

**SDS PODCAST
EPISODE 887:
MULTI-AGENT
TEAMS, QUANTUM
COMPUTING AND
THE FUTURE OF
WORK, WITH DELL'S
GLOBAL CTO
JOHN ROESE**



Jon Krohn:	00:00	This is episode number 887 with John Roese, global CTO and chief AI officer at Dell. Today's episode is brought to you by the Dell AI Factory with NVIDIA and by Adverity, the conversational analytics platform.
	00:21	Welcome to the SuperDataScience podcast, the most listened to podcast in the data science industry. Each week we bring you fun and inspiring people and ideas, exploring the cutting edge of machine learning, AI, and related technologies that are transforming our world for the better. I'm your host, Jon Krohn. Thanks for joining me today. And now let's make the complex simple.
	00:54	Welcome back to the SuperDataScience podcast. We've got an absolutely insane guest on the show today. John Roese is global CTO and chief AI officer at Dell Technologies, the giant Texas-based corporation with over 100,000 employees and \$88 billion of revenue in 2024. What a great guest to have on the show. John's responsible for Dell's future-looking technology strategy at accelerating AI adoption for Dell and its customers. With an unreal career stretching back several decades, John was previously global CTO at EMC, global CTO at Nortel and CTO at Broadcom, amongst many other top roles at world leading tech companies, board memberships, and deep involvement with the private equity and venture capital ecosystems. He holds a degree in electrical and computer engineering from the University of New Hampshire.
	01:44	Despite John being such a deep technical expert, today's episode stays relatively high level and so should be of great value to any listener. In today's episode, John details how Dell narrowed 800 generative AI ideas down to eight high impact projects, he tells us about proof of concept prison and his strategy for escaping it, he talks about where multi-agent teams will make the biggest impact in enterprises. First, the unexpected way AI is

creating more construction jobs than any other sector, as well as new careers that will emerge in the coming years because of AI and how quantum computing and AI advances are entangled in a way that will dramatically change the future. All right, you ready for this invaluable episode? Let's go.

- | | | |
|--------------|-------|--|
| | 02:35 | John, welcome to the SuperDataScience podcast. It's an honor to have you here. Where are you calling in from? |
| John Roesse: | 02:40 | I am up in the mountains of New Hampshire at my house here. I just flew back from Austin yesterday evening, so I'm here for a day or two before I head off to my next trip. |
| Jon Krohn: | 02:52 | Dell headquarters, I imagine, in Austin in the Round Rock area there? |
| John Roesse: | 02:56 | Yes, I have been there. I'm there quite a bit, but that and Silicon Valley. I've been back and forth to California quite a bit because we're doing a lot of work around trying to make agents work and a few other things that need the industry to work together. So it's a day in the life of a CTO in tech. |
| Jon Krohn: | 03:11 | Yeah, and so we will be talking about agents a fair bit in this episode. I've got some questions for you on that. For people who are watching the YouTube version of this, they get to see... They're actually inside of John Roesse's personal dojo, which is really cool. There's swords. Are those swords? |
| John Roesse: | 03:30 | There's a few swords, a bunch of shinai, which are kind of bamboo swords for kendo and a few other things. So yes, yes, I've done martial arts my whole life and so it's nice to be able to do that in a orderly way in your house. |
| Jon Krohn: | 03:44 | Yeah, very cool. So beyond your martial arts skills, you are also the global CTO and chief AI officer at Dell. You're |



responsible for establishing the company's future-looking tech strategy and accelerating AI adoption for Dell and its customers. In a recent Fortune article you said that ROI, return on investment, is the first and most important question before funding an AI project. Do you want to talk a bit more about why that is so paramount?

John Roesse: 04:16 Yeah, in the early days of GenAI, let's say two and a half years ago, we got very excited about everything that you could do with it and the things were kind of abstract. They were things without context. "I can access the entire internet and ask any question I want," and all of that's great, but at the end of the day, if you're in a business, the things you do probably ought to be connected to the desired outcome of your business, which usually, if you're a commercial entity, it's about profit, revenue, margin, cost reduction. Things of that nature are kind of important to you. And so what we've learned is while there's a lot of enthusiasm about coming up with many theoretical uses of the technology, technology is only useful if it actually does something that has meaning to the entity you belong to which, in the case of Dell, we very much care about the commercial success of the company. If you're a university, I was just talking to a bunch of university CIOs this morning, you care about educational outcomes.

05:13 So at the core of every technology, AI or not, there's got to be a purpose of doing it. And so what I said in that article is, look, it's great to understand the technology, it's great to see the art of the possible, but at the end of the day, the decisions you have to make about what you actually do. Does the rubber meets the road? Do you apply resources, should be that the technology is actually being applied in service of an outcome. And that outcome is usually very much correlated to a process that you could improve that will make your business better in a measurable way. If you have a sales force doing AI that

makes your sales force spend more time with customers, which is something we did, is a very good idea. Doing AI that makes your sales force slightly happier and more engaged is meaningless unless you can measure it. So this connection between material ROI and your AI activities is actually essential if you really want your AI strategy to be meaningful.

- Jon Krohn: 06:10 Yeah, we dug up in our research, our researcher, Serg Masis, he gets really deep into some things that you've said or written in the past. He said that there was an instance where... This is probably the same time period, but I'm going to put some quantities to this. You discussed, when GenAI first started to become really powerful two or three years ago, you received 800 generative AI ideas from Dell employees and you narrowed it down to just eight. So you took 1% of those ideas. Do you want to fill us more in on that process?
- John Roesse: 06:42 Yeah. There's a lot to that story. So here's what happened. So GenAI occurs and really the ChatGPT moment. You have this new tool that honestly, I had been working with large language models before that and I knew what they could do, but when that came out... And my head of research sent it to me as it was released. Said, "You got to look at this thing," before it was even in the mainstream. And I'm like, "This is really interesting. This is better than I've ever seen with Roberta and Bert and earlier tools." And so it came out and then I work in a company that there's a guy whose name's on the building who's very engaged and excited about technology and he sent a note to the whole company and said, "This is important," which is absolutely true. And then very quickly about 800 ideas showed up about, "This is all the stuff we could do with it."
- 07:29 And I have kind of a bit of a running joke that... I'll apologize for non-technical people, they won't get this

maybe as much as US geeks. When we went and looked at those ideas, what we concluded is a bunch of people got together in groups, maybe individually, but probably in groups to ideate about what you could do with this, and the only real qualification to be in that meeting was that they probably all saw at least one episode of Star Trek because the ideas were interesting, but they didn't align to the actual technology in most cases. "I want to build the holodeck," "I want an AI that will replace salespeople," and there's nothing wrong with that. It created the art of the possible. That was the unlocking of AI that people started to realize this could be meaningful. But if you start with that, if you have 800 projects that are every idea you could imagine, completely unvetted, not grounded in reality, where do you go?

08:20 And so the journey we went on wasn't... Initially we tried to take 800 and find the ones that would matter, and we concluded you couldn't. It was too hard because you just didn't have context. And so we actually ended up flipping the model. We didn't throw away the 800, but we asked a different question. We said, "Where should we apply this?" Not, "Where could we apply it?" And that flipped the model to go to that ROI discussion we just had, which we said, "Well, why are we doing this in the first place? We're doing this to make Dell a more successful company, and how do we measure that? We measure that in profit and revenue and cost and regulatory risk. Okay, let's focus on those things." Then we said, "Well, where should we target?" And we picked these core four areas of supply chain, sales, services and engineering.

09:00 And then we said, "Okay, within that, what is it about those that we could make better, like make our salespeople more productive by freeing up time that they spend preparing or make our engineers code better?" And that led us to connect the two dots because what we probably found in most cases was there were ideas in

there about how we could improve the sellers content preparation phase, which is really the biggest impact we could have, or where we could focus in engineering. Could we do QA or product management or core coding? It led us back to core coding. Every company I talked to, every customer I talked to has exactly the same scenario. They have this abundance of ideas and the real challenge is, "Okay, where do you figure out where to start?," 'cause you can't do 800 of these things.

09:45 If you'd done 800, you would still be debating 800 ideas and have nothing in the production. Today we have things in production, they impact our business in a positive way. We got over the finish line. But yeah, it was a fun journey and I will tell you, whether the number's 800 or 500 or 1,000, every single customer I talked to went on that journey and is probably still kind of stuck in the process of trying to figure out how do you extract or find the place to start? Sounds really simple, but with infinite surface area, finding the actual place to begin when every idea is probably pretty good, is incredibly hard. But you can't do 800 concurrent AI projects. It's just not possible for even the biggest companies in the world.

Jon Krohn: 10:23 And to borrow some terminology that you've used previously, this is escaping the proof-of-concept prison that everyone's stuck in, right? So do you want to tell us more about POC prison and maybe what makes a company ready to transition from AI experimentation to scale production?

John Roesse: 10:40 Yeah, it's funny. I am a big fan of experimentation with technology, broadly, inside of companies. In fact, our AI journey started about eight or nine years ago. I actually started that process and me and the former CTO of VMware went to Michael and Jeff and Pat Gelsinger and a bunch of people and said, "This thing's kind of

important." And we made a decision, and this was way before GenAI, that we didn't know what was important about it, but we actually gave permission to the entire company to start experimenting. We didn't do any top-down, we just bottom-up. Said, "This is important. If you're a business unit, you should think about this. If you're building a product, you should think about it. If you're developing platforms, you should consider this," and generally way before ChatGPT happened, on a typical year, we had anywhere between 500 and 1,000 AI projects going on. 80% of them absolutely zero impact to the business, but we didn't have a problem with that. They weren't occupying that much time.

11:35 But what was happening is people were getting comfortable with the technology, they were starting to learn about it, and by the time we got to ChatGPT, we weren't starting flat. We had people that had kicked the tires and people had kind of understood it. It accelerated dramatically past that. So if you haven't done that and you're starting right at ChatGPT, it's still important to do experimentation. But there is a difference between an experiment and production, and we've created a bright line between those. Production is when you choose to actually put this into production at scale, that you're putting significant resources in it and you're actually betting the company on it. You're choosing this will be a foundational piece of your enterprise going forward.

12:14 And so we have this process that we allow a lot of experimentation and we actually encourage it. But the way that you tip over into production at Dell, and we think this is something other people should do, is there's just a series of things that have to be true. The very first one is, "Do you have an ROI?" It can be a great idea, but if it can produce no material impact to the business, I'm not putting it into production. I'm not interested in that. The second one, which was an interesting learning is, "Does

this AI project actually build on the way we want to run the business in the future?" Or which is very common, "Is it a big blanket that we're throwing over a giant mess to hide bad processes, bad structures?" And we call that modern Dell. If it is not about the modern way we want to do it, we will not do the AI project.

12:59 So even a great tool that somebody can prove to me will save us a lot of money, but it'll do it by hiding a bunch of structurally unsound practices, we will not implement AI there because that's just a crutch. It's not going to be sustainable. And then beyond that, then you have discussions around, "Is it technically viable, does it meet our security and regulatory compliance obligations?," and then it goes into production. But that front end is so important because it says, you escape from POC prison not by finding cool technology, you have plenty of that. You escape from POC prison by figuring out which cool technology projects actually are going to create value to the company in a material way at a priority level and are not taking you backwards or hiding the sins of the past. They're actually about the future. It is all about the future.

13:42 You don't want to apply it to the past. You want to apply it to the things going forward. And so if you get those two right within.... If you have a hundred experiments going on, I bet you can go through them and find the top three that have the highest ROI and the biggest alignment to your future strategy and objectives. Those are the ones that move and then you move them into production. And once they're in production, that's where you scale them, that's where the investments come, that's where you measure them. And in our experience, if you pick the right ones, they actually produce a lot of ROI and they get the flywheel going, which is pretty exciting.



- Jon Krohn: 14:10 That's exactly the thing that I wanted to talk about next, was the ROI flywheel. So you've talked previously about the importance of choosing high impact projects that trigger this flywheel of AI success. Tell us about the flywheel and how to get one going or maybe the kinds of missteps that prevent one from happening.
- John Roesse: 14:27 Yeah, so the flywheel concept came out of... Some people have different ways of thinking. I'm a visual thinker, a pattern guy. I'm very good at connecting dots. And as we started to do this, we could see that if you are able to really understand what matters, what's going to move the needle for your company and if you are able to connect that to the technology that will do that, you're not just doing a one-off. What you're doing is you're creating effectively a flywheel, because if you put the right projects into that process of getting AI into production, the net effect, the thing that the flywheel will produce, if the input is a fantastically properly vetted, high priority, high ROI idea, and the flywheel works about getting it into production, the output is ROI. It's actually cost savings, profitability, revenue, risk reduction, things that matter to you.
- 15:24 But it turns out that because it's a flywheel and the reason it's a flywheel is that your first project might be a novel set of technology to play, but it turns out there aren't that many ways to do AI stuff. There's foundational technology that we can talk about. And once you get the first one going, the second one isn't another snowflake. If you do it right, it actually uses much of the same technology as the first one. And so the cost to do it is lower. The speed to do it is faster, and you can imagine that you get this thing going and it starts just shedding just a huge amount of impact to your business. That's if you do it right. Your question is what about if you do it wrong? And so the biggest mistake people make is right

now, I will guess that in most enterprises, their flywheel is not even moving, it's not producing any ROI.

- 16:06 Is your board happy with the ROI impact of your AI efforts? If the answer is no, then your flywheel is not moving. If your answer is yes, then it is working. And if it's not moving, the wrong way to start it is to throw a really cool project into it that produces no ROI. And I don't want to pick on specific examples, but I will be somewhat specific. There are places in companies where there are good things to do with AI, but the effect of it will at best be happiness, goodwill. And while those are good things to have in general, those are not the things that get the flywheel moving. If it costs money to create a slightly happier workforce or a slightly more comfortable work environment, while those are good things to do later, you won't be able to afford to do them if you don't get some ROI moving.
- 16:52 And so we tended to stay away from those. We went right after the areas where money and revenue and profitability and cost lived. Sales, services, supply chain, engineering, those are the core. The ones we didn't go after were more of the G&A functions where honestly, even if we had the best in class in some of those functions, nobody's going to pay us for that and we're not going to make any money and we're not going to really reduce cost dramatically. And while we are now in the position because the flywheel is going, in fact, we can now go after them because once you have something moving with a lot of inertia, throwing in an occasional one that doesn't produce a lot of ROI but creates a lot of goodwill, you can afford to do. But trying to start a flywheel with something that actually doesn't provide any fuel for the next project is a bad idea.
- 17:37 So, that visual has been really helpful to us to explain to people why their particular project, which looks good on face value, is the wrong project to get the flywheel moving



and it needs to come later after we've got the flywheel moving to produce the thing that the board, Michael, and everybody wants us to produce, which is material impact to the business.

- Jon Krohn: 17:57 This episode of SuperDataScience is brought to you by the Dell AI Factory with NVIDIA, two trusted technology leaders united to deliver a comprehensive and secure AI solution. Dell Technologies and NVIDIA can help you leverage AI to drive innovation and achieve your business goals. The Dell AI Factory with NVIDIA is the industry's first and only end-to-end enterprise AI solution, designed to speed AI adoption by delivering integrated Dell and NVIDIA capabilities to accelerate your AI-powered use cases, integrate your data and workflows, and enable you to design your own AI journey for repeatable, scalable outcomes. Learn more at www.Dell.com/superdatascience. That's Dell.com/superdatascience.
- 18:45 It's uncanny how the next topic that you've gone into three times in a row now is exactly the topic that I had lined up. Although I think in this case you've actually covered all the questions I have, but my very next question... It's like you're sitting reading my notes with me and actually nobody has seen the exact ordering that I have them in except me. So that's wild. The very next thing was I was going to say you emphasize that Dell focuses its AI efforts on four strategic pillars, which you mentioned there, engineering, supply chain, services and sales to get that ROI flywheel moving, which makes a huge amount of sense to me, but it is so easy to see how it could be overlooked, how you could end up prioritizing projects that are really cool, that make some employees lives easier in some way, but if that's the first project, if it doesn't deliver ROI, then you might not get authorization to do further AI projects.

- John Roesse: 19:39 Or you might not have any budget. Dirty little secret, IT budgets aren't growing dramatically and if you want to do this stuff, you have to create value before you do the kind of things you want to do. But yeah, absolutely, at Dell we picked those four areas because that's really where the bulk of the things that we can move the needle on exist. It was interesting because early on when you do that, you create somewhat a culture of abundance of AI and a culture of starvation of AI in certain places and even the order in which you do them, like for instance, our sales force was the last one we turned on, and it wasn't because we didn't want to, it was just that it turns out, we'll probably talk about it later, data matters in this thing. And the data underneath the things we wanted to do for the sales force wasn't in the best shape, so we had to fix the data stuff and work through some issues and then ultimately we were able to stand up for 20,000 sellers, a thing called Dell Sales Chat, that is now profoundly changing the way they work and improving their effectiveness on levels we didn't even anticipate. It's better than we thought it would be at our scale.
- 20:44 Picking those four things in any company, it will be different, but really what they are is the thing that makes you special. I used to tell people, the one question you have to answer before you start any of the technical dialogue is, "Do you know what makes you special as an organization? What is your core source of differentiation? What is it that if you improved in some way you would win?" And not to pick on, I love my HR friends, but having the best HR organization in the world is not the core source of differentiation for Dell. I want to have one of those. If I have that, but I have a lousy product and a bad sales force and a weak supply chain, I'm out of business. So there's definitely a tiering here.
- 21:24 And so going through that exercise of just saying, "What is it that makes you special?" And by the way, it's

different. Like I said, I was just talking to a bunch of education CIOs. They have a very different center of the universe. Yes, they want an efficient university, but the primary goals are things like attracting the best faculty, producing the best graduates that have the best attach rates into industry and have the best careers. Those are their strategic priorities. By the way. You should use that as the litmus test of what you do for AI. If you have five choices and two of them move that needle, go do those first. But you start with this very non-technical discussion of what is it about your organization that differentiates it? And if AI is a tool that can make your organization better, connecting those knots by having an understanding of differentiation and a tool that is aligned to that differentiation is critical.

22:17 And we went through that exercise and like I said, I have some [inaudible] on because you can imagine certain groups are, "We're not happy with that," because if you had 800 projects, lots of people thinking about it. But we have a culture that says, "Look, we're all here to win." In fact, I'll give you a story. Every quarter, I do a three-minute thing on our quarterly review broadcast about the state of AI because I want to keep everybody on the journey. We have a very bought-in population, people at Dell really care about this stuff. The previous quarters, it was all kind of status update. "This is what we're doing, this is the new stuff." The last one we did about a month ago, right after the fourth quarter, because we had just finished Dell sales chat, we had put the fourth one into production. We said there were four and we have all four groups now running and doing stuff and having an impact.

23:02 My message wasn't a status update, it was a thank you and I thanked every single person in the company, and there were three groups. There were the people that actually built and implemented these four things. There

were the users of them that bought in and had the impact. And the third group I thanked was everybody else for working with us to allow us to focus and get these done because if we had tried to do 800, we would still be an inch deep and a mile wide and have made no progress. And so that's tough because some people aren't going to get to do the project they want and some groups are going to go second and if you want to do this right, that's the only way you can actually get it to move fast because at the end of the day, the only AI project that's an absolute failure is the one that never goes into production. You never get it to do anything. It's still a concept. That's the POC prison thing.

23:52 Being in POC prison is not a good thing. It means you haven't escaped into production and if you haven't escaped into production, you haven't actually created any value and no one likes stuff that doesn't produce value.

Jon Krohn: 24:05 POC prison doesn't sound to me like a good thing. Those were great anecdotes, super helpful for any enterprise organization that's trying to make the most of AI. I love that. In our discussion of AI, the example so far have been around generative AI in this conversation. Let's talk about the natural next step that has emerged after generative AI, which is agentic systems because as generative AI has become powerful enough, as LLMs have become reliable enough, we've started to be able to rely on them more and more on their own. Do you have, John, your own definition of what an agent is?

John Roesse: 24:49 Yeah, I'm going to give you a bigger picture view and then I'll define an agent. So AI attached to the enterprise, applying AI to the enterprise, actually has two different parts to it, of which only one we've done so far. Agents are the second one. And the reason for that is the source of differentiation of an enterprise. A lot of us in the industry have said this over the last couple of years, even though

people weren't necessarily paying attention, but there were two parts that make an enterprise an enterprise, the real core source of differentiation. The first is your proprietary data. You know things other people don't know. That's actually very powerful. That's why you don't share your proprietary data with people. My customer list is very valuable. My source code is very valuable. And those are a sustainable source of differentiation. Even if the people change, the brand changes, the world changes, having proprietary data is very, very important.

25:40 The second source of differentiation is the unique skills in your organization, that you have people that can do things better than other people. At Dell, we have the best thermal and cooling people in the world, the best client developers in the world, the best storage software developers in the world. And the result of that is that translates into better products, interesting innovation, patents. And so if those are the two sources of differentiation, and the journey we're on is to apply AI to an enterprise and those are the two things that matter, it's interesting because for the first couple of years at GenAI, we actually went after the first one. A chatbot, a rag system, all of these things are just tools that allow us to unlock and create value from our proprietary data. What is a rag-based chatbot?

26:26 It is a tool that takes proprietary data and makes it generative. You could take all of your service information and if I gave it all to you in raw format, it would be of no value. If I embed it into a vector database and present it to you through a generative interface, you can ask and answer any question on anything I know, anywhere. That is incredibly powerful, and we have been doing that now for about a year at scale in the industry and it's transforming everything. We're getting huge value out of this. In fact, almost all of our projects that are in production are just that. They're a generative capability to

unlock our proprietary data in novel ways that just changes the curve in terms of productivity. That's great.

27:08 Agents are not that. Agents go after the second one. They are about the digitization of a skill. They're about saying, "I'm not just interested in unlocking the data. I'm interested in distributing the work. I actually want an AI that doesn't even require me to do a task, that it can actually operate autonomously. It can operate without human intervention. In fact, I'm not even going to tell it how to do the job. I'm just going to give it an objective and let it go, and I'm doing this aligned to the skills that I need it to do." So for instance, when we think about agents in the enterprise, now there's two views of this in the current thinking. One thinking out there is that agents will be replacements for multi-dimensional humans that can do everything. That's AGI and ASI. We're a long ways away from that. The reality of agents is that they are actually the digitization of more narrow skills.

27:59 I use the self-driving car example. I do not have a self-driving car today that can drive anywhere in any situation and navigate it successfully. What we do have is self-driving cars. They've been in San Francisco and other places, where if you geo fence it, if you narrow the scope, we see this in the trains and airports, there's no driver on them because it has one job. It moves from terminal to terminal without a human intervention. Well, that's what's going on with agents. The first generation of agents are saying, "Could I take a task, a skill and could I move it into AI not as a tool that a person uses, but as a manifestation of that skill autonomously, that I can just tell it to do something. I can give it an objective and it's smart enough to figure out how to reason through that objective. It has access to a set of data and it can deliver an outcome equivalent or better than what a human would've done for that particular skill."

- 28:51 And yeah, there might actually be humans doing those specific jobs that might not do them anymore because agents can absorb them. But what you don't have is a fully well-rounded entity that is the equivalent of a full human being that can do lots of different things. Think about in your life, how many different things can you do? Well today the manifestation of agents can probably pick off a few of those, but what they can't do is pick off all of them and create a completely equivalent of your whole well-rounded human being, including your ethics, your morality. That's a really hard problem. That's AGI and ASI, a different journey. And so bottom line is you take these two technologies, first gen GenAI, which is what we call reactive AI, that a human is in the loop and the human asks the AI to do something and it gives it an immediate response, but ultimately the human is the doer of the work and these are tools around the human.
- 29:41 And then you move over to this kind of second generation of agentic AI, which are complementary, and now you have a situation where the human is on the loop, they're the supervisor and all they're doing is creating objectives and delegating work. And now the AI independently is able to take that task, figure it out, run with it, and even run with it in perpetuity that it may never go back to the human being because it's been delegated below the machine line. The reason it's so important to distinguish these is that one, they aren't even the same technology. Well, this one, the center of the universe is a large language model with some data around it. It's a very static data set. An agentic environment has large language models, but they're used for part of the equation. They act as somewhat of its brain, but it has a body, it has a knowledge graph where it creates its own representation of data that it represents what it's learned and its memories and its evolution of skills.

30:31 It has interfaces around it that allow it to reach out into the real world, something called tool use and function serving, where it can actually go and activate a tool and interact with the world and perceive things. Very different technical architecture and quite frankly, appropriately so because it's solving a different problem. Now fast-forward into the future of an enterprise. Well, yep, still got proprietary data and still got unique skills, except now I have a path to digitize both of them. And that's the thing that's going to profoundly change most enterprises.

Jon Krohn: 30:59 Very nicely said. That was an amazing explanation of agentic systems and how they evolved out of the reactive systems as you described them. Something that you haven't touched on yet, though you probably have anticipated as my next topic already, is that so far everything you've been describing has been agents acting on their own really, or we haven't talked about them working in concert. So let's talk about teams of AI agents. What kind of governance and orchestration frameworks do you foresee emerging to manage ensembles of agents responsibly at scale?

John Roesse: 31:36 Yeah, it's funny, we built our first autonomous agents over a year ago now. We built a two-agent system to write research reports way before this was cool and probably less than a year ago, I showed those agents to the leadership team and that kind of got us all thinking about it. And now we built lots and lots of agents that we've been able to build them that run CNC machines and do all kinds of things, but they're not necessarily fully in production. But we've been working with this for a long time and what we learned is the real value of an agent is not an agent in isolation. It's just like the real value of a person is not an individual. It's a collective. We do much better when we have multiple people working together on complex tasks. Turns out agents follow the same pattern.

- 32:21 Now we proved you could do that. In fact, every one of our agentic systems, from the day we started, had at least two agents and a human being involved. And eventually we had one system that's 1600 agents working on a problem at one point because they can flex in and out. They actually have ability to grow and they have this concept of being able to hire additional agents. If you need a skill, just hire another agent. You tell them they can do that, they do it. The bottom line though is that as we went on that journey, one of the things that we realized... As you are in the front end of technology, you realize what gaps are. And the gap we have right now, which is an industry level gap, is we have no real framework or agreement on the interworking between agents.
- 32:59 We've agreed on the communication protocol. It's JSON. It's basically this idea of it's clear text over a digital interface in a messaging format. And that's actually really cool because you can actually watch agents interact with each other in human language and you can be a participant, which is pretty powerful. But then whole bunch of other problems show up, like, "How do I authenticate an agent? How do I authorize it? How do I share knowledge between agents that aren't working for the same company? How do I do a job prompt, which is how you start these things, but do it in a way that a Dell agent and one of our partner's agents can actually work together? How do we talk to both of them?" That's not even clear. And so there's this long list of things that we have to work out. And the good news is, that's why I've been out in Silicon Valley recently a lot, is we're working with our technology partners and a lot of the ISVs and we're all the same opinion this needs to be solved and we're going to go solve it now.
- 33:53 We're not going to solve it probably in a standards development organization over five years and probably won't even get solved as a pure open source project. It will

become a set of industry activities that become consensus. And in fact, there's one protocol called model context protocol that Anthropic came out, which is not the solution to the total problem, but it's actually a very good way to have a model talk to data. And it does it in a way that seems well-thought-out for that particular part of the problem. There's another one, funny enough we're talking today, that literally yesterday Google announced something called Agent2Agent, which looks promising. And we know a bit about that and we think it's an interesting approach to solve maybe some of the authentication, authorization, interworking problems, but we're not there yet.

34:35 And this is the nature of these technologies that even though the vision... The vision of an agentic environment in the enterprise is not, "I have standalone agents doing tasks in isolation." It's this vision of, "I'm a human and I'm responsible for something very complex and I'm going to break that down into the functions or the jobs that I need to be done to accomplish that complex task. But because they're agents, they're going to work together as a collective to do that work for me." And by the way, if that sounds familiar, that's exactly how you build human teams. That's exactly how we have always done it, except now part of that team are a set of agents. Some of them may still be people, but because we know that that's really where the value is created, and then extend it even further. The real value of human collaboration is not collaboration in your silo. It's collaboration across your enterprise or across your ecosystem, and that requires interworking. And we don't have those standards in place. We don't have them well-defined.

35:30 Now, like everything in AI, considering the term agentic wasn't even really well understood in December. I did my end of year predictions and I predicted that agentic would

be the word of the year in 2025. Every conversation I had, I had to explain what agentic was.

- Jon Krohn: 35:46 This episode is sponsored by Adverity, an integrated data platform for connecting, managing, and using your data at scale. Imagine being able to ask your data a question, just like you would a colleague, and getting an answer instantly. No more digging through dashboards, waiting on reports, or dealing with complex BI tools. Just the insights you need - right when you need them. With Adverity's AI-powered Data Conversations, marketers will finally talk to their data in plain English. Get instant answers, make smarter decisions, collaborate more easily—and cut reporting time in half. What questions will you ask? To learn more, check out the show notes or visit www.adverity.com.
- 00:36:30 That's my next topic, John. This is weird.
- John Roesse: 00:36:34 But anyway, the interworking stuff, we're working on it and it's moving really fast. The Google announcement yesterday, good progress. We'll see if that carves off more of it. I am 100% confident that before the end of this year, we will at least have de facto approaches to build trustworthy interaction between agents in a reasonable way. It will still be level four autonomy that we will [inaudible] it. It will not be infinitely flexible. It'll not deal with all the corner cases. But the bottom line is I don't need that. I just need my ecosystem to work together in a collaborative way that I trust and then I can get huge value out of this. So it's a journey, but agents are skills.
- 00:37:10 Skills ultimately are interesting by themselves, but way more interesting when you combine them. They get even more interesting when you combine them across administrative domains and organizations. Agents are following the same path. We're just going to have to invent the way that they actually do that securely and

trustworthily, but it will move fast because there's a huge value to doing it and there's a technical appetite to go solve the problem.

- Jon Krohn: 00:37:31 There's a flywheel that can emerge.
- John Roesse: 00:37:32 I know. And it's amazing how fast you're producing a lot of ROI and there's a lot of value. People move fast in the world. And by the way, a lot of the things that we're going to do are things we've done before. We don't have to invent an entirely new way to authenticate an agent. We just have to decide which way to use a tool that we already have. Authorization, same thing. Knowledge sharing, we have knowledge graphs. There's a lot of people that are talking about using things like confidential compute and a technology I really like called partially homomorphic encryption or homomorphic encryption, which is a cool tool to be able to process data without seeing it. And these things have actually really interesting applicability to things like multi-agent ensembles. So we don't have to invent everything. We just have to take the things we've figured out that maybe could be applied somewhere else and use them the right way to achieve the goal of having a trustworthy collection of agents being able to accomplish a task.
- Jon Krohn: 00:38:21 Very cool. I like how you're touching on homomorphic encryption there. It's something that I'd love to dig into in more detail, but we might not have time in this conversation 'cause I have lots of exciting topics still to get through. The next one that I was going to talk about was how, as you just said, at the end of 2024, you said 2025 was going to be the year of agentic AI. In that same conversation, you also predicted new jobs like software composer, AI interpreter and thermal plumber.
- John Roesse: 00:38:51 Thermal plumber, yes. Exactly. Yeah. This is a fun but necessary experiment you have to do. Any of us in the

industry, anybody in a leadership position, the number one source of angst in AI in general is this general fear of displacement of humans that we are going to shift a bunch of jobs to machines. And on a very personal level, if you're a human being involved in this world right now and you're seeing things like agentic and generative AI systems and all of the things that we're talking about here, which are very exciting and very real, you contextualize it to yourself saying, "Does this impact me? Could I potentially not have a job? Will my job change? Will my company exist? Will my world get changed?" There's a cartoon I love that somebody sent me a million years ago, and it's a professor in a classroom and he asks the question, "Who in this room likes change?" And every hand goes up and then he says, "Who in this room wants to change?" And no hands go up.

00:39:48 We're kind of opposed to changing ourself. We might like change, but as long it doesn't impact us. When you start thinking about AI, you start to realize very quickly that change is inevitable, that every big technology inflection, you don't want to be the last farrier when the internal combustion engine came out. There's a really important need to understand that. The problem is it's happening really fast. And so I don't think we collectively are spending enough time in, I'll call them deeply intellectual conversations about really thinking about what the real jobs are. We can talk at high levels and say, "Oh, every technology has always created jobs." That's true. Data shows that. Probably going to happen this time. But if you're a person who has a job and you think your job's going to go away and nobody's told you what the future jobs are, that's a very awkward situation.

00:40:34 So I actually took some time last year and we started to think. I actually have a much longer list. We picked out a few and put them in that blog to say, "Okay, if you really start working with this stuff, you realize that the human's

role does change, but there's a whole bunch of new jobs that need to exist for this thing to work because of the technology inflection." And so the ones you mentioned are good examples, like imagine a world where all the software writes itself. Well, kind of a problem there because the active writing software only happens once you know what program you're trying to create, what problem you're trying to solve. It also requires judgment about there are many different ways to build a software program and it depends on how you're going to use it. Do you create microservices or monolithic software? Is it got to be cloud native or not? Is it 12 factor or not? These are decisions that an AI by themselves can't really make.

00:41:22 And so we started to say that one of the roles that will absolutely exist for a very long time is some human being has to be the composer of the software. They don't actually play the musical instrument, but they decide what this system should be, what is good. Because without that, you can't even give a prompt to an AI. It doesn't know what to do until you tell it. And if you just tell it, "Write software," that's not good enough. And if you tell it to, "Solve my sales problem," it won't know what to do. And so you're playing this role of composition of leader, of decision maker. And so actually I was just talking to some universities again about this, and I said, "You know, I need you to produce people that might have some code proficiency. Hopefully they know how to use coding assistance, but I really need them to understand good software architecture and how to build a system and what it needs to look like without necessarily having to write the code themselves," because that's the skill you're going to need.

00:42:12 Second one that we talked about, it was thermal plumbers. It sounds great, it gets people thinking, but it turns out the skill set necessary to make a GPU cluster work are this composition of skills that don't typically

intersect. To be technical, I need someone who understands computer hardware engineering and thermodynamics. Now I'm an electrical engineer with a computer engineering option. I know a lot about computer architecture. I only know something about thermodynamics because it was an optional elective that I happened to have taken. Nobody taught me about fluid dynamics to be an electrical engineer. But it turns out if you want to make a GPU cluster work, it's direct liquid cooling and you have to understand the intricacies of how thermodynamics and fluid dynamics work and you have to understand how GPUs work. And what you're really doing is managing this thermal envelope, the place where the GPU runs its best, without collapsing and without being inefficient.

00:43:08 That is a very specialized skill. But if you look at the kind of academic disciplines necessary to achieve it, they don't really usually intersect. One's a mechanical engineering problem, one's an electrical engineering problem. Well, this is a both problem. And so there aren't going to be a lot of thermal plumbers, but without them, we're not going to be able to run these clusters. And so you're already seeing that job form inside of the big clusters because it's a super important piece of this future architecture. And then the third one, which I really liked because those first two are pretty specialized, you got to be like a really good thoughtful computer science person or a really, really smart engineer. The one in the middle is the fascinating one, which is what we call an AI explainer, and it basically says, "Look, we are going to more and more produce data and insights using AIs," and that's great. We should do that, "but the way we deliver it to humanity is equally important."

00:43:57 And so we already have some examples today in things like genomics where you have this technology mining through your genome and discovering that you have

certain attributes, some of them good, some of them less good. And so what you find in most cases, if you have a marker for a whole bunch of really bad things, like let's say it comes back and you are likely to have Alzheimer's, Parkinson's and something else bad, it would be unconscionable to send an email to you. It needs a human being to empathetically explain that to you and to make sure that you understand what to do with it. And the person that's doing that is not just a technologist. They're not even the clinician. They need to understand the data set. They need to understand how the AI came to that conclusion, but they also need to empathetically explain it to you.

00:44:42 Now, that's a very specific example that's already happening, but take it into any number of other examples where you're doing a performance review. Okay, I have seen, "And we are going to build and use technology that automates that entire process." Should the performance review be an email or a portal or a text or a chatbot, or should it be the manager having a conversation with you? So you become an AI explainer, but you're explaining not just your opinion, you're explaining what the data told us in a way that a human being could understand. Even in academia, I would give an example this morning of when you have a performance issue. Let's say there's a student who is... The data is showing they're going to fail out. They're not doing well. Funny enough, we have tools that are going to emerge that will tell us how to fix that, that we can actually get them back on track.

00:45:28 Do we just send them a bunch of emails and hope they figure it out, or do we have a responsibility to put a human being right in the middle of that that can translate what the information is and what the plan is and connect that human being back into the right track? And so everywhere that you have machine generated data, sometimes it's benign and you can just deliver it and it's

great. But there are more and more places as we use this in medical, in performance management, even in social services, where the need for the interface is still humanity because we're dealing with humans, but the skill we need is not just someone who can talk to a human, it's someone who can bridge that gap. So they have to have a kind of new literacy about why the technology did what it did, what it's telling you.

00:46:12 I think that's an enormous job. If you wonder what's the call center of the future, it's that. It's a much more maybe sophisticated job, but it's one that it biases not towards the technical skills of humanity, but towards the other skills in humanity. It's the BA path where the other ones were the PhD engineering path. And so we just went through that exercise and honestly, we found a dozen of them that were very interesting and they all seemed very reasonable, and we're actually seeing them happen within our own company that these jobs are starting to form organically. I think we owe it to ourself, our population, society, because this is moving so fast, to spend quality time as we deploy these technologies, looking for what happens to humanity, what jobs emerge. When I go through this narrative with a lot of people, they get a lot more comfortable. They feel like, this isn't net-zero that I'm just going to wipe out humanity, and there will be no jobs. Yes, things will change, but there will be new jobs that are created and new jobs that are impacted.

00:47:09 By the way, the one thing I will give you as a caveat on that is, ignore everything I just said. The single biggest job creation of the AI cycle is actually construction. It's construction workers, plumbers, electricians. The amount of infrastructure that is being built and will be built to power this AI transformation is bigger than the infrastructure bill that was passed in the United States several years ago. It is a gigantic public works project, but it's run by the private sector and it's going to employ a lot



of people, and it's going to do that for a very long time. So there is definitely a job creation angle, but there is also, unfortunately, a very fast moving disruption happening. And so we're going to have to be really thoughtful about what the future jobs are and help people get there because this does change who does work and how work is done, especially as we move into things like agentic.

- Jon Krohn: 00:47:54 SuperDataScience Community turns 10! Watch your email inbox this month for exclusive offers and surprises to celebrate this milestone. Having successfully trained thousands of professionals worldwide, SDS is the perfect platform to take your AI and Data Science skills to the next level. Expect new content, behind-the-scenes access, and one-time bonuses to accelerate your career. Whether you're a longtime learner or just starting, this celebration is for you. Not on the email list yet? Start your free trial at [SuperDataScience.com](https://superdatascience.com) to be first in line when anniversary deals drop. Don't miss joining the next decade of learning and innovation.
- 00:48:39 Fantastic perspective. I love the way that you brought all of that together and this forward-thinking that you've been doing about how AI will impact jobs. Something else, and we're going to have to try to squeeze this in quickly, given the time constraints that you have, but something really fascinating that you've talked about that's very forward-looking or seems very forward-looking is the connection between AI and quantum computing. So specifically you've said that AI is the thread connecting all modern technologies and that quantum computing and GenAI are two parts of the same story. Do you want to tell us more about that?
- John Roesse: 00:49:10 Yeah, absolutely. So for those of you aren't that familiar with quantum computing, it is basically a different way to do math. It's a computer that does math in a different way. Instead of having binary where everything is a one or

a zero and you try to convert that into the application of math, it is able to simultaneously look at any particular value between one and zero at the same time in something called a qubit. The qubit is the atomic unit as opposed to the bit. The bit can only be one or zero. A qubit can be any value between one and zero. It's really not even one or zero. It's any value. And the result of that is that if you have this system that can work in qubits and each qubit can be almost any value, it allows you to, with a very limited number of qubits versus traditional systems, look at probability. Basically look at almost every permutation of a particular answer simultaneously, where conventional computers will have to look at them one at a time.

00:50:05 And so it turns out that that breaks a bunch of math. There's a bunch of things like cryptography, specifically things called asymmetric key management protocols that bank on the fact that the ability to factor prime numbers, which would require you to figure out one at a time, what the answer is, is so hard that with a big enough key, it would take forever for a computer to do this. It turns out quantum computers can look at that simultaneously and instantaneously get to an answer. Now the ones that can do that don't exist yet, but they're coming. So think of it as a quantum computer is a tool that can do math in a different way, and the kind of math it does is really good at looking at lots of things simultaneously and coming up with the best answer.

00:50:48 Well, turns out the intersection with AI is pretty interesting because if we think about things like training AI systems, what you do is look at lots of information and you try to convert it into a mathematical representation of lots of information, the entire internet, every piece of data you have. And there are definitely thoughts that if we had that type of computer, the process of training things may become significantly faster and better. Even on the

inference side, the ability for you to make a decision more quickly in something like an agent, how fast it could decide and reason across something, if it could look at everything, every possible option instantaneously, would improve dramatically. And so while there's still a lot of work to figure out exactly which algorithms, all of us do believe that given this additional capability, this new way of doing certain kinds of math... By the way, quantum computers don't do everything. They only do certain kinds of math really well.

00:51:41 There are enough early indications and theories that say this would be incredibly disruptive. And what I've said is the day viable quantum computing is available at scale is a bigger disruption than the day that ChatGPT came out, because whatever the state of the art is of AI at that moment in time will suddenly become three, four or five orders of magnitude faster and better. That's a gigantic thing that will come. It's just a question of figuring out when. Interestingly enough, the two are even more related because the path to get there has been a very slow path. But now what we're finding is the application of AI to build the quantum computers, to run the quantum computers is accelerating the quantum cycle that we're actually figuring out how to do containment theory better. We're learning how to interface with them better. We can program them easier.

00:52:24 And so there's this mutual beneficial cycle between the two of them that as quantum evolves, AI is accelerating the path to make them viable. And as quantum computers become viable, they will inevitably create a computing infrastructure that will make AI significantly better. And so I don't know what happens after that, but we're heading towards that date at some point. It's not tomorrow. It's probably not for a few more years, but it's also not decades from now. And so I think we're going to see quantum utility and then quantum supremacy. And

one of the big impacts is it's absolutely going to touch and impact the way that AI works. And so I always tell people, "Yeah, the lay person doesn't have to worry about this quite yet. We in enterprise absolutely have to and we have to pay attention to it," because imagine if you could replay November of a couple of years ago and you knew what was going to happen. You knew in advance that there was going to be this disruption in November and it was going to change everything.

00:53:21 Well, I'm telling you right now, there's going to be a disruption in the future that's going to change everything. I can't tell you the exact date, but I can tell you to prepare for it because it's important and it's going to be another one of these quantum leaps forward, no pun intended.

Jon Krohn: 00:53:33 Yeah, as another pun, I guess we could say that the future of quantum and AI are entangled.

John Ruese: 00:53:37 Are entangled.

Jon Krohn: 00:53:38 [Inaudible]

John Ruese: 00:53:40 Exactly.

Jon Krohn: 00:53:41 We probably got that from something you said. It's in my research notes for this. We need to start wrapping up unfortunately, because this has been a fascinating conversation. We could have spoken for hours and maybe someday we'll have that opportunity to do that, but for now, we need to start winding down. I always ask my guests for a book recommendation, John.

John Ruese: 00:54:01 Yeah, I've given this answer a few times. Not recently, but there's a book, I have some attachment to it. I don't know if you know Stanley McChrystal. He ran the Special Forces, very interesting guy. I know Stan pretty well and I think he's a really smart guy in the sense that he

understands some big picture things. He wrote a book called Risk, which I really like. And what it is, it's a narrative. I actually got interviewed for it and we had some interesting conversations and he talked to lots of people. It's not a tech book. It talks about military, industrial. And the whole point of it is, it helps you think through how do people handle risk and change and the disruptions that are happening around you, because fundamentally, if you can't manage risk... Everything I just talked about. Picking the right project is a risk management exercise because if you pick the wrong one, you'd go out of business. If you pick the right one, people are going to get irritated that you didn't pick their thing.

00:54:57 And so being able to work through that scenario continues to be a theme that I think people are struggling with. So it's a good book. Like I said, I have some connection to it, that he interviewed me for it, but I like Stan. I've recommended that to lots of people as a good way to take a step back from your world and look at dealing with risk in all kinds of different scenarios. And you find these patterns inside of it that help people quantify risk, understand, make it data-driven, don't make it emotional, all of these things that help you navigate risk because risk and change are kind of the same thing in many cases because change introduces risk. And if you're not willing to take risk, you won't change. And in the AI cycle, it is incredibly important that we are comfortable changing, which means we are comfortable managing and selecting the right path, which is really about managing the risk. So anyway, Stan will love that I gave him another pitch to his book, but I really do like the book and I have recommended it to lots of people.

Jon Krohn: 00:55:50 For sure. It sounds like a great recommendation. Everyone in the class raise their hands if they like

reward. Everyone in the class, raise their hands if they like risk.

John Roesse: 00:55:57 Exactly.

Jon Krohn: 00:55:59 Nice. And very, very final question is clearly you are a tremendously intellectual individual with a huge breadth of knowledge. How can people, after this episode, continue to get your thoughts, say on social media or something like that?

John Roesse: 00:56:13 Yeah, yeah. Funny enough, I have a YouTube series now that is just really... And it was really driven by the fact that this is moving so fast that conventional marketing doesn't work. I'm glad we're doing this because honestly we have to use other tools because what we're talking about today... I mean, we talked about a thing Google announced yesterday. If that went through a traditional marketing process, nothing against marketing process, but it might take a month to get that out. So I have a YouTube series. I'm on LinkedIn very heavily. We're doing these kinds of things. It's great to have these conversations. My advice to people is, I've said this to governments, I've said this to industry, engage with other people. People are talking about this. Find the channels in social media. Find the channels in other spaces. The more you just hear what people are thinking about, don't blindly follow them. Don't blindly follow me. It's just data, but be exposed to it 'cause this is moving very fast. There are a lot of thoughtful things happening. There's a lot of learnings around it.

00:57:03 The only mistake that you'll make in that journey is to be disconnected from it and flat-footed and not know anything. But there's just really good vehicles today that we just didn't have before. And I think there's a lot of people like me that are talking a lot about what we're learning and shamelessly, copy what other people do,

learn what they're accomplishing and that will help you navigate this forward. So yeah, glad to be here for that.

- Jon Krohn: 00:57:25 Yeah, fantastic. We will have a link to your YouTube work in the show notes for listeners. John, thank you so much for taking some valuable time out of your valuable schedule for us. Really appreciate it and hopefully get you on air again sometime in the future.
- John Roesse: 00:57:41 Great. Glad to be here. Anytime.
- Jon Krohn: 00:57:49 What an honor to have John Roesse on the show. In today's episode he covered the importance of ROI as the primary factor in AI project selection, focusing on areas that impact business outcomes rather than just interesting technology applications. He talked about Dell's strategic focus on four key pillars for AI implementation, engineering, supply chain, services and sales. He talked about the AI ROI flywheel concept where initial high impact projects generate results that fund future AI development, the distinction between reactive AI tools humans use and agentic AI, autonomous systems that complete tasks independently. He talked about how teams of AI agents will work together requiring new standards for authentication, authorization and knowledge sharing, the critical link between quantum computing and AI advancement with each technology accelerating the other's development, and emerging careers created by AI adoption, including software composers who design systems without writing code, thermal plumbers who manage cooling for GPU clusters and AI explainers who translate AI outputs into human terms.
- 00:58:56 As always, you can get all the show notes including the transcript for this episode, the video recording, any materials mentioned on the show, the URLs for John's social media profiles, as well as my own at



superdatascience.com/887. All right, thanks of course to everyone on the Super Data Science podcast team. We've got our podcast manager, Sonja Brajovic, media editor, Mario Pombo, Nathan Daly, and Natalie Ziajski who are on partnerships. Our researcher, Serge Masis, our writer, Dr. Zara Karschay, and our founder Kirill Eremenko. Thanks to all of them for producing another invaluable episode for us today. For enabling that super team to create this free podcast for you, we are deeply grateful to our sponsors. You can support this show by checking out our sponsors links, which are in the show notes. And if you yourself are ever interested in sponsoring an episode, you can do that. Just go to jonkrohn.com/podcast to learn more.

00:59:50 All right. Otherwise, share this episode with people who'd love to have it shared with them who might enjoy it. Review the episode. I think that helps. Get the word out in podcasting platforms, YouTube, subscribe if you're not already a subscriber, but most importantly, just keep on tuning in. I'm so grateful to have you listening and I hope I can continue to make episodes you love for years and years to come. Till next time, keep on rocking it out there and I'm looking forward to enjoying another round of the SuperDataScience podcast with you very soon.