SDS PODCAST

EPISODE 1042: BUILDING FOR THE AGENT-NATIVE APP STORE, WITH SALESFORCE'S TYLER CARLSON

Jon Krohn:	00:00	SaaStr, the best known conference for SaaS businesses used to have dozens of employees. Today it runs on three executives and an army of AI agents. My guest today explains what that shift means for how all software gets built. Welcome to another episode of the Super Data Science Podcast. I'm your host, Jon Krohn. Today's returning guest is Tyler Carlson, SVP and head of product for Agent Exchange at Salesforce. That's the marketplace that now brings Salesforce and Slack apps, agents, and AI plugins all together in one place. We've partnered with Salesforce on this episode in which Tyler details their new headless strategy, why plugins are the new apps, and how AI practitioners can get what they build in front of Salesforce's vast customer base. Enjoy. Tyler Carlson, welcome back to the Super Data Science podcast. How's it going, man?
Tyler Carlson:	00:55	It's going really well. Thanks for having me back.
Jon Krohn:	00:57	Yeah, you were on the show in January, episode 956, for those people who are keeping track. And when we spoke in January, Agentforce 360 had just opened up to what you guys call ISV partners, which are independent software vendors. But now I understand things have opened up a lot more. What's changed since January?
Tyler Carlson:	01:19	Earlier this year, we made a really strategic shift at Salesforce and announced this strategy, which we call headless Salesforce, which is really all about taking all of the investment we've made in our platform, the APIs, now the MCP servers, and making that available to customers to use as infrastructure to build their own amazing experiences. What these AI tools and harnesses have enabled you to do is build these really interesting personalized artifacts that reference data from solutions like a Salesforce. And more and more, we're seeing that the interface that we have spent the last 20 years building is, for many of our customers, are a good jumping off place to start, but it's the lowest common

denominator. It's the thing that sort of solves the needs of the most customers horizontally. And for many, many years, we've enabled these frameworks for customers to customize the UI and tweak the way that it's laid out just for them.

02:37 But with the advent of these new tools, you can create personalized user interfaces on top of the data and workflow and permissions model that we've shipped with our platform for the past 25 years. So it's a really interesting transition in the way that software is delivered. And I think we leaned into that wholeheartedly from a strategy perspective and said, whether it's your interface, someone else's interface, a custom interface you've built, we want to take all of that investment in your customer data, in your workflow processes, in the security and trust layer and all of those things you've built, who can see what data where and make it available as a platform for you to build these amazing experiences on. So we did that at the early part of this year, and then we've continued to evolve that strategy. We launched Claudeforce Partnership, which is basically taking that idea, which is you can build amazing custom experiences using CloudCode and Cloude Cowork on top of the Salesforce data workflow and platform to do some pretty amazing stuff.

03:52 And we've now opened that entire platform up to anyone to build those kinds of experiences, whether you're an individual at a company who just wants to build your specific workflow dashboard for supervising support cases or how you want to manage your team's pipeline and forecasting process. You can do that now, but also all of the partners that have been on this 25-year journey with us as a company can also now take all of that expertise, vertical expertise, industry expertise, and build their own expressions of what these experiences should look like. So it's a really exciting time to be a builder working in and around the Salesforce technology because

I think we're really embracing this idea that you can build these incredible bespoke experiences on top of our platform layer while also keeping the value of what you've invested in the data process and security layer of Salesforce.

Jon Krohn: 04:58 Yeah, there's a huge amount that you can do, as you say, with that kind of security guarantee that you have of a big enterprise platform like Salesforce, so you can get going very quickly with powerful agents in your organization. And so I was at Dreamforce in San Francisco a few weeks ago, as I'm sure you were as well. And I went to, I'm probably going to butcher the naming of exactly what this was like Agentic AI City. What was that called?

Tyler Carlson: 05:25 I don't remember exactly, but it sounds about right.

Jon Krohn: 05:27 You know what I'm talking about. So in the basement at the conference, they had this huge city of vendors showing off how Agentforce is being used live today in organizations. And a lot of these are huge businesses that everyone has heard of like AT&T, Formula One, but the coolest demo that I saw there was from a business called SASTR. So like software as a service, TR, SASTR. So I knew them as a conference and that is mostly what SASTR does, but they're also a platform. And SASTR is a big deal. My previous company that I co-founded was a SAS platform and it was a big deal for us to go and was one of the best opportunities, one of the best conferences we could go to, to get SAS investment, to get SAS clients, to meet other SAS companies. So yeah, really cool conference. And SASTr used to have dozens of employees and now it's three executives and an army of Agentforce 360 agents, which is pretty wild.

Tyler Carlson: 06:35 It is. What they have done, and we've had their team come talk on our earnings report about what they're building, they've shared at our conferences, they are really embracing what it looks like to build sort of an

agentic platform for what they're doing. But also I think it really highlights what we are trying to accomplish with moving Salesforce from a predominantly user interface oriented SaaS platform to really a layer of infrastructure for your customer workflows, for your business process workflows that is for humans and agents. And I think there's really kind of two versions of that world. You have humans interacting with your platform through what is now becoming the new version of an operating system, which is an agentic harness. And you are building these ephemeral artifacts that express business process and logic, you share them with other people, you build and iterate on your process.

07:49 And then you have companies like SaaStr that are actually building effectively an enterprise agent that does the work of what would be a workforce of humans. And so those two patterns are really emerging as the way in which I think people are going to consume software in the future. It's built for a human that is AI assisted in everything they're doing, and it's built for an agent that is sort of operating a business on your behalf through that software. But importantly, I think those two things are going to have to have some shared foundations. And our strategy is that we can be much of that shared data process and security infrastructure that supports both of those personas working against the same sort of enterprise goals. So I think what they're doing is incredibly powerful. They're really leading, showing us what is possible with our platform in many ways.

08:52 So it's really exciting.

Jon Krohn: 08:53 Yeah. And something to highlight in what you're saying there and something that I think the SaaStr example really highlights is that when people think of Salesforce and then you think, okay, Agentforce 360 is part of Salesforce, you first think, what sales processes can I be automating? Or how can I be automating things related to

customer data? But the SaaStr example shows how it can actually be any kind of what you call the business workflow. It doesn't have to have anything to do with sales or customers and Agentforce 360 can still be effective there.

Tyler Carlson: 09:24 Yeah. The things that people have been doing with our platform, I think have even astounded us over the years. I mean, their entire business is built around completely different processes than a sales and service process that run and operate on top of the Salesforce platform. So I think what we're trying to do is take all of the innovation that we're building to support those sort of canonical use cases that Salesforce builds for, the sales, the service, the marketing, the commerce personas, but make those primitives that we're using to build that also available on the platform so that a developer or another company who wants to build a business that sort of relies on Salesforce as some of that infrastructure can do that, which is really exciting.

Jon Krohn: 10:17 Yeah, it is exciting. Another thing that has changed since you were last on the show is that Agent Exchange has now brought everything into one unified marketplace. So Slack marketplace apps, if people aren't aware, Slack was acquired by Salesforce some years ago. And so Slack Marketplace apps are now available alongside Salesforce Ones in the Agent Exchange Unified Marketplace. So first of all, you might need to define for our listeners who aren't familiar what the Agent Exchange is, how that relates to Agentforce 360, and then tell us about the unification.

Tyler Carlson: 10:50 We have had for the past 20 years, I think we launched it in 2005, so 21 years, an app marketplace for Salesforce. It was called the App Exchange. It was a place you could go to get apps and extensions and capabilities that extended the Salesforce platform. It was a thriving ecosystem of places for startups to build specialized

applications for certain industries or workflows, and whole companies were built on the back of this ecosystem and marketplace. About a year ago, we started to really realize a couple of different things were changing. One, the boundary and definition of what an app or an extension for Salesforce is was widening. Some of these types of experiences like what we've built with Claude are not traditionally the kind of app or experience that you would expect to be listed on what was then called the App Exchange. So we kind of wanted to widen the definition of what we were doing.

11:59 We also had this amazing ecosystem from Slack. There's all of these agents and apps that people have been loving using in the Slack platform for a very long time. And so our mission is really to figure out how to make this a very, very capable discovery engine for customers to figure out what they need to help get the work done that they want to get done, whether that's an agent, whether that's now a skill or a plugin for one of these AI harnesses or experiences, whether that's data asset that they need. So the aperture of what is necessary in sort of an enterprise SaaS marketplace is widening. So that required some really deep thought about how we change our marketplace mechanics. Our marketplace was built for Salesforce apps. It had keyword search. It didn't really have a lot of business context about what that app actually was or what it did.

13:03 And it really was operating like a marketplace of the mid to early 2000s. And so we needed to really evolve our thinking about this in addition to one, just building a better marketplace for the assets that extend Salesforce. Two, we realized there was this really important dynamic that was shifting in the world, which was it wasn't just humans that were discovering capabilities anymore. I talked about that sort of dichotomy of a human with agent assistance and an enterprise agent acting on behalf of a company. You can imagine a world where I, as a

procurement or IT leader at a company have my own agent harness and I set a task for my assistant to look at the business requirements I have, build a set of evaluation criteria, go do a bunch of research on what solutions in the market fit this criteria and can help me compare it against other solutions, issue an RFP to the vendors, collate the responses, stack rank who does best against all of these different capabilities, and then present me with a buying recommendation.

14:21 And that workflow was something that our platform was not set up to support really in any way. We had really not thought through agentic optimization for how is our content served to agents. We hadn't really even though about how do you actually give better business recommendations. So if you looked at a listing on our platform and you searched, we were doing keyword search across the words that were in the listing. So we needed to do something more. We launched semantic search, which enables us to be a little bit more thoughtful about what's there, but even that isn't far enough. What we're building now is how do you actually help someone make a decision? If I go to the marketplace and I ask a question like what are the best CPQ solutions, quoting solutions for a manufacturing customer? That question has embedded in it a bunch of depth, what is a CPQ solution?

15:25 What are the qualities that make one good? What are the capabilities that CPQ solutions need to have for manufacturing customers? What are the follow-up questions that you would ask the buyer to say, well, what are you most interested in? Is it this capability or that capability? And I think trying to figure out how we make our platform better at helping that kind of discovery is a huge strategy for us. So to get back to the point of your question, why we had to bring all of these things together, well, we had to bring all of these things together to build the foundation for what we want to build in the future,

which is really a capability reasoning engine that enables you to find the things you're looking for to solve the business problems that you have, whether they're built by Salesforce or one of our many amazing partners.

16:17 We want you to be able to solve your problems using our platform capabilities.

Jon Krohn: 16:24 Yeah. And when you say it allows you to find a solution, you includes agents, of course.

Tyler Carlson: 16:32 Yes, correct. Your agents, you, your agents' agents, whoever is looking.

Jon Krohn: 16:39 Exactly, exactly. What a world we are now in very much so. And in this world of agents, we have things like skills. So there's Salesforce specific skills that are available in agent exchange. There are plugins into AI tools, third party MCP servers, and then agents themselves. So yeah, at least these four different kinds of things, agents, MCP servers, plugins into AI tools and skills, which probably a lot of people are familiar with from their kind of experiences with cloud skills and that kind of thing. But when should my listeners who are hands-on AI engineers, data scientists, they're building solutions with agentic AI, when should they be going straight for a prebuilt agent, going for an MCP server, going for a plugin to an AI tool, or just going for skills?

Tyler Carlson: 17:29 It's a good question. I mean, I think the thing that I'm seeing start to play out is I think the way software works for certain personas is changing pretty drastically. I think you kind of have a pattern emerging where the large language models kind of represent what would traditionally be considered compute. They're the reasoning and computational layer. And then you have these agentic AI harnesses like a CloudCode or CloudCowork, ChatGPT work, news, dots, all of these things that sort of sit at this layer that would've

traditionally been occupied, I think, by the operating system. This is sort of the iOS or Windows sort of layer where you're orchestrating the compute across different tasks. And then you have sort of the plugins, skills, MCP servers, which I think occupy the area in the stack that really would've been the app layer. So it's like these are the different applications and extensions that you can plug into the operating system that do domain specific work are opinionated about how certain things are done.

18:39 And so as I think about that architecture, I think really deeply about how we help people that are trying to build something in that world explore the things that they need across the different layers of that stack. So our customers, what are the different models that they might want to be using across their stack? What are the different AI harnesses that have connectivity to Salesforce? So when we talked about the Claude experience we have, I want to be able to find like, hey, if I use Claude Cowork, it works really well with the Salesforce platform and I can build these amazing experiences there. And then at the top layer, we need to do the extra work to make these experiences really work well for the types of things that people are doing. I think this concept that you can kind of just take a really, really intelligent model and throw it sort of like a undefined problem is going to continue to sort of diminish in the way that people are using the models.

19:44 It's why more of the value is actually accruing to the harness and plugin app layer because once you get that part right, you can start doing model routing, you can pick the right compute for the right task is effectively what you're doing when you're picking which model runs under the hood. The thing that actually gives you the good output is what is the context that it has for that particular task? What are the sort of evals that are running against it to make sure that the output is the way that you would expect it to be? And so more and

more of our partners and developers are moving from developing the UI to developing the skills, which are the instructions for that model, the domain specific context that's required to complete the task, and the evals that are required to determine whether the output the model created actually accomplishes the goal that you wanted it to do and is it high quality and all of that stuff.

20:47 And so that combination of assets, the sort of skills instructions, the model sort of instructions, the evals to figure it out, and then the connectivity layer, what MCP tools do I have access to to accomplish the goal? That bundle of things, which is sort of trending towards a definition that maybe we all have a standard on as a plugin, but the different surfaces all have their own sort of variance on that becomes the asset that I think we should all be aspiring to build for our own products. So that's what we've done with the Claudeforce example is we're sort of starting to think what would Salesforce look like expressed through that new application paradigm? So I think that's kind of how I would think about it. If you're a developer building something, I think you should be thinking about what does that look like expressed as a plugin in these surfaces because that's effectively the new app.

21:45 And then how do you find customers to consume that and how do customers find the right ones? That's kind of the intersection of what we want the agent exchange to be is for our Salesforce customers, how do you find the right skills and plugins to do the thing you want? How do you enable your agents to find the right skills and plugins when you've given them a task to do and they don't have the right tool to do it? How do you govern and manage the use of those skills and plugins across your enterprise? So you have a bunch of different teams using different ones. How do you make sure that there's some standardization across the way they interact with your enterprise

surfaces? So there's a lot of really interesting problems to solve, I think, in that space.

Jon Krohn: 22:31 Really nicely said. I particularly like the analogy you made of LLMs being the compute layer and then agents or harnesses, sorry, being like the OS layer that we have in traditional software. And then everything else that you were talking about, skills, plugins, MCP servers being different apps that that operating system can work with. That made a lot of sense to me and I've never heard anybody talk about it like that before. So you're doing some good product level thinking over there, Tyler. I love it.

Tyler Carlson: 23:01 Yeah, appreciate it.

Jon Krohn: 23:03 You mentioned Claudeforce there, and that was another big announcement at the Dreamforce Conference in San Francisco a little while ago. And that ClaudeForce integration works both ways. So it means that Claude is available within the traditional Salesforce app, but then likewise, Salesforce integrations functionality are now natively available within the Claude interface. But you were mentioning evals in your last answer as well. You guys have obviously spent a lot of time on getting evaluations right with Salesforce skills and plugins because they don't just work on Claude, they also work on OpenAI's codex, they work in Cursor. So how do you set up your evals so that you can be sure that say a skill behaves reliably across all of these different AI tools that people use?

Tyler Carlson: 23:55 It's an interesting question. I think we're like everyone sort of grappling with this task. And I think we have some scars from early iterations of what those skills look like and were they working well or not? How predictably can you get the right output from them? And I think what helps is sort of two things. One, I think understanding the domain problem you're trying to solve is a really

important step. So there needs to be some level of taste in the output of the thing. And I think one of the things we've lost a little bit in the current sort of zeitgeist around models and using them for these tasks is I think people are so impressed with its ability to get to 80% that they forget about that last 20%.

24:44 And I think there's some element of this too where software in its current iteration, because it had to sort of solve for the needs of many with different use cases, getting to that 80% was sort of good enough. If you got to 80%, you could solve for the needs of most of your customers and you gave them the ability to sort of customize that last mile 20% of what they needed. I think with the right instructions and evaluations, you can actually get to a higher output for that. And so working with our engineering teams to figure out how do we actually test, evaluate the outputs that are happening. I think there's also some interesting things that we've done with the platform itself. I think trying to help the sort of harness understand what tools to use at what time is sort of embedded in the skills that helps with a little bit of this.

25:40 Two, I think there's an element of making sure that we're sort of testing these things across all of the different surfaces so that some of the nuance differences between them are accounted for. And you can build some of that logic into the layer that you're offering. And so I think what we really want to get to is a place where our customers can sort of manage how Salesforce is expressed across all of these surfaces in one central place. They can build it one time and have it sort of control how it expresses across these different surfaces. One thing that we built for that is what we call our headless experience layer, which is effectively a translation framework for UI componentry. So one of the things that you start to get into as you start to build these

sort of expressive software experiences in AI harnesses is you kind of get fatigued with a wall of text.

26:39 So everybody gets this sort of like wall of text back and I can ask it for this, that and the other. And you can start to do things like ask it to build HTML artifacts to sort of express this back to you. But then you get this drift where everybody's HTML artifacts are drifting away from the way that maybe the best practice for looking at that information would be. So we built this framework that allows us to basically build these UI primitives and tell the model how to express that in its own sort of UI language. It's going to feel native to the interface that you're in, but it sort of has this opinionated expression of how the data should be laid out that helps the model do the right thing.

27:37 And so I think more and more what we're going to see is there will start to be these new primitives and protocols that are created that help manage and mitigate the differences between the surfaces so that you can kind of have a managed layer like Salesforce handling how it is expressed across all of these surfaces and give the enterprise control over that so they can manage what is effectively a new explosion of personalized software as people build these artifacts themselves. And so how do you find the right balance between us giving them an opinion on how something should look, but not going all the way back to this world where we give you this very rigid UI that is not very expressive. So that's how I would think about it. I think one is, I think there's emerging capabilities for how to build and manage these skills plugins, the evals associated with it that we're all learning at the same time.

28:41 But I think the second part is we're going to continue to build these new protocols and assets that make managing this better because we're sort of really at the early stages of what building for this type of paradigm looks like. I

think if you would make the comparison to the transition to mobile, we're at that 2007 timeframe where the iPhone just launched, the app store just launched, people are figuring out what apps look like. And if you remember back to that time, there were a bunch of really cool apps that were created, but a lot of them were very simple ideas. And then over time, from where we were then to where we are now, the sophistication and the tools that were built for delivering those applications has just exploded. And I think we're going to get to a similar place with the new world of plugin skills, et cetera.

Jon Krohn: 29:38 For sure. Things are moving fast and things are moving faster than then for sure. And so there's lots of opportunity, more opportunity than ever, I would argue, for people who are the listeners to my podcast who are hands-on AI practitioners. And so actually on that note, for our listeners who have deep domain expertise in AI. But maybe they've never built their own app or they've never built their own software company, or maybe they've never built their own item for being included in something like Agent Exchange. So some kind of AI tool for agents. What are your recommendations for my technical listeners who haven't gotten started before with these kinds of initiatives of their own, but now can do it easier than ever because of the Agent Exchange?

Tyler Carlson: 30:33 Yeah. I mean, I think what I've always said on this is I think you need to be really laser focused on what problem you're trying to solve. So pick the domain, understand where the pain is, create a solution that uniquely attacks that problem, and really zone in on doing that well for a specific problem set or domain. And I think if you can do that, it will become clear what the target audience for that solution is. And then you have to think about, well, how do I get distribution? How do I get this solution that I've built into the hands of these people? Whether you're looking to do something that's just open source, you're looking to do something that's going to make you money,

that part is where we are trying to enable people building for the Salesforce ecosystem to rely on us to help them with that.

31:30 So the agent exchange marketplace is really leaning into how do we help builders and publishers get their assets into the hands of as many of our customers as possible by helping to match customers to solutions that they have. And so I think my advice would be, as we continue to iterate on what that is, the distribution sort of rewards, the access to distribution will accumulate to the developers and publishers that are building solutions that actually solve real customer needs in a meaningful way. And so that's what I would think about. And then I would think about what's changing. I would think about that new technology architecture that I talked about and what that metaphor means for how you build solutions. And if you're thinking about building the old way, challenge yourself to think about what would this problem that I'm solving look like in this AI native world where it's expressed as a plugin for an AI harness?

32:41 And if you build with that sort of mentality at the start, I think you'll be well set up for the future that's coming.

Jon Krohn: 32:49 I think a lot of your answer there showed how you have become and are such a prominent product leader at such a prominent company like Salesforce. Really great advice there. I love the way that you're thinking about problems that need to be solved in this new paradigm with agents and all the plugins that they have available to them, including through environments like the Agent Exchange. Thank you so much for taking valuable time out of your day for us, Tyler. Before I let my guests go, I always ask for a book recommendation, as you may recall from when you were here in January. Do you have a new book recommendation for us?

Tyler Carlson:	33:24	Yeah, I think last time I was... I'm a big sort of sci-fi fantasy novel nerd, and so I read business books in my spare time, but I think things that I really get excited about personally is just what I'm reading for pleasure. So I'm reading an amazing sci-fi series called The Sun Eater, I think it's Sun Eater Chronicles. I'm not exactly sure. I don't remember the exact name, but it's a great sort of expansive space opera style sci-fi book series, quite long. I like the long ones. It's very fun to get immersed in a new world. I think it's just a really exciting story and narrative, and that's the kind of thing I like to read. So that would be my recommendation if you're into that kind of thing.
Jon Krohn:	34:11	Nice. Thanks, Tyler. And then my final question, as always, how should listeners be following you or following Salesforce after this episode?
Tyler Carlson:	34:18	Yeah, follow along with me. You can reach out to me on LinkedIn. Would love to connect with all of your listeners. I'm trying to be a little bit more active, posting about things and stuff that's going on with the Agent Exchange in that environment. Also, if you want to join, join us on the journey. If you're building a solution and you want to get it listed and get what you're building in front of the Salesforce ecosystem of customers, we'd love to have you all working with us on the Agent Exchange. So yeah, please take a look at the Agent Exchange. Also, give us your product feedback. I think there's a lot more that we could be doing on this product. We have a long and expansive roadmap of things that we want to deliver, and if anyone's out there with ideas, we would love to hear them.
Jon Krohn:	35:07	Awesome. Thanks for that offer, Tyler. And yeah, maybe catch you again on the show soon. Hopefully this is the kind of cadence we're going to have.

Tyler Carlson:	35:15	I would love to. Happy to be back anytime. And yeah, of course, thanks for having me, Jon.
Jon Krohn:	35:21	Tyler Carlson certainly knows his stuff. In today's episode, he covered how the headless Salesforce strategy turns 25 years of data, workflow, and permissions infrastructure into a foundation you can build your own personalized interfaces on. He provided his analogy for the new software stack where LLMs are the compute layer, agent harnesses like Claude Code are the operating system and skills plugins and MCP servers are the apps. He also left us with his advice for builders. He said to zero in on one painful problem in one domain and design your solution from day one as a plugin for an AI harness. I hope you enjoyed the conversation. To be sure not to miss any of our exciting upcoming episodes. Subscribe to this podcast if you haven't already. But most importantly, I hope you'll just keep on listening. Until next time, keep on rocking it out there and I'm looking forward to enjoying another round of the Super Data Science podcast with you very soon.