**SINA**

When you think about machine learning or AI using databases, one of the most obvious questions is what's in those databases? And do we want to expose that to in the context of our cyber unaffiliated nodes in this context to machine learning, you can use encryption, to make data unavailable to the general public that may try to look at that data. But at the same time, if you put it on something like Aleo, you can prove things about that data that are true or false, or that fall within certain parameters. And leverage that data without actually giving away identifying information.

**CRAIG**

Data privacy continues to be a challenge for companies deploying AI, particularly the privacy of training data, which constitutes some of the most valuable IP for companies as algorithms and compute become increasingly commoditized.

**CRAIG**

This week, I talk with Sina Kian, chief operating officer of Aleo, a platform for building private blockchain based applications. Aleo supports zero knowledge proofs, a powerful tool to verify the correctness of a computation without revealing the input data, or the computation itself. This makes Aleo particularly useful in machine learning, where sensitive data such as medical records or financial information needs to be kept private but has tremendous potential for training AI systems.

**CRAIG**

Before we begin, I want to mention our sponsor NetSuite, Oracle's cloud-based enterprise resource planning software to help any business manage their financials operations and customer relationships in a single platform. For the first time in NetSuite 22 years as the number one cloud financial system, you can defer payment on a full implementation for six months; no payment, no interest for six months for a full implementation of Oracle's NetSuite. To take advantage of this unprecedented financing offer, go to www.netsuite.com/eyeonai. That's EYEONAI all run together.

Now here’s Sina.

**CRAIG**

Let’s talk about Aleo. I've had a little bit of a look, but why don't we start by having you introduce yourself and then introduce Aleo. And then I'll get into some questions.

**SINA**

Okay, great. So, my name is Sina Kian. I'm General Counsel and COO here at Aleo. I graduated from law school, after law school, I clerked on the DC circuit for two years and then the Supreme Court for Chief Justice Roberts. I practice law at a firm called Wilmer Hale here in DC for a couple of years. Then I changed gears, actually went into the finance world, worked at Blackstone for about four years. And during that time, in an investment capacity, I got very interested in the technology behind blockchain generally and I met the folks over here at Aleo and got very interested in what they're doing. I joined initially as VP of strategy and have since been promoted to COO and General Counsel. So happy to talk about Aleo itself.

As background, I was also an advisor to the Privacy and Civil Liberties Oversight Board, so long been very interested in issues around privacy and how they work consistent with sort of national security and other public policy goals that we have. So Aleo. If you take a step back, if you think about blockchains, generally, if the internet allowed unaffiliated notes, which is to say computers, phone phones, other nodes like that, to communicate with each other, even though they're unaffiliated, and have no reason to trust each other, blockchain sort of took that a step further and allow those unaffiliated nodes to cooperate on a task and specifically on updating a ledger.

Now, the mechanism by which they did that is basically yelling the answer out loud, which is one node will say, hey, Craig sent Sina one Bitcoin or one ether, one, whatever, and all the other nodes write it down. And so that's how they update the ledger in terms of like data. Now, if you and I crack started talking about what are use cases for ledgers, we'd start talking about, obviously, the financial use cases, bank accounts, that sort of thing. Maybe your real estate titles, we start talking about healthcare records, you know, vaccine records, maybe voting databases. And for almost all those use cases, if not all of them, you want some level of data confidentiality and privacy. Which is to say you don't want to yell the answer to the entire world of unaffiliated nodes that you don't know and maybe don't have great reason to trust. And so, what we're doing at Aleo is we're allowing the same thing to happen. But for the answers to be encrypted, which is to say we yell encrypted answers out loud. And you can still verify that everything being said is true, you just don't know what the actual content is, as a completely unaffiliated node as a user as a program, you may. But in terms of what you broadcast to the entire world that's encrypted. And that unlocks a lot more use cases. So, we're very excited about that. So that's very long answer. To help explain what we're doing.

**CRAIG**

Yeah, yeah. Well, that's interesting. And you're specifically targeting machine learning solutions. Is that right?

**SINA**

The overlap between what we're doing and machine learning, machine learning and AI is really, really interesting. Because there are a couple of there are a couple of interfaces here. The first is training data. When you think about machine learning or AI using databases, one of the most obvious questions is what's in those databases? And do we want to expose that to in the context of something about an affiliated nodes in this context to machine learning. And so, one technique that you could use is something called homomorphic encryption. Another one is federated learning. Another one isn't actually exactly what we're doing, which is using a kind of cryptography called zero knowledge proofs to prove information from data without actually revealing the underlying data, which would allow machine learning to train without necessarily having exposure to the kinds of data we don't want, machines we don't quite understand yet to have access to.

**CRAIG**

Well, just to slow down a little bit. So, I'm familiar with homomorphic encryption. And I'm, I'm familiar with federated learning. How is this different from either of those.

**SINA**

So, zero knowledge proof is basically an application of cryptography that allows a user to prove something about a set of facts without actually sharing that underlying set of facts. So, it's a particular and as a statement has three features. First, the proof has to be complete. Meaning if someone provides the proof, we know with certainty that the underlying statement being proven as true. The second is, it must be impossible for someone to provide a zero-knowledge proof of a particular statement if that statement is false. And then finally, and this is the key one that distinguishes it from some of the other techniques we're talking about, is the proof must be zero knowledge, meaning the proof must not reveal anything about the statement other than the fact that the statement is true. And so homomorphic encryption overlaps with that very significantly, but it's its own distinct field.

**CRAIG**

And in a in a in a machine learning training scenario, how does that work? I mean, federated learning, you're sending the model out to different data pools, that various owners have, and you’re training and then coming back and updating models so that that data never has to leave the possession of the owner. And the only thing that that goes back are the weights from the training exercise. In this case, where is the model? If you're training a model. And where's the data? And, yeah, how does it work?

**SINA**

Okay, let's take a very simple example. Imagine that we are looking at certain kinds of health data. And we want to know, if the health data shows correlate, or there's something we can learn about it based on age, which is people under 21. They have certain traits over 21, they have certain traits over 65, do they have certain traits. And in that context, you could actually look at, so I can hear baby in the background.

**CRAIG**

That's okay, that's okay.

**SINA**

So, in that context, you can actually look at the health traits you care about and get proofs about the age of the person without actually knowing anything else about the person. So, if you imagine a human doing it, they may pull out files, they may see, for example, things like the race, the name, the ethnicity, gender, and they may allow that to affect their thinking. But in a, if you're having like a machine learning context, where the only thing you want to do is isolate, sort of a scan result and the date of birth, then you can actually have it only prove that okay, here are the people who are over 21. And here's what's happening and is there a correlation or is there not without actually ever having to look at anything else? And so, in a way, you can sort of think of it as helping to blind the learning, so that it doesn't consider factors we don't want it to consider.

**CRAIG**

Yeah. And the data in that case resides, say in our hospital database.

**SINA**

Well, so if you're talking about a blockchain, you can actually put it literally on a blockchain and then it can be it can be leveraged by hospitals, insurance companies, anything except they won't actually have access to the personal underlying information. So, it's a matter of how you structure the data. You can have it the traditionally sort of be on a database where it's like managed. In that case, you don't need a blockchain to accomplish anything that I'm talking about at all. You can leverage something like zero knowledge proofs without a blockchain. Or you can look at a world where you put it on a blockchain, it's completely encrypted, which means that it's, no one can go to the blockchain and see your personal information or this health risks are these health records. But you can't, if you opt in, provide proofs of certain characteristics about it, so that it can be studied. As a way of allowing your users to have a lot more sovereignty over their data, what can you post, how can it be exposed, and then you can, you can combine that with machine learning to be more precise about what you're studying, and what you're looking at, when you're looking at data sets, as opposed to allowing machine learning to just go at it, come back with whatever it may come back with. Because what we know already, is the data sets aren't totally neutral, they reflect human biases. And we may allow, we may allow or permit machine learning algorithms to adopt those human biases without even understanding that we're doing that. So, this allows us to adopt more safeguards against something like that.

**SINA**

Think of it as using encryption to obscure what you're putting onto a database the same way your passwords are obscured. Same way, all sorts of sensitive information is obscure, like your credit card information, when you send it to the internet, you're not sending your literal credit card information. It’s hashes so that if your data packet is intercepted along the way, people don't actually have your credit card number. So similarly, you can use encryption, to make data unavailable to the general public that may try to look at that data. But at the same time, if you put it on something like a Leo, you can prove things about that data that are true or false, or that fall within certain parameters. And leverage that data without actually giving away identifying information. And you can tailor the parameters of that however you like. It's totally an open design space. And so, this is very new, it's very, very, very early. And it's kind of ripe for creative thinking, given that you have this tool that allows you to upload information, protect that information such that as with encryption, and yet be able to draw inferences from it and understand facts and correlations from it.

**SINA**

How do you best leverage this? One of the interesting things? So, there are two things going on here. There's the machine learning aspect, and there's a blockchain aspect. What's interesting about the blockchain aspect is how it allows you to leverage something that doesn't have to live on one platform. So, it's not just one hospital doing this. Every hospital can put their data together in something like this. And they can all benefit from all of the data that they're cumulatively providing, without having to reveal anything about their patients the world without having to compromise their patients privacy and getting in with getting their patients consensus. And they can allow their patients to retain ongoing sovereignty over this, which is the withdrawal to do whatever they want to own their data. And so, thinking about how a system like that enables better forms of machine learning, I think is just a totally ripe design space not only for products just fill us perfectly at the computer science level, at the at the level of theory, and that level of like actual practice. It's just totally new. And I think it's very, very exciting. That's just on the training part. There's a lot more that helps with effort hinted at this with actually like the output, how you can help control what factors are being looked at.

**SINA**

And also, so take one of the products we're working on that we think is a good use case for our blockchain, which is digital identity. We allow through this product, folks to be able to upload facts about their digital identity can be their passport and scan your passport onto a blockchain. And then you can leverage that passport going forward across platforms completely agnostic to the platform, which is say, right now we all take for granted the status quo, if you go to Amazon, you give them your information, you go to Walmart, you give them your information, you go to some random website, you give them your information, but you could actually just have passport grade, digital identification, that you leverage at each of these websites without actually having to reveal much more about yourself, depending on the website, depending on the nature of your relationship with that. Now, combine that with what's going on in something like machine learning, it doesn't have to just be your passport, it could be your vaccine records, it could be anything. And then you can allow machine learning to, to you can allow a more competitive machine learning landscape, which is to say you can take that same data to multiple different algorithms and allow them all to train and try to come up with something from it. As opposed to having the data live on just Google or just Microsoft, we're one company, we're all sort of relying on one company to do its best, but that, and it's a much more open-source view of the world. It's a much more competitive view of the world. You know, one of the troubling things when you look at AI is how singular, it seems to feel, you know, one or two, or just a handful of companies potentially can own the space because of their access to data. But if you can open up that data in a way that is privacy preserving for the people who whose data it is, then you can allow a lot more competition. And if you sort of believe classical theory, competition is going to result in better innovation, better products, and more responsiveness to the to the policy concerns that many have raised around AI.

**CRAIG**

Are you familiar with Oasis Labs?

**SINA**

I've heard of them. Yeah. Yeah.

**CRAIG**

Because it sounds similar. They have a combination of secure hardware and cryptography, cryptographic techniques to allow privacy preserving computation on the blockchain. But they have kind of a blend of, of strategies. So, from for Aleo’s users, all of this happens through a web interface or through an API. And does. Yeah, explain how it works.

**SINA**

We're building an open-source protocol that we're hoping to launch later this year. When this open-source protocol launches, are our audience for that it's actually developers, it's people who build the applications we're talking about. We're not, for example, building a healthcare application at the moment. But I believe as people learn about what we're building, they'll start to see the potential for things, the things that we're talking about, which is okay, you can leverage this open source blockchain, and you can build an application on top of it, that uses that, for example, funnels certain kinds of data, puts it on the blockchain in a highly secure way, and then allows it to be leveraged going forward. And one of the reasons I came up with healthcare is everyone kind of intuitively understands there's something wrong with the way healthcare records work. You have your primary care physician, when you're younger, maybe you move maybe they retire, it's very hard to get your records. A lot of people don't know if they were vaccinated when they were vaccinated, what they're vaccinated against. In that environment to actually get together, okay, who is everyone who's had a screen for colon cancer? You know, what are accuracy rates? Who's doing it better? It's very hard, but because the data is so fragmented, and one of the reasons it's good for the data to be fragmented is because it's privacy preserving, if at all, the doctor just said, Hey, everyone, here's the scan for my latest patient, then then your information would be sort of revealed to the whole world So that that's not an attractive sort of way of creating machine learning, enabling that data. But what I'm talking about, if there are developers out there listening to this, and they're thinking, Oh, that's my wheelhouse. They're our audience because they should be building experimenting with how to do this. And when I say this, I mean, create an application that allows you to upload this specific kind of data, or any kind of health data allows you to leverage it, and, and gives you the benefit of access to the data without actually compromising any individual's personal information.

**SINA**

In our laws aren't completely foreign to this. Our census law, for example, allows the Census Bureau to analyze data and to publish reports on that data, so long as it's functionally anonymizing, which is to say, the nature of the publication doesn't give away, it's not too so specific that it would give away who it's talking about. Right. So, collecting data and preserving its privacy while benefiting from it is not foreign to us. This is just a totally new tool that I think is going to be really enhancing for the for the AI and machine learning space.

**CRAIG**

Yeah. And it's open source. So, the business plan for Aleo is to open source it, get it widely adopted, and then offer sort of expert services or consulting on specific applications.

**SINA**

Exactly. Right. Yeah.

**CRAIG**

Yeah. Who's the competition? I mean, are there other open-source projects like this? I know, this has been an active space. I mentioned Oasis Labs. I've had Dawn Song, who's the Berkeley professor behind it on the podcast. But I, frankly, haven't looked at any of this for a long time. Is there a competition?

**SINA**

I first of all, I think I would do terrible injustice to all the really talented people out there working on this from one angle or another. So, I would be reluctant sort of have a list right now I don't think of it as a competitive space. What do I mean by that? I sort of think of it as an innovative space where everyone's trying to build out what the technology can enable at an infrastructure level. And then what's enabled on top of that is really, really incredible. And then that sort of when competition kicks in. So now when I see other people working in this space, I sort of view them as really important in a sense of partners, because they're doing R&D, they're doing the marketing. And the reality is like, there are computer scientists, very smart computer science, like Dan Boneh over at Stanford, who know this inside and out. But the world of developers, they're still they're still learning about this. And so anyone kind of introducing these tools to the world explaining these are tools you can use, I view as sort of a partner, it's just too early to think of it sort of being like, who are the competitors, you know, around the internet and the 1980s, like you could conceptualize of competitors, but just a little too early to think of the space that way.

**CRAIG**

What is ZK ML? I see on your, on your website.

**SINA**

Yes, this z k in our space is a common way of saying zero knowledge, zero knowledge proof. So, this is your knowledge of machine learning, which is a machine learning based on an algorithm that does not have access to the underlying data can only draw inferences from it.

**CRAIG**

Yeah. So where are you in the project right now you say that it has not been launched are? Yeah, yeah,

**SINA**

We're in the midst of our testnet three. And so, we just opened up the ability to deploy applications. And we've had many hundreds of applications deployed on our testnet. So far, it's working really well. The point of our testnet is to identify bugs to invite an open-source community to sort of begin the project of making this a global open-source project as opposed to something that we own are particularly important for. Because as you said, we'd like to pivot to applications and enabling use cases on top of it.

**CRAIG**

Yeah. And the is how wide is or how large is your developer base right now, people that are participating?

**SINA**

Yeah, it's, you know, there are lots of different ways to measure it. There's like the programs that are being deployed. There are the nodes that are being run so many 1000s. Right now, what I find heartening is that At, there's a very enthusiastic core group of folks thinking about zero knowledge on our testnet on other in other projects, and that group seems to be getting larger, and the contributors are increasing in what feels to be exponential with exponential numbers. So, we're sort of very heartened by all the trends we're seeing there. And then the number of people who are not computer scientists who are familiar with zero knowledge proofs is increasing, the Treasury Department released a report referencing them, the Office of Science and Technology Policy at the White House is, is getting familiar with this technology. So, the audience for this is growing really, really fast. And we're, we're very excited about that. And I think the machine is time for the machine learning world to recognize the Venn diagram overlap here, because it's very interesting.

**SINA**

I mean, let me give you another way that this can be leveraged that I think intersects with machine learning and, and the user experience online. So right now, if you go to Amazon or Walmart from the United States, it through your IP address roughly knows not only you’re in the United States, but like maybe what city you're in. And maybe even more specific than that, and you don't think too much about it. But the products that it shows you are a function of that. So, what's available to you, given where you are. And if you went to France, you have a slightly different experience. And if you went to another country, maybe you have a very different experience or no experience at all. So that's the sad theater currently.

**SINA**

And by extension, any machine learning model on top of that, understands where you are, and gives you an experience that's somewhat tailored to that. We can also build it, so that let's just take the digital identity concept. Amazon or Walmart, or the next website would know that you're under 12, or under 18, or over 21. And you wouldn't think too much about it. But what's available changes, because alcohol on the website is certain adult content on the website, or dangerous things on the website, you know, things that 12-year-old shouldn't be playing with on the website. And as a result of machine, machine learning that's built on top of that could also be leveraging that information. Again, without knowing name, date of birth, anything specific to the individual, just that, hey, it's maybe certain things are age appropriate, maybe we don't want machine learning for children of a certain age. Or maybe if we do want it, we want to tailor it a certain way we have certain public policy goals that matter.

**SINA**

So having credentials like this that can be leveraged and that the entire internet would be responsive to has the trickle-down effect of any sort of any sort of what I'll probably call machine learning on top of that, sort of watching the user experience, seeing how the user interacts, would be sensitive to that particular criterion. And so, you can imagine when you're thinking about TikTok or social media, this could be very important, especially to parents.

**CRAIG**

Yeah, it's, so that application relies on kind of the passport idea, the ID, the ID, for an individual that then is accessed by whoever wants to access that individual. I mean, how would that work?

**SINA**

Yeah, so like, think about how it currently works, which is you go to buy wine or website, they ask you are you 21, you click yes, that's the end of the, you know, at the end of the questioning, which probably good enough security for wine, we probably don't have an epidemic of minors ordering wine on websites, but where you have greater concerns, maybe you want something more than just data that the individual can arbitrarily input. You know, they know they just need to pick something that says they're over 18 or 21. We'll just pick it. And so, you can have a lot of fake identities. I mean, this applies in so many contexts. By the way, Craig, if you look over at like ticket sales, something like 40% of ticket sales are sold to bots. So put aside a human lying about itself. It's just like, something that's trained to lie to the actual, to the actual merchants, you can replace that with something where you have to get passport grade identification, which can’t be faked, digitally signed by the US government can be authenticated against the US government database. And that would be the identifying credential. And so, the ability of a child to suggest or over 21 or the ability of someone to suggest their US citizen when in fact they're, you know, a North Korean citizen. I think it limits that very significantly because as you're talking about passport grade security, where some of these websites just have, you know, click through your 21.

**CRAIG**

And to develop that. And that that there would be a key or something that someone would use

**SINA**

maybe a combination of a private key, which is specifically the individual. And LiveScan, which is to say something to so if I had your private key, you know, the LiveScan would show actually, it's not Craig, even though he has Craig's private key. And in the case of parents, you can imagine introducing multi party computation or other things, so that maybe there even two or three keys are necessary for login, it's just a matter of how you again, it's open design space, it's how you want to design it. But this is introducing the ability to have one, digital identification that is passed per grade across all of these websites, as opposed to you go to TikTok, you share your passport information with them, you go to, you know, Instagram, you share your password, sharing your passport information with everyone. And by the way, none of these websites really want to be responsible for collecting that kind of information. Right.

**SINA**

Now, on the other side of it, what's I think particularly interesting for policymakers and for and for a lot of parents out there is that once you introduce technology like this, the excuse for, for example, a website that has like adult content to not be screening out minors goes down very, very significantly, they can't say, oh, we didn't know they tricked us, it's, you have a way to figure this out. Right. And that also applies to anyone collecting data, which is to say their excuse for holding your data in this honeypot that hackers find very attractive goes down. Once you introduce technology like this, that's no longer necessary. Now let's go over to like AI what we're specifically talking about. Right now, when we talk about data privacy, it's typically sort of privacy policy, you know, long scroll down thing about what they're going to do with your data that many people don't read the click through, they move on. But you can start introducing data minimization, again, not foreign to the Census Bureau. And it's just a concept that says, you have to use a reasonably available technologies to minimize your, your what you're doing with people's privacy, right. And in many contexts, I think that's going to be very attractive, it's going to be attractive from a product design perspective, which is to say you even if policymaker said nothing about it, users may have preferences. But from the perspective of policymakers who, you know, in at least western values, we care a lot about people's privacy, what that means for their dignity. And so, we are not going to have like our government go and collect everyone's data and like, feed it into some sort of machine learning exercise. We were maybe other countries would do that, but we wouldn't. Now, this is a very interesting way of saying, okay, well, we can continue with the machine learning, we can actually pioneer this technology, but we don't have to invade people's privacy to do it.

**CRAIG**

What is it going to take for, for widespread adoption, which is always the case with these things? I mean, I've, I've listened to so many ideas or proposals that sound great, but until you reach the tipping point where it's widespread, and generally accepted, it remains kind of a niche.

**SINA**

Yeah, I think that's completely right. And there's a difference between having theoretical conversations and comprehend, and actual solutions that are adopted in practice. If you look at zero knowledge, proof, there were invented as a concept in the 1980s, five, six years ago, the time it took to do one transaction for zero knowledge proof orders of magnitude slower than today. That's a result of investment in R&D that's happening in the United States and happening globally. And that time is going down. And it's scaling better and better and better and faster. And what happens with these systems is as we get better at doing it on the software side, we can actually specialize hardware to do it faster and faster and optimized at the hardware side. And that's what that's the transition from having something like a CPU to the GPU to even in certain circumstances, an application specific integrated circuit, which does one thing but does it very, very fast and optimized, right.

**SINA**

So that starts to make this practical for large scale uses, like whether it's credit card data, or health records, that kind of thing. So, we're getting closer and closer to technologically of being there. So, if you think about the internet, when it first starts and it sort of dial up and it's very, very slow, and we could talk about video streaming in that in that in those days, but it's you have to get faster and the way it gets faster is more investment. And the way to get more investment is to get more excited about what it's actually capable of. So, I think what to answer your question, what it takes is more people learning about the potential here, more people experimenting in the design space, support from the government in terms of R&D, and around privacy enhancing technologies as a way to enhance how we do AI, AI and machine learning. And just more experimentation, as I said, that will, if this design space proves itself out, and I think it will, encourage more investment, and you'll see it scale better and faster.

**SINA**

So, from my perspective, when people talk about something like blockchain, we've long been in that like 1970s, to 1980s era for the Internet, which is the concepts have been introduced. There's this new and interesting thing we can do. But it's not quite practical yet. And we're getting closer. And the only distinction between in my mind between the timeline for the Internet and the timeline for something like blockchain technology is open-source systems tend to have a lot more eyes on them, and they tend to move a little bit faster, right. So, it's not a small group of resources are doing it anyone can open up the coins code or the next projects code and learn from it and build on it.

**CRAIG**

Yeah. Has the government expressed interest in this? Or are you working with any government agencies to include them in the open-source project?

**SINA**

Yeah, so the White House Office of Science and Technology Policy invited comment letters earlier this year on privacy enhancing technologies generally. We issued a comment letter in response to that. There's a lot of interest on the Hill in terms of privacy enhancing technologies, especially as it relates to machine learning. There's a lot of interest around digital identity, I think people understand that we can do identity better, we can do it in a way that increases compliance that increases reliability and also increases privacy, which is kind of an interesting combination. Usually, those concepts are seen as in tension as it relates to blockchain. Generally, I think we've been a little bit tripped up on threshold questions about, you know, whether a particular token is a security or not a security, there are these open regulatory questions that I have confidence will resolve over time.

**SINA**

And I think, especially as it relates to crypto, there's been a lot of focus on the asset itself, because it's invited this speculative, like speculative frenzy around the value of something, everyone's heard some story about how somebody got very rich off of a Bitcoin or the, you know, Dogecoin, or the next thing. And so, a lot of the social commentary has been around those asset values. But one of the insights that we had is, if you look at Bitcoin, for example, it's not just a ledger, for Bitcoin, it's a ledger for identity, each one of those public addresses is an implied identity. Now, it's transparent, which means once they once anyone figures out, you know, that it's your address, don't know, your whole transaction history is not really workable for a lot of the ledger uses we've talked about. But if you add a little bit of privacy to that, you can actually add a lot of meat to what the identity holds, and what it says and how it can be leveraged. And so those are the insights, I think that will take this technology to the next level. And I think everyone's very interested in that. And they should be.

**CRAIG**

how does this project intersect with the web 3.0? Or web three?

**SINA**

Yeah. If I if I understand the term web three, charitably, I think what it's getting at is the ability for nodes, i.e., computers on the internet, to hold and transfer digital assets. In the past, when we think about that, we sort of had this copy paste problem that's required an intermediary which just say, take music. If I have a song, and I sell you the song, I can just copy paste it locally and sell you the song. And so there now there's two of them, and then the value of the music plummets because there's infinite supply. And one of the things that Bitcoin did is actually come up with a way to enforce scarcity without requiring an intermediary, kind of to own all of that, because once you introduce an intermediary to own all that, if you take, for example, health records, they become a very powerful, almost monopoly. That is the owner of records that you don't want them necessarily to write. So insofar as web three references the ability of like nodes, to own assets transfer assets, were very much sort of within that definition, in that this would be a decentralized open-source project that allows hospitals or anyone to build on top of it something where credentials and assets can be moved around and leveraged without necessarily needing any particular intermediary.

**CRAIG**

So, let's presume that this becomes widely adopted. From a consumers point of view for the ID, what would I have to do to get my ID key and then use it on various websites?

**SINA**

Yeah, so, if this is adopted, the consumer could use their phone to scan their passport. Your passport has a chip that allows it to be to be scanned, so you can scan and upload all your information onto one for like, let's just say there's a project emerging. It's called like Aleo ID. You can upload it through Aleo ID, and then that project Aleo ID would have to go to, you know, the ecommerce websites, the wherever and have them integrate so that when the user shows up to the website, that underlying ID is leveraged, for example, to make sure that content is age appropriate.

**CRAIG**

So that would require the open-source project and I presume there's a there's a foundation in addition to the private company, is that right?

**SINA**

Yeah, yeah, so imagine there's the open-source project, that's just the layer that exists, then a team would come a company would come, it could be for profit, it could be just, it could be do nonprofit could be the government. And they would build an interface that allows people to say, okay, we're going to scan our passports onto the blockchain. And then that company would go to whoever, Amazon or a gaming company, right, these are the gaming companies and say, okay, the amount of bloodshed or violence or whatever should be limited for this age group, or certain games can't be downloaded by this age group. And so when you go to that website the same way, for example, you have single sign on with like Google or anything, or just you can even do it as like a the way IP addresses work, which is, the user doesn't need to experience anything, you can just go to the website, leverage that those credentials, and then get the appropriate experience based on his credentials without ever touching the website, again, hey, I'm Craig, I'm this age, or, Hey, I'm Sina, you know, this gender, or this country of origin, the other things that are on your passport. So, like, just to take us like a simple analogy, sort of visualize what this looks like, in the real world, you go to an event, and you're on the you're on the guest list, and they give you a wristband, and the wristband indicates, for example, you're over 21, and you're on the, on the guest list. The bartender never asked for your ID in that context, they just know from your wristband that you belong there, and that you're over 21, without the wristband, you would show the bartender who you don't know, your name, your address, your date of birth, the issuance, issuing state of your ID, all this all this information has nothing to do with you getting a drink at an event. And so, this is kind of like a digital wristband, where when you show up, the experience is tailored based on who you are without actually having to reveal who you are to your counterparty.

**CRAIG**

Right. And the and when you say you show up in the physical case, you're wearing a wristband, in this case, it's this key that represents your identity on the blockchain, you would enter that when you're registering with the website, or would that be tied to your I don't know your IP address?

**SINA**

Three are a couple of ways to do this. I mean, it could be tied to your actual particular device. So, you sign into your device, right with your with your face scan, and it could automatically leverage those credentials. It could be on a per website basis, the way you do Google single sign on. There are a lot of them again, it's very early as to I think it's too early to predict what the winning product looks like. But from the user's experience, I would hope that product is seamless. I would hope it protects their privacy, and I hope that it would help them have an experience online that's appropriate. especially for their age, you know, look, again, nine-year-olds on the internet, different experience and 40-year-olds on the internet. And I think those are things that we can build in reliably now. And the old excuses about well, it's impossible to verify age, or it's prohibitive to those I think are going to fade away.

**SINA**

And I also hope that what it means is, the days of us giving away our addresses and our dates of birth and other personal information to hundreds of websites, none of which exists because they're particularly good at security, they're ecommerce, they have different missions, their mission is not security, and exposing ourselves in these centralized honey pots all across the internet, just to have an internet experience, I hope that we can look back soon. And see that for what it is, which is a security nightmare that compromises the user experience in totally unacceptable ways. I mean, when is the last time anyone even had an emotional response to a major hack? We've had OPM hacked we've had those who keep our credit scores hacked, you know, that says, as much data as you could possibly want from the perspective of a hacker, and much of this is needlessly kept.

**CRAIG**

Yeah, yeah. And then from the point of view of like a hospital that has a lot of data that wants to make it available to a machine-to-machine learning, training exercise. They would, there would be some,

**SINA**

You can build a mechanism that allows you to do that. And what I really like about this is that the mechanism can be agnostic to the machine learning counterparty, right now, we sort of built a system that rewards monopolistic or oligopolistic and says that they can go, and they can go and do the paperwork of collecting the data. And then they're the only ones who have a stab at this. But all of us know intuitively, if you stipulate the patient data is protected, it's better to have 10, 20, 30 companies competing to do this better. Right. And we, we can more enable that as we learn to play with encryption as a design tool. So that we can actually invite competition into machine learning instead of taking for granted as many do now that the biggest companies are inherently already the winners.

**CRAIG**

Now, the what's the next step for you guys that I mean, if presumably, you say there's a foundation and then the company. I mean if people want to explore this, certainly, they can go to aleo.org. But I would guess there is a whole GitHub page or repository or something where people can get involved in the open-source project,

**SINA**

Open source, there's GitHub, there's a GitHub repo. We are testing that is currently live, they can actually go and deploy applications, you can simple, very simple applications, or we welcome more complex applications. Just to sort of learn how this works. We have a language called Leo, which abstracts away a lot of the hard cryptography and makes it accessible to developers more generally. They can experiment with Leo. Learn to use that. We welcome feedback. We welcome any sort of any sort of feedback, whether it's, we're looking for bugs, we have a bug bounty, but we also, what makes it better from a developer's perspective.

**SINA**

Because the next steps, our priorities are to take this from the realm of computer science and make it available in the realm of engineering and developers and people who are just typical coders, right? Because then once you open up that audience, it can go outside of my theorizing about what's interesting, and they can actually put rubber to the road and figure out what works from a product perspective. And I think when it really gets interesting is when you start having product managers think about who are our end users, what is the best experience for them? What facilities distribution, the questions that make this a practical reality, as opposed to a very neat computer science breakthrough that's theoretically interesting.

**CRAIG**

You know, as I said, I've talked to people about this over the last five years, similar kinds of things. Do you do have a sense it's a little bit like autonomous vehicles? We've been talking about it for a long time, and we're still waiting. Do you have a sense when this, whether it's Aleo’s solution or Oasis Labs solution or somebody else's solution? How long it will be before we're actually using this stuff in our daily lives?

**SINA**

Yeah, look, I think it's up to the open-source community who work on this, it's up to us, it's up to all the other projects that are working on this, to make this more like the internet in the early 1990s, than the internet in the late 1970s. Which is to say, educate people, show them the tools, show them the possibilities, publish blog posts, thinking about different solutions. You know, we had one of our computer scientists, Ian Myers, published a paper on identity, we took that paper and used it to build out a proof of concept around the digital identity product, which is looking really attractive already. So, in some sense, these products are I think, you know, on a 12-to-18-month timeline in terms of like being practically available. Now, whether teams execute on 12-to-18-month timeline is a different question. But I think something like digital identity is currently plausible, and is very, very interesting. And it's something that should be, I think, a very high priority for governments in particular, because it's a really attractive way of making the mess of the internet in terms of like, hackers, all the stuff, we've been talking about the data that's available privacy, a lot less messy and a lot more user friendly.

**SINA**

The more complicated things that we've discussed, which is making healthcare data broadly available to multiple teams of scientists, you know, the issue, there is no less a technological issue, as is the case with electronic health records, and more, getting the institutions that are currently incumbents and how they own data and how they monetize data. Sometimes the incentives aren't perfectly aligned there. And that could be a longer process. So, there's the technology, which is like closing in on, on availability. And from here on out, I think the main thing is educating people and increasing its ability to scale. And then there's just sort of the real-world dynamics of what's an easier place to disrupt? Where are there incumbents who are going to be very resistant? That sort of thing.

**CRAIG**

Yeah. Are you talking to any of the big social media platforms? You know, one of my pet peeves is, is Twitter because of the enormous amount of sort of toxic speech that goes on there. And you know, I wish Elon Musk would require people to register their Twitter accounts under their own names. I think it would, it would, you know, you would see a lot of that toxic speech disappear.

**SINA**

It's a great point. I mean, it's amazing the extent to which the social media experience has disappointed many, many people. And it's also amazing the extent to which not having control and sovereignty over our data has been used against us. If you mentioned in 2010, for example, that social media would result in data leakage such that foreign adversaries can use that data and leverage it to make us mad at our neighbors and US yelling about more things, and the blood pressure of our country to go up. It just would have seemed a little bit implausible. How would they do that? Why would we respond that way? It seems a little silly. And yet, that's sort of what's happened. Our data has been rather than like something we own, it's been turned against us. The on the pro side, it's like maybe the advertisements you see are a little bit more relevant. And on the con side, you know, the country is more divided, because they're spiraling with like plots, and they don't know they're interacting with bots, and there's a lot less accountability. So, I think it's a really, really wise thing to point out.

**SINA**

And from our perspective, it would be very attractive to start leveraging things like identity in a privacy preserving way that authenticate what's happening. I mean, one of the things credit we didn't talk about, it's a near term risk of what people broadly call AI is the ability to fake content. We're all familiar with faking an identity in the sense that like, there are all these bots that are clearly not humans. But there's also increasingly where we know what we understand in fake news, but fake quotes, fake audio files. Over time, there’s going to be fake visuals, fake video, and the ability to authenticate data content by linking it back to users without exposing the privacy of those users is a very attractive proposition. Because you can say, okay, well, you know, that actually did not come from the phone they're suggesting it came from. And we don't need to actually reveal anyone's data, we just know that it's a non-authenticated image, and therefore highly, much more structured, much more suspect. So I think not just social media, but the ability to authenticate in a courtroom, actual content, like did the person say that. You may know as well as I would if you go and look for, like videos that are faking what politicians are saying you're getting increasingly persuasive looking and sounding stuff, and so authenticating while preserving privacy, highly attractive, particularly in context of social media.

**CRAIG**

Yeah, but are there conversations going on with any of those social media platforms about this technology?

**SINA**

No, we haven't had direct conversations. But I think that would be really attractive. Our focus over the past two or three years has been completely on the computer science and converting it into sort of an engineering reality. But now's the time, you know, and I think we'd be very interested in talking to Elon over at Twitter, or any of these platforms about integrating this to provide their, their use of better experience. I mean, there's just no way they're monetizing data. And that is their business model in many ways. But there are things happening on social media, that is not the intent of the owners of these social media platforms, and that they would acknowledge as bad and they don't know, they don't know what to do with it. And they're not particularly incentivized to know what to do with it. So, we'd be very happy to talk about solutions on that front.

**CRAIG**

Okay, so that's it for this week. I want to thank Sina for his time. And I'd like you again to consider visiting netsuite.com/eyeonai if you're interested in a full implementation of Oracle's software with no down payment, and no interest for six months. As always, you can find a transcript of this episode on our website, eye-on.ai. I find that you see a lot of things and understand things reading a transcript that you miss with purely audio or visual.

And remember, the singularity may not be near, but AI is changing your world. So, pay attention