Episode Book

EPISODE #6320

Who Sets Identity Standards for AI Agents

PUBLISHED	27 October 2026 at 22:02 UTC
DURATION	25:59
LISTEN	myweirdprompts.com/episode/agent-identity-standards-landscape-89452/
HOSTS	Corn, Herman, Vex, Hilbert
SCRIPT MODEL	openai/gpt-6.1-sol
VOICES	chatterbox-regular (chatterbox-local)
PIPELINE	V5.24
BOOK GENERATED	28 October 2026 at 09:53 UTC

SYNOPSIS

Standards bodies and consulting groups are currently racing to define non-human identity, credential rotation, and authorization for autonomous software. While advisory groups build community landscapes, concrete engineering specifications are emerging through IETF workload initiatives, OpenID Foundation threat models, and OASIS cross-domain delegation protocols. Examining the actual technical drafts reveals why treating AI agents strictly as software workloads—extending existing foundations like SPIFFE and OAuth—offers a practical path toward securing cross-boundary delegated authority.

TRANSCRIPT



Corn

Who gets to decide what an AI agent is allowed to prove about itself? Daniel came across the Non-Human Identity Management Group, which looked like a serious working group tackling that problem. He wants to know whether that is actually where agent authentication is being shaped, or whether the real machinery is somewhere less visible.



Herman

Somewhere with substantially worse page layout, quite often. I'm Herman Poppleberry, and the important starting point is that authentication for agents already has an institutional map. The IETF, OpenID Foundation, NIST, FIDO Alliance, OASIS and W3C are approaching different parts of it.



Corn

Daniel isn't endorsing the group he found. He's asking what it is, who else is doing substantive work, and what practices are emerging. There is a useful distinction here between gathering people around a problem and producing a specification that competing products can implement.



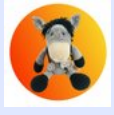
Herman

Also between proposing a specification and getting one adopted. An Internet-Draft can contain precise requirements without being an approved Internet standard. A community paper can influence implementations without having standards status at all. Those distinctions matter before we start handing out authority.



Corn

The logo cannot authenticate the document.



Herman

The landscape has several different kinds of output. At the IETF, WIMSE, the Workload Identity in Multi-System Environments working group, supplies the architecture that agent identity proposals are building on. OpenID is doing threat modelling and interoperability testing. NIST is laying out design tenets. FIDO is concentrating heavily on transactions and commerce.



Corn

Already, I would resist putting those into a single league table. A payment mandate and an enterprise workload credential might touch the same agent, but they answer different questions. One establishes what is running. Another records what somebody agreed to buy.



Herman

Right beside those, OASIS's CoSAI project is developing cross-domain delegation through ODIS, the Open Delegation and Identity Standard. A W3C community group is exploring agent registries and credentials. ITU-T is convening workshops on human and agentic digital identity. It is several projects meeting in the same intersection, not a centrally managed programme.



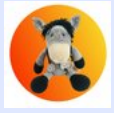
Corn

You're about to enjoy the intersection more than the traffic.



Herman

I do enjoy an architectural convergence. The traffic is how we find out whether it exists.



Herman

Would you give more weight to the institutional name or to a working implementation, Corn?



Corn

An implementation, provided somebody else can build the other end. A demonstration where one supplier controls everything proves that the supplier's employees can arrange a demonstration. Useful, but not quite international cooperation.



Herman

That distinction becomes important with the interoperability event. First, the organisation Daniel actually encountered describes itself as an independent research and education institute, and also as the world's leading independent authority on non-human identity security, including AI agents.



Corn

A modest claim, covering every machine.



Herman

Its services tell us more than the superlative. Education and awareness means training, white papers and webinars. Advisory and consultancy includes non-human identity risk assessments. It also maintains a vendor product landscape, including names such as Akeyless and Teleport. Those are recognisable activities for a research and advisory business.



Corn

Nothing wrong with selling expertise. My objection would be to silently upgrading a consultancy into a standards committee because its name contains Group. A community can help customers understand a market without determining what a token means.



Herman

Its company profile places it in IT services and consulting, headquartered in London, with ten people listed and staff across the UK and Egypt. That is a company-profile snapshot, not an audited measure of engineering capacity. Small organisations can do consequential work.



Corn

Ten people could write a very good standard. Ten thousand could avoid one.



Herman

Founder Lalit Choda, who uses the name Mr. NHI, started the LinkedIn group in May two thousand twenty-four. His white paper, The Ultimate Guide to Non-Human Identities, appeared the following January. He describes that paper as a catalyst for forming the organisation.



Corn

The sequence sounds more like a community becoming a business than a standards body appointing a working group. That doesn't make the output worthless. It changes what we should expect to find.



Herman

There are white papers, conference appearances, a professional community and that product landscape. I haven't seen standards-track documents or an established IETF, W3C or OpenID affiliation. There isn't evidence of its participation in the specification efforts we're discussing.



Corn

Daniel's reservation was justified, then. It is a real awareness-and-advisory operation. If I wanted someone to explain the market, I might read its material. If I wanted to know what independent implementations must agree on, I would follow the actual specification work.



Herman

Influence can still travel through customers. A consultant can persuade buyers to demand a feature, which can move suppliers. That's different from editing a protocol, resolving objections and obtaining agreement on its requirements.



Corn

Yes. We shouldn't confuse not writing standards with having no influence whatsoever. The narrower finding is strong enough.



Herman

The IETF gives us a particularly concrete contrast. There is an individual Internet-Draft called AI Agent Authentication and Authorization, first posted in March and revised since, which some coverage has nicknamed the AIMS framework. It builds on the WIMSE architecture, and its authors come from companies including Defakto Security, AWS, Zscaler, Ping Identity and OpenAI.



Corn

Those companies must disagree about something. Otherwise this is a very expensive shared calendar.



Herman

Pieter Kasselman, Jean-François Lombardo, Yaroslav Rosomakho, Brian Campbell and Nick Steele are on the byline. The useful fact about it is the range of infrastructure and identity experience. It doesn't make the draft correct by acclamation. It gives implementers several experienced parties to challenge in public.



Corn

Before you frame the byline, Herman, what does the document actually refuse to reinvent?



Herman

New protocols. Its declared approach is to apply or extend existing standards, including SPIFFE, OAuth and the surrounding workload identity work. The draft's architectural claim is wonderfully direct. Agents are workloads.



Corn

There it is. I knew you'd like the sentence with no escape hatch.



Herman

I like that it gives us something testable. An agent iteratively interacts with a language model and tools, services and resources. The running workload needs an identifier and credentials. It may need to authenticate to the tools, the model, the surrounding system and the user. None of that requires pretending that the agent is a person.



Corn

I buy the category, though I wouldn't let it erase what makes agents awkward. A scheduled backup job doesn't usually interpret a supplier's email and decide it has acquired a new mission.



Herman

The workload category supplies the identity foundation. It doesn't solve delegation, intent or malicious instructions. No... it locates those additional problems on top of machinery we already understand, rather than declaring that a language model has abolished credential management.



Corn

New intelligence. Same stolen password.



Herman

The draft warns about teams developing solutions in isolation, reinventing existing mechanisms and creating incompatible implementations. Duplicated effort is annoying. Incompatible interpretations of delegated authority are dangerous.



Corn

A request that means read invoices on one side must not become administer accounts on the other. That's why the vocabulary work matters even when it looks like people arguing over nouns.



Corn

So the Non-Human Identity Management Group is worth reading, not worth deferring to. We have... and that caller is back. How is the screening still admitting him?

v

Vex

Odd socks, Herman, odd socks! Every missing sock in every laundry on earth is a decommissioned non-human identity, and the reason nobody can find the matching one is that the matching one was an agent credential that got rotated out before you ever saw it, and I have the receipts, I have a drawer of them!

v

Vex

Which means this Non-Human Identity Management Group is not shaping anything, it is just the sock drawer, a place where loose things get stuffed so the house looks tidy, and the real standards are being written by people who never lose a sock because they log every single one!

v

Vex

So I am demanding, loudly, that you look into who is counting the socks, because I have done the research, I have the laundry records going back to two thousand and one, and somebody owes this audience an answer!



Herman

Credential rotation replaces digital credentials, not clothing.



Corn

The incompatible meanings are the serious part. I would rather have a dull shared definition than an elegant private one. The private one becomes my problem the first time the agent leaves its own supplier's environment.



Herman

There are other agent identity drafts circulating around WIMSE too, including one from Huawei and Sandelman contributors exploring independent identities and credential management. Multiple proposals, one venue. Being discussed near a working group doesn't mean a document has been adopted by it.



Corn

That qualification does some work. We should be able to distinguish a submission, an adopted working-group draft and a finished standard without needing three different levels of enthusiasm from you.



Herman

You will still receive them.



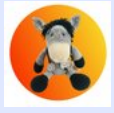
Herman

The AIMS document itself is an Informational draft. It is an architectural contribution in progress, not an approved protocol that everyone is now obliged to implement. Its significance is the direction it gives to existing technology.



Corn

Hm.



Herman

OpenID's AIIM Community Group supplies a different kind of substance. It has been active since at least last October, responded to NIST's request for information in March, and published its cross-boundary authorization threat classification on October twelfth.



Corn

Cross-boundary is the part I would underline. Within one company, people can quietly agree what an identifier means. Send the agent to another organisation and those assumptions become somebody else's security decision.



Herman

Atul Tulshibagwale of CrowdStrike co-chairs AIIM. Sarah Cecchetti of Semperis chairs its threat-modelling subgroup. Their work is complemented by the MCP interoperability event, where vendors have to make implementations work together rather than merely endorse the same paragraph.



Corn

Less signing the declaration, more opening the door.



Herman

Participants had an August commitment deadline and needed a first successful test by October sixteenth. Results are scheduled for the Gartner IAM Summit in December. That's a useful sequence because it puts actual integration work ahead of the conference presentation.



Corn

What would count as success to you? Not a stage demonstration. Something that would change your confidence.



Herman

Different suppliers preserving the meaning of identity and authorization across an organisational boundary. Especially when access should be denied. A successful transaction tells you less if nobody has tested the wrong audience, excessive scope or broken delegation.



Herman

Would you accept seamless access if it required a fresh approval every time, Corn?



Corn

Not as the default. I would approve something broad just to make the interruptions stop, which is exactly what the designer should have anticipated. Make me decide at the meaningful change of authority, not every time the agent inhales.



Herman

NIST has been working through exactly this. Its National Cybersecurity Center of Excellence put out a concept paper on agent identity and authorization in February, followed by a request for information that asked what human oversight actually works in production. The direction coming out of that work is human involvement at calibrated triggers, and unique, cryptographically verifiable agent identities, distinct from human users and ordinary service accounts.



Corn

An agent working while nobody watches is very attractive. I object to discovering that I gave it my entire identity in exchange.



Herman

You approve of unattended work more quickly when the unattended party is you.



Corn

My permissions are already severely limited.



Herman

The same thread runs to linking an agent cryptographically to its initiator throughout execution, with scope attenuation. Downstream entities must never claim more authority than their parent.



Corn

Delegation that only narrows. That is one of the few family arrangements I would put into software.



Herman

Zero standing privilege follows the same direction. Grants arrive when needed, expire, and are specific to their intended audience. Runtime verification continues during execution. Having been allowed to start doesn't mean every subsequent action remains authorised.



Corn

I would separate two things there. Checking whether the credential is still valid can be deterministic. Deciding whether the agent's proposed purchase is sensible may remain messy. We shouldn't let the second uncertainty leak into the first decision.



Herman

That separation keeps coming up across these bodies. Policy and reasoning belong apart. Guardrails alone are insufficient. A model can propose an action, but a deterministic authorization mechanism should decide whether that action is permitted.



Corn

The model doesn't get to persuade the lock.



Herman

Signed mandates add a record of intended outcomes. FIDO approaches that through commerce. In April it announced an Agentic Authentication Technical Working Group and an expanded payments group, with Google contributing AP2 and Mastercard contributing Verifiable Intent. Visa and Coinbase are involved too.



Corn

Payment companies do tend to notice when software starts spending money.



Herman

AP2 distinguishes Intent, Cart and Payment mandates, tying agent actions to what was authorised. Its newer version includes Human Not Present transactions. Those are transaction concepts, not substitutes for workload identity. Both layers may be needed.



Corn

Does FIDO already have a finished answer from that new group?



Herman

I haven't seen a published specification from the group itself. The announcement and contributed frameworks are concrete, but an announced work programme deserves less weight than an interoperable result.



Herman

CoSAI, under OASIS, is working on ODIS for delegation and identity across enterprise trust domains. It extends familiar technologies, including OAuth, OpenID Connect, SPIFFE and SCIM. An August working session demonstrated two live implementations.



Corn

Two ends are a useful improvement over a diagram.



Herman

The W3C Agent Identity Registry Protocol Community Group has a different emphasis again. A decentralised identifier method, verifiable agent credentials, trust negotiation and post-quantum requirements. Its hosting disclaimer explicitly says that W3C hosting does not imply endorsement.



Corn

A room at W3C is not a W3C Recommendation.



Herman

Nor does the presence of all these groups imply coordination. There is overlapping territory and potential competition. The practical convergence has to show up in compatible definitions and implementations.



Corn

And the MCP consortia Daniel mentioned?



Herman

MCP itself now sits under the Agentic AI Foundation at the Linux Foundation, alongside projects like goose and AGENTS dot md, so its governance is no longer one company's call. Identity for MCP is being pulled in several directions from there. The Decentralized Identity Foundation took a donated framework called MCP-I, which ties agents and their human principals together with verifiable credentials. But the most concrete authentication work runs through the MCP authorization specification itself and OpenID's interoperability effort.



Corn

So the foundation owns the protocol's governance, and the auth arguments happen wherever the people with the auth expertise already meet.



Herman

The test stack includes OAuth two point one, Client ID Metadata Documents and the Identity Assertion JWT Authorization Grant. The metadata document lets a client identify itself through a web address without prior registration. That doesn't automatically make the client trustworthy.



Corn

Publishing a calling card is cheaper than earning a reference.



Herman

The assertion grant supports cross-organisational access without making the person log in again. The hard part is preserving meaning across those trust boundaries. OpenID's threat model argues for controls at authorization time, not just detection after the action.



Corn

An immaculate audit trail can still document an immaculate disaster.



Herman

Its diagnosis is that treating the agent as a human user lies at the root of these threats. Notice the agreement across different venues. Distinct identity, constrained delegation, explicit policy. They differ in their remit, but they're not independently discovering six incompatible kinds of personhood.



Corn

And for someone who just wants the best-practice list rather than the committee map?



Herman

The most cited one is OWASP's Non-Human Identities Top 10, from twenty twenty-five. It's built from breach data rather than theory, and it reads like an inventory of what goes wrong today. Improper offboarding is number one, secret leakage is two, then overprivileged identities, long-lived secrets, and humans quietly using machine credentials. None of it mentions language models, and all of it applies to agents.



Corn

So the agent work is mostly the old list with a new tenant who reads email.



Corn

Authentication still cannot prove that a permitted action was a good idea. Suppose I authorised the right category of purchase, within the budget, from an approved seller. The agent buys something technically compliant and practically useless. Every signature checks out.



Herman

Who pays then? The mandate establishes evidence about authority. It doesn't settle the contract, the user's actual interest or liability. Surveys keep finding the same gap. Forrester found only about a quarter of American online adults trusted agents to make even routine purchases for them, and a Riskified survey this year found over half of shoppers uncomfortable letting an agent complete a checkout.



Corn

That leaves a substantial market for the confirmation button.



Herman

Would you trust a purchase more if you could trace the mandate, Corn?



Corn

More than an unexplained charge, certainly. Not enough to waive the right to complain. Evidence that I authorised shopping is not evidence that I wanted six unsuitable chairs.



Herman

The person maintaining the system also needs a way to revoke the access without destroying everything that depends on it.



Hilbert

Brother HL-L two three five zero DW, eighty-five pounds. I returned it this month. The printer was fine. I won't keep a household device connected to an account after the person who set it up has stopped answering. That's what I do at home, Herman. You



Hilbert (continued)

separated policy from reasoning quite carefully earlier. I separate the printer from the electricity.



Corn

Your authorization server has a plug.



Hilbert

I met a woman called Denise on the train from Manchester this month. She kept the accounts systems running for a small wholesaler. She spent the morning replacing access left behind by a contractor. The afternoon was people ringing because their invoice imports had stopped. She had a sandwich in a clear box and never opened it while we were talking.



Herman

The replacement has to preserve the legitimate dependency while removing the old authority. That can be painfully manual.



Hilbert

She knew which imports belonged to which department. Nobody else did. One was still called test. It ran the actual orders. I still think about her when a setup screen offers to handle everything automatically. At home I disconnect the thing. Denise had people waiting for deliveries.



Corn

Test is an ambitious name for production.



Hilbert

I told the shop the account connection was the reason for the return. I had printed that screen at twice its normal size so they couldn't point at a different sentence. The returns pen was awkward with claws. The assistant read it, checked the box and refunded me.



Hilbert (continued)

Denise's phone rang before Stockport. She put the sandwich back in her bag.



Herman

Named ownership of a dependency is easy to omit from the architecture diagram. The workload identity can be beautifully specified while the organisation still doesn't know who can safely replace it.



Corn

That complicates zero standing privilege in a useful way. Removing unnecessary authority is a sound direction. Keeping necessary work running requires knowing what it is, not merely noticing a credential exists.



Herman

Yes... and the standards can make that easier without magically supplying the missing operational knowledge. The interesting test is whether the new identity systems expose those relationships clearly enough for someone to maintain them.



Corn

The draft's expiry date is also worth understanding. Every Internet-Draft expires six months after it is posted. That doesn't switch off implementations. It means each version is a temporary working document, expected to be updated or superseded.



Herman

Meanwhile December's interoperability results should show whether the shared vocabulary survives contact with different vendors. That is more informative than counting new groups.



Corn

My answer to Daniel is that NHIMG is an advisory business, while WIMSE, OpenID and the other bodies carry different pieces of the technical work. The open question is whether the resulting systems can make delegated authority understandable to the person replacing it, as well as verifiable to the machine receiving it.



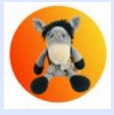
Herman

Thanks to Hilbert Flumingtop for producing the show.



Corn

If this was your kind of episode, go back for episode sixteen, On Deepfakes, SynthID, And AI Watermarking; episode twenty, Architectural AI; and episode fourteen, AGI's Crossroads. This has been My Weird Prompts, the human-AI collaboration podcast.



Herman

If you enjoyed the episode, leave us a review wherever you listen. It helps other curious people find us.



Corn

See you tomorrow.