



Transcript provided as a free resource by



Data Mesh Radio Episode #162: Creating Data FOMO and Keeping Close to the Business

Interview with Dacil Hernandez

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/>

0:00:00 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.



Transcript provided as a free resource by



0:01:48 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 Dacil Hernandez, who's the Director of data and AI for Northwest Europe at Nagarro. Some key takeaways or thoughts from Dacil's point of view. Number one, at the end of the day, tech is the easy part in data. Creating value through data is the hard part, and why do work if it's not to create value, right? So focus on creating that value. Number two, it's incredibly easy and pretty common for IT in the business to get disconnected or never be connected in the first place. Focus on creating and maintaining relationships with a steady flow, a constant flow of context between both sides. Number three, dig into if your business partners actually understand what data ownership means, what it entails, they may be willing but not capable to own data at the start. Work with them to up their capabilities and understanding. Alex Bross, the previous interview episode talked a lot about that as well. Number four, data ownership should not be treated like a hot potato being passed to anyone other than me, right?

Number five, tell your business counterparts, I need your help to help you. Can you unlock... You can unlock far more value through collaboration than waiting for requests. Number six, your data strategy should give people FOMO, that fear of missing out. Give them incentives that make it feel like if they can't miss out on the value you're creating. Number seven, an interesting idea. Use gamification to find data quality issues. Don't make it a shameful thing. Find the issues, get them remediated and celebrate the people who found them. Look to drive positive energy around your data. Number eight, talk to potential data consumers before creating a data quantum or anything similar in data. You won't know what you can offer that will be valuable to them until you know what they want. So have the conversation. Number nine, bad requirements and or bad requests lead to bad results and bad data. Ask why someone wants work done. What are you trying to accomplish? Let's focus on the outcome. Number 10, it's easy to lose sight that optimizing to turnaround time doesn't typically optimize time to actual value and you create hard to support data assets, right? So this is about rushing to be too quick and so you often don't end up getting the value that you really wanted out of it.

Number 11, dig into what data consumers are trying to achieve and why their idea of what they want may not be the most complete or best way to structure the data or



Transcript provided as a free resource by



insights. Again, have the conversation. Number 12, when looking at data quality, tie your metrics and measurement to value what aspects of data quality matter to that use case. Number 13, data trust is obviously crucial. Data quality issues will happen when they do remediate and put rules in place to prevent the same issue from happening again. And show business partners that you fixed the cause. It was a data downtime incident. Treat it like you would a software incident. Number 14, to achieve actual sustainable change, you don't have to make big sudden shifts changes. You can and should break change into much smaller pieces. Finally, number 15, centralized governance might not be your data bottleneck or even if it is, if you aren't mature enough to do federation and decentralization right, Data Mesh might put you in a worse spot. Really consider your challenges and maturity level before jumping into something like Data Mesh. With that bottom line up front done, let's jump into the interview.

Okay. Very, very excited for today's episode. I've got Dacil Hernandez here who is the Director of Data and AI for Northwest Europe for the consulting company, Nagarro. And we're going to be talking about a lot of different topics here, but like a lot of what Dacil has seen, of where these types of things, whether that's Data Mesh specifically or similar types of implementation tries going wrong. So we're gonna talk about how important it is to involve the business and how we can't just have IT run with things. A big issue that IT has had historically as the point by point data pipeline type thinking of do this, then this, then this. Instead of like, what is the overall thing that we're looking to do? How do we actually go in and teach the domains? There's so much that we could be doing with your data. Help us understand what you could be doing.

And then talking about how important it is to understand and have the conversation around what data quality actually means. That there isn't a overarching thing of, this is what is quality data and this is not quality data. And then we'll talk about how Data Mesh implementations actually could happen and how we have to be focused on keeping the change moving forward, but also not having chaotic change. 'Cause it's very, very easy to slip into that. So with that, is the backdrop to what we're gonna discuss. Dacil, if you could give people a bit of an introduction to yourself and then we can jump into that conversation at hand.

0:08:48 Dacil Hernandez

Yeah. Thank you Scott. Like you mentioned, I'm really happy to be here. Who I am. I'm actually... That's a good question to ask. I started working with data for... It wasn't my plan. It was not that you say when you're a kid, "Oh. I want be a data manager." There's not nothing that you say as a dream. But I love the purpose of collecting data to create insights. For me, it's always have to be a purpose of why we do what we do. I've been working with data analytics for a long time. Longer than I want to say aloud.



Transcript provided as a free resource by



And what I've been seeing, I'm a little between tech and business. I see myself like... I do love technology. I'm heavy in that, but my heart is in the business and the value of the data that is great. So I'm a little bit of... A purple people we're normally called. I'm a blue and I'm a red and something in between.

0:09:54 Scott Hirlleman

Yeah. And I think it was surprising to me when I really started to dig into a lot of the data space. How many people that are in senior positions that are so focused on the technology. And I think we need more people to... It is tangible. It's, it's a little... It's not easy, but it's easier to not have the additional context of the business. So we need to make sure that we don't fall down that same route. 'Cause it's tempting to go down that 'cause it's a little less ambiguous. But I think that the business is really, really crucial.

0:10:36 Dacil Hernandez

Yeah. And I think it's a little bit of what you were discussing in... Tech is easy. There is things you can comprehend. You take a data and you move it around and then you manipulate and then that's in another place. Creating that value is where there's more uncertainty. You don't know... Something that you thought you had collected for one purpose out of the blue is something else. And I think this is a little bit of why sometimes business and IT are so disconnected because we are in... The piece is a little bit different, I believe. And this is something that will bring us to the topic that we will discuss about how really, really you need to have the business or the operations or whatever part of the organization in your journey when you start a Data Mesh journey. The whole point of the Data Mesh from the beginning was that you have that kind of domain. Big companies will have decentralization no matter what. So either you do it with control or you let it go like it normally does.

So, it's a little bit of the... It's super important to have that and the understanding of business driving that discussion. It would be much easier to go faster if it is happening instead of... And everybody... I've been in a lot of events and I've been talking to a lot of my peers, a lot of clients and everybody say the same. "Business first." Focus on business, focus on people. And actually, I had a reflection when I was listening to all these experts saying business first I was like, "Our data's kind of a teenager." You hear your parents saying something, but you still do whatever. Something totally different because everybody's saying ensure that business is first, people first. But we still focus on tech. So I don't know either we are in the teenager phase where you don't listen to whatever is being said to you or we have some... We need to change the narrative of what we say... When we say business or people first, what do we mean?

And I think this, finding your stakeholders in this journey, because if you start from



Transcript provided as a free resource by



the tech, if tech start driving this, find the area where you feel this is a winning maybe a specific domain like in marketing or finance, and drive it together with them, and I think it would be much easier. And then you will have this kind of success story that everybody wants.

0:13:23 Scott Hirleman

Well, how would you say, so it can be different in all sorts of different organizations, but what are the learning points for people to actually go and collaborate with those stakeholders? Because you know what I've been hearing from a lot of people is, okay, data is IT's purview, it's IT's domain. So anything data related, you just take care of it. Even the business folks are saying, "I will tell you what I want, you just go do it." Instead of like, "We need to collaborate together." So how, it's almost like it is that the data the data side is being like a teenager, but the business side isn't being a parent. The business side is like, you just go do your thing. Like you, you do what I want you to do and I'm not actually gonna work with you on this. So like.

0:14:21 Dacil Hernandez

What I think is about, we have, this is an interesting topic because one of the things I've seen and I've actually seen by myself because I've been part of large corporates where we have the same challenge is that data have been always a part of like IT, it belongs to IT. And in IT in the past, if you need a computer, you need a software, you send a service desk request and it's fixed for you, IT hasn't, it's not something that you do yourself. And I think for business data's the same. If I want something or I need something from data, if it's IT. And then I just, I don't need to get involved in that matter. So I think in my head it's a little bit of... In a lot of organizations you don't know where to put the data organizations team. That's a little about the topic, but it's interesting because either in my whole career I've been part, either reporting to the finance people, either a finance manager or an operation manager. Or at some point I think we became IT but data itself doesn't have yet a place on the, by itself. So I think as long as business things... IT runs data or is like an IT thing, this will be the challenge because they don't understand that they own the data.

And when you say to business... The first thing when you do Data Mesh and you start looking up your domains and you try to find the data owners and people say, yeah, but you know, I can own the data, but they don't understand what it means. And I think I can speak with a lot of my peers that finding the data ownership within business is difficult because I think there have been a disconnect. Like, it's not, and it's interesting where, I remember the first time we had, I was working in a company where my data people didn't really understand where the numbers were coming from. So I actually forced my data engineers to be where the data started and the call center, listen to incoming calls. And they were shocked because they never would be that close of the source of the data.



Transcript provided as a free resource by



It wasn't as serious as once it was things behind the numbers. And I think for them was like an eye opening that, aha, this is what we're collecting or storing or analyzing. So I think, it's both and how if tech, if you're driving this as a CDO or you know, your Data Mesh journey is starting of any kind of data journey from tech that almost, always is like that you need to yourself and find the domain in your company that is more mature. Normally it's finance because they like numbers and they like quality and they have it. They love to collect and do massive report in aggregated data so find your early adopter, whatever you call it, and start from start small. Everybody said the same. So it's nothing new that I'm saying. And yeah.

0:17:36 Scott Hirleman

One question I would have there is this I think is where I'm hearing of people having issues in going and having these conversations, is that you are going to them and saying, you are the owner and you're going to them saying, I'm giving you responsibilities, and so it feels like I'm no longer responsible for this. You're now responsible instead of going to them with, "Here's the value of this and here's the value to your domain and here's the value to the greater organization." But most people at most come in with, "Here's the value to the greater organization." And some people have said, "Well, this gives you more freedom and this gives you," But a lot of times they're like, "Well, I don't have the people to handle this, or I don't want to handle this."

0:18:21 Dacil Hernandez

Exactly. Want, Scott, that's the issue. They don't want to do it. And it's like how somebody mentioned yesterday was talking to, it's like a hot potato. Everybody is running around because nobody wants to catch it. And I think, one of the things I always say is, let, I need your help to help you. That's something I always say, if you don't help me, I cannot help you back. So I remember the first discussion about you had a data set that have to be, you have different teams trying to get, create some kind data product or a place where that data should be put. And we were like if you are not there and you don't try to help us, it's impossible for me to help you back and get your data, the report you want. So I think, like you say it takes time and it's a little bit of, like you say, what is the story that you're telling to the business to get involved? I'm totally with you. You cannot go say, "This is not your possibility. Bye bye. See you only once a month when we do the scrum..." What do you call it? The demo session and what we were doing.

0:19:32 Scott Hirleman

But, it sounds like from both sides we're saying that initially the conversation is... Both sides are saying, "IT, go run with it. We don't wanna collaborate with you." So IT has to get out of their own way and say, "Okay, instead of doing something and delivering



Transcript provided as a free resource by



something, we have to focus on delivering what is actually necessary in the mid to long run. And so we have to kind of break through these barriers," but, what is the point that you're finding? What is the conversation point that you're finding, you said, "Help me help you," but, is there anything specific that somebody can go in and say, "Hey, we get that we're asking you to own the data, but the first consumer of this data is typically the people who own it themselves." So let's talk about what you are not doing with your own data that you could be, or how do you get them over that hump about... Like you said, they're finding the data owners is far different than finding data ownership, right? You find the people and they'll go, "Yeah, yeah, I'll own it." And it's like, "Okay, but do you actually know what this means?" How can the IT team go in and actually help?

0:20:52 Dacil Hernandez

And the ownership is also something in person. I remember a case where we have, we are collecting data in the company I was working on, and the data was collected and stored for one purpose, and it was meant to be able to reach somebody. So the data was used to be able to call somebody. So it was like for contact purposes. My data scientist team figured out that some of the parameters were important for the defining a center pattern. So if you have this value, you will response in that way. And that wasn't a data perfect force. So when we supposed to do this, the idea was identify there was a mobile number or a landline. And when we went to the database, they didn't have any mobile numbers whatsoever. Zero. And we were like, "What? You don't have mobiles in this country?" Was like, we couldn't understand what happened that this both data sources were put in the same column because the people looking at this were saying in an interface where you could see like a box and you couldn't have the numbers and somebody died.

So it doesn't, the need of that data was high quality and collected in a way for the purpose of calling, but not for the purpose of doing advanced analytics. So at that point, when we, as a new domain that we want to use that data, we said, "But data is not good enough for us. So then you need to restart." And oftenly it's like we say, maybe for the call center is fantastic, but for us it wasn't good enough. So then you need to sit down again and say, "Okay, back up. We need to redo this and try to find, create a new data product on that." And that's what is important to have a platform that allows you to do this fast moving. So it's not a big process to move because then we will start touch up this kind of quality and when the data used from somebody else.

So just because you think the data... You create your beautiful data product or you expose your data and you think it's perfect, but then it's not for somebody else. And I think the important thing is that, you have that speed to be able to change how that data is going to be used. So how is the classification of quality and what's kind of...



Transcript provided as a free resource by



And I think it does back to, how do you do that is to identify that kind of use case when you feel that data have a purpose that is... Maybe we have lucky and it's collected in the way that it's supposed to be used for, but if it's not, you need to be able to do this fast and act on that change, for example.

0:23:38 Scott Hirleman

So when you're going into a domain, you're talking about, let me help you, right? Help me help you as well. How do you actually work with them to see what unlocking this domain or this data could unlock for the organization, but especially for themselves? Because what I'm finding the pushback is, and I don't know if it's really the actual pushback or it's just people are meeting some resistance and they're interpreting this pushback differently, but it's, they go to a domain and they say, "You have this data. We want this data for the organization." And the domain goes, "What's in it for me?"

0:24:20 Dacil Hernandez

Of course.

0:24:20 Scott Hirleman

And so, how do we go into there and partner with them to see what is in it for them? Or do we have to go and we have to realign their incentives? Or how do you get over that hump and how do you get them to see what could be possible if they're actually playing in this, right? They're looking at it as a prisoner's dilemma and they say, "Well, if I roll on these other people and don't... By not providing my data and they provide their data to me, then I get all the benefit. And I don't have to have any of the expense, any of the cost," but, how can this participation net them more? Or do we again, have to go in and either kind of have somebody above them beat them down and say, "Go do this," or we have to realign their incentives? How are you finding that working?

0:25:19 Dacil Hernandez

And I think that... Now we'll sound like... But a good data strategy is the key to... And I know that sounds like, yeah, yeah, yeah. But you'll have to have a strategy that focus on data and data driven... What it means to be... If that company wants to become data driven and you have to... You need to have a data strategy and you need to make sure that kind of everybody understand that this is going to happen. And a little bit of, what is the rule engagement? Why we are expecting to do our data, and collaborate?

If there's no good strategy, it will be more difficult, because then, domains needs to, like you said, "I need your data, please give me your data." And what we want is people saying, "Hi, here's my data, check it out, we have created this amazing exports



Transcript provided as a free resource by



of our data. Please come and see if we can collaborate." That's what you want. This proactive way, not the reactive, "I will only act if somebody ask me or force me to do it." So for me, a good strategy from top level and a good understanding of what do we mean to be data driven. Because otherwise you will still fight, and that will be your reality, and you will not able to change it. And so you will... So some companies are amazing, they're super data driven, and they don't have any issues at all. Other companies are not that mature. And for me, I sound like everything will solve with strategy, at least I believe that it have to be in place.

0:27:06 Scott Hirleman

Well, but it is... It depends on the organization, but exactly what you're saying. I think this isn't something a lot of people have said, which is that your data strategy has to include not just the vision, but it needs to include the incentives, right? 'Cause just getting people bought in on a vision isn't going to actually... There are people who are like, "Oh, wow, this would be so awesome for the organization." But people are like, "I need to focus on... If I'm getting ahead based on A, B, and C and our data strategy is only aligned to X, Y, and Z, I'm going to focus on A, B, and C." Most people are, right? There is this thing of, "Okay, well, shouldn't everybody play together and shouldn't everybody collaborate?" And it's like, "That's not the way human nature honestly works."

0:27:57 Dacil Hernandez

No, exactly.

0:28:00 Scott Hirleman

How do we align...

0:28:01 Dacil Hernandez

I heard, I've actually talked to a peer of mine and she mentioned that in her company they have done some kind of gamification to find quality issues. So you get points every time you find a quality issue in the data. Actually it sounds super, I was like, "You need to tell me more about that one." But for her was to involve business too. And for them, they work, this kind of gamification to have... Or you can do, I remember we had like a competition between two departments to do the data cataloging. We're supposed to do a lot of mapping of a mass of data and it was difficult to... We did a game about that. And so we... And everybody wants to win. So it became... Everybody knows about you doing this because we need to do it. But if you make it a little bit fun, for some company it works, for other, maybe not.

But that is something that you need to think about it, how to make this kind of exercise fun, and this little bit of "What's in it for me". You get your... It's like when you run and you get in and start a new app, they say, "Oh, you are faster than the last



Transcript provided as a free resource by



time." You know that the system telling you that, but you are happy to see that start on the literary now, "Congratulations, you are faster." Everybody likes that feedback. So why not create it with data collaboration or data cataloging or whatever issues with data. Our brains are triggered that way. We need to have that confirmation.

0:29:46 Scott Hirleman

Gamification and being seen and things like this, at NAV they talked about, they literally anytime somebody brings a new data product onto their Mesh, that team gets sent a cake. Right? A literal cake.

0:30:04 Dacil Hernandez

Yeah, that's nice.

0:30:06 Scott Hirleman

And so then people take a picture with the cake, and then they'll put in the little blurb from the data catalog as to what the data product is about. And then they can have a, okay, we're gonna have our cake celebration. That's also a show and tell, not of just the cake, but of the actual data product. And you can...

0:30:26 Dacil Hernandez

Yeah, yeah, yeah, yeah. That's fun.

0:30:28 Scott Hirleman

Just little things like that. But the incentivization... So, most business leaders are probably not gonna go, "Yay, I get a cake," or, "Yay, my team gets a cake."

0:30:47 Dacil Hernandez

No. But you create this kind of positive energy, where it becomes that... Because at the end of the day, everybody wants to use data to extend kind of value. To create, either you get better business, or you can create new products or new things that you can actually, depending what industry you are. Or you get faster implementation or you get better understanding of how your system works. So, in the end of the day, everything is here because you need to have an input to the company value, or whatever they're doing. So of course this is a little bit of a... Just to make it... Maybe in a period of time it makes sense and then later on when becomes not natural you don't need that things anymore.

But, again, if you need to have... And another thing that I think is interesting is that a lot of... You don't know what you can offer because you don't know what the other person wants. So have that kind of discussion, creating that kind of forum where you can say, "Okay, we have brainstorm or innovation room data, so we have this data set that we are thinking to create, expose it because we believe it makes sense." And



Transcript provided as a free resource by



then you can maybe for your organization, this could be really useful for marketing purpose, for example. And you talk to them and you have that conversation, "What do you think, guys?" And come to that proactive... But for this you need to have these kinda of people that are between these kind of... I don't know if you call it business analyst but this kind of people that can connect these dots on the data site and have because if you are in one domain, you will never see outside that, you will never see that your data can be used for something else because it's natural. It's not all over you work up and say, "Oh, I can use this data for anything else." It's, I don't know if it's nothing to do with Data Mesh but something with the use of the data.

I bought an oven that there was a WiFi communicator so I could get signals from my oven and then my friends are like, "Why you buy this? This is stupid." Like a linearity thing, I like technology, let me buy my oven that would talk to me when I'm not home. And what happened is that one day I got a signal from the oven that the pizza function was on in the middle of the day. My son's supposed to be in school, so I called my son and I said, "What are you doing home having pizza?" And he's like, "How do you know that?" "The oven told me!" So outta the blue, my oven became a spy. And I don't think that was the purpose of that data but now you could actually, maybe they start doing a new product, see what your kids are eating when you're not home. So again, there is a little bit that data is, that's the beautiful about data and things that we collect. It can be not always is super fancy and you can do really amazing things but just make it right and accessible and have this kind of that you can... I think trust is that one of the things that we discuss often, when you work with data, you need to trust the data you have.

And that's a thing it's much better very important than quality. We were supposed to be discussing that quality is subjective, is dependent who is going to use that data for. For some people, they have to be zero errors because it's life and death data. But for some other people that data can be like, "Yeah, I'm 80% quality. I'm happy with that. I don't need more." But whatever is there, you need to trust that the data you're getting is what you're supposed to be getting, good about quality. But it's a matter of ensuring from the people that you are collecting this data, making sure that it's compliant, and everything, that you can trust it. I think that's the main focus that these kind of data products have to ensure trust more than quality measurement.

0:35:01 Scott Hirleman

Yeah. So what I was seeing in a lot of what you were saying is that when you're going into these conversations, you need to find mutually beneficial points. You need to go in and actually say, okay, again, it's the help me help you. But a lot of this seems to be that we don't have high context exchanges around what data do we have? What data could it be used, the art of the possible, all of these and that IT is trained to listen for the requirements instead of the pain. And if we actually go and listen to the pain



Transcript provided as a free resource by



and say what challenges are you actually having? Not immediately what should I, okay, boom, I'm gonna write down, file your Jira ticket from those pain points versus, what are you actually trying to accomplish?

How do we get people to one, share with the IT teams to do that and how do we get the IT teams, how do we get data people to kind of relax a little bit and say, "Okay, do you just wanna vent or do you want that solved?" And this raps into that quality thing of, oh you're really frustrated by the quality but what does that actually mean? How do we get past the ways of working that we've always done? The patterns and everything that we've done has been this call and response this request and do work.

0:36:42 Dacil Hernandez

Yeah. But they have been, this is something that they have been forever like supply and demand, you ask and I supply. And one of the things that we know in the past regarding bad requirements mean bad delivery, that's one that we know. But I think it's a matter that, don't know, I ask why a lot. Every time somebody asking me for some kind of piece of data or something they want to, I will try always try to ask and then try to understand why and I ask why a lot of times and then I try to explain back to the person this, you mean so I try to repeat what they were asking me for if I understood correct.

And I remember one time that we had, somebody asked me, we supposed to do some actions when something happened trigger an action and that action made that we have to do something. The business owner came to me and said, We need this as soon as possible. As soon as this click, it have to be like real time. And then I said, "Okay, fine." And I was looking and I said, But wait a minute, what happened when you click on that? Yeah, yeah, we create and he said, Now we create this was to be able to send a letter. Okay. And I'm like but how often do we send letters? Once a day? So what you mean is that you need to have that data stored before the end of the day so we can get the letter to sent because it's not real time. So we are agreeing this... And for me it was like, wait a minute, why do you need it in real time?

Because I knew that... I don't get... If it was an email, by all means. But that was a physical letter that have to be sent. So I say, Okay, so what do you mean is you don't want to miss the deadline of the day or the whatever this file have to be sent to the post office. So the thing is to ask a lot of whys and a little bit of curiosity on... Be brave to ask if it doesn't make sense for you, why do you want this data, why? Ask, don't be afraid of, sounds like you're stupid and you don't understand what you're saying. Because I often ask why so many times. It's like I learn from my kids when you say do something like why? They don't get... They're not happy with the first answer you give them. So it's like why and why.



Transcript provided as a free resource by



So ask why and also... Sounds... But more interactions you will request. It's something like this you like and you show it. No, no, but it's not that. And sometimes it's like in the opposite. I said I just want to have a new line in a dashboard. That line we started, we didn't have the data in the system, so for him he was just, "yeah, but you can do a new line." And we are like "okay," so I have to actually show that person in the system what it could mean that line, because he didn't understand. For him it was just a line. But for us first we need to have the data because we didn't have it and the data was in a data warehouse, so we have to back the source. So we do this kind of, what do you call it, the path where all these that we have to do to able to get that line because business didn't know.

So this is a little bit of both trying to explain to somebody that's understand and what it means. And another thing that also we often talk about is when you had business people asking for data, they see, say that you have a system and the system has names. So it's like name, address, that's what they see. But in your data warehouse name is maybe 340_DSLT, something else. And when business ask for a specific for data set, get lost in all the transformations you do for your data. So at some point it's like, what is that that we're supposed... And happy... In real good companies you have your data lineage and a lot of companies don't have that.

So it's like talking the domain language. So what is that people want and ask them to show you send a print screen, do, you know... This is the data I'm aiming for that is coming from here and be part of more acting in the conversation to have. And it take longer time because you have to spend time with the... But the thing is, and more you talk the more, you learn and you learn how the understanding is mature. And in my career, never seen anybody in my opinion, everybody's there to help. Like you say, everybody wants to help. Is not people like, "Oh I don't want to help you, I don't like you." So I think everybody's super professional and everybody, but it takes time and if you open that conversation then it gets easier in time. That's what I believe. But I maybe naive, but I think people are good in nature.

0:42:09 Scott Hirlleman

I think what you said is it takes a longer time to get to from step A to step B but it probably takes less time to get from step A to step Z because you're iterating and you're getting that feedback and that you also aren't delivering things that people don't want and then having to go back and redo the work and things like that, you know, you have to balance. Is this perfect? Is this perfect? Is this perfect? Along the lines versus are we heading where you wanna go? Alla Hale on her episode mentioned. What would having this unlock for you? That question instead of why really trying to push back and say, why? Why do you want this versus what is this actually gonna do instead of the why some people bristle at, which I think is silly to me because if you don't know what you want from this, then why are you asking for



Transcript provided as a free resource by



it?

What would having this unlock for you then opens up a bigger door for a conversation so that you can get some more context and then you say, Oh you're trying to accomplish this. Well, I actually know about this other data source that's probably better aligned to what you are doing and you're able to add more value to them more quickly and kind of iterate towards that, what they're actually trying to achieve instead of what they told you they want.

0:43:39 Dacil Hernandez

And the thing is a maturity level, and before I became... For a short term in my career, I was a product manager for, with people, .NET developers. And they were doing our website and I remember one we had a bunch of data that was supposed to be displayed on the website, I thought was too messy. And I asked the developer, Can you please reduce the number of words you can see and you can put an expandable button so if you click and then you get more data. That's what I said. And this was a stressful, we did it like the day before we went live. So we tested on the fly and I opened that data set and then I asked her how do I close it? And he said, You didn't ask for that because I assumed that he figured out that I want to go back to where I was from the. So instead of saying to him that I need to have less data in the landing page and they'd be able to see more, which was by me because I was too specific what I asked for. And in the beginning I was upset because don't you get it, I need to close it again. So that was my first reaction. But then when I start thinking about what did I ask for, I got what I asked for. That's it. No more than this. So whose fault was? I guess both but. So now every time I ask for something, I think really good about what is the value that, what is it that I want to achieve and I'll give details on how to going to do it. And this apply for data as well.

0:45:28 Scott Hirleman

I struggle with that because I'll tell somebody, here's what I'm trying to achieve here might be a way to do it. And then they lock onto that, oh, he wants it done this way. And I'll even say I do not care under any circumstances as to how this is achieved. And they still lock on too. I'm just trying to make it easy. I'm trying to make it easy as to say here might be, here, might be something that works and it just doesn't.

So, we were kind of heading towards talking about data quality, and I think this ties in well to what you're even talking about of like, what are we trying to achieve in having that quality conversation with people as to what actually matters and that there's so many different aspects to quality that, is it the completeness? Is it the timeliness? Is it the correctness? Is it, that this thing is... I talk about within Kafka, in the Kafka book by Gwen Shapira. There's an entire chapter on what exactly once means and how to do it. When you think about how much time is spent explaining



Transcript provided as a free resource by



what exactly once means, when you know that... It means only once, you know at least once, and no more than once, only once. It's, every single thing is only done one time and it's all encompassing. But, it's much more complicated. So like talking about this...

0:47:00 Dacil Hernandez

Yeah. Super complicated.

0:47:02 Scott Hirleman

How do you extract the actual needs and then share with them as well the incremental costs? Because again, we've taken in requirements when we think about data instead of, "Okay, you want this to be... You want your completeness to be five 9s? Okay, that's going to mean an extra six weeks of development time and it's going to cost us an extra \$500,000 a year, versus if we do it at three 9s, it's gonna take two more weeks and, it's gonna cost us an extra \$100,000, versus two 9s." Is that really... two 9s, we can get it to you in two weeks. Like, what do you want?

0:47:48 Dacil Hernandez

Yeah. And, that's something... I remember the... I was working in a company where they had this beautiful dashboard about data quality, and they had green numbers and red numbers. It was really nice, but I realized that what they were saying, if how good... If the cell was fit in, they didn't know if actually it was real data in that field. It was like back to the telephone number. They knew that it was something there, but if there were good numbers, or bad numbers, that was a measure. "Oh, we have 500 records." "Okay. Good for you." Doesn't say anything to me. Can I use this for 500 records? And then we started discussion, okay, how do we measure... Because that's a little bit of another kind of quality. It's like, we're exporting 500 rows and we get 400 rows in. So there's no issues with that, no. That's one piece that you can measure.

And I think that I always introduce, we call it like how valuable the data is. So if it's high quality, but the value is fully low, maybe we don't care that much in making sure that everything... depend, like you mentioned, how much it cost to bring that 100% right. So you need to have your value measurement on, how valuable is the data. 'Cause if you have something that's low quality, but it's high value, then you need to work really hard to make that quality better. Otherwise if it's like quality low and the value is low, then you can just have it there because you don't need to have it more quality than that. So I think that is important, the balance between... And the quality depends on what they're going to use it for. So if I need that data to be able to make a call, I really need good numbers. So a real number, it have to be a telephone number, not something else, for example, to make it superconcrete.

And I think that that hit me a lot when we talk about quality is when you say that



Transcript provided as a free resource by



your data scientist team is creating an algorithm to predict something. And then, it have to be like... You need to have like... What do you call it? Precision that is super, super high. That means that your quality... The data have to be precise to be able to run this. So some things like predicting, like if you have cancer in the picture is, you cannot be wrong. So you need... Your error margin have to be superlow. So then, then you say you get a lot of energy and then you make more parameters to quantify the images. Everything that they have to be around is only one measurement of quality. But for that topic it's like, you don't need to go... Everybody thinks that, "Oh, I need 100% quality," because this is how we are. But really? And for a lot of, in the... When you do a lot of the modern, you do, you don't need to have 100% quality in a lot of things you do. If it's good enough with all 80 or 90, that means that you have good data, then you have some data that is noise. Yeah. That you don't noise, what do you call it? Noise. It's like noisy data. They're like, "Yeah."

0:51:07 Scott Hirleman

Yeah. How do we actually communicate that though? Because people are... This is something Zhamak and I discussed recently about how frustrated we both are by binary thinking and like it's... Is this a one or is this a zero? Is this high quality data or is it not high quality data? Versus what do we need? What is the use? How do we communicate that to people where we can say, "Hey, here is how far..." Like, "Here's what the metrics are showing us," right? As to the different quality metrics, right? Okay. So this thing is within this amount of time, we're populating it within three minutes of this thing happening, but our accuracy is not all that high precision or whatever. Like how do we get it to a place where... Because so much of our time... When people talk about how much time is spent cleaning the data, it's mostly about trying to figure out how much can I actually... It's not even like the prep work as much as how much can I actually trust this? Like, you have to go and you have to do all this work to verify and reverify and all this stuff, and it's just not... It's not value add work. So how do we make it so that people can inherently understand how much they can trust?

0:52:32 Dacil Hernandez

That's a good question. And I think, let me go back into my thinking, because often when somebody realizes the data is bad or is something wrong with the data, normally it's business, business either... Let us super simplify because it's... Later you see a dashboard and you get a number, the revenue for last month, and the numbers are crazy. So business is almost the first one to react something wrong. And then they go back to maybe the BI team and say, "There's something wrong with the data." And then the team says, "But we don't see any mistake here." And then you go back to the data team and say, "There's something wrong with the data." And then the data team figure out that maybe was some kind of... During the process that when the information was something, that get a mistake. And I think the issue is



Transcript provided as a free resource by



that often, in some organizations, they have really good... Make it where every time the data pass any transformation, you check that they're actually seem good.

But when business is the one all the time discovering bad quality, you don't trust them, anything that you see. So... what we have have done in the past is that when you find a mistake like that, for example, you can actually automate it with some rules so that that mistake doesn't happen again. New mistakes will happen, but that one should never happen again. So if you know that one, you get a data set where you have commas instead of dots from a client, then you know that, "Okay, I need to ensure that dots are comma or whatever issue there was." And so focusing when a mistake on your final quality issue, make sure that you have rules and processes where that will not happen again.

So you reduce and as... Trust takes time and you know that. If some... If you're used not to trust your data, you need to prove that you know. It takes... Trust is hard to build. When you had it and you lose it, then it... Ensure that... Let me see... If there's some data quality issue, if you have knowledge about that, make sure that didn't happen again. And put controls in the way, so business... It's better that you and data team see it before anybody else sees it. I think is... So the best thing will be that all data in the business is the quality that I expected to have. But it's still... A lot of the issues with data and the qualities, every time we transform the data from the raw data that has been indexed to the end users, see it. You can have... And sometimes, even three or four different transformation before the data comes out. And that's... Or you have different applications, have the same data treat in different ways. So you have one number, one application, another number, another application. Which one is right?

Good question. Decide, maybe both are right. Just pick out one. I always say, "Just, pick out one." But asking business to quantify what I mean with quality for them, I guess it's like when you see you need to trust, doesn't make sense that we have... In Sweden, we are around 10 million people and then we have somebody saying that 20 million of Swedish people are doing something, we don't have that much people here. So try to be... And in other things it's like...I've worked with finance a lot, so my main experience is with finance and financial data, so that's why all my examples have to be... I guess, in other industries, it's more and more difficult. And it's more important with Mesh... More industry heavy that the machines have to be running a certain speed or it's important the air conditioner works in a certain way. And for some places it's crucial. But back to your question, how! If I knew I will be rich.

0:57:00 Scott Hirleman

If you have observability is trying to do it, data observability is trying to show trust, but it's still exactly what you talked about of, you can't necessarily show is this correct



Transcript provided as a free resource by



versus is this filled in, do we have a result? And we can look for external validation and things, but yeah... So we've covered a lot here. I did wanna wrap up a little bit around what we were gonna talk about of, how do we look at positive change instead of merely chaotic change, because Data Mesh is gonna create a lot of change. So how do we really look to... Just what would your tips be when talking to people about keeping the chaos out and keeping it so that the change is headed in the right direction instead of just a direction?

0:57:54 Dacil Hernandez

In my opinion, important thing is to do is like... Companies are thinking about what we have today doesn't work. I don't know what else is out there. The only thing I know that, I have low usage or I have people that are now exposing the data, we have domains. So if you are at a point where you figure out that what is happening at the moment, is nothing that I want to continue to have, and you start looking, what is out there? What I can do? And transition that change is like... You always said, you don't need to do big change, you have to ensure that you break your challenge in small pieces where you can say, "Okay, but my major issue is that we need to have..."

Ensure that when more sets have... We need to have the business user can extract the data when they want. That's a particular piece that you need to address and think about... Focusing that and don't go then with something else because, so have a good priority on what is the challenge that you want to solve because you cannot do everything at once. And in going in a Data Mesh approach, again, maybe this centralization is not for you guys. Maybe something else that is behind this mess. So be sure that you ask two experts, have somebody that can help you in this journey and say, Okay, is this something for us? So every time you want to do a change, you need to realize is this is for me or for my organization or whatever we are? Are we mature enough to do this jump or do we need to wait?

Or maybe a centralized approach with the good data governance can make happen... Can make a success. So in my opinion, this big change transition, you need to... There's a lot of peers... There's a lot of people out there that where you can... You have your Slack channel in the Data Mesh where you can actually ask questions. Does it make sense? So you're not... Nobody's alone in this journey. There are a lot of companies, and I think one of the things that when you come to a level of seniority is that it's okay to say, "I don't know what I'm going to do. I don't know what I'm doing. Help me," and it's okay to say it loud. Doesn't matter how long you work in the industry, I say okay, what I know that it doesn't work today, but I dunno how to solve it yet. So for me ask questions and talk to your peers or a vendor or whoever thinks they can help you totally.

1:00:54 Scott Hirleman



Transcript provided as a free resource by



Yeah, Zhamak and, she was talking about... Well, I can't just say that there's a problem, so I have to come up with the solution as well, which is the thing behind it. But I don't think we all have to have that every single time. I think we can have open and honest conversations. And I think that when you were saying measure and really figure out if this is your problem, I think a lot of people are saying, we have data challenges and therefore we should do Data Mesh. And I think that a lot of us are saying that's your... If somebody has brain tumors, yes, brain surgery is the right call, but if somebody has a sprained ankle, brain surgery is not the right call. So think about what... Figuring out the centralization is your actual challenge, if that's the thing that's causing you to not be able to do what you wanna do.

1:01:54 Dacil Hernandez

Exactly.

1:01:55 Scott Hirleman

Well, Dacil, this has been such a great conversation. Thank you so much for sharing your knowledge and your wisdom and your insights here. So is there anything we didn't cover that you think we should or any way that you'd wanna wrap up the episode? Any point you wanna hit on?

1:02:12 Dacil Hernandez

Yeah, and back to if you are thinking about Data Mesh and you listen to... And you don't get wiser because when you're in more and more, you get more and more confused. If you're interested in more Data Mesh either reach through... I'm recommending this Slack channel that's the Mesh channel that Scott has or reach out to me or, and we can discuss like you said, are you ready for this? What is actually your... You need a brain surgery, you don't need somebody else to help you. So I think and it's because I love... Don't forget it's a lot of money involved in doing this transition. It's not cost free and you already invest a lot of your infrastructure and your data. So is this something that, in my opinion, don't rush it. So is...

1:03:06 Scott Hirleman

I think that's really important. I think a lot of people are saying, Oh, I need to keep up with everybody, the whole myth of if we're not a 100% data driven by two years from now, we're out of business. And it's like, if you're going at a pace that's not actually sustainable for your organization, then you're setting yourself up for failure anyway and you're throwing money at it instead of throwing the care and, the really careful consideration that it needs, you're trying to do the thing simply to do the thing instead of you're trying to do it to accomplish what you need. So and I'm sure there's gonna be a lot of people that would love to chat with you and follow up. Where's the best place? Is it LinkedIn or.



Transcript provided as a free resource by



1:03:54 Dacil Hernandez

LinkedIn? Yeah. Nice easy to find me. We're not many that have the same name that I have. Thanks, Scott, for this. But that's another story I can tell you.

1:04:06 Scott Hirleman

Yeah, I'm the only person as far as I'm aware with my actual name in the world. So...

1:04:11 Dacil Hernandez

That's it, it's that's easy to do, but it was really fun, Scott, to talk to you like always, and I hope that the listeners have a little bit of... We didn't get them a headache instead of solving things.

1:04:28 Scott Hirleman

I'm sure it was very useful. I think it was a very good conversation. I wanna thank you for your time today and I'd like to thank as well, everyone out there for listening.

1:04:40 Dacil Hernandez

Thank you. Bye.

1:04:42 Scott Hirleman

I'd again like to thank my guest today, Dacil Hernandez, who's the Director of Data and AI for Northwest Europe at Nagarro. You can find a link to her LinkedIn in these 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 fulltime to help out the Data Mesh Community. If you're looking for a scalable, extremely cost efficient, multi datacenter, 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.