Data Chaos to Clarity for Successful AI Adoption

Transcript | auto-generated and lightly edited

Craig Thielen, Chief Innovation & Product Officer

The first topic that we want to hit today is around data monetization. And the first question I'd like to always look at is why does data matter so much? And one way to look at that is if you look at the top 10. Companies in the US from a market cap perspective, one thing that jumps out right away is you could argue at least eight of these organizations, if not nine of them, are data centric organizations in which they have really mastered. Data and the topics that we're here to talk about and the results show for themselves. So that's one very strong indicator of how organizations are able to get significant value.

A term that I think we've all heard on this call is that data is the new oil and it's a it's a catchy phrase. But one thing that a lot of organizations have is they have an oil field. You know, full, full access to an oil field. But that doesn't necessarily mean that they're able to utilize it. That doesn't mean, you know, oil doesn't do much good if it's sitting, you know, in the ground a mile deep or 10 miles deep or under the ocean. It becomes usable once we extract it. Once we refine it, once we turn it into many, many, many, you know, thousands and thousands of products. And so that's really sort of maybe a better analogy when it comes to data. We believe every organization is on it on their own oil field of data. And what we're going to talk a lot about today is how do we extract that? How do we leverage it in ways that perhaps we've tried before, but we can do much better, much faster? And what are some of those techniques?

So with that, I think it's really important to talk more about defining data modernization. A lot of people, when we first start talking with them, the thing that jumps out is selling data. To clients or to partners or to others. And that's certainly one way to monetize data. But there are many, many, many ways. Some of these ways have been around for a while and some are relatively new. And then we'll talk about why things are different today, but everything from how you can optimize sales, marketing and growth. Again, that's been around operational, you know, optimization, better, faster processes, more automation if you think about technology every type of technology that we have today, the advanced technology, whether it be robotics, whether it be mobilization, virtualization, decentralized, you know, like Bitcoin and currencies or whether it be artificial intelligence.

Whatever the digital twin, whatever the technology is, one thing that it relies upon is a mastery of data. And so that's again what we're going to kind of go through in four different segments today and the speed at which we can gain efficiencies through data. We'll talk about that as well. Data-driven products and services. A lot of organizations that are traditional asset intensive organizations that make products, that build products, that engineer products are finding that they need to include data-driven products along with their to stay competitive. So that's certainly one, data subscriptions and platforms is a newer one that is an opportunity for a lot of organizations. And so you can see on the top half is really more internally focused data modernization opportunities and the bottom half is more external. Innovation is a critical one and again with AI we can just go at light speed, faster than we could even three or four or five years ago, again with access to data and we'll talk about what that means.

Decision intelligence again has been around for a long time, but a lot of organizations we see are still working with data that's 30 days old, 90 days old, they, you know, quarterly results, annual results, even monthly results and the flow of data is just much, much too slow versus some real-time decision intelligence that could be based on customer sentiment, could be based on social media, could be based on a lot of data that we now have available and we can process much faster. How we can even leverage out external data with benchmarks and intelligence industry insights. Again, we can do this much, much faster than we ever could previously and then rethinking ecosystems and partnerships. So that's I think a good breakdown of the different types of monetization and the idea is this isn't necessarily a comprehensive list, I think we can think of more, and the list can be somewhat endless, but these are good categories to say, what are we actually doing? How are we engaging with these? Is there anyone driving that? Is there anyone responsible for that? What tools are we providing them? What data do they have? Is a really good start to understanding what's possible with monetization. And then the next question is how do we prioritize, how do we implement some of that.

One of the things that we've done is we've applied some of this thinking to ourselves. And so about a year and a half ago, we looked at all of our internal processes and services and said by using AI in every single thing that we do, we found immediately a 30% efficiency gain and that we pass that right through to our customers. So the same engagement from a year ago or two years ago or three years ago is at a minimum 30% less even we found even some of them are 5070% less. And so that's nice and we need to do that to stay competitive and we are doing that.

But what's also important about that is how we can help our clients do that in every part of their business, every department, every function that they have that we've proven it on ourselves and we certainly have done it with many of our clients. So that's just one example.

So one of the things that's very different about where we are now versus you know two or four or six years ago is that the speed is significantly different and AI is really driving that and so if you if you have the assumption that you have data, you have access to it, and we'll talk more about what does that mean, data readiness. Then you can go through and you can execute on some of those data monetization opportunities much, much faster than you ever could before. So for example, we have implemented and found that that in the AI world we can go much faster than we could ever conceive of before. And so we have a three five day Sprint process where we can identify an opportunity which could be a process efficiency, it could be a product or service where we can identify a current state, define what a future state looks like. We can immediately build a prototype that we can get in the hands of users. They can test it, they can provide feedback, and then we can in the third Sprint already implement that and put that into an MVP use and then lay out what is a 30 day improvement and expansion scalability look like. So that's how fast some of these opportunities can go and we really try to apply that to as many of those opportunities as possible. So you get wins, you get value. It no longer takes months, quarters, years. Of course, this doesn't apply to every business case, but it does apply to many of them when you have data readiness when you have Data and AI assurance, when you have data as a product that foundation that we're talking about here today, then that allows you to go at the speed of this kind of an engagement.

So really to sort of summarize a data monetization and this is really a good lead into the other areas. The keys would be one is first of all to really view your data as an asset and we started the segment with talking about oil like an asset and.

To do that, we call that a data centric organization and there's really 3 pillars of that and we can maybe go through that at another date. But we've been helping organizations shift to a data centric, which is really data is part of your business strategy. It is your business strategy and it's really valued as an asset as much or more than any other asset you have. Of course, you have to have data readiness. When you have data readiness for AI, you really have data readiness for monetization. Building out data as a product is a really core foundation that can help enable some of the value realization and then finally establishing data and AI assurance. This is really important as we proliferate some of these technologies to really make sure we understand what's happening, why and there's governance in place. So with that, I believe I'm going to be handing things over to Matt Bryant to talk about data readiness.

Matt Bryant, Director of Data & Analytics

Hey, thanks, Craig. And as Craig talked about, right, there's no doubt that AI is very much a multiplier for the business, but that's assuming you can kind of trust the foundation data that it's built upon, right. And a lot of times what our AI opportunities are is they tend to be automating the manual processes that already existed along with that same underlying data. And the challenge is if the underlying data has quality issues, then from an AI perspective, it's going to lead to unpredictable results or hallucinations, right, or other negative effects, and one recent example actually is with Amazon. In December of last year, Amazon actually had a 13 hour outage over in China and it was all because of data and the quality of the data and not having the right safeguards kind of in place. So therefore, addressing that foundational data piece before you kind of get into the AI automation aspect is very crucial to your success.

So what are really the challenges with data and AI? And surprise, surprise, it's no different than what it is for people who are interacting with data, right? The data's not understandable, it's not accessible, or sometimes kind of the reverse. It's overly accessible from a data perspective and it's not necessarily accurate. A lot of times when we kind of get into client engagements, we often hear, hey, it takes X amount of months to get someone up to speed on our data. And a lot of times, just because the data's not documented, it's not understandable.

I was recently working on a client where we were talking about the acronym MSA and there were four different definitions of MSA. It turns out that the real definition was actually a fifth one from a third-party vendor that they were using outside, but no one had that context, right? Especially, we talk a lot about the technical aspects of the data, but we don't tend to talk about the business context and the relationships. And that's really what AI kind of needs. You may have a data catalog that's out there, keeping track of some of this information, but it's very much focused on kind of and compliance and less unnecessarily on your ability to kind of find and use the data. We talk about what kind of data is not accessible and available.

There are a lot of companies out there that kind of talk about data democratization, right? Saying, hey, we want to make our data available for the business users, but that's really not usually the case, right? Because there's access questions. Who can, who can be trusted to access that data? Literacy, right? Who can be trusted to analyze it correctly? Who can make decisions with that data? And unfortunately, that has negative impact from an AI perspective, right? The AI may not be training on the right data, incomplete data, stale data. It may be using outstanding facts because it couldn't get to the data. And sometimes that's because the business areas, right, are creating shadow data pipelines, which kind of bypasses some of this governance, but they need the data in order to kind of make their use cases work. The third challenge obviously relates to kind of data accuracy, right? And obviously it doesn't it reflect what you're expecting? It's typically that. Unfortunately, the data accuracy issues are typically the costliest to both kind of detect, identify and resolve, and they're usually someone's called the silent issues because they're the ones that it's your customers that are telling you about right from an AI perspective, what that means is false confidence, right? You're getting the wrong answers from AI or the erosion of trust. After a while, people aren't going to trust that the AI is giving you the right decision.

So, in today's world, what can you do to kind of make a I ready, right. And it's really we've got a couple of recommendations to kind of address the data challenges that we referred to earlier, right. The first recommendation is really to think about accessibility, right, and creating that unified view of your data and from our perspective we recommend the implementation of a kind of a business semantic layer that kind of acts as the guardrails for your AI, right? It allows you to create those consumable tables kind of with predefined and verified queries. So that when your AI needs to answer the questions, it's got reliable guardrails to keep it kind of on the path and answering the right questions with the right data, right?

So very much it's about, you know, reducing prompt ambiguity for Gen. AI systems in particular, but it's just making it more explainable for not just a I, but also for your business users.

The next recommendation is leverage metadata to drive as much automation as possible. One of the challenges today, right, is that there is so much data coming in that it's there's too much of it to really be able to do all the manual processes that we've typically done with data today. And so as you get to this more and more data, your manual processes just can't keep up with it. And so that's where you really have to kind of flip it and say, look, all of our data is gonna have metadata attached to it, right? I'm gonna classify it, I'm gonna tag it, right? What this does is it really starts to also change your governance processes from being manual processes where I go to get approval to get access to data to really much more of a people get access to it and now it's an exception only process that puts in place and that's one of the things that we found that's huge for being able to roll it out at a much broader, faster pace and getting the people access to the data kind of from a quicker perspective.

The third recommendation we have is around kind of ownership and trusted data, right? And we really recommend the creation of a data marketplace with a data with data contracts, honestly for both internal and external data sharing. And most people think a data contract is more for external data sharing, but you really need them internally as well, right? And that's really what allows your data consumers internally to have confidence that the data that they're seeing, right, it's certified, it needs a certain level of quality, it's got the right service level agreements for it, right?

From an AI perspective, you want that because you want your AI to be able to choose what data it needs to use, but know that it's going to be in compliance right with let's say AI regulations that are out there and guarantee that the data is going to be fit for purpose for what it wants. So that's a very different way of thinking about data.

So this kind of leads to the question of what is a data product, right? And there is no one formal definition and the definition I've got up there is just the Wikipedia one that's out there. But I like a different variation of it, which is I've seen out there, which is referred to as a data product delivers a high quality, ready to use set of data that people across an organization can easily access and apply to different business challenges, right? And so the reason I like that is I think it touches on what really makes an effective data product, right? It solves a very specific problem or a set of problems. It's reusable, it's actively used and being tracked that way. But most, most important kind of at the end of the day, it's really truly delivering.

Business value. I think the other thing with a data product is it's not just the data that's getting delivered with a data product, your data product is a solution really, right? And it's going to be include other things such as the documentation, the other or practices that you would have associated with products in other industries, right? So if you think through other industries, when they think about products, they think about usability, support, right, usage tracking and those are all the same kind of.

Aspects that you want to apply to a data product, right? And from an AI perspective, one way to think about it, to go back to that example earlier, AI is really just a consumer of these data products and you have to treat it AI consumer.

A little bit different than you may kind of treat some of your more broader users because a is can sometimes be if you give them too much permission, right, that's what kind of can happen right when they can go out and have a much and it can be a very negative impact on your environment, and so you really have to think about how do you define your AI user? What access rights does it need to that broader set of data?

So one of the questions that we sometimes get is I've got a data warehouse, why don't, why do I need data products? And I think it's interesting to remember what data warehouses were created for originally, right? They were created to answer the what happened question, right? That's a very different perspective in the AI world, because in the in the AI world, you're moving much more to a what now? OK. And when you answer a what now question, you need to know the business context, you may need to know how a metric is going to be consumed. Let's say how the logic is used across the different environments. In the data warehouse world, you had that human interpretation at the end of it to look at it and say well. Hey, what do I do now with that? In the AI world that that decision is really now getting occurring is occurring on a micro level and so it's no longer being taken by humans. Humans are still setting the strategy, but it's the execution that becomes automated, right? And so that tribal knowledge and that context that maybe you got to go to Geary down the corridor and say, hey, what does this mean and why would I use that? You no longer have that capability because now you have to define that much more concretely, right? So this really means moving from teams that provide numbers, right, to teams that build systems that turn those numbers in the next actions. And that's really your AI.

So what's the process for kind of creating usable and effective data products, right? And the general process I think everyone understands is the same, right? I have to get the raw data, I got to clean it up somehow and then I'm going to kind of combine and aggregate it for the different people that are going to need it. One of the things that you have to do in data differently than maybe how data has been handled in the past is we have to treat the building of data pipelines with the same rigor that software engineering teams have been doing using for complex software components. And so we have to start thinking about these data pipelines more as tasks that have an input, they're going to do some transformation and they're going to do some output and you have to kind of build everything around that aspect.

Everything has to be automated on an ongoing basis, right? Whether you're masking data, handling missing values, et cetera. That's the other benefit of starting to think about the pipeline differently is that you're testing really starts to get built into the data pipelines, right? Where you've got your test kind of predefined, oh, it's in production and it's constantly monitoring that the testing's working. So testing no longer is just kind of handled as part of a project delivery step. You know, we've worked with multiple kind of healthcare and financial service companies to design, structure and implement these steps, you know, with the proven implementation best practices that are needed for that. And with that, I'm going to turn it over to Lyndon, who's going to talk about AI assurance.

Lyndon Carlson, Director of AI

Hi everyone. Good afternoon. Perfect. Just to recap before I jump in, Craig started this conversation by talking about monetizing data, turning data into real business value. Then Matt just walked us through how organizations are preparing their data and really treating it as a product. So where we're going to go from here, now we're going to take the next step in that journey. And we're going to look at what happens when you take those data products and start using them to power AI. So in these scenarios, AI is going to be generating recommendations, predictions for automation and this really plays into what Matt was talking about earlier on the what now with the data rather than how we used to look at it. So in order for AI to be capable of these tasks of recommendations, predictions or automations, AI can't just analyze the data anymore. It needs to be able to make decisions with that data. And when AI becomes the decision engine, trust becomes the final critical component of truly monetizing and leveraging your data. However, that trust component isn't super trivial. Many people are coming to enterprise AI for the first time. Maybe they haven't even used generative AI in their personal use, and they're coming with a healthy level of skepticism. And truthfully, it's usually for good reason. There's been concerns around accuracy, privacy and data protection, and then the ethical use of AI.

And then again from that, moreover, from that enterprise perspective, that skepticism is healthy. So what we've seen in multiple studies is more than 80% of enterprise AI initiatives fail to deliver their intended business value.

And what's important to understand with these failures is they're not usually around how the model themselves are performing or working. The problems typically show up somewhere else, somewhere around trust, governance and operational control.

So what organizations are doing when they see problems like this or kind of what they're noticing is models producing answers that sound confident that are actually very wrong. Sensitive data is leaking through prompts and integrations and then those organizations will see automated decisions that nobody can clearly explain. And what this will lead to and why we have this triangle with trust at the top is if leaders start seeing those kinds of issues, trust is going to disappear very quickly. And then if leaders don't trust the AI, they're simply just not going to use it. And then if AI is not being used, that data is never actually getting realized and we never really monetized that data. We never made a data product out of it. So this is where my song and dance comes in.

This is where AI Assurance comes in. AI assurance is really about making sure that the AI systems that are built on enterprise data remain reliable, secure, explainable and properly governed. Because when organizations have that layer of trust at the top in place, they can confidently scale AI, and that's what ultimately allows them to unlock the real value of their data.

All right. So one thing we see consistently is that AI systems, they work beautifully in demos, but then the real issues start to come when you know things have gone to production and some things are kind of popping up as issues.

Some issues are models are beginning to drift because they have underlying data changes. Outputs are becoming less consistent. Teams are struggling to explain why certain decisions are being made, or they're discovering that sensitive data is leaking through prompts and integrations and more recently, one emerging challenge that we've seen is something called shadow AI. And this is where employees start using AI tools that haven't been improved or governed by their organization. And this is becoming pretty prevalent and it's going to have some consequences. So analysts predict that 40% of enterprises will experience some security or compliance incidents related to shadow AI within the next few years. So these really aren't just theoretical risks, they're operational realities.

Organization space once AI system start scaling and they happen because AI is often treated as a one time product rather than a project rather than a system that requires ongoing oversight before we go to the next slide. This image here kind of shows some of the risks and problems around AI projects that need to be addressed with AI assurance and governance. A lot of these look very similar, familiar and similar to like the typical software development, but there's a couple on here that are unique and I just wanted to call them out. So prompt injection, model drift, integrated accuracy. For those who don't know what prompt injection is, its a vector for attack where people can come in and say that you have an AI agent that is hooked up to your SQL database to do queries. You might have a nefarious person come in and try to get that agent to run instead of a query, a command against your database.

Drop your whole database and then you're out of production. Or they might try to get a response from your AI that would hurt your branding and then use that for nefarious purposes. Model drift. So depending on what provider you're using.

They might issue result or different versions of the same model, and if you're on some sort of automatic upgrade, you'll be put with that new model. But with what comes with that is you might regress on certain attributes that you were hoping would stay consistent with your model. So it could progress on tone it.

To regress on, you know, accuracy and these are all things that with a proper evaluation framework would help assure that that doesn't happen. And then that kind of plays into degraded accuracy and how this is different from typical software products.

So when you develop programmatic software, you expect the same deterministic output every single time. AI is a little different, so you have to be able to create some governance and strategies around technology that is not deterministic, it's probabilistic.

OK, so now that we've talked about where AI systems tend to break, the next logical question becomes what do organizations actually need to do to validate and control? So there's a few foundational areas that consistently matter most.

And this is also what we advise our clients to focus on. The 1st is governance. Organizations need very clear policies and ownerships around AI systems. This means knowing what models exist, who's responsible for them, and then how are they going to be used. This sounds super simple, but this is actually something that many companies don't have visibility over. Second is model reliability and responsible AI. Generative AI systems can produce answers that sound extremely confident but are actually incorrect. You may have heard this term before, but this is called hallucinations when the model just tells you something that's not true. Makes it up. Third is AI security. AI introduces a totally new vectors of attack. This is something I talked about on the previous slide with prompt injection, sensitive data leakage, and then risks associated with partnering with external models or APIs.

This 4th is autonomous systems. So as organizations begin deploying their AI agents that can take actions, it sometimes becomes critical to find clear operational boundaries on what those systems are allowed and not allowed to do.

And you see a lot of AI providers actually try to focus on this as their main value offering is deploying solutions that strictly adhere to whatever previous governance and protections you've had around your organizational data. And finally, there's operational monitoring where AI systems change over time. You know, models can.

Drift data can evolve and performance can degrade. So organizations need mechanisms like these to continually monitor AI behavior and intervene quickly if something starts to go wrong. Together with these five areas, this forms the.

The core validation layer that organizations need to have in place before AI can safely scale.

One of the biggest misconceptions about AI assurance is that it really only applies to the models themselves. But in reality, AI risk starts much earlier than that, starts with the data. And that's why AI is going last in this presentation is because everything that we've talked up about up to this point really plays into.

How successful an AI adoption can be? So we talked about the data value loop, assessing data, preparing it, creating value, delivering solutions, and then optimizing AI assurance cannot exist as an extension of this data value loop. Instead, it has to be.

Existing and across the entire life cycle of that data value loop. Because when something goes wrong with AI, it's usually not the model. That certainly is a possibility, but it's most likely coming from data hygiene problems upstream.

So to address this, this again starts with data governance, everything we've talked about so far, making sure that the data is managed properly, ownership is clear. Then there's data consistency and quality because of inconsistent or incomplete data leads to unreliable AI outputs and then a couple of other things that we address in AI assurance that doesn't come up as often is data bias. Bias inputs will inevitably produce biased outputs. So make making sure that when we're using data for AI and it's focused on decisions that it has a diversity of data so it can understand various different scenarios. And then finally we need pipeline assurance across the full AI lifecycle from data preparation to data model development, deployment and then ongoing monitoring.

The key take away from this slide and from what I'm talking about today is trust in AI is not created at the model layer. It's built across the entire data to AI value chain.

All right. And then last slide here for me. When organizations implement strong AI assurance practices, something interesting actually happens. So AI adoption actually increases, and this goes against what many people assume that governance and assurance is going to slow innovation down. But when it comes to AI, it's usually the opposite effect, because without trust, AI initiatives tend to stall. Projects tend to stay stuck in pilots because organizations aren't confident enough to scale them into production. And then organizations that successfully scale AI typically.

They focus on building 3 core foundations and this is what we also focus on internally at Trissential when we're building AI systems for ourselves. First is trusted trusted data, so ensuring that information feeding the AI system is accurate, consistent and reliable.

2nd is trusted models, validating that AI systems perform well, produce fair, explainable outcomes. And the third is trusted operations. So we're continuously monitoring AI systems to make sure they remain secure, reliable and aligned with business goals. AI Assurance connects these three layers together, and when organizations get that right, they're able to scale AI across their enterprise much faster, more safely, and with far greater confidence.

Q&A

I can certainly go through the chat and see what everyone's asking. And please do you utilize the chat if you have questions or even just comments. We'll start at the top. What's the difference between using data for analytics and actually monetizing data?

Yeah, I can take that one. It's not that data analytics is gonna go away. It's been around for decades. It's a it's a great step in the process. It allows you to get insight out of data. Many times data analytics is used for very precise questions, very precise purposes. Monetization is all about getting value through it, and that could be using some analytics to improve a process to identify trends in the market, but it's really about the action that it's going to cause.

And so monetization includes people, includes processes, includes structure, includes even sometimes physical changes. So it's really about getting the output and the outcome. And the other thing about analytics that's shifted is that we now can.

Mine data much better than we ever could. We can use AI to not just ask questions to, but we can have AI analyze large volumes of data to determine what predictability is in the data, what trends are in the data that perhaps humans aren't even aware of. That may alert us to all of a sudden this market segment, this product is trending up, trending down, or there's a relation to other outside impacts or inside impacts of our organization to you know perhaps revenue generating or profit generating activities. So we can use analytics in a in a much different way and sometimes I think.

In in the AI world that we have now, we can get it closer to the user, get closer to the decision maker than we ever could before. But good question.

How do you balance going fast first to market with the risks associated with AI? Is drift inevitable? Do you address it all up front or build a strategy to adjust on the fly?

Yeah, I'll go 1st and then Lyndon, you can talk about the drift aspect. It's really what a lot of organizations are struggling with today frankly is going fast, meaning for many organization means putting.

Some versions of AI, in some cases many versions of AI, many tools in the hands of many people. So that could be, you know, going fast. For us, what we try to do is we'd really try to balance. So we don't think there's a right answer in terms of one goes before the other, but we think that parallel processing is probably the best that for most organizations. Again, we don't have one-size-fits-all solutions. We really look at the organization, the size. The complexity, the industry, the maturity that you have and say where are you now? Some organizations have very good policies in place, very good tool policies, very good assurances. Some have literally none. And there's of course wide spectrum in between. And then of course things are changing fast. So if you had a great AI policy and governance and assurance two years ago, it's probably very outdated right now. So what we are seeing with I would say the vast, vast majority of our clients is that they really need to be probably accelerating on 3 fronts. One of the fronts is the strategy, the governance, the policies, the literacy, sort of these big strategic, you know, what is our AI strategy? What is our data strategy? How does it tie to our business strategy all the way through education, all the way through sort of those strategic decisions on the opposite end of that spectrum is getting tools in the hands of people and getting use cases through and usable and adding value. And so you can do that very quickly. I showed an example of what we're doing, which is a 15 day cycle where we can get AI to be practical and valuable and in the middle is how you scale it. In the middle is how you know what is the architecture? How do you build data products? How do you scale AI? How do you allow multi language, multi LLM, multi platform models fit for purpose?

And so how do you provide that architectural runway? So those are sort of three big categories of what organizations can be doing. And we just believe that you should be probably be advancing all three of those. If you just focus on putting a tool or two in the in people's hands, you're going to fall behind on governance and assurance and perhaps even ideation of use cases, and if you focus on any one of those too much, the others will lag behind. So that's just a general comment. Lyndon, I'll let you answer the drift question.

Yeah, absolutely. I think Craig, you did a great job from the leadership perspective of like speed of the project. But I think one thing that I'll touch on is the perspective of the developer, the engineer that's implementing these projects. So John, to your question about going fast and getting to market from the engineer's point of view or from the developer's point of view, like setting up what's called evaluation frameworks to catch model drift is the same evaluation frameworks set up that the developer would use to define their definition of done. So when they're going through a Sprint for that AI application, they know exactly what metrics to hit, what accuracy levels to hit before they can move on. And before Tri Central, I've been a part of projects where the developer didn't get that asset before starting and it slows things down considerably because you can never clearly define what's done and you can never really move on to the next task. And then also that same evaluation framework becomes a governance and assurance asset later down the line to where, to your question, when models start to drift, new models come out, you can catch.

It because you're already using that for your development lifecycle.

But yeah, to the other point, does is model drift inevitable? I would say, yeah, I would consider it inevitable, especially if you're staying on top of the latest models that are out there. And drift doesn't always have to be centered around accuracy. I think that's usually the common metric talked about, but it could be around tone too. So I think everybody remembers when Open AI switched from 4.0 to their first version of five that had significant drift in how the tone was coming out and a lot of people didn't like that. But there was if it was a custom application that had a framework, it would have catched tone and then reverted back to the earlier model as an example.

Great. Now before I move on to the next question, I think Craig has a question for the attendees. This is just really understanding your pain. What's the biggest obstacle or challenge your organization faces with data in AI? Feel free to share your comments in the chat. We'd love to understand that.

Next question. As Agentic AI systems begin to automate analytics and decision-making, organizations are effectively allowing AI to interact autonomously with enterprise data. How should companies rethink data governance and data product architecture to ensure these systems remain trustworthy while still enabling innovation.

Yeah, I can definitely take that. From a data governance perspective, it's about reframing the question a little bit. So the question changes from who gets access, which is typically what it is to really what is allowed under what conditions, right? And so once you change it to that perspective, that changes then your data product architecture, right? So then from a data product perspective, what you're doing is defining again, what's the metadata that I need to collect about each thing to understand and that condition and apply it, right. I'm then putting a policy in place to say, hey, under these conditions, this is what I'm gonna allow, let's say an agentic AI type agent with this from this business area to use this type of data and then what it does. And then the other part of that is you changed on the consumption side to say, look from a consumption standpoint internally we have to be reading that metadata at the beginning and say, hey, am I allowed to use this data for this use case. And so in some ways your agentic AI agents are almost saying, hey I found this data, am I allowed to use it because I want to use it for this business use case? And then your business use case and those parameters are kind of being evaluated.

Yeah, I'll just a layer on top of it. It's such a great question and I think a lot of people are struggling thinking ahead. Most organizations are not anywhere near getting to an agentic environment, but that's clearly where we're all headed. And I might have a little different take on this, but first of all, I would just say this is a huge opportunity I think to govern and get visibility and traceability decisions like we've never had before. So that's it's not just a risk, but it's a huge opportunity. So the first thing I would say is kind of layering on what Matt said, treat data as a product. If you truly treat data as a product, then you will have clear ownership of that data product. You'll have quality standards, you'll have defined data contracts that Matt talked about. And so you have a definition. If we think about how organizations make decisions today, they don't have all of that in place.

We just spent last week two full days with two clients and they are using spreadsheets as, you know, as a basically an ERP and there's tons of data in all of these spreadsheets on the side of all of the other enterprise systems.

That's not being traced, that's not consistent, that doesn't have quality standards and they're making decisions that nobody knows why or how they're making those decisions. So as we enter this agentic age, we have an opportunity to really get visibility and traceability to all of that. Second is we can put in real-time controls.

Things like lineage, access policies, monitoring, auditability. Again, anytime we have data, which in AI we can track every step of the way we should have, we should have AI that is traceable. We shouldn't have AI that's in a black box. That would be that would be a risk factor for sure. And then lastly is that we want to expose the governed data products and APIs. We want to have visibility to those things and then we can use another layer that Lyndon talked about, which is a assurance to monitor AI and have dashboards and have visibility on what's happening in much more of a real-time way than we have today. So I would, I would just add a few of those comments on as well as let's, let's look at it as an opportunity, not just a risk.

Yeah, I would also say too that a lot of the challenges around this are not like totally new to AI and that we can build off of the same security frameworks that we were using before. So you know, the way that I think about it is you should either treat your agents as particular people within an organization that have a particular set of access levels to certain data or the agent should inherent inherit the access level of the user using it. But in either case either scenario you're leveraging the.

Security controls that were around before AI. So what documents get an access? What parts of the database get an access? It's just extending that into this new technology.

Assuming those things are known, Lyndon, right? So a lot of organizations, those things aren't known. They're not documented. They change person by person, day by day, week by week. So it's an opportunity to really define those things and then monitor and you can govern them in a much better way.

All right. So we have our last question and then we'll kind of wrap things up here. So this question is around a this case is for over 500 products. So keep that in mind when I'm asking this question if the current state has multiple data repositories to tell a story about a product. What are your thoughts on building an AI model to extract information on current state versus consolidating first versus some hybrid?

Yeah, I think this is a like a really good question that comes with its own trade-offs, right? So if you were to take the approach of let's have the AI go throughout 500 program products aggregated together for a response that could work, but you're also giving that AI a lot of tasks and it could breakdown and then also you're adding in a level of um probability where you don't need to. My first initial state is like you should probably consolidate that data into an aggregate so that the AI can access it and inference over it and have one clear task. Sometimes that's not possible and you could build out a multi-agent framework. So it really depends on the scenario, but at when it comes for AI performance and accuracy kind of at a raw starting point, I would say aggregating that data together is probably going to be easier for the AI.

Now, yeah, now the one thing I'll do is this is very similar to a use case that we actually had with the healthcare company a couple of years ago. One of the challenges though, right, when you're consolidating that data is everyone always underestimates how long that that's going to take, right? And part of our recommendation is to put that if you want to think about the business semantic layer that can sit across each of those repositories and then what it allows you to do is have your agentic AI leverage that or your BI Applications and then what you can do is behind the scenes start to migrate the data and consolidate it into the into the better quality established system. That way you can still get immediate value versus having to wait for the full consolidation to happen.

Yeah. And there's, you know, these kinds of this, it's a great question. I would just say we do get this often. We had another healthcare provider organization that had a vision to consolidate all patient data in the cloud and that overall road map was 10 years long and so that clearly is not a viable solution for all parts of the organization. They can't wait one year, much less 10 years.

So the answer, you know, there could be three or four different options. You know, one is let's take what we have and let's analyze it. The other on the opposite end is let's get it all consolidated, which could take a long time. And the third in the middle could be let's virtualize it as an interim solution, which is what we ended up helping healthcare, large healthcare system do is get short term momentum with virtualizing data and then long term they're moving to the cloud and there's likely some hybrids in between. What we just always would recommend is a very thoughtful business case, you know it, it matters the size of the organization, it matters the industry, it matters what you're going to do with it and what the value of that is. And so to have a thoughtful business case scenario is going to save a lot of time and effort. And the great thing is we can now use AI and we can do that analysis much faster than we ever could before, and we can also do proof of values and proof of concepts really, really quick on some of these concepts as well. So we try not to have one-size-fits-all answers for clients because we know every client's in a little bit different place and the answer is likely gonna be different and depending on all the details of the scenario and what the ultimate value of it is, but good question.

I think that segues really well into next steps, how people could maybe get started.

Yeah, it's the number one question we get when we first talk with clients is great. You know, we want to go faster. We want to get more value. All the things that we talked about with data readiness and monetization, it's not a lack of desire.

It's really of how do we prioritize and how do we get started. And so one of the fastest, quickest, easiest way to get started we found is by collecting more information about them and their environment. And so we've created a self-survey tool. It takes 15 minutes or less.

What we do is we take that data, we benchmark it against industry best practices and then we create a prioritized list of improvements that give opportunities for monetization, risk identification and then we sit down and discuss it and talk through it and see what makes sense for that particular client, that particular even department of a client or that particular person and what they're working with. So it's a quick way to gather more information, get more insights and then provide feedback and determine what next steps, might be most beneficial. Sometimes those next steps are another conversation, sometimes it's an assessment, sometimes it's taking one use case and saying if we can get one use case done in 15 days, then we can sell the next three use cases and the next six and then we can talk about some of the bigger topics like architectural runway and strategy, governance and things like that.

So that's available to everybody on this call and everybody that you know gets the results from this in April. I think you'll be finding a way to communicate that to everybody.