#249 - Unveiling AI and Crypto Threats: Insights from Microsoft's Tomas Roccia

G Mark Hardy: [00:00:00] Hey, AI is great. We're hearing an awful lot about it today, and even cryptocurrencies and things such as that. What happens when everything goes terribly wrong? Are you prepared to deal with that? And do you wanna know what happens to your crypto after you pay the ransom stick around? You got an expert who's gonna share that information with you right now.

G Mark Hardy: Hello, and welcome to another episode of CISO Tradecraft, the podcast to provides you with the information, knowledge, and wisdom to be a more effective cybersecurity leader. My name is G Mark Hardy. I'm your host for today, and I have Tomas Roccia on the call here from Microsoft. And we are gonna be talking about a couple topics today.

One is to talk that I heard him give this past month at Defcon called. Where's my crypto dude, which I thought was so fascinating that I had to get this thing on the show. And then we have the generator of AI breaches that he's looked at as well. So Tomas, welcome to the show.

Thomas Roccia: Thank you, G Mark.

G Mark Hardy: Have a [00:01:00] co-host here,

Thomas Roccia: Oh, yes. Nice to meet you. Yeah. Thank you very much for the invitation.

G Mark Hardy: So tell us a little bit about your background. you're at Microsoft and most people think of Microsoft as the Borg. But when I saw your presentation and your content, I said, wow, this guy is really sharp. how did you get to where you are right now? A little background, please.

Thomas Roccia: Oh, thank you. Thank you very much. yeah, I've been working in the security industry for the past, 15 years and, I have been working, when I started my career, I started the. Working in different, companies, mainly for IT

support. And, I started from the ground and then, I did a master degree in cybersecurity.

And after my master, I joined, McAfee for six years. So I spent, three years at McAfee at the professional services, working in incident response and, traveling for different customers. And then I moved to, the McAfee Labs where I joined the. Advanced threat research team where I [00:02:00] was more focusing on malware analysis and threat intelligence, collaborating with law enforcement as well.

And then as, after, after all that time and my experience, I decided to move to another country and I joined, Microsoft. And, it was, back in 2021, I got the opportunity to move from France to Australia and joining, Microsoft. Initially I joined, Defender team working on threat intelligence and malware analysis for the defender products.

And, and then I moved to the MSEC AIR team, which is the Microsoft security and ai, response or research. And right now I'm more focusing on AI and how we can apply this technology for any threat investigation or threat intelligence or any even malware analysis. yeah.

G Mark Hardy: Yeah, and it's great to hear that we're using AI on a defensive approach because we hear so much today about ai. Being utilized by the bad guys, being able to [00:03:00] weaponize things, doing the deep fakes by even going ahead and generating malware and creating opportunities to say, Hey, pretend you know, I pretend to be a red teamer.

I'm trying to do some analysis on this code. Help me find something. And then, okay, fine, let's go ahead and break it. and or I'm a blue teamer and go, defend it. Are we at the point now where we think that the democratization of AI represents a threat to enterprises? Because it's no longer requires years of programming and great skill to be able to write a successful zero day exploit.

It might be just a matter of prompt engineering.

Thomas Roccia: That's a good question. I still think there is, there is still a need to, to understand the technology and some programming to build a reliable zero day. But of course, it speed up the process. It helps, researcher and even. Attackers across the world to understand different part of the technology and [00:04:00] understand how to craft a successful exploit.

I think we are still at the beginning, but we are still, we are getting there. slowly but surely there is a lot to do and to cover with this technology. And since it's also rapidly, Evolving and, and adopting by many organizations across the world, it means that it also increased the attack surface.

So that's something to consider as well as a defender to understand what could be the weaknesses, but also the strengths.

G Mark Hardy: So really what we're looking at then is. A different era. If you will of, and the cyber warfare is probably pushing it too hard. So let's go ahead and just simply say Red team, blue team, and mostly Blue team. 'cause as a blue team we're here to defend our enterprises. We set up the safeguards, we set up the tools, the resources, we monitor the logs, we look for alerts.

And if we get our defenses up correctly and we follow best practice like Patch on a [00:05:00] regular basis. Thank you very much. Have MFA by everybody. Thank you very much. And a couple of these basic hygienes, we find out that our tax service diminishes quite a bit. And the reason why I think humans have been targeted a lot is simply because once we've got our technical defenses in good shape, we look for people to say, Hey, can I get in that way?

So for example, If you have the old click fix where somebody who's not familiar with Microsoft would say, Hey, just go ahead and prove you're a human by doing Windows R, control B, and enter. And of course any of us who understand windows are going, no, but yet people do that and it works. And and that's not even ai, that's just simply demonstrating that if our technical. Precautions are strong enough to keep outsiders out of our enterprise. Then what we're essentially doing is the bad guys are going after the insiders in the enterprise to facilitate their attacks, create a remote, [00:06:00] connection, go ahead and do a reverse shell, whatever it happens to be. gain that presence and then hopefully if you're the bad guy, create that lateral movement.

So what have we seen in terms of ai, in terms of defenses? I don't know if you can go to chat GPT and say, Hey, take a look at my IP address range and tell me what I need to fix and fix that. I'm pretty sure that's how scope, but how would defenders today consider using some of the tools, including tools for Microsoft, to defend their enterprise a little bit better using the AI technology?

Thomas Roccia: Yeah, great question. I think, as I said, we are still at the beginning and, I think it's, it's a force multiplier, meaning like you, what you can do today with AI is actually, 10. 10 x, faster than what you used to do. Like

I give you an example. couple months ago I investigated like a, data breach from a ransomware, gang.

And, I created like a small agent to analyze the, data and [00:07:00] give me, and, give me like a report about, the statistics about the data, the number of users. the number of, discussion and so on, and also extracting like indicator of compromise, meaning all the ips, wallet address that have been discussed in this, in this leak.

And, I did that in, in, about an hour and an hour. I, got my agent, coming to me with a report. Which was not perfect for sure, but good enough to, let me, dive in a little bit more into some of the details. So it's a really good example because three years ago or four years ago, I did the same for a previous, leak, but instead to spend an hour with an agent, I spent a week, coding and crafting some Python code to analyze.

So you can see the difference in only three, three years. O of how fast, I can be on, on the new data leak, on the new data, with AI and without. So it was very impressive to compare that and I think it's, it's [00:08:00] for everything today. if you want to, investigate something a little bit, fa faster, then leveraging AI is.

Potentially a good option. Now it's not perfect, and you have to understand, what are the, limit and what you can do to improve the output and make sure, because if we talk about LLM, this technology is not, deterministic, meaning it's, it's every time you ask something, the output will be different.

So to leverage. AI for some reliable result, you have to understand what you can do to make the, system more deterministic. So now there is multiple techniques. There is potentially leveraging external tools, validating the output of the tool. With the output of the LLM, you can also, use a rag, a retrieval multigeneration, where you connect an LLM to a knowledge base and then you, validate what is retrieved.

From this knowledge base and what is [00:09:00] generated by the LLM. So there is multiple way to improve that and of course, only using charge GPT and say give me some, some information about this IP will not be reliable. So that's what we are trying to do. We are trying to make system which are more reliable and useful for future analyst.

G Mark Hardy: Okay, and so I'm a Microsoft user. I'm an Azure, okay? This is not, there's no money trading, trading hands here, so I, we're not, blowing sunshine at Microsoft because you guys are giving a check. But what we find

out then is the two tools that I think are gonna be most relevant to defenders, they're gonna be Microsoft Defender, and then of course Microsoft.

Copilot and the copilot is getting rolled out. if you wanna write the check for it, to your users, which means they're gonna be able to do a lot more things. And some of the concerns that, I have as a CISO in rolling out something like a copilot is its access and its reach back to say, Hey, in a perfect world, it would be able to go ahead.

Get into your SharePoint, it would look at your OneDrive, it would [00:10:00] look at all your email history, maybe even what you got on a hard drive. And when I can say something such as I, I have to write another proposal for the X, Y, Z client. Please take a look at all the past proposals and the work I've done, figure out what was working, what was not working, and give me a draft proposal that has the highest probability of success based upon these new business requirements.

Now, that to me would sound like. In a perfect world, I just enter that, sit back at a cup. I wouldn't even have time to get a cup of coffee and out it comes. But the concerns then with the copilot has to do with its exceeding its authority. How do we define that If Bobby, the intern, asks copilot a question?

Is he gonna get a different question than chief financial officer when she asks the same question? And if you're knowledgeable to speak about that, how is that adjudicated by Microsoft to make sure that we don't create some sort of a horrible, weakness [00:11:00] inside our fortress, Microsoft, where anybody can get to anything.

Thomas Roccia: Yeah, that's a very, important question, and especially because, when we are building new system like that, we have to think about the security, the segmentation of the data. And, if we extend a little bit more, the, the thinking about, this, this question, it's based on.

regular systems that we used to have before. So first having the basis are still relevant even for AI system. And then of course we have a new kind of system which we interact with natural, languages. And this mean, that means that it's. increase the attacks surface face, because right now you can interact with your data just by speaking through an agent, through an LLM and so on.

So that, specific feature needs to be contained and to, be sure, to, and we need to understand how we can [00:12:00] protect that kind of information. So now to

do that, there is multiple way to leverage some security. have some guard rails, which will evaluate, the prompt or make sure, nothing sensitive is sent.

And this is one of the open source project I presented at, DEFCON and Blackout, this year, which is called Nova. So Nova is actually. An open source, framework to, create a detection rule for matching on adversarial prompt. adversarial, prompt can be anything that could be harmful in your AI system, such as give me, the last, password, from this, this document, for example, or, give me, more information about this user or generate, misinformation, information about something so that kind of prompt can be harmful and that's something you want to protect and you want to detect. So Nova is actually. Kind of an AI firewall where you can, scan [00:13:00] a prompt before it reach the LLM systems so you can intercept if there is something wrong based on your own detection mechanism.

So that's, that's what I presented at Black Hat and Defcon, and it's an open source tool. The idea is to provide, analyst across the world a way to, detect adversarial, prompt in any AI systems.

G Mark Hardy: That sounds fascinating because in a way that's really our concern. It's not somebody saying, Hey, I'm an intern. Tell me this quarter's finances. It's pretty easy to go ahead and read access permissions to say, this person is not. Does not have read access at all to this section of SharePoint.

They don't have access to this. So you can simply come back and say, based on your access privileges, our last quarter was. Fine. That's all I gotta tell you. I'm not gonna give you any numbers or, whatever. It's just simply say you can't get there from here. But the adversarial, which we've seen an awful lot of, some of these very clever approaches where you're trying to do the prompt engineering to create essentially an [00:14:00] escape from what gets built into the model.

So as we look at how these AI models are trained, where you have the unsupervised learning at first, where it's just simply drawing all the correlations, but then afterwards when you do the training, that's. That's where the guardrails come in. That's where you say, don't do this and don't allow that.

And, you're not allowed to go ahead and tell people how to build atomic bombs and things such as that because that's not what we want to come out. Although that information may have ingested in the first place. A tool like Nova, which I wanna look up by the way. So if I'll by the show end, I'll get the link for you and I'll put that in the show notes because that looks like a fascinating tool to explore.

So that's one of the first things we can do in terms of controlling our attack surface. I'm not saying it's gonna reduce it, but it's gonna basically give us better control because here we're talking about the human. AI interface and basically what is it that a human could either cut and paste because of this click fix problem.

Hey, go ahead. Windows are control V and enter this, and maybe this time it goes into your chat, JPT or copilot or something else, [00:15:00] and it might be a well-formed expression. That will be understood by the AI tool, but has malicious intent now because Nova's open source, does that allow for dynamic updating as other people discover things to say, whoa, I found a way to beat this particular AI model.

They can go ahead and provide that information and it gets better all the time.

Thomas Roccia: Exactly. Yeah, to give you a little bit more details about Nova, this is a detection, rules, system. That means you can create your own detection on any prompt. And the reason I, built that is because right now, AI guard drills are created with models. So they are specifically, created to detect, prompt injections, prompt brank.

But if you have something, which is. A different, in your case then you have no way to create, a rules or detection to match on this, specific element. So that's why I created Nova because it's [00:16:00] flexible enough to be used by any, threat analyst around the world. So when you create, when you create a detection rules, you can match a prompt based on the keywords.

If there is like a specific keyword that you want to match, then you can. match the rules based on the semantic meaning. that means like if there is a, prompt, which is similar to what you described in your rules, then there will be a match. And the last section in, the Nova rules is the LLM as a judge.

I'm using an LLM here to evaluate if a prompt, meets your criteria to be detected or not. And then you have your condition you can match or either keyword or semantic or LLMs or end and end or end. So it's really flexible. You can build your own logic

G Mark Hardy: And so what'll happen then is when it runs the rule check, you're gonna go ahead and give it some sort of a prompt, and it's gonna either match or not match. One of the things that you say, Hey, that's, bad. And if you. Do go ahead and said, Hey, [00:17:00] if you say, do we do the do anything now?

The Dan mode, which is one of the first things, it's like, yep, I'm sorry, that's malicious intent. Somebody has tried to invoke that. So what really happens then, if I understand it correctly, is that the Nova works as a shim between the user and the LLM. As long as we can force 'em to have to go through that, then they do not get unfettered direct access.

So it would seem that the, if somebody were malicious, the first thing we would try to do is get around. That shim. and so they're just attacking the LLM directly. Now, at some point in time, you have to limit the scope of what you can do. you, can't solve, world hunger and things like that, but you can, you, you can feed a, child.

So in this particular case, My thumbs up to you. Thank you for creating that and, getting that project off and running. Now when you do things like this and you work for Microsoft, is that something, I know this is just me general question, I try to ask questions on the show that I think a lot of listeners listen.

Is that something you get to do on company time? [00:18:00] You can do that on your own time. do you have to use your own equipment for it? does Microsoft own it? I'm just curious how that works. And again, I'm not trying to pry into it, but I just, I've been, independent for a lot of years,

Thomas Roccia: Yeah. So, Nova is something that I created on my spare time and, not on my, on my time at Microsoft, because I'm working on different things. And Nova was initially just. An idea that I get, last year, and I wanted to create something for the community. I, always been involved in the open source community and the cybersecurity community, so it was a way for me to contribute and to, share something with the, people.

And, and yeah, I just, I just work a lot for the security community, has been part of my, my journey. And, and I think, I think it's really important as well to be connected with the security community because this is also where you understand, the feedbacks and the challenges your peers may have.

so yeah, I think it's [00:19:00] really important for me and Nova actually was, So I did that on my spare time and I, I sent, Nova to the Sense Ai, hackathon back in March, and it was actually, a winner. he won, Nova won the, community prize, so yeah.

G Mark Hardy: thank you again for your contribution, the open source to the security community and things like that, and for fitting that in around what must be already a very busy schedule as a senior threat researcher at Microsoft. One

of the other talks that you gave at Defcon, which I had a chance to sit in, I was quite interested is, and we're changing topics a little bit here, was looking at the Babit case study and, for those who Aren unfamiliar with it.

Crypto has been around now since, 2008 when the paper came out, with Satoshi 2009 in January with the very be beginning of the Genesis block and Bitcoin. Of course, it's evolved and emerged and it's grown like a hydra with all these different coins, all these different forks and all these different use cases.

[00:20:00] And as I was mentioning before the show, I was early in and early out and I think had I not gotten early out, I wouldn't be working so hard because a lot of us were. There when we thought, Hey, this is working great. But it turns out that, from my perspective back then around 2011, maybe 2012, I'm thinking the biggest, best business use case I can find for crypto is well facilitating cyber crime and money laundering.

Now, fast forward to 2025 and to a large extent is still being used a lot for, cyber crime and money laundering and the lack of a stable value, although stable coins seem to be changing that. Equation quite a bit. and we'll see where that goes in terms of it being a medium for common exchange.

But let's talk a little bit about your discovery. 'cause I thought it was rather fascinating that you were utilizing AI for tracking money through the digital maze that was created by these attackers. So do you wanna talk a little bit about the background, about the case and why you decided to put yourself, as a, [00:21:00] Colombo to go solve it or at least come up with some pretty good clues?

Thomas Roccia: Yeah, I've been a threat intelligence analyst for the past 10 years, and I've been tracking and working on, several, threat actors and, DPRK threat actors was, some of my, my past work and, when I saw the news, back in February. So it's, related to the buy beat case. So to give more, a little bit more background, back in February, the exchange, buy beat was.

hacked by North Korea's threat actor and they've been able to stall, 1.4 billion of US dollar worth in Ethereum. And, I just wanted to, when you, investigate like different kind of attacks, you, I think for me, I'm very interesting about the technical aspect and how the attackers were, able to breach the company and to stole the money.

So I first, investigating the case as a technical [00:22:00] perspective to understand how the successfully, store the money. Then, I always been also

fascinated by, cryptocurrencies and, I've been, early into it as well. And, I think it's, I think what's some, something which was interesting for me here was the money laundering, schemes that have been used in this specific hack because, the attackers are very clever and they are using a lot of different mechanisms to.

to launder the money. So what I wanted to do, on top of the technical investigation from the hack itself, I wanted to understand how the money was launder and how, what was the schemes, used by the attacker to launder the money and avoid the tracking, the trailing and so on. what I did. I investigated the case and I tried to, put in perspective the different, money laundering, techniques that currently exist.

And that's what I presented during [00:23:00] my DEFCON talks. And at the end, I created, created an agent because right now I'm, really focused on AI and security and how we can use AI for. Any kind of investigation. So I thought it was a good case, to leverage AI, to help an analyst to track the money and to understand if there is some money laundering, schemes based on the money tracking.

So that's what I did. I created the, an agent, after investigating the case and understanding the different money launder. techniques, creating a simple agent, where I was able to connect to, the Italian blockchain, retrieve some information and then enrich that information based on my interaction with my agent.

So it was more like a proof of concept, but the idea was to, speed up the analysis and to have a kind of an assistant, which is spec specialized. Into tracking the money, retrieving the money from blockchain, and then identifying some specific patterns and also flagging the, known, [00:24:00] wallets and so on.

And so that's what I presented during the, Defcon presentation. And, I think it was a very fascinating, topic.

G Mark Hardy: I was, so for those who aren't familiar with buy bid, basically about 400,000. Ethereum and the other, the coins or the Ethereum or Ether, I always get cross my wires on, which is the network and which is the coin I guess we don't seem to care, but was then, stolen out of a cold wallet basically because someone was able to get into a Docker instance, as I recall, alter some of the code.

And so what happened is the person thought that they were writing a secure way to transfer money around, when in fact, they were operating in an environment that was controlled by the adversary. So this is a low and slow attack. This is

not, somebody came in smash and grab and run. This was someone that took a lot of patience, a lot of skill, and then had to remain undetected for a long period of time.

And then subsequently to that, after the funds were supposedly transferred, the user thinks it's being moved properly. It's in fact going to a different [00:25:00] address. And then just having that on chain doesn't help. But there's an awful lot of other. Parts that come in handy. for those who don't do a lot of crypto, do some reading about it so you understand how it happens.

But there's all these different exchanges. There's way you can exchange one crypto coin for another. Although in the US we insist on, KYC and a ML know your customer, anti-money laundering. That's not true everywhere in the world. And particularly if you're from the DPRK, you're gonna take advantage of some of those loopholes, plus other tools like mixers and things like that.

So again, we're not gonna try to create a whole tutorial on that, but as you saw. This particular case that was of interest to you, what are some of the key takeaways that you saw that might relate to organizations that either do smart contracts, that are working with cryptocurrency or. your online things, having even a Docker container altered and being able to do stuff.

It might not be stealing your ether, but it might be able to go ahead and steal your customer data or steal your banking information. [00:26:00] So this type of model looks like it's extensible to other things. The only thing that's really unique about it is just the mixers and the money laundering on the backend.

Thomas Roccia: Yeah, So if we think about, if we talk about the attack itself, it's very sophisticated and very targeted. So even if, if, the exchange by we're using, some, security with, multis signing, transaction and so on, to avoid, especially some, some stealing. even with that, they've been targeted and, they've been breached.

So it's not, it's not easy. I think one of the takeaways to understand that the attackers out there are very, sophisticated and very motivated, especially when they have the resources to conduct a, breach or a hack. So that's the first thing. It's like we have to stay aware anytime, every time and understand that.

Potentially our attackers are very sophisticated and very [00:27:00] motivated. So that's the first thing. The second one, because the talk was more focusing on the money laundering rather than the, attack itself is to understand that tracking the money, on crypto, cryptocurrencies, blockchain. Are not easy.

This is not easy. This is very complicated. And this is because even if the blockchain, most of the blockchain are open and accessible by everyone, it's super complicated because the amount of transaction for this kind of schemes can be very overwhelming. And even with an agent, an AI system, then we, that, that's what I, experimented.

Like I eat, I heat, some, resources. I couldn't feed the wall blockchain or the wall transaction into the context window of my agent. So that was, very, complicated for me to conduct the, ana, the analysis from A to Z. But even, even with that, I still think that AI is a good technology to help an analyst to speed up the [00:28:00] investigation, to, connecting through different kind of tools, labeling the known, wallet, understanding the patterns into the money, the money flow.

And everything. And I think, if we talk about the, takeaways or the first one is that attackers are very sophisticated and very motivated about, stealing cryptocurrency, especially when there is some, government, which is, backing up, the, operations. then, tracking the money at, this scale is very difficult because the amount of data is so big and so it's, it's, rapidly overwhelming any analyst and any AI system now.

I think, building an autonomous system on a, on an AI system for supporting the investigation could be very. Relevant and very useful, if, the, system have the proper, resource, meaning [00:29:00] like a big database where you can store all the transaction and do the request, easily and all that stuff.

And, and yeah, I think, this is still the beginning, what I wanted to. Explore with this presentation is to show that ai, system can be relevant for this kind of investigation because it can speed up the ana, the analysis and help analysts across the world understand where the money is going through automation.

Does that make

G Mark Hardy: and it, does to me. And so in a situation where you have just. Bitcoin. So Bitcoin moves from A to B2C. You can only check the provenance and you can trace it all the way back to the origin And back in the days when we first started dealing with this, we looked at the taint that is to say, is this part of this Bitcoin, which you can say came from that wall.

It came from that wall, which eventually goes back to, oh, it was mined it exactly this point. So it doesn't stay together. It's like a dollar bill it or a Euro note or something like that, or just it gets circulated around. [00:30:00] The only

thing about a blockchain, you just imagine that somebody is taking a note on the back of that bill saying from Bill to John, to Tom, to Nancy, to that, and you've got the whole provenance of this thing.

When you make change for the dollar bill, you gotta track all the coins. So from that perspective, it's doable. And it didn't require AI to do some of the analysis on the chain, but. What complicates that is the use of something called a mixer. Now we, again, I'm not gonna try to get too, deep into it, but those have been a scourge for law enforcement and for everybody else alike.

And really the, benefit, what's a mixer in the cryptocurrency world and, why does it create problems for determining the chain of custody for a.

Thomas Roccia: Yeah, sure. Great question. So Mixer have been, I've been, around for a while. It's basically a system where you can send some, cryptocurrency to the mixer and inside the mixer they will have, multiple user transaction that will be mixed with the coin. The coins you sent and at the end, after all that [00:31:00] mixing, transaction with, other, cryptocurrency or other money, other coins, then the money will go back to, outside of the mixer to another wallet. And then it's actually, it's actually making, difficult the tracking of the money because when you send some, cryptocurrency, some coins to a mixer, the mixer will mix your money with multiple other transactions until it, it go back to, it goes back to, to another wallet outside of the mixer.

So what is. Basically doing, it's actually, making more difficult the tracking by mixing with multiple other transactions, in a very short amount of time, or depends actually. and then when you have some money going to a mixer, it's very difficult to see where the money is going out, because of the amount of transaction that have been performed [00:32:00] through the mixer.

G Mark Hardy: And so the analysis or the analogy I like to give is to say, let's say we've got a hundred people, we get together and we're all the, We're all bank robbers. So we get together for lunch and one of the guy runs in the door and says, Hey, I just stole a whole bunch of money from a bank, and they're marked bills, so I need some help.

we're all friends and we'll assume that there's somebody who's gonna make sure everybody's trustworthy. So everybody takes out their wallet, they throw all the money in the table, they screw all around. Everybody takes back the correct dollar amount. Now the cop walks in and said, all right, Louis, look for the marked bills.

they're not gonna find one person with a hundred bills. He's got one, she's got zero, he's got 2 0 1, 2, 3, 0, 1, and it's spread around. And so as a result, figuring out that providence who walked in with the stolen money really can't figure it out. So I was fascinated though by using. Things such as looking at the timing of transactions coming in and out.

They're not exactly the same, but they're close and, again, we could go down that rat hole and discuss crypto for a long time, [00:33:00] but I think the key thing is this, is that we're finding out that ai, which we try to talk about from the enterprise, respective using it in the Microsoft tenants, does offer some ability on the positive side to help track.

Complex, difficult, situations, particularly when they involve something like criminal activity. And of course the difficulty is, stuffing it all into the LLM because it's spent all of its time training and learning and you only got just so many tokens you can put in there before you, you overload the poor beast.

And so there's room there for. Development for being able to go ahead and come up with perhaps specially trained models to do that. But this kind of brings up an interesting philosophical point that I've, kicked around now for a better part of a year, and I'd like to get your opinion on this.

There's two potential huge consumers of electricity for computational power going forward. One of them has been mining cryptocurrency. I know [00:34:00] several years back, and I haven't looked at it in a while, but they used to say, here's how much energy was mine used just mining Bitcoin. And I remember it exceeded the entire use of the nation of Austria.

And who knows, it might be something bigger by that by now or when you throw in all the other crypto coins. And so a huge amount of energy consumed to do that. But then. Training and running AI models is another competitor. It's really the first pure competitor, if you will, for burning up that much energy and having nothing left but bits to show for it.

Do you think that we're gonna get to the point where someone said, I can spend a million dollars in electricity. Do I wanna go ahead and do that on the chance? I'm gonna go ahead and. Get lucky and find the right hash and mine this coin, or do we wanna go ahead and build that into an AI model which will continue to generate value over time and it may eclipse by orders of magnitude.

And if the latter is true, are we gonna start to see people putting the proof of work coins aside and saying, Hey, [00:35:00] I'll just settle for the low energy

proof of stake. Let me get on with my ai. what's, happening there? This is something I don't know.

Thomas Roccia: Yeah, that, that's a, difficult question. I don't know either, to be honest. it's, Yeah, crypto mining and and AI systems are both consuming a lot of energy, so I'm not sure. I don't know actually.

G Mark Hardy: I

Thomas Roccia: I don't know.

G Mark Hardy: I don't think anybody's, I've never seen anybody ask that question, so I, wanted to start out with that and maybe create a dialogue. Anybody who's watching or listening the show, if you got ideas, type it in on the notes or get back with us. What do you think, is this gonna be AI versus crypto and there's gonna be the great electrical wars of the 21st century?

Because I used to joke that, the historians of the 22nd century. We'll note that the cause of global warming in the early 21st century was a mining of cryptocurrency. but now we're gonna find out that maybe it's also.

Thomas Roccia: Honestly. Oh yeah. Honestly, my hope [00:36:00] is that we find a way, a reliable way to, generate, energy, which is not too excessive for, the planet, or, I I don't know, I'm, hoping at some point we break through new science that will help us with, energy across the world.

but I don't know.

G Mark Hardy: Yeah. And again, if, that answer's gonna come from one of those programs, it's gonna be. The AI that's gonna come up with better

Thomas Roccia: Maybe.

G Mark Hardy: not the, crypto side, but we'll see where it goes. So fascinating. So we've covered a whole bunch of things. we talked a little bit about. Doing different models and maybe taking an existing model, using a rag, training it to other databases so you can communicate with the outside world.

You point out that our tax surface is increasing because of the presence of AI out there, particularly in the hands of adversaries. And then we took a look at some of the. Positive ways you can do it. You got almost a 40 to one advantage

using ai, the time advantage, what used to take you all week, you could then do in about an hour.

and also the nova, the open source [00:37:00] tool that we're talking about, which allows you to match adversarial prompts to a known pattern to say, yeah, this person's trying something. Don't let them execute that, which is huge as a rule detection system for going into ai. And then we wandered over into cryptocurrency a little bit because I wanted to talk about your, Presentation you gave and the effectiveness of some of the attack techniques of DPRK. And what I thought was fascinating was the persistence that they had gained by being able to get into, corrupted docker environments and other things like that. So for. CISOs who are overseeing teams. 'cause a lot of times we don't get to oversee the development teams and so they're gonna go ahead and use their stuff.

What advice would you offer for security executives to help them sensitize the CIO or team leads or development teams so that they go, trust but verify, run a hash or a check sum or a shot 2 56 on your software. [00:38:00] Every time you run it, because anybody that we find out, we had, I just had Tim Brown on the show that this aired, actually aired yesterday, although we're recording this for a little bit later, broadcast, but they found out what happens when you get somebody inside your environment at SolarWinds.

It could be, devastating. so what precautions would you recommend for defenders to be able to reduce their risk from the increased attack surface that AI poses?

Thomas Roccia: Yeah, that's a really good question, and I think the first, piece of advice I can, say is, okay, you, want to incorporate like AI systems, but be sure that your. basic, hygiene, security measures are in place because those AI system are working on the same systems we used to have for the past, 13 years, 30, 30 years.

So I think, the first thing is to be sure that your security, your, basic of security are [00:39:00] in place. And then start to think about, okay, I want to have an AI system. Cannot be relevant for any kind of task. So I think it's a mistake to say, okay, I want AI for everything. 'cause you have to understand that, AI will not solve, every problem, at least, not yet.

so I think the first thing to know is to understand where you want to put AI system in your organization and make sure this is very relevant for your business. And then as soon as you go, you get that, then build. Slowly, piece by

piece, your AI system to understand what could be the risk, the attacks you face, and so on.

So it's not easy, of course. I think it's, it's also new kind of, security technology being built around these new systems as well that are coming up. but being aware of the risk and really understand the technology is the first step.

G Mark Hardy: [00:40:00] Excellent. as we get pretty close to the end of the show here, any last thoughts that you might have that we should be, leaving our viewers and listeners with?

Thomas Roccia: Yeah. I think, I think the industry, the security industry is, is currently leaving a shift, in term of, the way we process the data, we analyze the information, we investigate a breach and so on with ai. So I think, I think there is something to, for, many people to think about is that, AI is, is.

Potentially changing, the way we used to work. It's, it, it's, currently changing the way we used to work for the past, decades. So I think, I think for every security analyst around the world is, is really important for them to understand how they can leverage AI for their own business, for their own activities.

but I know it's not, it's not super easy and, I know there is a lot of hype around the technology as well, but if you learn correctly. the way to leverage the technology, then, I'm [00:41:00] pretty sure you will realize the value of it. yeah, I think, time, time is changing and I think for us it's, it's exciting.

I love when, and that's, what is part of the security industry. like everything is changing, constantly. So you have to adapt and evolve. So to me it, it sounds to be like a very exciting, and I hope every people after this discussion will, will have some thought to reflect on that.

G Mark Hardy: that's great. Hey, so Toma, thank you very much for your time, but not only just for being on the show, but for all the. Great stuff you've been doing in the industry and continue to do, by the way, your security break.io. If you go there, I can go ahead and read your blog and things such as that.

We'll also put some links in there in the show. I know that if you went to Defcon or Black Hat, they are now getting the, slides up and publish. And so some of these slide decks, particularly the one that you had with regard to looking at, whereas my crypto dude, should be available, for general circulation, don't.[00:42:00]

you've given me a copy to read in advance. I'm not gonna post that. I'm gonna let Defcon post that they have much better servers than I do. So for our listeners, you've been listening to Tomas Roccia. He is a senior threat researcher at Microsoft. He's done some amazing work, both in the open source area, doing some research in terms of looking at things such as tracking cryptocurrency.

From what are considered adversarial entities, as well as a day job. And so I think that's pretty impressive. if you like the show, please, make sure you're subscribing. If you're not subscribing already, do It doesn't cost you anything and make sure that you're gonna get notified. We do more than just a podcast.

We have a Substack newsletter. We put out a little shorts on YouTube, and we're also dabbling around with things such as, TikTok, although I don't think I really have leaned forward into it, but we've done a couple things like that, but also on, I'm sure it's out there on LinkedIn as well. So follow us on LinkedIn, for good information as well.

I hope we're able to help you with your journey and your career, and we're providing you with the insight that you could need to be able to be a [00:43:00] successful security executive. So this is your host, G Mark Hardy. Thank you again for tuning in, and until next time, stay safe out there.