difficult with finding that career path, but I think it is more rewarding. So one thing we wanted to kind of start off with was the talking about what like what ethics actually has to mean when you're in that private sector versus the public sector, because in Singapore, a lot of the government entities are rather large. And so kind of looking at this from a different lens of, okay, if you're in the government, if you're not in the United States, the government in the United States doesn't have much ethics embedded into a lot of the things they do. They have some, but like how do you start to think about like, if you're not guided or

the government is guided by rules and regulations as to what they can do, but private companies aren't. So why don't we start about like, what are the rules and regulations that you think are useful that people could look at from the government side as to what they even can't do? And that can guide us a little bit towards what we shouldn't do and things like that.

0:13:21 Esther Tham
Sure. So maybe for the benefit of like listeners, if they're like worldwide, maybe just to set some context as to like, I guess the environment that Singapore plays in. So as everyone knows, we are kind of like one of the most safe cities in the world. Like I grew up literally, I could walk out in the night and no one's going to really come mug me and stuff like that. So a lot of us, we kind of grew up in this environment where we feel like really safe, really secure. And unless we've gone overseas and experienced very different countries and different climates. So we really take our safety for granted. And this kind of like physical safety translates to like, I guess the mindset of how safe like online is and how safe our data is. And I guess maybe to set more context, like in Singapore, we have our national identification number that actually is correlating to our birth certificate number. And it also used to be also doubled up as our passport number. So essentially, we have like one set of national ID that is used across like three or more different identifier documents.

And it wasn't until like, I think 2006 that the government really like enforced and say, hey, we actually need to delink the national ID number from passport numbers because it used to be if I had to renew my passport, I would actually get the same exact serial number back. Now we don't. We have like randomized and unique numbers for every single passport that's issued just to cut like all these stolen passports, fake passports, etcetera. So it's kind of like I realized that even for myself until I went over to the US to study, I kind of really didn't know or had the concept of what identity theft, identity fraud like is and how to really like protect ourselves. And also it used to be up until 2016, companies like any companies can freely ask for your national ID number, the full string, and people just willingly give it out like for free. And we don't really realize like the repercussions of us like doing so because that like string of numbers like you can use it and you can look up like people and their data and whatever data that's actually stored in like the government databases you hack into that.

So I think for us, we are really very, I guess, privileged in a sense that it's not so like up until recently, there hasn't been a lot of I guess very malicious like acts and like stolen identities, but we're kind of seeing it like more and more right now in Singapore as well with regards to like identity theft and like just hacking of bank accounts and so on and so forth. So I guess just a bit of context as to maybe why like for us in particular, it's becoming more and more important that we actually take steps to

safeguard our personal data if we know that other people might not be able to do it for us or do it like well like enough for us. So I guess to circle back to the government. So the government wise like because they are held like accountable and Singapore kind of like prides itself on being like not as corrupt as you would if you compare it to like other countries around the world. So they kind of take steps, I guess like good steps, I would say at least to really like protect citizen data and to make sure that it doesn't get like hacked or stolen like as far as possible.

And for one, I know like we are moving towards like having a digital identity, it being like more safe and secure and like not easily accessible or not easily, I guess, counterfeited in most cases. And for agencies, they do like they are very careful in handling like our like personal and private data. And even for agencies that do contractually like have people participate in longitudinal research, they do protect the data like properly to make sure that it's not distributed, even though like there have been cases of I guess civil servants actually accessing people's like private data for their own like personal gains, whether it's like personal personal gains or for other like monetary reasons or what not. But there have been like really bad repercussions for people who have done so like accessing like government data for their own personal gains. So I think for the government wise, they are very careful. But for private sector wise, like we can't be like really certain, right? Like how they really like use our data. We know that a lot of companies collect a lot of data, maybe more data than they might need.

And for customers, like I guess in a sense, we rarely question why they might need like all sorts of data. They ask for email, they ask for a phone number, they ask for like your phone number. Like your address, your credit card number. Like is it really necessary to give it all to them? Like for me, I would question. I rarely give up my credit card information for one, like to let them store it. Even if it's transaction, I kind of let them transact it, but don't store it. So I don't know, Scott, like do you like how much data do you actually let companies store?

0:19:14 Scott Hirleman
Well, and I think that's the question, right? Of well, why do you need this data? Why do you need... If I am going to be transacting through this company very often, and I don't want to have to necessarily put my credit card in every single time if I'm going to be doing a lot of things. But if it is a one off transaction, yes, I don't want to give them that. Why do you need my address if it's a digital transaction only? Why do you need this thing? Maybe you need my billing address to get it through from a credit card standpoint. But if not, if I'm using PayPal or something like that, why do you need these things and what are you going to use them for? It's kind of the same thing why Apple has started to restrict and show like on the iPhone stuff, Android's doing it a little bit too. But of this app wants access to this, this app wants access to

this and showing kind of fine granularity. And they like... Are you going to... One, do I trust you? Right? Like I have everything. I have a Raspberry Pi Pi-hole, which is like a DNS router. And I block anything from any Facebook entity. So like Instagram, Facebook, Giphy, like any WhatsApp, any of those things.

So but like you said, we're more aware of this. Are people generally aware of this? I think that's the good question of, I don't think that Singapore is that different in that you feel safe. Most people haven't thought about what are my threat factors? What is my threat surface area around my digital information? And so I think you've got something where your government is heavily scrutinized where a lot of other governments kind of aren't. And I've had on the folks from NAV in Norway and their government is kind of proving that they're a government service and most government entities want to be a service. But you look at the CSA, the FBI that are collecting this information, even oftentimes illegally. And so it becomes difficult to trust that government.

But one question I would have about what you were talking about there is when you don't have the rules and regulations, how do you think about a company having... How would they guide themselves as to what is good and what is not good, right? And how do you exchange that information with people about why am I collecting this information? Do you have things that you've seen that you're an experienced designer. You know, everybody, at least in the States, has those EULAs, those end user license agreements or those whatever that pop up and you just click, yes, I've read this and move through, you know. So have you found a good way? Let's not even talk about the ethics of what it is, but have you found a good way to communicate, like, what are we going to use this for? Why do we want this?

0:22:46 Esther Tham
I think that's a dichotomy. In terms of UX, user experience, what we aim to do is to really ease and create almost frictionless experiences. So for instance, like you mentioned. If I transact with this site or this retailer like many times frequently, I would store my credit card information so my checkout is like quicker every time. They remember my information, I can just like one click checkout or whatever, like Amazon does. But if you think about it, like, how should I put it? In terms of like security, like how secure it is, do you trust the retailer to be able to protect your data? So collection is one thing, usage is another thing, protection is the third thing. Nowadays, there's a lot of instances of people like hacking into systems. We've had a lot of high profile security breaches and stolen like data from like hundreds of customers. We had a situation where like Razor, a company in Singapore, they had a security breach because of some mishandling of code from the vendor. And they actually brought this vendor to Singapore High Court. They settled it out of court, but it was a very high profile case a couple months back.

Singtel, our telecommunications company, there was also like security breach. And I think employee data was stolen, along with like credit card information, PII data was stolen. Most recently, Carousell, an online like secondhand retailer, I was part of like the customer data that was actually like stolen or hacked into. And the government like sort of like tracked it down. And that data set is being sold like on the black market to who knows where. So I guess it's like, how much do you trust these companies to actually protect your data as well? The government has like tried like really hard to make sure that they don't get hacked. But there are also so many companies out there that we give out our data to. I don't even know how many like companies I have accounts with, but probably more than like 20, 25. So yeah, like it's kind of like one thing that I've been thinking of like more and more as I kind of read all these like news reports, even just local news about like all these like security hacks and breaches. I always, I will kind of like, well, it's down to like, how protected am I with my data? Not even talking about usage and what kind of targeted ads they throw at you, but is my data even safe anymore?

0:25:46 Scott Hirleman
Yeah, well, that makes me think of in the US there are these three kind of credit agencies that... Experian, Equifax, and I can't remember the third one right now, but that they track all this information and they store all this information about US citizens without anyone's consent. They grab all this information, all this information is pumped to them and we have no ability to opt out as a US citizen. And then there was that big, I think it was Equifax got breached. And so everybody's basically personal information and all of like this credit information and things were stolen. And then I locked down my credit from that just so nobody could open up a credit card or anything like that. And then when I tried to move my official address with the US Social Security office, I literally couldn't because Equifax own my address. I have to be able to have my credit unlocked for the US Social Security office because they use this company that didn't handle our data well. All of people in the United States data well and it is super crucial of like our financial well being, our social security in our retirement is controlled by this company that had almost zero repercussions from leaking everybody in the United States data.

So I'm jealous of you in certain aspects, but like, I still think when you're thinking about that experience design, are you working with clients, are clients wanting to provide reassurance to somebody that they can understand what is being collected, how it's going to be used and how it's going to be protected? Or are they just trying to not even have anybody think about that because even bringing that up then gives people hesitation, right? When you're thinking about moving through checkout and then you see this long list of here's how we will or will not use your data, it makes you start to think, well, do I want to give this company my data? And you start to do that

of it can give pause as to the additional information. Have you seen that your clients are good with sharing that or like being open and honest about that information that they're collecting and how they're using it and how they're protecting it?

0:28:35 Esther Tham
I can tell you for one that my government clients are definitely very transparent about like data that they're collecting and what use it is for. And they do actually make you read or at least do the conscious like checkbox to say like, hey, I've read these pages of like EULA and agreeing to it before I like proceed with any transaction or agreeing to use certain services. For private companies, I would say not so much like even for maybe like checkout on transactional types of service. It's very, I guess, minimal compared to like how verbose and how detailed the government's like terms of use of like certain systems are. But the thing is that even if we provide like EULA, the problem is getting people to actually read it, writing it in the language that people understand it and getting them to really consider like what are the consequences of me basically providing this set of data to whichever service or company that I want to transact with. Like mainly for me, I guess the biggest issue that I feel is actually getting each and every one of us individually to really read and understand and be conscious and aware of what we are giving up, like what kind of rights we're giving up when we transact and like use services.

Because there's only so much like we can write it out, make it as easy for you or as hard for you to access it or transact without like going through and reading it. You know how like when Apple shows you the EULA, you have to scroll through the entire thing before you can click like agree or disagree, right? So they kind of force you to go through the whole spiel. And even so, who reads it? I kind of just read like the first few paragraphs and just skip the rest. I'm pretty sure you probably have not read every single word of these kind of like terms and conditions, right? Let's be honest here.

0:31:09 Scott Hirleman
Yeah, well, and I think that kind of brings up the question of do people really want to do they just want to use the service and they just want to use the service and they just want to use the service. Yeah, do they just want to use the service and they don't want to care or right? And is this like the people have talked about the digital information wallet or things like that where you say, hey, I'm going to give hey, okay, I want to start to use this service. Okay, here are the five things, the five concerns that we if you're the digital information wallet holder or whatever you want to call it, that here are the five concerns that we have about this company, right? Here are the five issues that we think could be that you should be aware of. And do you want to give them that? And so that there is something that is a barrier, but like like you said, nobody really reads through the EULA and they kind of do that on purpose, right? I

don't know if you ever heard that story about the guy who won $10,000 because he said like, and we will give you your, we now own your immortal soul or something like that is part of their EULA.

And when he brought it up, there was a bounty on it that the first person to bring it up got $10,000. And it was out there for like years and years before somebody did it. And so it was like, do you think that people want to care or do people just want to feel safe and then be, and then they get taken advantage of, or like, is this both sides don't really want to do this. And so the companies just want to be able to collect what they want to collect and monetize however they can. And the users are just like, fine, fine, whatever, I get it. And I don't want to have to care about that.

0:33:08 Esther Tham
I kind of have to go with like a bit of both, like, because like you said, the user, the companies, they just want to make profits, right? The AARR framework of acquisition, like they want to acquire like users, which means that they want people to sign up, create accounts, and then like repeatedly like come back and give them more business. And to get accounts, you need like the user data. And then we've always heard like these, I guess, horror stories of companies exchanging like customer data with other companies for like other benefits that they might have or like partnerships that they have. And then they kind of like push extra product onto you or like what not that we at that point don't have control over how our data is shared, like among like other from companies to companies. But on the other hand, like for consumers, we like, oh, we just really want to use this service or buy this product at this like point in time. I just want to transact as quickly as possible and just like pay and then just check out and just like use the product.

So I think it comes down to like unless we actually are at the receiving end of some scam or some like a malicious act that's done to us. It's like I don't want to think about it until it happens to me. And people have that like notion that, oh, it's one in a million chances of like me getting scammed. So until it really happens, then it's a once bitten, twice shy thing, right? I don't think about all the extra safety and precautions that I actually should take with regards to like protecting my own data, because unless it's out there, if it's not even out there for the world to like consume or it's not made public, then people will know that I have this set of data that you can take from me. But once it's out there, like there's no way of putting the genie back in the bottle. So I think it's kind of both ways, right? We need to really start thinking about what kind of data we're putting out there, because it used to be we didn't really have to provide so much data. When we go to buy things, we go to the store, we pay cash and then we get our product and we leave. They don't take our phone number.

They don't take our name. They don't take our credit card. But now we do. Now they take everything from us. So there is the difference, right? Like it's so much easier for them to just like do what they want with like us. And because we are also in the end like commodities to these other companies that they might decide to monetize us to some other company to get even more profit or benefits from like the other companies as well. So we're just giving up all control.

0:36:18 Scott Hirleman
But I mean, you kind of touched on it in there is do people care about that control. Is it that until they get hit? And so if you're say you're a company and you want to act ethically, so you want to disclose all of this information. Not only like I mean, do you see that as purely something that is of benefit or like it's like click here for our read through the EULA and then click here for our inhuman language EULA, like because we have to use all the legal language, but click here to actually understand what we're doing. I mean, again, if it prevents you from doing business with a lot of people, simply because you are giving them the actual information? Let's say you are using the data ethically, but you're being ethical in how you share that you're going to use the data ethically. But you're giving people more information. You're giving people more potential points of friction versus that hundred page EULA that nobody's going to read through and just go, Okay, okay.

I mean, is it something where have you seen? I mean, maybe we haven't even seen companies really doing this very well of really explaining what they're doing, what they're using the information for. So have you seen anybody that is doing this really well outside of maybe the Singapore government, but like and putting it in their language and then that's had a positive impact because if it if it has a double negative impact, right, of then you have fewer people signing up and fewer people being comfortable with what they're doing, like then it's no company in the world is going to say, “Okay, yeah, we want to do that, even if it is ethical.”

0:38:11 Esther Tham
Yeah, I don't know. I don't know if I've ever seen like companies that will like risk, I guess, a profit margin to be completely like transparent about, hey, like there's potential that this data might be used here or there or some harm might befall you because we store your data. You may get hacked like one day, like who knows, like just listing down like all these like things that could happen just because you decide to like sign up with them. I don't know. But on the other hand, wouldn't it make these companies like really much more trustworthy? Like maybe trust is not a value we put like on enough. Like do I should I transact with a company that I absolutely trust or should I transact just because it's convenient, like for me and I can get like what I want at a good price? I guess what what do we really value right in this like very capitalistic world? Do we just want our commodities and we don't really care?

0:39:14 Scott Hirleman
Yeah. Or have we been trained to not care and that people actually do want to care? It's but like, I mean, if I were to talk with my parents about this, right they're in their 70s. And so they just would not want to have this conversation because it's just, I don't want to even think about it. And so is this something that that as people are becoming more and more digital native and things like that, that this does become a thing again, but that when we think about the mass market is it that 50% of the people don't really care? And so, okay, as we start to move forward and forward I mean, Katharine Jarmul, K-Jams, who's also at Thoughtworks was was on recently and was talking about kind of privacy in that that aspect. And people having Finsta, which was fake Instagram accounts. And so they would but that was not really to protect their information from the company. It was to protect information from each other. As to, hey, I can't fine grained controlled my my privacy settings. So I'm going to create different layers of accounts and be like, hey, here's my like main main account. And I don't publish or I don't post much on there. But then I've got my fake account one that is for kind of my circle of friends and their circle of friends and that and then I've got ones that's just for my very specific circle of friends and all of that.

So I mean, I guess what I'm trying to ask here is what benefit do you think? I mean, is do we have proof yet that higher trust drives business value or do we need to have, like if somebody is listening to this and they're like, hey, I'm going to do this and I'm listening to this and saying I want to move more towards ethical data usage. I want to use more towards ethical collection, usage and protection. How can they sell that internally? How can they do that? Is it that they just kind of we need some brave souls to go and test it and say, did this add more value? Did we test this out? And we saw that this led to better results. Maybe AB testing a user experience of, hey, when we really explain like, what is this going to do? What are these settings do? Does that lead to better engagement, better trust and things like that? Or what have you seen anything that's proved out that people should be testing this out rather other than, hey, it's the ethical thing to do. It makes us feel good about ourselves. But companies aren't about feeling good about themselves, unfortunately, in most cases.

0:42:06 Esther Tham
Well, the researcher part of me would say, yes, we need to test this because it's hypothetical. We just assume that, hey, maybe people are maybe part ignorant, part like apathetic. They need to see that, hey, it's actually going to be beneficial for me, for my family, for everybody if we actually do practice a lot of ethics. And I think people maybe want to care, but maybe they just don't have enough knowledge about, hey, what is all this digital, where is it headed? What exactly is data being used for? I think not a lot of people really understand what data actually is used for, what it means when you say, oh, I'm a data company. I don't think people understand. I don't

think people understand what that truly means. So if people don't understand, then it's hard to make them understand. We probably need some sort of proof that is easy for them to, I guess, see that it is working, that ethical data and ethical data usage is working and is better for them as well. I don't know if there are companies that actually do do that.

0:43:31 Scott Hirleman
Yeah. I think this is one of those where it's kind of a pie in the sky right now until we have somebody that's actually doing it. And then, because we keep seeing the things about Facebook and TikTok or ByteDance or whatever and all these unethical uses of data and that we're trying to figure out, okay, but are the ones that are really doing ethical use, we never hear about them because they're never the ones that become successful? Or is it just that nobody's really tried it yet or things like that? So I'm hoping we can get people to go out on that limb, but it's difficult. So we've been talking a lot about the communication. What have you seen that when you're looking at interacting with the government, do they have these crazy long EULAs or do they have things? If somebody wanted to start to communicate better about, hey, here is what our collection, our usage and our protection are going to be, is there anywhere where you would recommend people that are interested in doing this to start to do it?

0:44:49 Esther Tham
Where to recommend? I actually have not seen anybody do EULAs well. It is exhausting to read them. And as much as we try, as much as the government tries, they put it easily accessible and you can pull it up as and when you need to, and they really iterate, hey, you need to really read this. There are repercussions if you don't abide by certain rules and regulations. But I don't know, I still find it's hard. It's just it's not in our nature to want to read legal documents, unfortunately. And it is a legal document unless we can find it a way to make it more easily consumed or just ingrain in us the knowledge of how to know what's good for us. I guess it's in part, maybe it's also something for me to take back, how can I encourage this kind of sharing and ethics for business? And also, in part, how can we encourage everyone, even for your elderly parents, to really understand and safeguard themselves?

0:46:17 Scott Hirleman
And I think that's one of the big issues is, again, we're kind of in a... I don't even know if it's a chicken and egg scenario, which comes first, but are we putting the cart before the horse of saying, we need to do this, but if there isn't demand for this, or that we have to test that there is demand for this and that there is value to both sides, and, and, and. Is that going to really inspire anybody to go out there and do it because we don't have any proof that this is of any value to them? So why would they do that? Why would companies be okay with somebody going out and trying to do

this when there isn't any proof yet that this is of value? And so, but I do think... I think that we can start to have...EULAs have to have legal language. So they're always going to be terrible or I believe so, because you can't clearly explain something and still have legal protections, which is very unfortunate. But I mean, I think having a very guided setup around, hey, here's what we're, here's what we're asking for here. Here is the additional value.

You talked about this in the precall of how do we communicate to people, we're not requiring that you give us this extra bit of information. We need your email because we need to give you the ability to log in. We do need that. But if you give us your phone number, what we then can do is we can work to provide this additional value or we can provide text alerts if you want those text alerts around, Okay, we're an ecommerce site. And we think that you might want like once a day we're going to send you a text that says here's a hot item that you now are the only one who's going to have access to it. Anybody who signed up for text alerts are the only ones who do. Or we, if you give us X amount of information, this is the value that you get from it. And that you make it a somewhat of a transaction and you say and here's how we're going to use it, right? We're only going to use it for these purposes.

I keep getting these things from banks where they tell us exactly how they're going to use all of our data because it's required in the state of California. But you literally can't opt out. They're like, we're going to share our data with our third parties in this way and this way. And it's like, and supposedly you can opt out. But you go onto the site and they're like, no, you can't opt out of it. We're just telling you that this is what we're going to do. We have your information, so we get to do what we want with it. So if somebody were to head down that path, you're an experienced designer. Is that something that you would say should be part of the sign up? Is that something that you should because again, it's additional friction, right? When you think about a sign up funnel, you think about where do we lose people? You want to make sure that you get people signed up. And then do you start to ask for the additional information or do you say like, how do you think about that kind of aspect to it?

0:49:53 Esther Tham
That's a tough choice, right? Do you want to please the client? Do you want to, like, make sure that, like, people know, like, what they're doing? Like, I mean, I guess, or at least simplistically for, like, experienced designers, you want to advocate first and foremost for the users, right? I mean, obviously, you have to balance it with, like, what our client's needs are because they still need to, like, make money. But, I mean, as part of best practices, like, I would say yes. Like, let people know, like, as much as possible. Like, hey, I'm giving you this piece of information, this piece of data. At least, let them know, what is it going to be used for? If anything, it's a courtesy to people, right? I'm giving you something in exchange for what? What's in it for them if I give

you my phone number? Like you said to receive some text messages or, like, hey, the alerts for... I don't know, early bird signups for certain, like, deals or what not. So that's that. And then obviously, if it's not required for doing a business please just don't let people even have the option to provide it for them if necessary.

If you have never needed their phone numbers why even ask, like, for them in the first place? Sometimes people just blindly fill out forms even though you tell them that, hey, this is optional. They just give it up. So, if you don't let them enter it, then you would never have to collect it and they don't even need to know, hey, it's actually just information that they don't need. But, I mean, nowadays we don't think twice, right? Companies just want to, I guess, they just want to collect. So, maybe they was like, oh, maybe down the road, I might need this data. Because what are you going to do? You're going to make the users, like, go through and read another, like, set of EULA after they've agreed to the first one just because you want to change, like, oh, I want to collect an additional piece of information from you. But I think certain companies do do that, right? They keep updating their privacy statements or what not, like, hey, this has changed. This situation has changed. Now, moving forward, we're going to do this or that or what situation, like, your data is now being useful and stuff like that. I've seen it from sites like PayPal. I think I've also seen it from Etsy previously with regards to privacy statements being updated and changed. So, I think there are companies that do try to be good about it and try to stay or keep us updated about stuff like that. But still, it's also up to us whether we read it.

0:52:49 Scott Hirleman
Well, and see, I'm exactly opposite when I get those PayPal ones. I'm just, like, I'm cynical about it because I'm just, like, what now are you trying to do to sell my data? What now are you... What more are you allowing yourself to do since you already have my data? And they do probably, I don't know, 15 or 20 updates a year as far as I've seen at least US ones. And it's just, like, I'm just done with this. Like, I'm not... I haven't logged into my PayPal forever because it's just, I don't... I don't think that I can trust it simply because you're constantly updating this. And when you're constantly updating it, I haven't seen anybody that does it from a trustworthy standpoint. I've only seen it that you're now giving yourself permission to do additional shady things and that this is a legal protection instead of a value add of, hey, you know... And maybe that's just how they communicate it. Of, hey, we've updated our privacy statement and we're adding additional protections around ethical use or things like that that we have to communicate in the legal aspect but can we move more towards as well communicating in a lay person aspect?

And I just don't... I'm hoping we can get there. But I think when we think about exactly what you talked about, if you don't know why you're going to use it, why collect it, and you go, well, it might have value. I talked to a bunch of people early on

in the Data Mesh conversation, and I was, like, you've got all this stuff in your data lake. You just kind of need to get rid of it because that thing from five years ago that you're never using, and it's, like, well, what if it might have value? And it's like, well, the cost of cleaning it up, the cost of everything, and exactly what you were talking about of the protection threat cost, right? Like, if you leak people's data, it can lead to big fines unless you're a US company because the US just doesn't have any teeth in fining people.

But, yeah, I guess, again, I'm kind of circling back to... Let's say somebody listening out there has bought in, right? They go, okay, Esther, I want to listen to you. I want to start to do this. Where do they start to look? Do they start to look at, hey, let's look at ethical collection first? Is that probably the easiest thing to say? Like, if we're not using this, why collect? Or is it protection? And then ethical usage is so nebulous, but where would you recommend somebody start? If you want to tie that into experience, great. If you don't want to tie that into experience, great. But if somebody wants to start looking at this it's a pretty big topic that feels like it's not, but it really is underneath. So where would somebody start?

0:55:54 Esther Tham
Well, you put it in the first place like the data lake, it stores, like, information and sometimes it just sits there because it's not used. So, like you pointed out, first step, if you don't need it don't collect it because you have to treat the lay person as if they don't even understand or know what's going on. They're just there to give you data. If you only need one piece of information from them out of five, just collect that one because if you are then going to store it in a data lake or wherever, then you're only storing that one piece of information out of five pieces that that person could give you. And if there is a security breach or a leak, you're only leaking that one piece of information of that person. So, by virtue of that fact less is more, you have sort of helped to at least give that person a bit more protection that they didn't even know was coming, right? If that is, I guess, a very simple way of putting it, less is more. So, if you don't need it, like, don't collect it, like, don't use it, then you won't have the additional, like, headache of wondering, okay, now there's just extra, loss I need to take because so much more PII data is being leaked.

0:57:29 Scott Hirleman
Yeah, and I think it is something where people are just like, well, I could leverage this or especially I get really nervous when people talk about data marketplaces. If it's an internal marketplace, fine. But oh, well, this value, this data would have value if we were to sell it externally. And it's like, okay, but one is that the business you want to be in? And two, are you there? Like you said, people aren't there just to provide you their data. This has to have a mutual exchange, but we kind of have to have people get to a level where there is a benefit or there is a cost to collecting this information

that you shouldn't be collecting. And we're just not there yet. And are you seeing that consumers or even businesses or whatever are becoming more aware of this? Or because I just I know there are sites where people are talking about this. And constantly people are like, why do they need this? Why do they need this?

But I'm just not seeing it from enough of a mass market appeal partially because the companies that want to prevent you from doing that control a lot of the internet conversation, right? If there's an article that's extremely critical of collecting this information, I don't think, or a site or anything, I don't think Google is going to be super happy to have it rank very highly or Microsoft or all these other companies that collect way too much information.

0:59:03 Esther Tham
Google knows everything about you.

0:59:06 Scott Hirleman
So yeah, yeah, for sure. But like, are you seeing that there's a rise in demand for this? Or is this kind of, again, is this a unicorn world, right? Like a thing that wouldn't be nice, but there isn't enough demand. So companies aren't going to get any value from doing this yet.

0:59:34 Esther Tham
I know for a fact, like companies are probably more cognizant of the fact that any misuse of data would result in like repercussions or like fines, or at least in Singapore, like they do, like they fine them for like bad behavior and like misuse of like data and what not. I think for like individuals, we are getting a bit more aware, or at least in Singapore, due to the rise of like people being scammed of their like personal data, the government keeps like sending us all these like PSA about like, hey, don't freely give out your information, especially for people who are operating scams that impersonate government agencies, which also bear a lot of like consequences, like the fines and arrests and like prison sentences are like out there and like reported and stuff. But I think people kind of just equate that like, I need to protect my data to not get scammed, but I think they haven't really clicked in with like, hey, that also extends out to like, maybe this company could be sharing my data or could be selling my data. I think that connection is not there yet.

They know about protecting themselves from immediate impact. If I get scammed, like if my bank account takes a hit, it's an immediate repercussion that I feel. But with regards to like, hey, if I transact with company A, but they might be selling like my data to company B, I don't see that like connection with regards to like me being an individual. And I think that's kind of where it stops for people. Like they don't see like these secondary, tertiary connections about how data is being like shared and

spread like outside of their control.

1:01:29 Scott Hirleman
Yeah. And I think it's coming up a little bit more in the US with like something like a period tracking app where there are... That in these states that have outlawed abortion or have very restrictive laws that there's it's now an attack factor. It's now a thing where law enforcement can subpoena these things. And if your government is working against you, which is happening quite a bit in the US and in a lot of countries, obviously, then you're probably a little more cognizant of it, but it's still kind of a scary world out there that you can have a lot of this stuff kind of going after you.

So, but yeah, so I think we've covered a whole heck of a lot of things here. But is there anything that we didn't cover that you wanted to or any way you'd kind of want to wrap up the episode in general?

1:02:33 Esther Tham
I think we've covered like quite a bit, Scott, and I guess like at least my perspective, at least my personal takeaway is that you need to understand the value of your own personal data and then what you are in turn exchanging for when you give up that data, right? Because that's what data is. It's in the value of like the information that we're exchanging in return for like what we think we're getting. So unless you really think that, hey, this company actually is going to value like my data and my privacy and they actually stated in the agreements that we sign, then we really need to hold them accountable for how they collect and use our data and how they protect it.

1:03:19 Scott Hirleman
Yeah, very much agree. So, well, I'm sure there's going to be a lot of people that would love to follow up with you. Where's kind of the best place to do that? Anything specific you'd like them following up about?

1:03:32 Esther Tham
I think the best place would be to follow me on LinkedIn. I guess that's the safest professional platform right now to be sharing data.

1:03:44 Scott Hirleman
Awesome. Well, I'll drop a link to that in the show notes. And again, Esther, thank you so much for taking the time today and as well, thank you everyone out there for listening.

1:03:54 Esther Tham
Well, thank you, Scott, for having me. It's a great fun talking to you.

1:03:57 Scott Hirleman
I'd again like to thank my guest today, Esther Tham, Experience Designer at Thoughtworks. You can find a link to her LinkedIn in the show notes as per usual. Thank you.

Thanks everyone for listening to another great guest on the Data Mesh Learning Podcast. Thanks again to our sponsors, especially DataStax, who actually pays for me full time to help out the Data Mesh Community. If you're looking for a scalable, extremely cost efficient, multi data center, multi cloud database offering and/or an easy to scale data streaming offering, check DataStax out. There's a link in the show notes. If you wanna get in touch with me, there's links in the show notes to go ahead and reach out. I would love to hear more about what you're doing with Data Mesh and how I can be helpful. So please do reach out and let me know, as well as if you'd like to be a guest. Check out the show notes for more information. Thanks so much.