diffuse*tap*
Virtual Event Series

Will AI Agents Rule the World?

Guest Speaker:

Jason Sosa
Founder & CEO
Azara.AI

Hosts:

Kenny Estes
CEO & Founder
Diffuse

Ayla Kremb
COO & Co-Founder
Diffuse

DiffuseTap: Will AI Agents Rule the World?

Last time on DiffuseTap, **Jason Sosa, Founder and CEO of Azara.AI,** talked to us about how close AI agents are to shaping the future of work, why businesses still rely on spreadsheets despite cool new AI solutions being available, and the worst-case scenario of adopting these new AI applications.

Want to make friends from the Diffuse Fund Ecosystem? Email contact@diffusefunds.com.

DiffuseTap: Where Crypto and Technology Shape Finance

DiffuseTap takes you to the heart of finance and digital assets. See how changes in crypto and tech are transforming the industry, with insights from top leaders and experts.

Meet the Speaker

JASON SOSA is an entrepreneur, keynote speaker, and global business advisor who has led teams in AI, blockchain, and VR, raising over $100M for projects. He founded IMRSV, advised a $20M VC fund, and is now the Founder and CEO of Azara.AI, making AI accessible for business. Jason is also a Forbes contributor and has spoken at major events like TEDx and the Singularity University Global Summit. LinkedIn: @jasonsosa

About Diffuse®

We are an alternative fund platform offering differentiated investment products. From digital assets to VC funds and beyond, we identify green field investment opportunities we feel will have market beating returns and turn them into professionally managed funds. For more information, visit www.diffusefunds.com.

KENNY ESTES: Mr. Sosa, thank you so much for joining us here today. Do you want to unmute and tell the good folks a little bit about your background and what you're up to over at Azara?

JASON SOSA: Great. Well, thanks to everybody for having me here. So, a bit about me. I started a computer vision company back in 2009. I have been involved in the AI space for quite a while. We were focusing on age, gender, and emotion classification using web cameras at the edge. It was like a Minority Report scenario where you would look at advertisements and they would change based on your age or gender.

That was our mission back then. We were in the TechStars incubator. I sold that company in 2015, worked at an early-stage venture fund in San Francisco, and have been doing AI business management consulting for a number of years. I've worked on everything from AI, nuclear medicine, and theranostics to VR headsets, wearable technologies, and 3D body scanning, just to name a few.

Over the last year, I've been working with some really talented people from OpenAI, AWS, IBM, and former TechStars folks on a new solution that's really solving the problem around generative workflows. The promise of science fiction has always been that AI would do the work for you. Right now, we are at the stage where it has been conversational.

The next wave we see coming is this wave of delegation where you can push a button and say, "Create an HR onboarding workflow for me." It asks you all the relevant questions. There is no code, no configuration. It simply assembles itself, writes all the code in the background, and routes itself to all the right plugins that businesses already use.

From there, it builds off the workflow you can reuse over and over again. So, that's a bit about what we've been working on over the last year. We've raised a little over a million dollars from some notable investors like Dave McClure, Yohei Nakajima, angels from Meta AI, Airbnb, and some institutions from Singapore, such as Redbadge Pacific. So, that's just a bit about us.

KENNY: Okay, amazing. We share some investors, which is exciting. Regarding chatbots, we all kind of know what that is—chatting and all that stuff. You are working with AI agents. Can you talk a little more about the state of play for AI agents? How far are we from realizing your goal of delegating all of HR to them? Where are we at? Who are the main players, and where are you trying to fit in?

JASON: Yeah, it's very confusing these days. Right now, everyone is familiar with chatbots. There's talk of digital employees, digital twins, AI employees, AI agents. These terms are often used interchangeably. The way I define it, AI chatbots are conversational. It's a one-to-one interaction. I ask a question, it replies, and we are all familiar with that from GPT and other solutions. What is coming next are AI agents that will be more autonomous. You will be able to make a command, walk away from the keyboard, and have it continue to operate behind the scenes.

The foundational layer, which includes models like GPT, OpenAI, and Anthropic, has everyone excited. Some smaller language models and open-source things like Llama are out there. We were just discussing in the breakout session that this is akin to the early days of crypto—everyone is excited about the possibilities with all these new ICOs coming out, but we are still in the building stage.

The orchestration layer on top of that foundational model is still being developed. Some amazing tools, like LangChain, Llama Index, AutoGen, and CrewAI, have emerged, but they're largely engineers building tools for other engineers. We saw a need to make this accessible for the 99% who don't code.

Many solutions out there are great in a proof of concept or demo, but when you want to put them into production, where your business relies on them and your customers depend on them, you need more than just a tool.

Currently, there are three main philosophies around AI agents. The first is the idea of drag-and-drop configuration, like Zapier or Make.com. There are powerful tools on the AI engine front, such as AgentHub, where you can mix and match models, data modifiers, and data loaders. However, most businesses would not know what to do with these tools. They are still operating largely on spreadsheets and emails. We saw a gap in the marketplace for those who want to leverage AI for business but feel intimidated by the technical overhead.

The second philosophy involves an AI taking over your browser, doing repetitive tasks. The problem is that browsers are intended for humans, and websites change frequently, requiring constant maintenance. The language of AI, especially regarding connecting with other services or applications, is largely around APIs. We see a way to make this more tangible.

We are building an interesting technology called Scenarios, where you can use prompts to define the context of your task. For example, with Claude 3.5, you can use a feature called Artifacts to generate a pricing calculator quickly, but it often makes assumptions. It would be great if the AI could state its assumptions and ask relevant questions to better understand the context, enabling it to build better code and workflows.

This is where we see things heading in the near term. Over the next six months, we will see technology emerge that allows you to push a button, give a command, and automate large portions of your business, doing approximately 80% of an employee's tasks at about 10% of the cost.

KENNY: I think we've already touched on this, but when you're talking about enterprise clients, a major concern is avoiding mistakes, right? If an AI misidentifies a role, for example, treats someone in the system as a janitor who doesn't exist, it's not ideal.

My understanding is that a lot of hallucination-solving involves quality checking the output and ensuring it's accurate, creating a hybrid solution. How are you dealing with that? Is there a solution where you have people doing QA on top of the AI's output?

JASON: There are a few different ways. At the foundational model, we are looking at a blended approach. Instead of using the most expensive model for everything, from long-term and short-term memory to planning and tool usage, we can apply the right model for certain tasks.

For example, if we're using a different language, we might choose GPT-4 for its robust language support. For a large context length, we might use Anthropic. We may fine-tune certain sub-agents, or a mixture of experts, which we call sub-agents. These act as individual critics, evaluators, and supervisors within the persona of the agent.

Another approach is to set up effective guardrails with deterministic rules, like tracking and order numbers, which cannot be hallucinated. On the other side, we can be generative with tone and language to bring personality to the user experience.

These guardrails and triggering mechanisms ensure that the AI checks back in with the human in the loop, especially for quality control, such as with agentic regulation. However, there is no defined orchestration layer or rules for anyone to follow, so we are still figuring out the steps and processes.

KENNY: Okay, I guess that makes sense. Let's take a couple of steps back. Right now, I think it's a good time to be a founder. My armchair quarterbacking opinion is that valuations are kind of ridiculous, and there is a massive amount of hype—probably overhyped to a good degree. Do you agree with that or strongly disagree? What's your take?

JASON: Kenya and I went on a 42-day trip across Europe. We heard a number of different pitches, and a lot of them are all about AI. I think we've reached a point of AI fatigue that is palpable, from both investors and the average consumer.

It's moving at such a pace that it's really difficult to keep up with. From a business perspective, they're almost like, "Let's let the dust settle, and then we'll figure out which one is actually useful." At a certain point, I think some investors are also in that mindset—waiting to see which one emerges with decent ARR and some product-market fit before deciding where to dive in.

But there are a couple of challenges. From an investor's perspective, if you are looking to invest in this space and want ARR, you could invest in an AI girlfriend application. There are plenty of those. There are plenty of apps that show you what a couple's baby might look like.

But there is no defensibility in any of these because the orchestration layer hasn't yet been built, and the application layer or AI-enabled services on top of that haven't been built either. It's like the ICO market in crypto. We were all like, "Yeah, the market is going to be awesome, the future is here," but here we are, eight years later, and we're still waiting for actual applications to emerge. The layers are still being built—it's very similar to that market.

So, it makes a lot of sense to really look at the jockey—what's the team and the long-term vision? Consider the data moat; where are they gathering this information? Ultimately, all these models will go to zero; it's a race to the bottom, as everything becomes commoditized. What's really interesting is the private data being mined, because all the public data has been used across the internet, sometimes illegally.

The private information is a valuable source that hasn't yet been tapped into by some of these models. Especially when it comes to providing leverage in work, the ultimate dream for a company is to either reduce the number of employees or make existing employees ten times more productive to grow revenues. That's the holy grail of where this whole thing is headed.

KENNY: A couple of things you mentioned resonated with me. One is the valid criticism of blockchain, where we have a lot of layer ones and layer twos, but no applications. I hope AI doesn't end up in the same spot—eight years in, and we're still waiting. You also talked about product-market fit, which is a good segue into Henry's question in the chat.

There are a lot of AI agents trying to solve different problems. Do you think we will end up with a bunch of domain-specific agents that solve particular problems, looking like companies with an AI model? Or will we see something more generalist and industry-agnostic doing all these tasks?

JASON: A lot of the AI solutions I'm seeing now are really just chat GPT wrappers around professional services. If you're building a specific workflow solution, it usually means it's very inflexible. For example, if you are building a workflow solution around customer service, sales, or support, you have to find other businesses with that exact same workflow.

But it's difficult because businesses are messy and change course all the time. The reasoning is not quite there yet, but it's improving quickly—we're seeing three or four major releases a year, while most companies have a procurement cycle of one year. It's moving incredibly fast.

I think we will probably see narrow verticals around very specific niches before we see any kind of generalized solution. If we do reach that point, we're entering into AGI or artificial superintelligence, and then it's anybody's game.

KENNY: Niches to start, and then, as they expand, hopefully, we find something more comprehensive. I guess that makes sense. From previous conversations, I know you are, unfortunately for you, closer to being a Bitcoin Maximalist than I am.

My question is, there is a lot of talk about the convergence of AI and crypto. On the face of it, maybe there's not much common ground—they might just be shiny new technologies appearing at the same time. Do you see these as two completely separate areas, or is there quite a lot of overlap?

JASON: There are a couple of interesting areas of overlap. We touched on the ICO market and why applications have not emerged. I think it's largely because they're still built by engineers, and the average person—the larger, broader market—cannot adopt them because the processes are complicated and intimidating, like long addresses and using MetaMask.

One area where AI could help is solving the user experience problem, either in a conversational way or in some other yet-to-be-defined process that simplifies it for everyday people. Another area is in verifying information.

We live in an era where we cannot trust our eyes or ears. Deep fakes are happening all the time, even in financial institutions, where people are losing tens of millions of dollars because someone spoofed a CFO on a Zoom call. We don't have a real way of identifying what is true and what is not. For example, no one knows Satoshi Nakamoto's real identity, but we would know it's him if he did a digital signature.

Digital signatures and other solutions like MicroStrategy's Orange Check, which verifies email authenticity, are areas where AI can complement blockchain. Lastly, AI agents equipped with wallets could be very powerful, especially if they cannot work well within traditional payment rails. The natural choice for AI agents in finance might be cryptocurrencies and Bitcoin.

KENNY: That sounds a bit scary to me—having an AI with a wallet means they control funds, which could potentially go wrong. In the last 30 seconds or so, what is the worst-case outcome here? We've talked about the best case—how AI can help with workflows and efficiencies. What about the worst-case scenario with AI agents in the future? Where could this go wrong?

JASON: We just saw a good example of what could go wrong with CrowdStrike and the Windows driver incident. Imagine AI agents acting as an attention router, filtering all the information coming to us based on our interests. I have an app and I go through and I tell it what categories I'm interested in, and it filters out all the things I'm not interested in.

So now, I'm getting this bubble effect of information, even more so than what I had before. Because it's all being filtered through this AI agent. I give it tasks to do, and now that AI agent is communicating with other AI agents, and they're sharing and passing off workflows to one another.

That sounds like a dream. I think that's the ultimate application that would be beneficial for everyone from personal to business. But it also can go extremely wrong. It could end up like a “1984” scenario, where the information controls us. There is a lot of potential for great, amazing things AI could do to benefit people, but it's also a double-edged sword with potential dystopian outcomes.

Thank you for downloading this DiffuseTap event transcript.

Sign up for upcoming sessions and check out past features and event transcripts.

Chok Ooi

Senior Partner at AgilityIO

DiffuseTap: Asia's Best Investment Opportunities

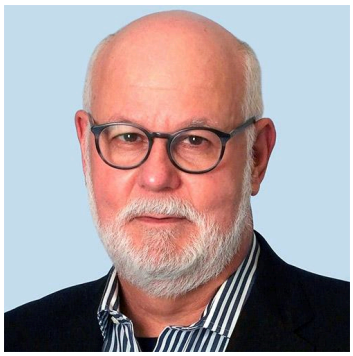
What does it take to seize the next big wave in Asian investments? Chok discusses the pivotal changes unfolding in investing in Asia, the trailblazers seizing the region's best opportunities, and why others might be missing out. Read on

Stevie Cline

Managing Director at Vol. 1 Ventures

DiffuseTap: Why Will 2024 Be a Great Year for Crypto?

Will 2024 be the year when crypto products rule the industry growth story again? Stevie talks about crypto projects poised to disrupt the market, factors causing the downfall of certain protocols, and the evolving role of NFTs today. Read on



Pat Daugherty

Partner at Foley & Lardner LLP

DiffuseTap: Who Decides Which Coin is a Security?

Who really determines which asset is a security and which is not? Pat explores the legal and regulatory dynamics of digital assets classification, the broader implications for the SEC's role in crypto, and the potential impact of crypto-related cases under judicial scrutiny. Read on

JOIN US