



AI LEADERS PODCAST #48: FROM DATA TO VALUE: HARNESSING THE POTENTIAL OF GENERATIVE AI AUDIO TRANSCRIPT

Erwan Menard [00:00:00] We need humans to make these technologies respond better to business problems.

Chetna Sehgal [00:00:14] Hello and welcome. My name is Chetna Sehgal. I'm a senior managing director with Accenture and the Global Practice Lead for the Accenture Google Business Group. Now, we are here to talk about something that gives a whole new meaning to the word disruptive: generative AI. The impact is unprecedented whether we love its promise or, you know, a bit apprehensive. But most of us do agree it will forever change our lives. Now we know we can't change the direction of the wind, Gen AI is here, but we can certainly adjust our sails to better control our destiny. And that's what this conversation is about, to understand those nuances, the power, the tradeoffs of these winds of Gen AI so we can adjust our sails. And I'm really excited that I get to have this conversation with Erwan Menard. He's the director of product management for Google Cloud AI and ironically, just got back from a sailing adventure. So, Erwan, welcome. First off, how was your trip?

Erwan Menard [00:01:19] Thank you. Thank you. Chetna and thanks for having me. That adventure sailing along the coast of Alaska was energizing and humbling. Putting yourself in the hands of Mother Nature in a place where there are very few human beings and a lot of wood logs floating on the water, that was incredibly refreshing for someone who spends a lot of time

in gen AI and technology. So its great remembering that there is Mother Nature, there is Planet Earth. I think we could do an entire podcast on that, but I'm not sure that's a topic you want to talk about today, correct?

Chetna Sehgal [00:01:56] That sounds amazing, and so we have a topic for our next podcast now. Wonderful. So now should we go on to the adventure of Gen AI? So first, let's talk about data. Now, we've heard this many times, for AI to be the differentiator, cloud has to be the enabler and data the driver. Now, as an enterprise, I hold my data on near and dear to me. So the question becomes, with generative AI tools and LLMS, is my data going to become everyone's data?

Erwan Menard [00:02:27] So I think a question like this deserves a very short answer. We need to be super clear about something like this when it comes to enterprises. And the short answer, at least as far as Gen AI and Google Cloud is concerned, is that your data is your data. Now, let's double click on that, because I think there are a few aspects here that have nuances that are important for everybody to comprehend, be comfortable with, to make the right decisions for your business. So I just want to touch a few things. Like, first of all, as far as Google is concerned, we are here building on years of research for responsible and secure AI. We've invented many of the technologies on which generative AI is based. In 2018, we published our AI principles, which I think is something that



a lot of companies could consider as they look at ethical aspects of AI, and this framework to guide our decisions on how we research, we design products, we develop, we roll out products. These principles are here every day in what we do. And so when we take the research and those principles and we look at building enterprise grade products, we really want to make sure that our commitment to data privacy is very easy to understand and absolutely unambiguous. So let's be concrete. If you're a Google Cloud customer, the way you're going to use your data in the context of generative AI is in the AI platform. We call that product Vertex. And so what happens is that you discover a number of models in what we call the Vertex model garden, and you build on these models, adding your data to optimize the behavior of the model to your business. And so in simplest terms, you use a foundational model that is frozen, that works in a Google tenant, that is the same for every customer touching it. And then you use your data in your tenant, in your perimeter of compliance and security and privacy in order to build what we call an adapter model that allows to optimize the behavior of the combined model to your business. And that optimization, that adapter model, the queries you put to the model, the data corpus you bring to do incremental training, the output, all of this is yours and yours only. And it's super important from a data privacy point of view, but it's also important from a cost point of view because you're leveraging this foundational model that saves you years of R&D that your company may not want to entertain on your own. You're going to build on top of that. It's very efficient, but that optimization has to remain yours and yours only, and your competitors should not be able to access it and so on. So this is actually really core to what we do at Google Cloud, and I think it has to be core to the way enterprises think of rolling out these technologies. Yeah.

Chetna Sehgal [00:05:30] Yeah, I love that because I think that's got to give a lot of confidence to enterprises, right? Let me pivot to another hot topic, people. One of the big areas of impact of gen AI is the workforce. So what does this all mean in terms of retooling, reskilling, re-energizing our workforce?

Erwan Menard [00:05:51] Yeah, I think. I think it

means opportunity. You know, you're talking about a change of the magnitude of cellular communications way back when, when the Web changed the way, we share info, and the cloud changed the way we think of scaling Internet services. So it's a huge wave for our society. We can all sit at the dinner table at home as much as at the office, and essentially time to results, and skills entry barriers are shrinking. I think this is the point of inflection that will bring, you know, the efficiency we've been looking for, like how do we unlock data for massive benefits to humans, to planet Earth? And so back to the question on skills. You know, if I'm a business analyst, I can see a way to improve time to value if I'm a customer service person or contact center agent. I can be simply more efficient at my job, focusing more on the customer, relying on the AI to help me with some of the administrative part of my work, or helping me finding an answer faster. If I'm creative, like I'm an architect, what I love is designing houses, looking for that picture to illustrate my work. If I can get that done by creating a picture in a dialog with AI, that saves me a lot of time. So I think these examples show a lot of opportunities in terms of new business outcomes and productivity. But what is also interesting is that none of these examples require you to have a machine learning background and certainly not a machine learning Ph.D. So that is huge. You can now, as a C-suite member, you can now think of equipping line of business people or app developers with technologies that are going to enhance their work, enhance the customer experiences they work on, and you can do that beyond your scarce machine learning resources. Do you still need machine learning resources? You do. But for example, take prompt engineering, which is the most simplest way to optimize the behavior of foundational model for your business. You can accomplish that with people who have no machine learning background. It's actually a new field that's developing. There's some very interesting discussions in the industry on the emergence of that role, how attractive it is to people of many different backgrounds. I think the pure machine learning people will also see amazing opportunities to grow their career and their skills because now, instead of thinking in a linear fashion about what problem do I want to solve, what data do I need? How do I build the model? How do I prove the model works well?



You start from an amazing foundation that others have built, and you can bring your touch on top, like in the Google Cloud Vertex model account. And you have 70 models right now that you can pick and choose depending on the modality and what problem you're trying to solve, and then you bring your touch on top as a data scientist. So I think every role in an enterprise in some fashion can be impacted by this. It's a huge opportunity and it needs to be organized. People can be scared about learning new things. You need to think about enablement. How do I train people? Hackathons? How do I raise the awareness? And then I know how do I help people really acquiring this incremental skillset that can help them bring that technology in there in their daily life? We've seen that inside Google, by the way, over the last year. The way we've been scaling out the enablement of Googlers on these technologies is a very rich experience that we're happy to share with customers as they go through that similar cultural evolution in their in their organizations.

Chetna Sehgal [00:09:37] Right. Right. And so it's about the reskilling, right? And so this notion that human innovation is optional is really a fallacy and proves that we will have a need for ongoing human skills in the world of AI. So that really should re-energize our workforce, given that the skills will be different and hopefully more exciting. And I think right, you talked about prompt engineers, data annotators, etc...

Erwan Menard [00:10:03] Absolutely.

Chetna Sehgal [00:10:06] So we've been talking about humans and many of us go through FOMO, the fear of missing out, and we hear that CXOs are going through that. They want to be in the game, but there's a cost to entry. And I know you touched upon costs a little bit, but let's talk about that. How should CXOs look at this cost of entry? Prohibitive? Worth it? Be measured about it? What are your thoughts on that?

Erwan Menard [00:10:32] So I think the short answer is we've reached the moment when you can prototype and iterate quickly on the AI. And it's a major shift in paradigm because if you think about your conventional AI projects so far, they were 9 to 18 month long journeys, and you were spending a lot of attention in data cleaning and

data preparation because you were...you are taking a corpus of data. You are building a model yourself. Garbage in, garbage out. If you didn't have the right data set, you may not get the proper business outcome and you had a relatively long cycle between the moment you formulate your thesis. I'm going to solve this problem that way and the moment when you could verify that. So why did the paradigm shift? Because those models are multitask by nature. They're going to give you an answer out of the box. May not be the best answer, but you are starting somewhere as opposed to starting from scratch. Then if you want to make that answer work better for your business, you have some relatively lightweight good efforts to do such as prompt engineering, as we talked about. Or you may have a corpus of data that you want to bring to further tune the model that takes us back to the adopter model we were talking about before. And that effort can be relatively quick and iterative to validate the business outcome and dial the proper TCO. What we see increasingly in conversations with customers is that once they pass the excitement of having a demo by using whatever foundation model, there is a real question around, okay, if I put that bot together on my website and it does ABCD for the customer, what is the cost and is that worth the incremental conversion rate I expect to get from that visitor on my website? And that dial between investment and outcome, we are very used to it, right? The industry is really good at that, and I think the opportunity we have is to address it in much quicker cycles. Thanks to gen AI you use less skilled personnel than you used to, and you compress the timeline because you start from a foundation baseline that you can optimize at your pace, at your cost. I think that's pretty essential in order to address the question of the cost to entry. Just go, iterate, prototype. Now we see customers doing large hackathons. We support many of them every week. They're a fantastic way to ignite innovation in your organization. At the same time, I think you want to think as a CXO, how do I organize that innovation? How do I pick the project that going to make a difference in employee productivity and customer experience in the right timelines at the right cost? You need that effervescent innovation needs to be given try and there's nothing magical here. At the end of the day, you have a budget. You need to decide what to do.



But I think the fact that the cycles are shorter and the entry barrier in terms of skills is lower allows you to move much faster through that innovation cycle. If I may...If you don't do it, your competitor is going to do it. So I think it's important to go into this with a conscious around the cost of entry, the need to select the right technology, but don't over plan by, get on it, get people on it. It's like at your home. People want to touch, prototype and iteratively confirm or downgrade their ideas. I think if we embrace that, we're going to see amazing things happening.

Chetna Sehgal [00:13:59] I love that it's about realizing value, right, and in your example, doing it quickly. So we had, you know, on this point, we had just recently a joint press release between Accenture and Google Cloud. So Julie Sweet and Thomas Kurian, the CEOs of our two powerhouses, they talked about how jointly we will help enterprises to reach new levels of performance and help transform the organizations. And this addresses the cost angle as it allows enterprises then to have the ability to truly realize the value from their generative AI deployments. So let's continue to talk about costs for a moment, but this time, the cost to our environment, sustainability, another big topic. Lots of perspectives out there and how Gen AI with its processing needs will impact the planet significantly over time. So Erwan, how is Google Cloud ensuring that your gen AI products are efficient in resource usage?

Erwan Menard [00:15:01] Yeah, I mean, here again, just like I was mentioning before, that we've been researching that very space of gen AI for a very long time before it became a topic in the in the area of carbon footprint, energy efficiency of our operations. We were actually we're the first major company to match 100% of the global electricity bill with renewable energy in 2017. And we've made that happen again and again. We're pursuing net zero emissions across all our operations and by 2030. So that's a very important topic for us. I'd like to bring that in context of gen AI in a very, very tangible way. Back in 2013, Google released the first version of a specialized processor, a computing unit for machine learning, which was the tensor processing unit, a.k.a. TPU. We had the V4, the generation four of this one, and it's amazing how

much efficiency we've been able to pack into these. We've actually extracted best practices of energy reduction for machine learning. They're packed into the product that they come also with design recommendations for customers. We've recently communicated on a what is probably the largest publicly available machine learning hub. It's a data center in Oklahoma, in the United States that has a cloud TPU v4, the generation four clusters, and it's 90% carbon free energy and it's five times more computing power than we had prior with the previous generations. And so we continue improving that hardware research. It's very important when we talk about these models to also talk about the research that happens at a silicon level, at a networking level, at a storage level. Now, above and beyond picking a cloud vendor that is conscious of that, another thing enterprises can do is to recognize that not everything is a nail that needs a giant model hammer. So the biggest model, the more versatile and multi-task it is, the more costly it is not only to train, which is probably not your concern, but to infer which is what enterprises will do in their services. So what we decided to do is to offer within our Vertex model garden dozens of models that you can pick based on your use case. We currently have more than 70 in the garden, which we just GA-ed a few weeks ago, and that number will keep growing. We think that getting choice is important to dial the right cost and consequently the right environmental footprint. The last thing I'd say is that we expect a lot of customers to start from the foundational model because this is a fantastic, fantastic baseline, as we discussed a few minutes ago, to optimize with that data and eventually decide to distill down the model to a smaller model once they know precisely what use case they want to solve. So think about it. The usual way of doing AI is I find a single product problem. I find the right data set; I build the model. I solve my problem with my model. We're reversing the product. You take a multitask model that is gigantic, you optimize it with your data to do a better job at your problem, and then you distill down to a small amount model and that takes your cost down and takes your environmental footprint down as well. So we think that the combination of the research we do on data center infrastructure, such as those TPUs and then the tools we're giving enterprises in platforms like Vertex so that they dial the right



model for the right use case. The combination of these two are the most practical ways to address the topic of optimizing resource usage as we apply this type of technology.

Chetna Sehgal [00:19:02] So basically right sizing the model to be used, correct? Right? So enterprises do want to do the right thing, talking about right sizing the right thing. You know, they say many work for the cause, not for the applause. But that does come with the burden, right? The burden to ensure your AI is ethical, that it's not biased. I know that's a topic near and dear to your heart. So what does CXOs need to put in place to address this piece proactively?

Erwan Menard [00:19:30] So one best practice that I experience as a governor is this very, very clear statement around the AI principles. So Google did publish them many years ago. We live by them. I can see it in the way we're developing products with the team. And just real quick, we're talking one, be socially beneficial. We developed AI considering social economic factors. Number two, avoid creating or reinforcing unfair bias. Number three, be built and tested for safety. Four be accountable to people. We designed systems with an opportunity for feedback and feedback loops to improve incorporate privacy design principles. We talked about your data, is your data at Google Cloud? Uphold high standards of scientific excellence and then be available for use consistent with these principles. So don't apply those principles only when you think and develop, really carry them all the way through the lifecycle of the technology. I would say to any CXO, think about making a version of this for your company. You could keep those seven or adopt them or augment them. And the second thing you can do, get yourself some tooling because those ethical topics are not only a matter of having a position and being able to articulate it and communicate it to your customers and your team members, it's also a matter of tooling and checks. So if I look at the Vertex platform, you have safety filters, you have a prompt library and from validation tools, you have reinforced learning with humans, as you were mentioning earlier. Yes, we need humans to make these technologies respond better to business problems. We give you safety scores that are off the shelf based on our experience,

but you can tune them to your business. If you use gen AI for coding, we give you recitation checks and we give you an evaluation framework for your models. I think principles in one hand, tools in the other hand are the things you just need to make very important items on your AI agenda, and that's how you're going to eventually drive your company and your customers in a very ethical roll out of these technologies.

Chetna Sehgal [00:21:53] I tell you, I love that advice. It's that combination of the softer piece of culture and principles with hard tools like don't forget either one of those, right? So AI has always been a controversial topic and with Gen AI, the promise of AI has increased, but so has the fear around it. So Erwan, what are your final thoughts around how one can look at this dilemma of AI?

Erwan Menard [00:22:20] I think you mentioned it at the beginning, right? This is a major milestone in recent history, like maybe like electricity at its own time. You know, it comes with possibilities and some concerns. So if I look at the way customers and partners are using this, I think there are two or three characteristics that are amazing. And each of them comes with, to your point, with this mix of a great opportunity and areas that need careful attention. Number one, we talked about it. Everyone can think of a use case. Isn't that amazing? This democratization? Entry barriers going down. Every employee, every group, a line of business group, anybody with no machine learning background can start thinking of a use case. A business analyst, a contact center agent can imagine ways to be more productive. Then, of course, the corollary topic is how do I organize this innovation beyond the excitement of a single hackathon? How do I pick the right projects? This needs to be done. This needs to be organized and responsibly. The second thing is that, collectively, we, not only the actors of the industry like Accenture and Google, but also governments. We have a new set of topics that we need to properly frame to ensure that the technology is used for good purposes. So we talked about data privacy today. We could have talked about security, like we're using machine learning, LLMs to improve the way you can prevent threats. These are the tools for



responsible AI that we talked about. So we take this responsibility very seriously as Google. But to be perfectly clear, that's probably the reason why we were not first to market a technology that we have led innovation on for decades. When we GA our products and our product right, we don't outsource development of Gen AI products. It should be known that we continue to make responsible AI our top priority. We put our principles in the center, and so we are the AI company in many ways, but we would not release something without putting safety and security around. I think that we're trying to role model that, but it's not only about Google, it's about Accenture. It's about every enterprise, it's about every government, and we collectively need to address those topics and frame them properly. The third thing that is kind of mind blowing is that, again, enterprises have different needs than consumers. Again, meaning we have a revolution that starts from the consumer point of view, because when the technology became available to every consumer, that's when this explosion of innovative ideas started. But enterprises have different needs. When they put their brand behind about answering questions, it comes with a certain level of expectations. It's also a great opportunity, right? I mean, you can bring enterprise grade products to market. So let's take an example. You can ground answers of your bot in your corpus of data. You can do that without developing your model from the ground up. You start from a foundational model. You optimized with your corpus of data, and then you bring an answer that is not only an answer that is conversational and much more friendly than browsing through an FAQ yourself on a website. But you can ground the answer. You can say, here is my answer, and that's where I did find it pointing the website visitor to a tangible resource on that same website that is verified and clearly grounded in factuality. That sort of opportunity is unique to enterprises because they have the data set, they have the databases, they have all the data sources with created facts that they can use to improve the quality of responses these technologies bring. So I think that's a great opportunity. And here again, it needs to be addressed. Being organized and professional about it, which data we use, do we have the skills to perform that model optimization? But if we do that well, the promise of the customer experiences we will

enable or the employees' experiences or employee productivity we can generate is massive. So yeah, you're right, there's a huge opportunity and there are new concerns that we need to collectively understand, address, organize. We feel really confident that the opportunity overcomes the challenge and let's go. Let's prototype together. Let's figure out the right tools, the right procedures, the right regulation to enable this. Because again, the experience for customers, productivity for employees, I mean, we have new frontiers that are opening up and that's truly, truly exciting.

Chetna Sehgal [00:27:10] You know, many enterprises have been successful over the years with disruptive technologies, but this one is like the mother of all disruption, right? You compared it to electricity a moment ago, and I totally agreed. And as Pollard said, the arrogance of success is to think what you did yesterday will be sufficient for tomorrow. Now, generative AI, we know, is going to be changing tomorrow in unimaginable ways. So I really enjoyed our conversation because as an enterprise, how do I welcome the wins of the generative AI by understanding its impact? And we talked about the areas of data, people, sustainability, ethics, gleaming value, and so I can adjust my sails to better control my destiny. So I really think the conversation, Erwan, we just had really sparked that thinking for me. It was an absolute pleasure talking with you. And by the way, your passion is very contagious. So thank you so much.

Erwan Menard [00:28:08] Thank you. Thanks for having me. And thanks for putting a little sprinkle of sailing in your final remarks, adjusting your sails. I really love that. Thank you for the opportunity.

Chetna Sehgal [00:28:18] Absolutely. And I would be remiss if I don't thank our listeners. So thank you all very much. And wherever you listen, don't forget to subscribe to our podcast. Be well.

Copyright © 2023 Accenture
All rights reserved.

Accenture and its logo
are registered trademarks
of Accenture.