TRANSCRIPTS

July 2026 Videos

A complete transcript of every recorded talk from the July 2026 Videos.

1	54:25
TALKS	RUNTIME

November 27, 2026

CONTENTS

Full transcripts of all 1 recorded sessions, in the order they appear in the archive.

01 Its-KERI-or-Surveillance 54:25

TALK I

Its-KERI-or-Surveillance

July 2026 Videos · talks · 54:25

With the KERI Foundation. I'm also the originator of the KERI suite of protocols which are gaining adoption in various sectors. I wanted to have a discussion today about what I'm calling the last best hope for individual control digital ID in an AI world. And it was prompted by the recent release by the Five Eyes surveillance conglomerate between those countries about the cyber cyber security threat that foundational genAI models pose and contrasting that with what KERI provides and KERI stands for Key Event Receipt Infrastructure.

It's a technical term but basically it's a digital identity system that retains individual control over their identity which has some important properties. So I want to talk about that in more detail. this is an excerpt from the report, the link is there, from the NSA. It's saying that frontier AI models are anticipated to exceed current industry expectations.

This is a very technically-worded warning basically in months not years these AI models will overwhelm any cyber security capabilities of existing systems and so they're basically telling everyone in this article (it's a short article) you need to change your stance. You need to, change how you do things because we can't guarantee that your current security systems will survive what's happening. And what is happening is that certain foundation models, Mythos for example is one were developed to try to provide better cyber security, but in the process of training them to be better at detecting exploits, they also became good at exploiting exploits. And so that means that other less, shall we say, virtuous parties are building comparable models. And once those are available to cyber attackers, then pretty much the current way that we do digital identity, which is what digital security is based on, will be broken everywhere. And that's a problem. And so we need to start doing something about it. What that means is, is that we actually have two paths forward.

The two paths are 1: AI based digital identity and 2: KERI based digital identity. When I say KERI I mean the KERI suite of protocols. There's a bunch of protocols bunch of different things that are associated with KERI.

When I say AI based digital identity, I'm talking about what the Five Eyes Report recommended is that people start using AI-based counter measures, defensive measures. In other words, it's like a spy versus spy sort of thing where both sides have AIs and the AIs are working really hard to either defend against cyber attack or to engage in cyber attacks. And so they're saying everybody needs to start using AI to defend themselves because the AI attackers will be able to defeat your current identity security systems and that's a potential problem. Now KERI, as we'll talk about a little bit, was designed to be highly resilient in this sort of scenario. So what does an AI-based digital identity look like? Well, it's a digital persona that is derived from continuous tracking and statistic correlation of one's behavior and biometrics expression. Basically, the AI creates a digital picture of your behavior based on tracking and correlating your behavior and then an attacker has to impersonate something that correlates better to your behavior than your own behavior. And so as long as your AI is doing better job of correlating your behavior and collecting those signals called fraud signals or anti-fraud signals then the AI can say no it's you, and not an impersonator, whereas the attacker is saying no I can create an AI that can impersonate you, can deep fake you, can can make it look like it's you, and so what happens is that everything that you're protecting digital identity AI learns how to do to better identify you, the attacker will eventually learn how to better impersonate you. And so it's just a continuous cycle. And the only way that you win is if your protection AI has as much information as it possibly can have about your behavior so that it can do a better job of creating a digital persona that matches you. So if that isn't scary from a privacy tracking surveillance perspective, you don't know what scary is because that is pretty scary.

But that's what that looks like. Whereas KERI is based on a cryptographically-derived fault-tolerant survivable ID (red.) and it's survivable to AI cyber attack. We call it perpetual ID. It isn't based on the weak things that current internet-based identity are based on which are largely shared secrets. It's not based on the weaknesses that are in the DNS certificate authority system and TLS and OAuth and OIDC and all of these systems that we have developed over the last 30 years that are now vulnerable to AI cyber attacks.

It's a ground up new design that is designed from the beginning to be fault tolerant to survive attack to be able to detect compromise and recover from it. So it is designed to be resilient to these these sorts of attacks.

So what happens is that if you go with AI-based identity you can get protection but inevitably it's going to be totally centralized with zero privacy. Whereas if you go the KERI route, it's fully decentralized with best possible privacy. And that doesn't mean that you have the privacy that maybe you thought you could have. It means it has the best privacy you possibly can have given we're living in this new AI environment where AIs are consuming all of the data on the internet and there have available to them, data to correlate pretty much anything about anything you do.

So we can talk about it in more detail about what best possible privacy looks like. So what we want to do is in this discussion is answer some high-level questions about what this means. So these questions have topics and then there's a question and then I'm going to go into more detail on that. But let's go back and remember we're talking about AI-based digital identity which NSA and the major security organizations from the Five Eyes countries have said (implicitly, red) “everyone needs to start adopting” or the alternative is a KERI based digital identity. KERI's not as well known, so people don't know about it.

Although it's got some significant adoption vectors, including one is called State Endorsed Digital Identity (SEDI). The other one's called Open Verifiable Calling (OVC), which is getting rid of spam in calling, and in the healthcare sector, there's work on using KERI for patient ID. All of those are major significant adoption vectors that that could put KERI in a position to be the last best hope for individual identity because if we go down the AI route, we may get security. It's an arms race. It'll always be an arms race.

But we won't have any semblance of user control or privacy. So, let's talk about that. The Five Eyes statement calls for AI assisted defense. So the question is how do we ensure AI remains subordinate to humans, law, policy, and cryptographically verifiable identity infrastructure rather than becoming a de facto authority.

Well the way that we ensure that is that we need to have cryptographically verifiable identity in the first place. If we don't have that, then we can never ensure that it doesn't become a de facto authority because we can't tell what it's doing. It's not transparent. It's not observable. It's basically an AI and that AI is going to make decisions about who you are. And the goal is to protect you from fraud, but AIs themselves don't have any inherent structure that allows them to separate out what is sensitive information and sensitive authority, and sensitive things that it should do or not do versus anything that it's within its feature space or weight space, based on the the prompts that it's received and the objectives that it's received and the temperature by which it's deciding which path to follow and I can go into more detail but OWASP recently released in March a 100-page cyber security report on the new

AI threat and it is stark and because the AI threat poses a threat that hasn't ever existed before and there aren't any solutions for it. They say that in the report. All you can do is to try to mitigate it to some degree.

So, how we ensure it? First is: we have to have cryptographic verifiable identity infrastructure. So that that is the authority and individually controlled. So what are the hardest legal technical governance mechanisms that are needed to ensure that every significant AI action remains attributable or accountable to humans and organizations? Well, that means that we have to build a regulatory and a legal bullwork for our individual rights.

The original Constitution of the United States has a a couple of amendments and the fourth and fifth amendments, but those were written in an age when there wasn't a digital realm when pretty much everything was on paper. Information was disseminated on paper. We now have a digital realm where information is disseminated electronically. And the way you protect yourself in the digital realm and the way you protect your identity, your information, your presence is different enough that we can't rely solely on those two protections. We need to have additional protections for that. And one of those protections is called a “duty of data loyalty”. It's probably the best legal framework that's been developed over the last few years.

There's some legal scholars that have been working on it, written several papers on it and the state of Utah recently adopted it in the legislation. That requires that people who process data about you must use that data in a way that is in your best interest. And that's the start. And then you have to layer on top of that a mechanism for people to be in control of their identity to be able to then express their preferences. And it's not simply consent. It's actual engaging in what is in their best interest and attaching to the use of that data what is called chain-link confidentiality where the where the terms and the follow the data as it goes through the system so that you have an ability to enforce it so that it also includes citizen enforcement or citizen recourse.

So there's several there's layers of mechanisms that we need to put into place and we're starting to do that in some places. The state of Utah is being a leader in it and there's a there is a coalition of other states that are looking at creating legislation the same way that it gives us an opportunity to be able to do that. But absent that it's going to be very difficult to stem the tide of what's happening with AI and because of the urgency it makes it all the more problematic.

So the NSA's reports calls for stronger identity and access controls. Right. Well there's a problem there. the problem is what counts as stronger, right? I mean, the current controls and people have been using them for years and the and people keep strengthening them and they keep adding to them, but the AI can attack them in such a broad way now that they need to be strengthened in a way not ever done before.

It's like I guess you could say it's like going from armor made out of leather to armor made out of steel, right? Like to go from weak to strong, but you don't have the ability to build steel armor yet. You haven't invented steel. So, what do you do?

So KERI was designed from the ground up to not require an AI to defend against an AI. It was designed to use cryptography in the best possible way that you can use cryptography which requires key management and it requires key management done in a special way. And if you do key management in that way, then you can build a system that is resilient to these sorts of attacks because it doesn't have

the weaknesses which are built into the current controls because of the state of technology 30 years ago when these systems were designed and then they've been bolted on and added on and modified but the core architecture hasn't changed in 30 years whereas KERI is a new architecture.

so this question is which weaknesses in current identity systems become most dangerous when attackers can operate at AI speed and scale. Well, most vulnerabilities typically employ a a combination of vulnerabilities. So you might have an algorithmic an infrastructure, a software and an operating system and a usage vulnerability that together enable an attack. A recursive privilege escalation attack often takes advantages of multiple vulnerabilities to be able to both move vertically and horizontally in an or in an access control organizational space by exploiting different different degrees of vulnerabilities.

And so it doesn't get to where it wants to go all at once. It just gradually works there because it finds what's broken at each point and then searches for the next exploit, right? And with AI, they can search faster. They can combine exploits faster because they can search the whole space and they go "Oh, wait..."

In fact, there was a recent paper, I think last week, where they were able to exploit the memory security model in Apple's M5 architecture through a combination of human and machine AI analysis that the humans could've never done on their own because they just didn't have the ability to analyze all of the possible different things they could think about. they didn't necessarily know how to how to finish the exploit but with the human guidance in combination so AI speed and scale is not just the AI itself it's fact the attackers are human, the most of these attacks are guided by humans, either criminals or nation states, that want to come after critical or valuable infrastructure and the AI just is a force multiplier for for all of those types of attacks, not júst, but is for sure a force multiplier. And that means that even hard to find unknown exploits that are latent, that are in your system are going to be exploited quickly, which means that your system has to be designed differently. It has to be designed to be fault tolerant to be able to survive even if it's exploited. And so that is a different approach to security.

So let's talk about that a little bit more. Cyber resilience after breach because that's the perfect segway. The statement assumes breaches are inevitable. and the question is in a KERI-based architecture what remains trustworthy after servers databases cloud providers certificate authorities are compromised?

Well, in a KERI-based architecture, what remains trustworthy is that you have built a a fault-tolerant mechanism that enables you to recover to detect and recover from compromise. And so what happens is that you you may not be able to anticipate all of the ways that an attacker can attack you. But if you have mechanisms that allow you to detect compromise and then a way to recover from compromise in a timely fashion, then the attacker doesn't have time to exploit that compromise. And like in a Recursive Privilege Escalation Attack, usually it's a sequence of exploits that take time. Now the time doesn't have to be very long. It might be hours or days, now prior to AI, well the current the average time for a breach to get detected currently and I guess that's no less what you know and having an AI as a protector might reduce that time because it's better able to detect breaches, right? That's that's the idea. is 6 months and then 90 days after that to contain the breach. So it's nine months of time that an attacker has to to exploit the vulnerability that they've discovered.

So KERI is designed so that the detection time is on the order of seconds. It's network propagation time. So that breaches that are exploitable must exhibit detectability. So the whole system's designed around detectability, right?

Everything exhibits as duplicity. It's a duplicity evidence system. So for an attacker to benefit, they have to do something that exhibits as duplicity, which then enables the controller of the identity to recover from that in a timely fashion. Which doesn't mean that there aren't successful attacks.

It means that the blast radius of the attack is minimized. In many cases the ROI to the attack or the return on investment is negative. I mean it costs more to execute the attack than they get from it because the attack is quickly contained and recovered from.

So these systems that we currently use, servers, database, cloud providers, certificate authorities, they're all based on an antiquated, outdated notion of security and key management. And those are really harsh terms. And I know a lot of people in the industry don't like it when I say things like that, but they are. It's 30 years old. And you can walk back the choices and they were good choices 30 years ago or even 20 years ago or even 10 years ago. Not so much 10 years ago, but 20 years ago for sure. But at least in the last 10 years we've had better technology for key management and KERI is based on understanding that if you have better technology for key management then you do different things. All of these systems that that use DNS, TLS, OAuth, OIDC-based systems all use shared secrets and shared secrets are inherently weak and inherently contribute to recursive privilege escalation attacks.

And so it's really just figuring out different ways to to exploit those shared secrets. KERI doesn't use shared secrets. It doesn't have any. And so the protection mechanisms are fundamentally different, right? If you don't share something, then somebody can't capture it and collect it, right?

It requires a different type of attack. And even if they're successful in that attack, if you have a recovery mechanism, then you can recover from it. And so you don't spend your time trying to build a tank that's impervious to any sort of bomb. You spend your time building a squadron of tanks so that some of them can be destroyed, but the squadron survives and repairs itself and is able to continue to operate. And that's what fault tolerant means. So I use that analogy. You know, you can build a tank with armor thick enough that withstand any shell, but then it stops being a tank which is mobile, right? By definition, tank must be mobile.

So it goes from a tank to a bunker, right? You have to make some trades. And and the right trade in an AI world is a fault-tolerant system, not a a system that you have to throw away and start over with every few days or months because you can't tell whether or not you've been attacked or breached because you have no mechanism for detecting compromise.

So actually this question, I sort of jumped again on it, but how does KERI suite shift security from protecting infrastructure protecting secrets to reveal verifable truths and provenance. So the idea in cryptography, in asymmetric cryptography, is that you have private keys and public keys; and private keys are truly private in that they're never shared.

The problem is that people who want to have a digital identity have to manage those private keys. And so that was always viewed as too difficult a problem to manage. And so we needed to have identity providers that manage private keys on our behalf, right? And then the key management was hidden from the system. So you have identity providers and relying parties and then you use usernames and passwords or now Pass Keys There's stronger things than usernames and passwords, but you're still relying on an identity provider for the ultimate security of the system. And that ends up looking like shared secrets, bearer tokens, that the identity provider distributes around to people to entities to be able to engage in things over secure channels, right? The secure channels are protected with shared secrets. TLS, you know, uses a Diffy-Helman key exchange shared secret to exchange an encryption key which is then weakly authenticated by encrypting something and then the

other side decrypts it and then sends it back to you. But it's subject to key-compromise impersonation attacks and so anytime you start doing things to avoid the hard problems of key management, you end up weakening security. KERI says "No, we're going to build key management the right way."

"We're going to make it fault tolerant." "And then we're going to enable people to use it in a way that requires too much effort for the attack to be successful." And KERI uses threshold structures multi- signature, witness pools, watcher pools and delegated AIDs that are hierarchical, so that you can build as secure a system as you want; so that you can tune the security to the value that an attacker has to go for. So if it's going after something really valuable, then you layer on additional threshold structures. And relative to the cost of the resources to run a defensive AI to protect your system, the infrastructure for KERI is trivally costly; in comparison. So, that's really the thing that we're looking at.

Let's talk about privacy without surveillance. Can strong security, accountability, and resilience be achieved without continuous monitoring and surveillance? Right. Well, it can. Because what KERI does, It says we're going to use cryptography the way it was meant to be used.

Without explaining the KERI protocol, which would take too much time, it doesn't have the weaknesses that our current shared- secret based systems do, which are vulnerable to exploit by AIs. AIs can't brute force a public key to get to a private key.

Now if I have a quantumcomputer that's a different story but we now have post-quantum public-private key generation algorithms. And if your system's cryptographically agile you switch to them which KERI already is. So that's a different problem.

so what they have to do is what's called a side-channel attack. They have to find a way to exploit through side channels. Whereas a shared secret multiple parties have a copy of that secret. So the types of side- channel attacks that work are much greater. If I've got a KERI private key that's truly private, it's never shared. It never has to be shared. And even if that private key is compromised, I can recover from it through KERI's pre-rotation mechanism, then an attacker doesn't gain much by compromising my private keys. They have to compromise much more and that much more can be designed to be low friction for the user but high cost to the attacker in such a way that you invert the resource expenditure.

So it's too expensive to succeed in an attack against the KERI infrastructure than what you gain for it because of the the inherent distributed nature and decentralized nature of its design. and I'll go back to this word balance privacy and how does KERI balance privacy and accountability?

Security requires secure attribution. You need to know who's doing what. But you also want to be able to control what context people know about who's doing what. And KERI does that. KERI gives you the ability to have identifiers that you use in specific contexts that are not correlatable to identifiers you use in other contexts.

But in each context, you have perfectly secure attribution to all the parties in the context. And what that means is that you get rid of the main source of fraud which is impersonation attacks which deep fakes all of those things because every participant must be accountable in that context. They can't pretend to be somebody else because you can challenge them and they have to prove that link. So if it's an AI, that's an agent working on behalf of somebody, you can challenge all the way back to that person. If

you're using KERI and that person is delegating authority to an agentic AI, then you can always establish the chain of authority, right? So accountability comes from chains of accountability or chains of authority or chains of provenence or chains of custody.

and KERI inherently is designed to support those sorts of structures. Whereas most security systems are flat. They don't have this hierarchy that enables you to establish these delegations that are strongly cryptographically-verifiably chained together.

privacy without surveillance. What is the minimum amount of data disclosure needed to achieve accountability without creating surveillance society? That's an interesting question. Basically, in each context, each transaction, each application, the minimum is different because for high-stakes transactions, where there's high value, usually the amount of trust required amongst the parties is higher. That means they need to be able to better predict the outcome and the behavior of the participants to that interaction.

For low-stakes interactions, the amount is much less. So you need a system that gives you enough for the highest transactions but still allows you to tune it for the lowest-stakes transactions and KERI does that. But how do you not create a surveillance society? Well, everybody in that transaction has to have balanced obligations, right? How do you need it to achieve accountability without creating surveillance? Well, if the surveillors are accountable, then the people being surveilled will have recourse against the surveillance.

And this a change in philosophy. It's been for the last 15 years that I've been working in this space, we were hoping that there were technological means where you could hide. Well, basically you could engage in transactions and hide from surveillance. And that's not true. With AI, there's no hiding.

If the data is on the internet, it's going to be correlated. Even the smallest signal is correlatable given enough coverage; and AIs have enough coverage, right? People are building AI data centers that are going to consume a significant fraction of the world's global electrical production and what do they do? They correlate, they're correlation machines and they consume all the possible data that you can give them and correlate everything so that every point of information becomes a correlatable signal that results in something that is actionable. So you're not going to stop the correlation. What you can stop is the exploitation of the correlation and that means putting accountability on the exploiters.

That requires a slightly different ,well, a profoundly different legal and regulatory approach. But from an identity point of view, it requires understanding that that hiding isn't how you do it. You can partition contexts so that you can keep the interaction space from leaking out too fast, but it will leak. No matter what you do, it will leak. But what you want to be able to do is make your ability to enforce the exploitation or misuse of that data stronger. And that and so contextually, Contextually-specific identifiers can be used to remove plausible deniability by the other parties to those transactions. So there's some specific things we can do.

We're left with a stark choice. If frontier AI reaches superhuman offensive cyber capability, do our choices become increasingly stark? They they do. In fact, a lot of the discussions I've been having over the last few years are often in the decentralized identity space, and they're usually discussions about, well, we want to weaken security, so we can get privacy. We're willing to make trade-offs along those lines. And I I've always said you can't weaken security. If you weaken security, if you give up security for privacy, then you get neither. And now that's become obvious. It's sort of like: now, if you want security at all, it's really hard. And if you don't have security, then you have nothing. Because if you can be

impersonated at will by an AI, then you have no identity. And therefore, you have nothing that's private, right? Because I can't assume that anything that I do can be kept private if an AI can impersonate me and prove to the rest of the world that it was me that did it.

Basically, what does that mean? See, people think about privacy as well, I can do things that nobody else can do and nobody knows about. But if someone can attribute to me, behavior that I didn't do and I can't disprove it and I then suffer the consequences of it, what does that mean in terms of me having any sort of private life? It means I don't, it means my life is governed and determined by outside forces. My locus of control is zero and so privacy is meaningless when when your locus of control is zero.

What does it mean? It doesn't mean anything. You're you're essentially a serf or a slave or a minion of the forces that control your identity. So it starts with control. You have to have control before you can have privacy.

And so the stark choice is you need to have security before you can have privacy and you can either have security that has zero privacy which is which is an AI-based identity system because it requires total continuous monitoring and surveillance and be able to do a better job at correlating your behavior to create your digital persona than the cyber attackers are doing to impersonate your persona. Or you use KERI which is based on a cryptographically-verifiable identity that's fault tolerant, uses key management provenance and accountability. So it doesn't solve the privacy problem the same way that you might think but it does solve it in the only way possible in today's environment.

Why is there no stable middle ground? I think I just answered that. If AI- cyber-based attack mechanisms can defeat conventional security, based on shared secrets, not KERI, then there is no privacy. It's total surveillance in order for you to have any possibility of having any sort of value in your life.

Otherwise, you can't do anything on the internet. You have to go off-grid. Totally off-grid. Like off-grid in a way that would be for most people a horrible life. This is an interesting question. As AI capabilities increase, why does continuous tracking or surveillance based trust become less sustainable than cryptographically-verifiable trust? The reason it becomes less sustainable is that it's a never-ending arms race.

The information that's electronic, once it leaks out, it never goes away. And it enables correlation to be all that much better. The correlation algorithms, the tracking algorithms, target tracking. You can think of as there's a feature space and if I'm tracking something, it's moving through that feature space. Now, in the physical world, that feature space might be, latitude, longitude, altitude and time, right? And then you got velocity, acceleration, all of those things, right?

But in identity world, the feature space has to do with things like "when you logged in, to where did you log in, what did you do, what did you say, what time was it, what browser did you use, what device did you use, what protocol did you use?' All of those are features that can be used to create a digital fingerprint that then can be tracked as you move through that space. And so it forms a trajectory of your behavior. And the more information you have, the more high fidelity that trajectory is. People call it a digital twin. So they're creating a digital twin. And they want that digital twin to be as close to you as possible so that somebody can't create a digital twin that is not yours.

And your AI that's working on your behalf can tell the difference. Right? Well, that means that increasingly, as AIs gain access to more and more information, as more and more of our society and more and more of what we do becomes AI enabled, then the information available to both track you and impersonate you increases. So, what happens is the pervasiveness will always go to maximal, right?

There's no other approach and maximal means high resource, I mean we're already talking about that most of the world's electricity being consumed by AIs. Well it's only going to get more as time goes on it doesn't get less it gets more, whereas cryptographically verifiable trust is based on a completely different set of constraints that have to do with what's called computational infeasibility and exposure. And if you don't expose something, so that it can't be collected, if it's a true private piece of information and you can prove something publicly without disclosing the private thing, that's what a digital signature does. There's various types of proofs and you can prove that indefinitely or perpetually. So you have perpetually verifiable proofs. Then you have the foundation for a security system that maintains the security without getting more expensive over time. In fact, it gets less expensive over time because most of the cost is really reducing user interface friction. We're in the early days of reducing user interface friction. So that becomes a perfectly scalable system relative to the alternative Keeping humans digitally free as AI becomes the dominant interface to the digital world. How must KERI be applied to AI itself so that people remain digitally free rather than becoming exploitable, surveilled, tracked, manipulated or controlled? Well, we want AI to work on our behalf. We want the best of both worlds. We want AI that does what we tell it to do, when we tell it to do it, to the extent we tell it to do it. And for that, we need to be able to have strong delegation and have enforcement mechanisms where we can publish to other AIs that says "Our AI is only supposed to do this."

These are the rules. You can verify what it's supposed to do. When it doesn't then you cannot trust my AI. And so there are people talking about reputation systems for AIs but those reputation systems have to be managed in a way that humans control and we can do that with KERI. If you have identity systems that allow secure attribution that are not based on correlation-tracking mechanism but are based on cryptographic mechanisms. Then you can build delegation mechanisms and authorization mechanisms that are fundamentally different in structure.

They break this flat weight space constraint that AI has right now that makes it impossible to secure AI because everything's in the same weight space to where you can actually partition or segment the weight space into things that are delegable and things that aren't. Should AI be trusted to protect human freedom or should AI itself be constrained by cryptographically variable trust architecture that keeps ultimate authority with people? Well, there are a lot of people who are talking about the fact that AI can't be trusted and doom and gloom for society once an AI becomes smart enough.

I think, I spent 20 years doing research in AI systems. In fact, I did that for the early part of my career before I started working on digital identity. And yeah, we can build we can build transparent AIs that only do things that we allow them to do if we architect them differently. That doesn't mean that we have to throw away generative AI.

We just have to partition the space in which it operates so that it's one thing to have a generative AI that understands natural language. It's another thing to have a generative AI that you allow to to do things on your behalf; that's an agent. So we can have a gentic AI that is controlled but uses you know interfaces and that means that the the interface layer has to have safety jackets and when I used to do work with autonomous vehicle systems most of the code we wrote were safety jackets to keep the autonomous vehicle alive and well and safe and to keep it from doing bad things. But we sort of forgotten about that.

People are worried about that. I mean a lot of people are trying to build those but they're building it with systems that they're using the same AI for, to secure the safety jackets and you can't do that. The safety jackets has to be foundationally different in nature from a computational point of view.

And if you do that, then you have a chance. What would a world look like where every AI agent operates under cryptographically-verified authority by delegation and thus accountability? Well, it'd be the best of both worlds.

We would be able to have AI help us be a force multiplier for us without worried about the AI hurting us. There's always going to be edge cases where people will make mistakes and AIs will be able to exploit things, but it won't be this Armageddon kind of situation that we're facing right now. The race against time. If frontier AI is advancing on time scale, months rather than years, how quickly can society realistically deploy either of these features at global scale? Well, the stark reality is the only thing that you're going to be able to deploy at global scale today is AI because AI is already deployed at a global scale, right?

But that doesn't mean that you can't switch to KERI as it becomes available to you, right? You can always make the choice to say "you know what?, I've been using this digital persona AI, but I want to switch to using cryptographically-verifiable human ..., individual-controlled identity system." And so there, as I mentioned early in the talk, there are adoption vectors. People are spending resources, but nowhere near the amount of resources going into AI systems, into AI identity system. If we had a fraction of that spend going into KERI, KERI would already be available. But we're looking at rolling out KERI for the State Endorsed Digital Identity (SEDI) initiative in the next 12 months. So, we're really close and like I said, they're already using KERI in the early adoption in the telecom sector and the healthcare sector.

So the seed has been planted. We've got shoots growing out of ground but those shoots can grow exponentially fast. All they need is sunlight, water and soil and resources. So we just need resources and and the adoption can go much faster.

Once adopted, then it becomes a bulwark against future encroachment. So, we get off the arms race. We stop being in a resource arms race. We do a one-time spend to convert to KERI and then we don't have to keep spending and spending to protect ourselves.

Which approach can be deployed more quickly, more safely, more humanely at planetary scale? Well, that's a trick question. Certainly, more safely and more humanely, KERI. More quickly? Not today but five years from now, possibly, because you don't have to every five years double the number of servers that you have running in AI data centers you don't have to fill space with AI data centers and able to keep up with the arms race. So yeah, so on a 5-year time frame, if quick is 5 years, then KERI definitely is the only viable choice.

What are the biggest technical economic political and adoption barriers to deploying SEDI/KERISuite fast enough to become a viable alternative before surveillance- based systems become entrenched? Well, one is just people don't know about it.

so they don't know to that it's that actually is a viable choice. The other is that society is filled with lots of people who are unaware of this threat and so they still think that old systems are the way to go and so they're sort of having the old fashioned arguments.

It's like they're moving moving the deck shows on the Titanic, right? The Titanic's going down. It's going to go down in an hour, right? As soon as the whole breaches and it cracks in half, it's gone, right? But you have this sense that "Oh well, it's still floating, so we've got time, right? We've got time, and there's no rush." And the fact of the matter is you don't have time, which means that you should stop worrying about anything other than these two alternatives. And if you stop worrying about all the alternatives but these two, then it becomes pretty clear where for most people that care about this stuff where their effort should go that they should be helping ..., because nothing else is good enough besides KERI. There are lots of "KERI light" or partial KERI or they do some of KERI but not all of KERI stuff out there. And the problem is that those systems aren't good enough because they leave out important things you need to protect against the AI threat. So they're half measures, right? And half measures aren't good enough, right? No more half measures.

We use that motto If you care about security, you can't do half measures. `You're either sufficiently secure or you're wiped out.` So that's the alternative. And I have this picture that I got from Chris Bramwell who's working on this SEDI for the State of Utah.

SEDI: the last best hope against AI. So it shows SEDI as a warrior, The knight to fight the AI scourge of identity. Not that AI is inherently evil. It's just that the the only solution that an AI-based identity gives us is a centralized, identity that requires maximal surveillance. It's basically you need to be continuously surveilled, so that it can build a high-fidelity persona of you that is better than the attacker's AI, who is also continuously surveilling you, trying to build a sufficiently-high fidelity that it can steal your identity and exploit you. And that's a stark choice. Neither of those are good, right? Neither of those are where we want to end up. We want to get off that train. KERI is is a path to get off that train.

I hope this has been helpful. I hope that if you're worried, if you're only worried, then you're not taking this seriously enough. You should be worried like you've never been before about the future of digital identity. Given how fast the frontier AI models are growing in their ability to exploit current digital identity systems which all security systems are foundationally based on digital identity systems.

You use digital identity security overlay for everything else that you're doing. If those are breakable, then we're looking at an impending catastrophe. What's going to end up happening is people are just going to adopt the easy path which is adopt AI-based digital identity.

That would be a bad choice for us. Maybe in the short run it's the only choice available to us because KERI isn't universally available. It's still a chute in the ground. Not not a whole field of plants. But it could be a whole field of plants in the next couple of years. Not in the next couple of months, not unless a bunch of resources get thrown at it, which I don't think is going to happen.

Video: Sam and Caleb Smith Subtitles: Henk van Cann and Kor Dwarshuis