TIME100 Talks: The Future of Enterprise AI

Words	5,082
Duration	00:29:10
Recorded on	Unknown date
Uploaded on	2024-09-18 21:30:42 UTC
At	Unknown location
Using	Uploaded to aliceapp.ai

Speakers:

Speaker A - 14.54%

Speaker B - 41.72%

Speaker C - 43.74%

Notes:

- I'm so pleased to be here with Parker and Daniela, who need no introduction. Both tremendous leaders in this industry and founders. Our conversation today, the future of enterprise AI. And we want to talk a little bit about trust and ethics.

- There's a healthy amount of questions and skepticism around what should I be using AI for. Quad highly recommend dipping your toe in first with a low stakes use case and building from there.

- Parker: Slack is a great example of how to bring AI into enterprise technology. In search, summarize conversations. Agent Forest slots right inside of slack. We're going to all be using AI more and more to get our work done.

- Much like time, Slack is how we communicate everything. It's often a, uh, really overwhelming amount of data and information. Claude can be an incredible, incredible partner in helping to summarize information. There's an immense amount of potential

across an application like Slack.

- All of Slack's data will be available in data cloud. Data cloud will also federate out to your snowflakes and all the other structured databases. We're building integrations to unstructured information. That's the power of the future.

- Daniela: Where are we today? What led us to the, and where do you think we're going in the future? Parker: I think we finally, in the enterprise, I would call it hitting the tipping point. I think it's the fastest technology has moved for me in a very long time.

Speaker A

00:00:00

I hope you're all doing well. I'm so pleased to be here with Parker and Daniela, who need no introduction. Uh, but I will give a small one. Both tremendous leaders in this industry and founders. I um, should say before we begin that, uh, time is owned by Mark Benioff, who we all know is the founder of Salesforce and co founder with Parker and that Salesforce is an investor in anthropic. Uh, our conversation today, the future of enterprise AI. Um, and I think we want to talk a little bit about trust and ethics and how these two leaders think about those two topics. Um, Daniella, let's start with you. Thank you for being here. How do you think about trust when it comes to your work?

Speaker B

00:00:46

Uh, well, first of all, Sam, thank you so much for having me. And it's great to be on the stage with both of you today. So anthropic was really founded with this principle of developing very powerful AI systems in a way that was robustly safe and ethical and trustworthy, uh, and fair. And really the way that we approach this is across a variety of different layers of how we actually train our models and then deploy them. The first is we absolutely believe that you cannot tack safety or ethics on after you've trained the models. At that point it's too late. So anthropic uses a technique called constitutional AI, which is essentially a set of guardrails that we bake into the technology from day one. This incorporates documents like the UN Declaration of Human Rights, but also industry terms of service to really help guide quad in how it communicates and works with others. We also do, uh, traditional trust and safety. And we work very closely with individual enterprises to understand what are their business needs and how can we provide value in a way that really takes into account privacy, security, trust.

Speaker A

00:02:03

Parker, I know trust is important to you. What role does it play in your vision for enterprise AI?

Speaker B

00:02:09

Parker?

Yeah, well, Salesforce started in 1999 and people didn't trust their credit cards on the Internet, much less their contact list or the customers they were selling to. So trust has always been our number one value. And when we launched into AI and we love being influenced by great leaders, uh, like Daniella and anthropic because of how they thought about AI. And so when we built Agentforce, we built a trust layer in. And so it does things like, I don't want my personal identifiable information going out of, of the trust boundary if I'm using some LLM that's out in the world that's maybe not an anthropic LLM, and I don't know what they're going to do with my data. I want to keep it in. I want to control hallucinations. Uh, so how do I scan things to see, was there a hallucination? Uh, I want to see if there's bias. Um, but also, these things are. This technology is not perfect. You know, it's not deterministic. So sometimes things are going to happen. And how are you going to deal with that? It's kind of like cyber security. You could say, like, we have the best cybersecurity group in the world. You're still going to have issues. So it's also about how you, uh, handle that and how you show up. So do you have auditing? Do you know what's going on? If the customer or your employees are saying, that didn't help me, or even worse, that was a hallucination, how quickly do you know about it? M how easily can you then go back and see what happened and resolve it?

Speaker A

00:03:50

Are there particular examples you have in mind where you've seen that play out so far, and how have you reacted to it?

Speaker C

00:03:56

Well, I've seen an example, actually, internally, uh, where we, uh, have summarization that's happening in slacken. Uh, and it's wonderful. There's been a couple of cases where the summarization picked the wrong person's name. It was like, it was the last person maybe the LLM, um, had thought about, or, and it said, this person just left the company. And that person let us know, hey, this isn't okay because I'm still here, you know? And, uh, so that would be an example where, like, it was a little emotional for someone, where the LLM, um, summarized something and said, oh, it's so sad that our head of employee success, Natalie Scardino, just left the company. Natalie's like, uh, I'm still here. Um, what's good, though, is you can have, um, references or annotations to go and say, like, well, where was that said? And you click on it as a human, and when you go look at it, you realize, oh, that's not actually what happened. As an end user, I wasn't stuck. You know, it's very important that you attribute in the AI. You don't just say, it's brilliant. Here's the answer. You say, well, how did the answer happen? And so when we saw that you click on it and you see, oh, it actually didn't say, uh, Natalie. So that's too bad. It didn't work in this one case, we're gonna fix it later. But as an end user, I wasn't that upset, because I was more upset at the AI, like, okay, I gotta go fix that. As a technologist, than I was about, you know, misconstruing something it had said to me.

Danielle, I'm curious. Um, many of us, probably all of us, watched Mark Benioff's keynote yesterday. The theme of this week on agents. Um, how does this evolution towards agents change how you test and build models like Claude?

Speaker B **00:05:51**

So I think the concept of an agent is such an important one, and not just because that's what this conference is really structured around, but I think, fundamentally, uh, AI systems of today are so capable in so many ways, and there are still limitations to what they can do in the real world. But I think in contrast to, say, a year ago even or a year and a half ago, models are capable of doing, uh, so many more things independently, right. Customers like United Airlines or Doordash are using quad to automate a huge amount of their customer service. Right. And probably everybody in the audience has called at and t at some point and gotten stuck in some kind of terrible loop with a traditional AI. And I think that has sort of led to, to a, uh, sense of fear around really taking these models and allowing them to have their kind of full agentic identity, uh, sort of revealed as the technology develops. But I think so much of what we're seeing today is when you build the right frameworks from day one and when you instruct the model correctly, how to engage, when to escalate the process of teaching these models when they don't have perfect information, when they do need to pull a human into the loop, I think that is something we will continue to see develop over the course of the next year, two years, three years. But I personally have felt incredibly astonished at the degree to which quad is capable of going and executing on a wide range of tasks in a row at an incredibly high level of accuracy. The last thing I'll say about this is, it is true that no model today is perfect. Right. Some models will hallucinate. They will consider so and so left the company accidentally. And I think an interesting thing to think about, depending on your business use case, is, is the accuracy of an LLM. Perfect is one way, but another way of thinking about it is, is this model more or less accurate than the human who would be doing this task? And I think over time, what we aim for is perfect accuracy on tasks that are easy to identify as having a yes or no answer. But in some cases there's also amount of subjectivity that comes in. Sometimes you might need a human to make those decisions, but increasingly an AI can.

Speaker A **00:08:16**

We started this conversation on trust and you just brought up fear. And I'm curious how you help guide partners, customers, other enterprises over that fear so they feel comfortable with this evolution.

Speaker B **00:08:29**

So first of all, it is not surprising to me that there is a huge amount of question, concern, right. Parker talked so nicely about for every new technology that emerges, folks are uncomfortable saying, how can I rely? How do I know I can rely on this? If you even think

back to 15 or 20 years ago, the concept of purchasing something on the Internet was not necessarily something that everybody was comfortable with, right? Over a certain age, folks might say, I really prefer to go buy something in a store. I do think there's a healthy amount of questions and skepticism around what should I be using AI for? The way that we usually counsel our customers who are getting started with Quad is start with a use case where you feel really comfortable, you fully know what success looks like, and deploy an LLM there will work closely with you and say, okay, what are the benchmarks that are important for you to hit, to use the customer service example, what's the level of accuracy that you're hoping for? Or what's the time to resolution that you want? And then just run a pilot. Right. With anything that is a new technology, we highly recommend dipping your toe in first with a low stakes use case and building from there.

Speaker A **00:09:40**

Let's get a little bit deeper into these applications. Parker, when we talked, uh, last week, one of the things you said was that up until the launch of agent force, that Slack was perhaps uh, the best example of AI at Salesforce. Parker wears many hats. He is the CTO of slack as well. In what ways is Slack? Ah, such a great example of how to bring AI into enterprise technology.

Speaker C **00:10:02**

Yeah, that's true. Before yesterday, um, one of the things we focused on with Slack is instead of having to go, uh, and address an agent or go and address the AI and ask it a question or ask it to do something, we just wanted it to be in the flow of work. So if you're doing a search and you're trying to find something in slack, we want to make that better. So you can ask it a question and it'll actually use vectorized conversations of all that unstructured information. And LLM will process it and it'll actually answer it in a human question. So I used it, for example, you know, uh, I joined six months ago, uh, to help bring Slack even closer to Salesforce and integrate it more deeply. But engineers make up code names, right? I'm sure, uh, you got a whole bunch of code names at anthropic. So I'm trying to learn very quickly. And all the knowledge was actually in all these unstructured conversations. And there was a project gondola. It was like some technology stack that they had, and I'm trying to learn all the technology. So I just said, what is project gondola? And processing. Processing and using anthropic did a great job. And it gave me this beautiful answer. It said, oh, uh, it's a technology for moving code to production safely in a reasonable way, etcetera. And by the way, if you want to learn more, here's where the references were. So it's a beautiful use case. So in search, summarize conversations. So I can't look at all these channels. Summarize them for me. Great workflows for me. You know, we're going to keep adding things. And so that's slack AI that complements agent for us because slack is the best interface. It's a conversational interface. What better interface for AI than slacken? Agent Forest slots right inside of slack. If you wrote the slack keynote, you saw that. So you can address Agent Forest. The agents that you're building

directly, they can be with you in channel. Uh, there's also third party AI's. I'd like to say that everybody should use agent force and that should be the only AI you use in the world. But maybe use Adobe Firefly because it's an amazing AI for building images that works right within Slacken. Maybe you're using workday and you want to use workday's AI. Why go to workday? Just use it in Slack. And so that's the beauty of, uh, Slack, both slack AI, but Slack also being a, uh, conversational interface. We're getting work done, and we're going to all be using AI more and more to get our work done.

Speaker A **00:12:40**

Daniella, those of us at time live inside of Slack. I'm sure the same thing at anthropic. You do a lot of work there. Have you learned lessons about how your interactions on Slack can be helpful in this conversation about enterprise AI?

Speaker B **00:12:54**

So Slack is one of Anthropoc's earliest and deepest partnerships. And part of why that is is we built a lot of originally, uh, native integrations into Slack with Claude and the Slack team was incredible at ah, partnering with us on this dimension. I really think in large part because like Parker said, the obvious use cases and applications are there for us. Much like time, Slack is how we communicate everything. So much of our work is done in Slack, and it's often a, uh, really overwhelming amount of data and information. I know when I leave, I'm going to go back and have 50 different slack messages to catch up on across different channels, different group threads. And what we've found is that Claude can be an incredible, incredible partner in helping to summarize information. Before we officially integrated Claude into Slack, we had a beta product that we called the anthropic times. And this was essentially if you were on vacation or you had just had a really busy week and you weren't able to keep up on all the messages that were happening, uh, in Slack. You could ask Claude to say, hey, what are the top threads that happened this week? What were any of the conclusions that came out of them? And you could really sort of hone the question to say in this giant body of information, especially at a company that's growing as quickly as anthropic, is, what did I miss? What do I most need to know? What are the key kind of takeaways? And I think there's an immense amount of potential across an application like Slack. But really, when you think more broadly, any sort of information set that's just very vast, where you need to pull out really the key pieces of information. I think LLMs have an incredible potential to really be able to do a lot of heavy lifting there.

Speaker A **00:14:43**

What we're touching on here, both of you, is this, this problem around unstructured data, which I think is key to the potential of enterprise AI. And Slack is obviously a great way to think about that problem. Parker, as you're thinking about that at Slack and at Salesforce, what is the potential for everyone in this room when it comes to the ways that we might

be able to harness unstructured data and how we work?

Speaker C **00:15:05**

Yeah, well, I think it's not about only accessing the unstructured data. That's what we have today. So in Slack right now, you can access all that unstructured data, but it's when the unstructured data and the structured data comes together and the LLM can use both together. So all of Slack's data will be available in data cloud. Some of that's available now, and we're building the architecture. So it's all available in data cloud. All the structured information is in data cloud. Data cloud will also federate out to your snowflakes and all the other structured databases. And we're building integrations to unstructured information. Email, your calendars, all, you know, g drive, you know, Onedrive, whatever, and then you want to. To understand something. So an example might be, how did, uh, so and so close that deal yesterday? Because that was a huge deal, and I would love to close a deal like that myself. Now, you might say, well, it's going to look at the structure information and say, well, they made these many calls. There was a. This discount. What were all the structured ways that that deal was closed. But then there's a whole team working around the deal. They're talking to the customer, they talk to each other. How do you bring all that together and bring a story to that, uh, question that says, well, this is actually how the deal was closed. It was because of that discount. It was because of the product fit. It was also because they had an objector in the account that they overcame because they talked to this other person and all those human interactions. So that's the power of the future. And we couldn't really tap into unstructured information like we can today without this new technology that I don't understand nearly as well as Daniela, but that really unlocked it. Like, wow, it's not just a, uh, keyword search under all this unstructured data, it's actually getting intelligence from it.

Speaker A **00:17:08**

Do you have anything to add on that, how you're thinking about unstructured data at anthropic today?

Speaker B **00:17:12**

I really think the more both the technology industry, but really just sort, uh, of the enterprise industry as a whole has grown in the digital age. There is so much data across these different companies platforms, ways of communicating, and if you talk to any executive of a Fortune 500 company, they'll say, it's hard to find things. I mean, that sounds so obvious. But the ability for these LLMs to essentially find a needle in a haystack, which is the name of an actual evaluation we do with Quad, we say, hey, if there's 500,000 words, can you go pick out the right word? And Claude's performance at that is like 99.6.

Speaker A **00:17:58**

What does that mean, the right word? Can you give me an example?

Speaker B **00:18:00**

For example, if you say, hey, we're going to give you a whole novel or something, but there's going to be a word that doesn't make sense in the middle of it. Can you find that word? You would think it would take a human, maybe depending on the length of, you know, a book, maybe a few days, to sort of comb through that, and an incredible amount of focus and attention to be able to say, why is the word hippopotamus? You know, in a story about, I don't know, in Virginia Woolf somewhere, Claude can do that in about 10 seconds. Can look through a book length and say, that word doesn't belong there. That's kind of a silly example, but if you imagine for business use cases, the incredible potential of a tool that is able to look across a vast amount of information and pull out something like that, obviously a book is structured, but imagine you have unstructured data, and you need to find a key piece of information across it. You're a doctor, and you say, hey, uh, I have 50 years of medical records about this patient. I remember there was, uh, something that was happening with them, um, in the 1990s, but I don't remember what it was. You imagine the power of this tool to be able to go and look across that type of information base and pull something out there. And additionally, something we're increasingly starting to see is that Claude can actually be transformative, not just at the individual task level, but at taking unstructured data and actually migrating it to be structured. There's a telecommunications company that we work with who is attempting to migrate a, uh, legacy code base into a modern code base. And the process of doing that would normally take thousands of developers many years. But with the aid of Claude, it's possible to reduce that timeframe to something like an order of months with a much smaller team.

Speaker A **00:19:50**

One of the key themes that we're covering at time today, and one of the reasons why I was so excited about this conversation is this year, seems to be defined by the interaction between these large technology incumbents and the startup labs and research institutions. And we're really seeing a flowering of partnership and collaboration this year. And so I would love to hear from both of you, starting with you, Parker. Are there lessons that you've learned, uh, from the. The partnership with Anthropoc, that you think are relevant to everyone in our audience, but also the industry as a whole?

Speaker C **00:20:22**

Well, you know, I think we share where you open is that value of trust. And we were just talking backstage a little bit about constitutional, uh, AI. And I still have more questions of, how does it really work? What does that mean? How is it in that LLM? We're building trust outside the llmdeh, so we're building a trust layer. And so how can that trust be even better where, you know, we are not? Well, actually, we are creating some LLMs. We have a

research group, but not on the scale of an anthropic m not on the scale of other providers that are building these massive LLMs. So we want to partner. And so what's the IP that's happening inside of anthropic that's getting better and better and how can we take more advantage of it? Obviously, the anthropic angle, one of the angles is around trust. And you look around this show, everyone's so excited, but they're also saying, can I trust it? Am I ready? And so that's where we have a huge opportunity to partner together and to learn from each other.

Speaker A **00:21:30**

Daniela, what lessons have you learned?

Speaker B **00:21:32**

Well, number one, we are so grateful to be able to partner with Salesforce because while anthropic is an expert and a leader in LLM development, Salesforce knows their customers incredibly well. And our ability to partner with them and help build a highly trusted product that is validated in the real world is so dependent on that knowledge and understanding of those business use cases. I also really agree with what Parker said just around the trustworthy, uh, alignment of how we both build the technology and then deploy it. I think without truly understanding what it is an individual business needs, it's very hard to build a bespoke solution for them. And so I think that combination of trustworthy technology that's developed ethically from day one with a deep ethical alignment to how we do business, I felt an incredible amount of synergy between the two companies.

Speaker A **00:22:28**

Uh, last year, time launched the time 100 AI. We did that because we thought this is really a crucial moment in our evolution as humans and in this technological moment. And Parker, you've been at the seat of a lot of change in this industry. I'm curious how you think about both from an enterprise perspective, but also just globally, from a technological perspective. Where are we today? What led us to the, and where do you think we're going in the future?

Speaker C **00:22:56**

Well, I mean, you saw in our keynote, uh, what Salesforce has done. Our journey, uh, overlaid on the evolution of AI. And AI has been around for I don't know how many decades, 50 years. And it's been around a long, long time. 2014, we really launched into it and it was all about, uh, predictive AI and machine learning. And we bought a bunch of companies and baked it into our technology. And it's still huge value in that. Everyone's talking about generative AI and it's phenomenal. But don't forget about good old machine learning. It's actually, and then m the generative AI could actually take advantage of that.

The workflows can take advantage of that. But then we hit that tipping point and some of us, not me, saw it way before, uh, you know, someone might say 2017, maybe it was 2018, maybe it was 2019. For me, it was like a year and a half ago or two years ago when, uh, OpenAI came out, it was public with like, wow, they hit that tipping point of the scale of the model and human feedback and I think we all were surprised. So we're trying to keep up with it right now. I would say that, uh, every day there is a new article, there's a new thing we have to read. I think it's the fastest technology has moved for me in a very long time. I agree with everyone, all the pundits out there. This is probably the biggest thing to happen in technology possibly ever. Uh, certainly, you know, Salesforce started on the back of the Internet and that invention of uh, this thing called the Internet, uh, that was phenomenal. Then the mobile phone came out, you know, the iPhone, and then social media came out, you know, for good and bad. Um, this is bigger than all of that. That being said, when it came out, we all scrambled. We were like, oh boy, Microsoft's going to win or my competitor is going to start doing AI before I do. They're going to transform their company. Some people are like, I'm going to fire all the people in my call center tomorrow. So everybody went really fast and that's why, uh, we talk about don't diy your AI. We all did. Even Salesforce, some people won't say that, but we did. We were like, everybody started building it. Every cloud build it, and our it department build it. Everybody go, uh, into AI because we were all anxious. It was like fear of missing out. I think many of you probably did that as well. And everybody went and we spent a ton of money as an industry, every company spent a ton of money and that benefited a few companies quite well. But in the enterprise it hasn't happened nearly as fast as everybody thought. And I think a lot of that is probably around trust is around accuracy. Getting it to be accurate and not hallucinate and be ethical, that is hard. And companies and brands, it's easy to trust it if you're an individual. And Salesforce started influenced by the consumer Internet and again, we're seeing the consumer or Internet, you know, when I can go to a chatbot myself and have it do things and teach me things and write shakespeare for me, whatever, that's really fun. But then taking it and applying it to the enterprise is what we've done for 25 years. And I think we finally, in the enterprise, I would call it hitting the tipping point. At this show with agent force, we have something that really works well and it's very salesforce. It's easy to use clicks, not code. But if you want to go deeper m and unlock it, you can, because it's an open platform, but you can be successful. You can go to Moscone west, first floor and learn about it. And so that's what we're so excited about. But it's just the beginning. To your point? It's just the beginning.

Speaker A **00:26:39**

Daniela Parker describes the scramble 18 months, two years ago where everyone went running towards this goal. You were one of the people who anticipated this happening. Was there a moment for you when you thought, aha, uh, here's the potential?

Speaker B **00:26:53**

I would say, rather than a particular moment, it has felt, probably unsurprisingly, at a place like anthropic. It's been more a set of moments. I think with every new model that we develop and deploy, we think there's no way the model could be that much better. Right? We've kind of picked so much of the low hanging fruit. It can write code better than the majority of developers in the world, write text better than a lot of writers. It can, um, answer questions at sub human speed. It can read faster than any person. And yet every time we continue to advance the state of the art, we find out that there's more that the models can do than we thought was possible. Sometimes we don't even know in advance what that's going to look like. Right. We'll say, wow, look, it's kind of able to do this next set of things that we weren't necessarily expecting. I also think building a little bit on what Parker talked about with just enterprise adoption and sort of the speed there, I think something that has been kind of interesting to see is because the technology is developing and advancing so quickly, it actually does take time for a business to get their head around it and then be able to deploy it, particularly at a large scale. When you also add in this question of trust, right, do I need to feel like this model is better than a human versus at human level to be able to deploy it? That can also really just sort of restrict the set of use cases that a business might initially start with. And so my expectation is really, you know, 20 242-025-2026 is all going to be about how businesses, particularly very large scale businesses, become comfortable with the technology, deploy it more broadly, understand its limitations, well, but it does just take time to really be able to take something from a research innovation three months before to something in full scale production across thousands of people around the world.

Speaker A **00:28:55**

Um, Daniella Parker, thank you for this incredible conversation. Thank you all for joining us. While you're here, grab a copy of the Time 100 AI magazine. You'll find it all over the place. Thank you so much.

Speaker B **00:29:06**

Thank you.

Speaker C **00:29:09**

Your agent.