



- Patrick Moorhead: Hi, this is Pat Moorhead, and we are back with The Six Five Summit 2023, day two for enterprise software and SaaS track opener. We might be talking a little bit about AI. I don't know, Dan. How are you doing, my friend?
- Daniel Newman: It's great to be here. Six Five Summit. We are in full effect.
- Patrick Moorhead: I know.
- Daniel Newman: It is such a big year for technology. We started this whole event all about navigating rough waters, but one of the ways that companies are going to be able to navigate rough waters, Pat, is going to be the investments they place in software, in AI, in those deflationary capabilities that enable companies to be more productive and drive towards growth.
- Patrick Moorhead: Right. And that's one thing I love about tech. It's one of the few industries, when you do commit to troubled waters, you power it down and you innovate your way out of the challenges. Doesn't happen in every instance, but I've seen a ton in 30 years.
- Daniel Newman: I call it "teching" our way out.
- Patrick Moorhead: I like that. I like that, Dan. I'd like introduce our guest. We can't chit chat this whole time. Kareem, how are you?
- Kareem Yusuf: I am well, thank you. How are you guys doing?
- Patrick Moorhead: Great. Thanks for coming on the show. First time Six Five, but we hope it's not the last. But we did have your fearless leader keynote the entire event last year, Arvind, and a lot of your senior leaders. We are excited to have you.
- Kareem Yusuf: Well, thanks for having me.
- Daniel Newman: Yeah, Kareem, it is great to have you here. You are kicking off the enterprise software track. IBM has a huge play in software, but sometimes because it's enterprise software and SaaS, people are like, "Oh, well, IBM's SaaS?" Well, you are making big moves. And, of course, the company over the years has been making big investments, but AI has been the thing of the year. It's the tech moment of the year. But you've been following this a long time. IBM's doing this a long time. I'd love for you to just give us a little color on what's been changing over the last couple of years and why is AI having such a big moment right now?
- Kareem Yusuf: 8, Well, when you think about it, if you step back to what it means to actually create AI, this notion of creating AI models, most recently the state of the art has been what I would call machine learning and variations on that called deep learning. And that's all been about taking a set of data, labeling the data, and then training if you like the model with that data. I like to say, "Here's a picture of an orange. You annotate it as an orange. Show enough pictures of oranges and hopefully the model, next time you show it an orange and an apple, will know it's an orange."



What you note in that, however, is that it's task-specific and, in its own right, labor-intensive because of the labeling you need to do. Now the advent of this techniques called foundation models, which leads you to what many people think of commonly as large language models, so I'll use that by way of describing, has been about how you do this training with vast amounts of data but less labeling. So, one, it begins with create this model that uses large amounts of data, so training on everything on the internet would be where people's minds commonly go. And then you begin to establish this base model or foundation model from which other models can be derived.

Simple analogy I like to use is think about teaching a child the alphabet. That suddenly unlocks many a use case from conversational speech to writing poetry to writing an essay, if you kind of follow that simple analogy. And so it's that power of foundation models and the very idea that you can use this base model to then create many more models quickly with just slightly more data that is driving this inflection point because what it fundamentally means is you can think very much about how you scale adoption of AI. And that is what we're focused on in applying those kind of techniques in the enterprise.

Patrick Moorhead: So as I learned at Think back in May, IBM was not new to foundational models. In fact, it's been a multi-year journey and, in fact, AI is not new to IBM either with machine learning and before that a lot of it based on analytics years before that. It's been a multi-year run. AI is getting better. It's also getting a little bit more complex.

But you made some very big announcements back in Think at May. Can you talk a little bit about how you're helping clients on this journey? You hear about hundreds and thousands of models, different models, the ability to put it together, the challenge with data that we see. Garbage in, garbage out. If the data prep is not there, if it's not protected, if it's not governed, how are you guiding folks through this?

Kareem Yusuf: So really two key vectors, if you like, when you think about this. One, how we've been using foundation models to drive new intelligence into applications or products that we provide today for various use cases. Think about products like Watson Orchestrate for Digital Labor or AI ops for IT automation-

Patrick Moorhead: For your own products.

Kareem Yusuf: ...for our own products. So that was the first step. You take what we can do in training these quite interesting models and embed them in. And we announced in particular a particularly exciting new offering called Watson Code Assistant that is all about having trained a foundation model on code being used for the Ansible community to help with code generation, code completion, code recommendation for those kind of automation tasks.

The second big vector, and in many ways for many the bigger news, was the announcement of the fact that we were now providing a distilled platform for AI and data to do all of that work. And our platform is branded What's Next. It's our What's Next platform for AI and data with three key components that kind of speak to how, as you say, you unlock this valley for others to build their own. First is What's Next for AI.



Think of it as a studio, an environment for the builder, to either focus on how they want to train up completely new models or, more likely, take foundation models that we provide or that others have provided via the open source community. You will recall we mentioned our partnership with Hugging Face to bring that community through and do the things you want to do on those kind of models. Fine tune them, prompt engineer them, at the same time still supporting the traditional machine learning techniques as well.

Second component was What's Next at data, which we showed was essentially a next generation data store for supporting the data you need for doing AI built upon an Open Lakehouse architecture and really bringing that kind of openness and modulation of various types of data into the fore. And then the last, you mentioned this, What's Next at governance.

How do you bring governance through this entire process? And when you think about governance around AI, you're thinking about do you understand where the data is being curated from? Do you have those, what I call, the data facts? Do you have your notions of your models and that audit trail? What have you trained? What are you expecting the model to do, the model facts?

And then how you run that all together in a cohesive way to observe how your model is performing in the wild, what corrective actions you choose to take or modifications along the time and actually bring in the rigors, if you like, of regulatory and compliance-style processes to this. Because within the enterprise, you ultimately need to be able to clearly articulate, and we're seeing this occur, why you've done what you've done and why it's doing what it does.

Daniel Newman:

So Kareem, it sounds like you're doing a lot for enterprise. And everything you just mentioned comes down to simplifying, scaling, even introduction for a lot of companies that are just trying to figure out how to get started and IBM having that really full stack approach from hardware to run this all the way out to consultants that can help you design, it's really important. But when I'm listening to you, it makes me think a lot about the digital transformation thesis of the past decade.

So it was all about how can you modernize processes to become more productive, to build scale in your business, to implement tech. And you and I always joke, "Well, there was no analog transformation, but what it was all about was processes that made your business better." But when I listen to this, this is going so much faster. We thought the last wave was fast.

This is happening at exponential speeds, meaning every enterprise is going to have risks of not getting on board quickly, but at the same time, there's plenty of ways they can mess it up. There's many pitfalls, many gaps. What are you thinking and how are you working alongside the enterprises to make sure they don't fall into these pitfalls and they don't miss these great opportunities because there's only one way to go. You got to get into this, but getting in wrong could probably be almost as bad as not getting in at all.

Kareem Yusuf:

So first thing as just as a bit of a clarification, when I use the word enterprise, I simply mean business. And so I'm actually not implying size of the business when I use the word enterprise, I just think of any business as running as enterprise. And so that's very important. And so what we are focused on scales and indeed my comments that follow in my mind scales from small to large. It's not about any particular size. But when I think about this notion of getting started in this whole world of AI and how we see it accelerating, it really comes back to me, and to my



mind to, the use cases. What is the use case? What's important about the enterprise in my opinion is it's all about domains of work, and as you mentioned, critical business operations. So where are we looking or where are we, one, as an enterprise looking for advantage?

It could be focusing on some customer service use cases and accelerating work there or internal employee productivity type use cases or even for the professional who's doing certain tasks to get them done more quickly. You often see that in the realms of digital labor. It could be for IT automation tasks or scaling more work that needs to be done on security or getting started with a smaller team. Why I pick those first and foremost is that I actually think, now when you think about this new world of generative AI and some of these new advancements that are occurring, for many it will actually start by consuming those capabilities embedded in tools that they need for their tasks every day.

It's kind of making those tasks better. And you will recall we actually had announced back at Think our partnership with SAP and, if you remember, they were talking about how they were using a natural language interface to make it easier for the SAP application user to use multiple applications together without actually having to think very deeply about them. "Give me my procurement order, give me that last transaction." That kind of stuff in natural language form.

So I think for a lot, it becomes that level of unlock. Then, to the point you make, for some they actually will want to do stuff with the underlying platform. Enhance things we've provided with their own data, build new models or new use cases from scratch. And as always, it begins with, as I said, that domain knowledge and use case.

Next, depending upon how you're entering this, do you know where your data came from? Do you know what you're using? It's so easy to just jump online and say, "Oh, I'm going to use that tool, that." Within business, we have to be a little bit more circumspect. You have to know. Accuracy becomes important. Predictability and behavior of any kind of AI model becomes critical. You have no rooms for hallucinations in a enterprise or a business or true business use case.

And so that becomes very important. And being able to effectively and easily track that. And the last thing I will speak to, which you actually mentioned at the beginning, it comes back down to how do you even consume this tech? This is why SaaS becomes very important. Because getting started, time to value, very often software delivered as a service, has that acceleration already built in. And so you are really talking about, in all that we are doing, providing environments that you can quickly log in to and start getting the work that you actually want to get done done.

Patrick Moorhead:

If nothing else, what you just talked about, hopefully, people were paying attention to that question, is also the striking difference between, I'll say, B2C, generative AI, and foundational models, and B2B. You don't have room for hallucinations. You really don't have room for a lot of mistakes. And I think if we've seen with prior technologies, if anything can stunt the growth of a technology, it's just coming back with something that, even if it's 90% effective, people are going to be a little bit tentative, even if there is a big warning there.

And I'm also very enthused to know that there are solutions and reminds me a lot of the old speech-to-text models, which when you narrow them down on a certain vertical, they worked great, like medicine. And it looks like that's a similar way that foundational models are winding up. If you narrow a little bit of the focus and you're using your enterprise data, not the entire



data of the entire Internet, your results are going to be more along the lines of what you expect in the first place.

And I think that should be a real eye-opener and a real motivator for companies to jump in here. So we can't go a day without some new breakthrough happening. One of the great things about being in tech, and I think we've all been in it pretty much all our careers, is that once the entire industry gets focused in on something, things happen on a daily basis. And keeping up at this amazing pace, but also keeping all of the important enterprise grade industrial strength principles in line that you have has got to be a challenge. How are you keeping up the pace as it relates to go to market? I really enjoyed the meeting that we had back at Think with you and Dario. Talk a little bit about how IBM is moving so quickly in this new age.

Kareem Yusuf:

Well, it's funny. When you step back and think a lot about pace, pace is actually, and I think this is a beautiful thing, being set in the open source community. There is so much going on there in the open source community. And as you well know, from our Red Hat business, Red Hat Linux, or OpenShift, which is key and essential to this, has been enhanced and we engage with various open source communities, to all that we do on the next level within Watson X and integrating nicely. That's how you begin to allow that pace to come through. And so I think that's really critical.

I think the other side of it is, of course, what we do in making sure that our capabilities are easily understood. Appropriate, to use the phrase nurturing through the digital experience. This is something that's essential for us to focus on and particularly critical in this time, more so than maybe in the past in the world of enterprise software.

And just to overload the term, there I'm using enterprise to refer to the bigger ones who often had large teams. I go back in time when we use the word enterprise. But for the business of today, the enterprise, the agile enterprise of today, big or small, nobody can afford or wants to spend their resources that way. So how do you allow your capabilities to be easily discovered, readily tried, easily transacted, them brought in and adopted all become important.

Side note, by the way, I do want to mention this. I think it's always important when we talk about foundation models to remind people, by the way, we're not just talking about large language models or language models. There's so much work we've been doing on code models, you might say programming language. I'll leave that alone for a minute. But think about geospatial data. We talked about the work we're doing with us-

Daniel Newman:

We brought out some models of things.

Kareem Yusuf:

Exactly. Yeah, the work we're doing with NASA on that one, to understand environmental change much more easily and support the geoscientists. Think about models you can build on IT events. That's how you begin to get into the whole cybersecurity side of things. Or think about the nomenclature of time series data. There, I'm thinking about things like machine performance and energy consumption and building those kind of models. This represent the entire spectrum of activity to which we're applying this and so, to your point, can be a lot more tuned and directed to the right business operations that matter.



Daniel Newman: The overall opportunity is to create what will be a new enterprise. And I think there's a little bit of an underlying theme in how you want to present the enterprise as not just large, but every company. And obviously, we are at this really interesting inflection where companies with AI will be able to get really big without having to get what was historically structurally really big because the way maybe only hedge funds or some things were able to do that, AI's going to enable.

And again, I think there's a whole world in the future of new kinds of jobs and roles that are going to be created that most of us can't even imagine yet. And so it's too early to say that we're going to have a displacement issue, but we're going to have a upskilling opportunity. Having said all this, I would love to get your take and something maybe prescriptive for our audience out there is about the mindset of this enterprise that we're talking about going from AI as a plus to their existing transformational efforts to an AI-led strategy. Talk a little bit about how they should think about this in order to make sure whether they are a five person company today or a massive global fortune enterprise, how do they get to that AI mentality in their business?

Kareem Yusuf: It's a really, really good question. And, for me, when I think about our purpose, how we define it, which is to unleash the intelligence in the enterprise with our software, I think very much about this notion of reimagining how work should be done. And so, to your point, you go, "Well, how do you begin to reimagine?"

And let's be clear, it's a little bit of a step ladder, if you like, or a scaled continuum as you work your way through. Often, for many, it literally begins with automation. That really is often the starting point. Automation, but no longer in the what I would call traditional sense of "I'm going to build all these business process automation stacks and the very IT programmatic centric." Automation that can be triggered by natural language conversation, which means it needs to be dynamically sequenced. That's a requirement as you build. But back to your point, how do you begin to think about it? It often starts with automation. And by the way, that's kind of easy in the sense that we all know things we'd like to do differently or a little better or a little easier.

Daniel Newman: Quite a long list.

Kareem Yusuf: Right. Isn't that why the classic summarization use case from large language models what everybody leads to? Because all of us in every day would just wish somebody would summarize that bit of data for us. I'm sure we go through it.

Daniel Newman: He feels us.

Kareem Yusuf: So you get that. You get that. And then I think it ultimately begins to evolve to that level of comfort where you begin to essentially let the AI do the tasks for you because, to your point, of upskilling or elevating from the bottom, it begins to take over some of those mundanes. And I think this is a very important point because it serves two key purposes.

One, to your point of upskilling, allows an individual to elevate to where we bring more unique value. Because all of us, if we can get that summary done and it's accurate and it's on point, that frees up time for us to apply our minds to other stuff. But, secondly, and even as, if not in some cases, more importantly, it allows you to bring in people from the bottom and get them more productive with less training time. Think about industries where you deal with a lot of churn and



you need to continue to retrain. By being able to accelerate folks up that curve with the right kind of AI-based assistance, you fundamentally transform how that enterprise works in its own right.

Patrick Moorhead: So, Kareem, this has been a great conversation. We've talked about history of IBM in AI. We've talked about the technology. We've talked about your strategy of how you're applying foundational models, a little bit about the products, but I know, because I've talked to a lot of your clients, and obviously you do because you work with clients, is a lot of them come to IBM to help them put their data to work. They might not have all the answers and the know-hows. How does IBM uniquely enable its clients to put their data to work in this new foundational model world?

Kareem Yusuf: Well, I think day in, you bring the last element to the equation, expertise. And we bring that expertise to bear in two very important ways. Within our technology-centric world, we have our client engineering teams. These are teams that go and co-create with the clients just as part of our overall process. Explore the use cases, build quick prototypes, carry out quick POCs, make sure that we can build things with the client that gets them going to rally.

And then, of course, we've got our consulting organization who have invested heavily in building expertise around what we're doing generative AI to also support this ongoing digital and business transformation. And this is stuff they know how to do and have done very well in the past, as you will well recall, as we drove out the hybrid cloud platform and OpenShift, it was the same model.

The client engineering team who knew how to work with the clients on container-based architectures and beginning to get the best of that from a product perspective and a consulting organization that really tripped down on leveraging hybrid cloud platform as a business transformation initiative and driving that out and supporting clients and adopting. And in that regards, that's something that I do truly believe we as IBM are quite uniquely positioned to bring that kind of combination value while being open and working with everybody else.

Daniel Newman: Yeah, I've always said that the full stack approach that IBM has is definitely one of the big opportunities for it to position because there are many offering different parts, meaning whether it's the SaaS, whether it's the consulting, whether it's the hardware and building out the infrastructure to do it. But you really need all those parts and it needs to be, it's continuous, it never ends, kind of like how we said transformation will never end. Well, this is going to be a transformation that's going to just accelerate and being able to do it fast is going to be incredibly important. Kareem, I want to thank you so much for being here with us to open up the enterprise software and SaaS track, day two here at the Six Five Summit.

Kareem Yusuf: Thank you ever so much for having me. It's been a pleasure, guys.

Daniel Newman: All right, everybody, stay with us. Stay tuned. We've got an action packed day two at the Six Five Summit ahead. Thanks for tuning in.