The State of Agentic AI Security

The Numbers Behind the Report: Fireside Chat Transcript

Speaker	Role	Organisation
Tom Eggleston	CEO	ProofID
Neil Evans	Chief AI Officer	ProofID

About this transcript

Organisations are under pressure to roll out agentic AI quickly, but governance and security are struggling to keep up. In this fireside chat, our CEO Tom Eggleston and Chief AI Officer Neil Evans talk through the findings of ProofID's State of Agentic AI Security report, and the practical steps that help organisations move faster with AI, not slow it down.

The report is based on independent research by Coleman Parkes with 300 business leaders. It looks at how far organisations have got with adopting AI agents, and whether their identity, governance and security controls are keeping pace.

Use this transcript to catch up on the conversation instead of watching the recording, or feed it to an AI assistant to create your own summary, explore a specific topic or pull out the insights most relevant to you. Every turn is timestamped and attributed to its speaker, so you can trace any point straight back to the original video.

00:00:06 **Tom Eggleston** · CEO, ProofID

Okay, so we're here today to talk about ProofID's new report into the state of agentic AI adoption. We're going to be releasing this very soon. We spoke to 300 business leaders across the UK, the US and the EU to ask them about where they're up to with their adoption of agentic AI. Really, really fascinating report.

Before we get into it, though, it'd be a great idea if we introduce ourselves. So maybe you introduce yourself, I'll introduce myself, and then maybe you can give us some of your initial takeaways.

00:00:35 **Neil Evans** · Chief AI Officer, ProofID

Yeah, sure. My name's Neil Evans. I'm Chief AI Officer for ProofID. I've been in the industry for over 30 years now, from an R&D background and product management, roadmap management, and I've been working with AI since the GPT-2 days. So I'm interested to see how this industry develops.

00:00:52 **Tom Eggleston**

Right. And my name is Tom Eggleston. I'm CEO of ProofID. So Neil, what were your initial takeaways when you read the report?

00:01:04 **Neil Evans**

I think what's not surprising is that there's a lot of pressure in the industry to roll out AI extremely quickly, and that is happening. But there's a tension there, with people struggling to follow up with governance and security around that, and that's causing a lot of friction within businesses.

00:01:22 **Tom Eggleston**

So let's start at the top with adoption. One of the first questions we asked the companies we spoke to for this report was where they are with adoption of agentic AI. And what we found was just under 1 in 5, so I think it was 19% of the companies we spoke to, have adopted AI agents into production.

And I've been trying to work out what I think about that stat. Do I think it's high, low or about right? And I suppose there's so much noise across the industry about AI, it's the hot topic that everybody wants to talk about. I suppose I thought it was a bit low. What did you think?

00:02:01 **Neil Evans**

Yeah, I think the key point for me is that agentic can mean a lot of different things to a lot of different people. On the one hand, you've got what I would call enterprise agentic services running on your network at speed, providing automation across all of your business, which I think is probably pretty unusual.

Maybe apart from some bleeding-edge Silicon Valley businesses. However, I think where a lot of people are on their journey is running Claude and Claude Code agents, those kinds of workloads, which are very different in their nature. The key point for me is that both of those workloads have got agency, whether you call them agentic or not.

The ability to call tools, to access files, to touch your core business systems: those are the things that are important to governance and control. So I think you need to build your policies around that and worry about whether it's agentic or not later.

00:02:58 **Tom Eggleston**

So let's talk a little bit about discovery. This is the idea of actually knowing what AI agents are operating in the organisation. I found it quite striking that only 2 in 5 of the companies we spoke to were confident that they actually knew what all the AI agents in their environment were.

How does that stat feel to you?

00:03:19 **Neil Evans**

I think it's a lot easier to track your agents in a more enterprise environment, if they're running on your network and in your data centre under the control of the IT team. I think what's really happening is people are running agents on their laptops, whether it's Claude Code, the Claude app or whatever else they've brought to work with them.

And it's a lot harder to track those. There are a number of tools and techniques emerging in the industry now to make that easier, but I think it's something that's growing and evolving quite quickly. So I'm not surprised by those numbers. It's just a sign of the lack of maturity in the market right now.

00:03:55 **Tom Eggleston**

Well, I suppose there's a bit of a terminology question there as well, right? So we're talking about agents, and then you've got AI tools like Claude or Copilot, which people are starting to talk about as shadow AI, very reminiscent obviously of the shadow IT we've had previously. And it just raises this really interesting question.

What actually is an agent? And we need to track all of it, I suppose.

00:04:19 **Neil Evans**

Yeah, absolutely. I think at the end of the day, anything that has agency on your network, whether it's a person using an AI tool or something that's more autonomous, is something that needs to be monitored and contained. I think agentic is a bit of a buzzword right now.

And it's more around how we build agents that can spin up and do the work of hundreds of humans. I think that's where that's come from. But I think you're right to lock that down a lot more from a security perspective. Anything that's got agency, that has access to tools and access to files, needs to be tracked and monitored very carefully.

00:04:57 **Tom Eggleston**

I suppose the fundamental point, actually, when you get down to it, is if you don't know about it, you can't control it.

00:05:03 **Neil Evans**

Exactly. It's all about visibility. How do you actually find what's happening in your network? Who's doing what? And that's a really difficult thing to do. I think there will be a range of techniques required to do that well. Obviously you've got things like network traffic monitoring, what's on laptops, what's installed, what your cloud estate looks like.

Ingress and egress firewalls. There's going to be a range of tools that make that really effective. And I think a lot of industries and a lot of companies are struggling with that.

00:05:37 **Tom Eggleston**

Is there anything specific around AI that makes this a really pressing problem? This idea that if you can't see it, you can't control it?

00:05:45 **Neil Evans**

Yeah. I'm from a cloud engineering background, and there's a lot of really good foundational technology out there, particularly from the microservices world, where there are ephemeral services: services that spin up and spin down really quickly. And there's a lot of technology out there.

It's really good at tracking that: observability, telemetry, workload IDs, all that kind of stuff. I think what's new about AI is that it's just another layer of complexity on top. AI services are a little bit more unpredictable. It's hard to define what they should do. You've got things like intent-based monitoring, watching what an agent is doing as opposed to what it's thinking.

It's quite hard, so it's a complex area. And I think it's just something that needs to evolve to catch up with the technology itself. Okay, Tom, let's talk about governance. One stat that really stuck out to me was that only 14% of the people in the report were assigning human owners to their agents.

And again, there were similar stats around managing lifecycle and identity for agents. How do you think that presents a challenge to businesses, and how they manage their agent landscape?

00:06:58 **Tom Eggleston**

Yeah, I agree, I think that's really striking. Another way of saying it is that six out of seven of the organisations we spoke to are not assigning a human as accountable for their AI agents. And you wouldn't do that with people. We don't have people coming into an organisation without a responsible manager. When we talk about people, there's always someone who's accountable for what's going on in a department or a team, and for the output being produced, someone's responsible for that.

And I think not doing that, particularly when we're talking about agentic AI agents, is a real blind spot. You've already mentioned this, but the way that AI works, and AI agents work, they're non-deterministic. They work incredibly quickly, at machine speed.

We don't always know exactly what they're going to do. If we're going to be putting them to work on real company data, maybe customer data, it's really, really important that we've got someone who's ultimately accountable for that, because things might go wrong. And we need to understand how that happened.

And who's going to be accountable for dealing with the fallout.

00:08:10 **Neil Evans**

Yeah, absolutely. I think it's a very natural response for IT teams to treat agents in the same way as cloud services and traditional services in the IT landscape, but it's just not the case. It reminds me of a very famous IBM quote from the 1970s, which said never let a machine make a business decision, because it hasn't got any accountability.

And I think we need to think of agents that way. We've got a responsibility to manage this technology. It's new and it's emerging, and we can't just apply our previous techniques to the AI landscape.

00:08:44 **Tom Eggleston**

I think it's a great quote, and it really brings it home. But I also think you can look at it a different way and say, well, AI agents are people too, right? Obviously they're not. But in terms of thinking about how we manage them, I find that quite a useful phrase to actually think about.

If we would do this for people, we should also be putting governance in place for agents in just the same way. That's having an assigned owner, but also having standard policies that govern onboarding and offboarding, just like we do for people. We also need to do this for agents because they're performing an important role within the organisation.

They need to be managed carefully.

00:09:21 **Neil Evans**

Yeah. And it's a really interesting topic. I've been listening to podcasts talking about AIs having personhood and having cash to spend and rights and all this kind of stuff. So it's an interesting way to think of it. Maybe that's going a bit too far, but if you frame it that way, I think it gives you the right way to approach the problem.

00:09:38 **Tom Eggleston**

Yeah, I agree with that. Another stat that jumped out to me from the report, and maybe this is a bit biased because we work in the industry, was that only 12% of the organisations we spoke to are actually routing all of their agent traffic through a gateway. That jumped out to me. What did you think?

00:09:58 **Neil Evans**

In some ways surprising, in some ways not. If you look back at the evolution of IT services, middleware sprung up very quickly to manage this problem in the previous incarnation of this technology. And I think what we're seeing here is adoption at a rate that's just getting ahead of that in most businesses.

But I think you're right. It's really important that we catch up with that and get some more control over our agent landscape. I think putting AI gateways in is probably an easier thing to do when you're running in the enterprise space. Maybe you're running agents in the cloud and the IT department is deploying them.

But when you've got people running apps and agents on their laptops, it's a lot harder.

00:10:43 **Tom Eggleston**

The thing I always think of when we talk about AI gateways is APIs. When APIs were invented, everyone got really overexcited, and all of a sudden development teams were connecting everything to everything. Which was really powerful. We were able to do a lot of things, but you very rapidly ended up with this unmanageable spaghetti junction of connections and different integration patterns.

It feels a bit to me like we're making the same mistake again.

00:11:13 **Neil Evans**

Yeah, I think you're right. The different types of AI gateways out there are becoming more and more wide and diverse. You've got AI gateways that are designed to interface between the agents themselves and the underlying models. You've got gateways that are designed to sit between the agent's tooling and your internal services.

And I think that's evolving. I think we will see some standardisation as the industry matures. But you're right to call out that we shouldn't make the same mistakes as in the past.

00:11:47 **Tom Eggleston**

Yeah. Just thinking about how long it took to retrofit API gateways to deal with that problem. Honestly, if I was sat with someone who was looking at building agents and thinking, what would be a really great first step to take?

I think an AI gateway is a great place to start, and it's not just about security and governance. I think it's a real enabler. What's your view on that, as someone who's worked in R&D teams?

00:12:14 **Neil Evans**

Yeah, I think you're spot on. Even if you're only running one or two agents, it's the easiest time to get a gateway in place, right? If you've got hundreds of them, it's a lot harder to get everything connected. So having that single control point gives you somewhere to do your auditing and your policymaking.

So I think you're absolutely right to make that recommendation. Okay, Tom, let's talk about trust and how we build trust around our agents. One of the stats that really stuck out to me was that only 15% of the people we asked had put short-lived, narrowly scoped access around their agents. What do you think about that, and how could we change things?

00:12:49 **Tom Eggleston**

It's really interesting. Another one that jumped out to me was that only 17% were actually giving every agent its own unique identity. And when you flip those and think about what everybody else is doing: we're using shared identities for our agents, and potentially giving quite a lot of broad standing access to our AI agents. Maybe that's something we've done with people historically, because we've not always been brilliant at managing this stuff with people.

But when you're talking about agents, it's a whole different ball game. We've already talked a little bit about the fact that agents work at machine speed, but there are other things as well. You don't exactly know what an agent is going to do. AI is not going to be bound by ethics in the same way that a person is, or by a job role or whatever it may be.

So I think it becomes really, really important that you've got a lot of control over what that agent has access to when it's carrying out a task. And I suppose what the report tells us is that currently that's not really happening. So thinking about the nature of AI and an AI agent and how they work, what we really want is to know that every single one has a unique identity.

So we can always tell which agent it was that did a thing. And at the point when they come to actually carry out a task, we want to give them the absolute minimum access they need to carry out just that transaction. We'll hear the industry talk about this variously as least privilege, zero standing privilege or just-in-time, all different ways of saying the same thing.

But we need to really be thinking about the minimum access we can give this agent right now to do the task it's trying to do, rather than just giving it standing access to do things whenever it needs to.

00:14:47 **Neil Evans**

No, I think that's a really good way of framing it. This all comes back to the differences between traditional microservices and software services compared to agents. It was easy to define least privilege

and minimum access rights for a well-defined software service. But when you've got an agent with unpredictable, non-deterministic behaviour, where its intent can result in dynamic actions,

that's when you really need a very different approach: shrinking its standing access rights down to the minimum, and that's where just-in-time, dynamic access comes in. The more intelligence you can build around that, the better. And I think observability feeds into this a little bit too.

If you can do it in real time, the more you can control it. Ultimately it's about minimising the blast radius of that agent. We like to talk about entitlement radius and zero standing privilege. All those things go together to minimise the damage that can be done in any given moment.

00:15:49 **Tom Eggleston**

Yeah, absolutely. I think when an agent is looking to carry out a task, there are really three questions you need to ask. How broad is the access we're going to give, and how long should the access be granted for? But then there's an even more fundamental question, which is: should we even allow this transaction to happen at all?

And if not, then we're not going to grant any kind of access. Now, that's actually quite a challenging problem, because we need to be looking at each transaction and making a qualitative decision on whether this agent is doing the right thing. Is this transaction or action it's taking appropriate

for the objective it's trying to get to? And I think that's a bit of a developing landscape. A lot of the technology to do this already exists. We're talking about policy-based authorisation, with a PDP framework behind it. An example might be: if an agent is trying to do a transaction in the middle of the night, we might look at that and say, that's a bit weird, that probably shouldn't be happening, and we'll stop it.

A very simplistic example, but it's the idea that even before we get anywhere near granting access, we should be asking the question: should this even be happening?

00:17:00 **Neil Evans**

Yeah, I think you're right. And there's so much terminology here, it can be quite confusing, but I think you summed it up really well. Should it be allowed to do it? If so, for how long, and how far should its reach be? If you can think about those three principles, you've got a really good starting point for building some policy around this.

00:17:16 **Tom Eggleston**

Let's move on to have a look at observability. There are a whole lot of eye-catching statistics throughout the report, but one that really jumped out was that only 1 in 5 of the companies we spoke to who've actually deployed AI agents can terminate them in real time if something goes wrong.

I guess the dramatic way of thinking about that is we've got a rogue agent, we need to do something: kill switch. These are the terms people talk about. It's a bit shocking to me that we have this situation where companies are deploying agents but don't actually have a way to stop them when things aren't working as they should.

00:17:56 **Neil Evans**

Yeah, I was quite surprised by that too, because when you think about observability and telemetry, those technologies are really mature, thanks to all the work that was done in the cloud engineering space. So I'm surprised people aren't leveraging that for their agent landscape. However, if you think about all the things you've covered so far, do you really need a kill switch?

If you've got zero standing privilege and you're able to do just-in-time access granting, it feels like a dramatic thing to want. Whether you really need it is another matter. When I think of observability, I think about what you need to collect in real time to build actionable insights.

What intervention do you want to carry out on your agent landscape? Is it trend analysis? Risk analysis? Is it a kill switch? I'm not so sure. So I think it's all about collecting the information you need to take the actions your business wants and needs to carry out on your agentic landscape.

I think it could be very different for different industries. So it's about building those data models, and I think it's a real-time thing, not a historical reporting thing. You really need to be able to respond in real time to those metrics and those signals.

00:19:08 **Tom Eggleston**

So I think what you're saying, maybe not in so many words, is that if you get to the point of a kill switch, it's kind of too late. And what you've learned from your time in the cloud industry is that this is something that happens before you get to that stage. You should be able to see that something's moving away from where it should be, and take action responsibly at that point.

00:19:29 **Neil Evans**

Absolutely. And it's a really unique problem for AI, because in cloud services and software services it's matured a lot, to the point where you have separate data planes and control planes. Having that control plane really gives you fine-grained control over those services and how they're behaving.

That's not the case with agents, because it's all going through one prompt. The control plane and data plane are effectively one channel. So there's probably a lot of work to be done there in terms of understanding that and building a more mature observability architecture.

00:20:04 **Tom Eggleston**

All right. So I suppose we're getting towards the wrap-up. At this point, it's probably worth mentioning to anybody watching this video that if you want a full copy of the report, just head across to ProofID's website and it's available there for download. But just before we wrap up, I'm going to ask a totally unfair question.

If you could get your crystal ball out, and we did this report again in 12 months, where do you think the industry is going to be? What will have changed? What will we have got right? What will we still be getting wrong?

00:20:39 **Neil Evans**

Oh, that's a really good question. What we're seeing already is the adoption and emergence of more standards around this. With many emerging technologies, certain vendors pave the way. New techniques come out, become popular and then fade away. But I think we're seeing increased maturity.

And I think that's going to accelerate, particularly as the scale of AI rollout speeds up, and we're going to need more standardisation and more best practices around this. If you look back at the cloud engineering world, we've got things like well-architected frameworks. We've got lots of standards around best practice for running a cloud landscape.

I think we'll see that for AI too, as the industry gets to grips with it. In terms of stats, I'd probably stop short of trying to predict the percentages, but I think the trend's in that direction. One thing I would say, though, is that security and identity will underpin a lot of that, because that's really the enabler for allowing that acceleration to continue.

I think the industry's got a little bit ahead of itself with rollout. So we need to bring in that governance and that security to act as an enabler for it to continue to speed up.

00:21:52 **Tom Eggleston**

Okay. I'm going to have a go at a stat. Why not? My prediction is that 12 months from now, if we ask the gateway question again, I hope we'd get 30 or 35%. And the reason for that is I just hope we can see that this is a pattern we've seen before. We've talked about this before, Neil.

None of this is necessarily brand new. It's obviously a different application of technology, but we have seen a lot of this before, and there are patterns we've used before. You've talked about that with what's been going on with cloud services, API gateways, all of that.

We've seen this before. So I would hope that when people start deploying this, they see that this is a simple step they can take, one that's going to really enable us and prevent us from making some of the mistakes of the past.

00:22:38 **Neil Evans**

Yeah. I hope that increases too, because I think it'll affect all the other stats. If you get into the habit of pushing AI through your gateway, you can do a better job of discoverability. You can do a better job of governance. There's lots more you can do once you've got the gateway in place.

So I hope you're right on that point.

00:22:56 **Tom Eggleston**

So you heard it here first. Neil reckons 35% for AI gateways 12 months from now. Okay, let's wrap up with that. Thanks, Neil.

00:23:04 **Neil Evans**

Yeah. Thanks, Tom.

Continue exploring agentic AI security

If you'd like to explore the findings in more detail, you can read the full State of Agentic AI Security report here: https://proofid.com/report/state-of-agentic-ai-security

Want to go further? Our Agentic AI hub brings together more practical guidance and resources on securing agentic AI: https://proofid.com/solution/agentic-ai