



Yes. Exactly.

Description: A disgruntled developer discloses multiple Windows zero days. Microsoft purchases its own bugs in massive campaign. VeraCrypt and Wireshark suddenly lost their dev accounts. A serious problem with recaptured domain names. How might AI help to secure open source repositories. A listener wonders what we thought of Project Hail Mary. Cyber security professionals tell us What Mythos Means.

High quality (64 kbps) mp3 audio file URL: <http://media.GRC.com/sn/SN-1075.mp3>

Quarter size (16 kbps) mp3 audio file URL: <http://media.GRC.com/sn/sn-1075-lq.mp3>

SHOW TEASE: It's time for Security Now!. Steve Gibson is here. He's got some more thoughts on Project Mythos, the new super-smart AI that can find security flaws. He's saying, you know what, this isn't hype. And he's got some real evidence behind it. We'll talk about Microsoft buying its own bugs back and ignoring other ones. And what does Steve think of "Project Hail Mary"? Yeah, even that, coming up next on Security Now!.

Leo Laporte: This is Security Now! with Steve Gibson, Episode 1075, recorded Tuesday, April 21st, 2026: "Yes. Exactly."

It's time for Security Now!. Yay, Tuesday has come. You've been waiting all week long to hear from this guy right here, Steve Gibson, the man in charge at Security Now!. Hey, Steve. Good afternoon to you.

Steve Gibson: Leo, great to be with you again, as always. I got some feedback from the 20,000-plus listeners who joined my email list to receive the show notes. As a consequence of my own schedule over the weekend, I needed to start work on this Friday. I finished Saturday afternoon and immediately sent out the show notes. So the feedback was, wait, is it Sunday? No.

Leo: People are paying attention, Steve. You can't pull one over on them.

Steve: Yeah. Unfortunately, they're paying very close attention because I also, because my emailing system assumed that I would be doing the emailing in the morning of the podcast, it autofills the day's date at the top of the email. But since I'm sending the mailing out typically on Sundays, and occasionally on a Saturday, I have to go in and remove the placeholder for the autofill and replace it with the actual date. Unfortunately, something is wrong with me because this is not the first time I've put in 2024 as the

year. And so I sent it out as 4/21/2024. And so several of our sharp-eyed listeners said, I don't think so.

Leo: Have we traveled in time?

Steve: And now I'm thinking I'm just going to leave it with autofill, and it'll date the email on the send date rather than the podcast's date, since that would work, too. Anyway, we're going to have some fun. I hope that today's podcast will put to rest any question about what Mythos means because two days after last Tuesday's podcast, the entire industry of security professionals - Bruce Schneier, who we know well; Google's CISO; I mean, a who's who - all cosigned and authored and produced a document that is intended to get the industry's attention because they all agree with me. So I titled today's podcast, "Yes. Exactly." Which is meant...

Leo: I love the name. "Yes. Exactly."

Steve: Yes. It is what I said last week, but we've got - I want to share, just to really, you know, this is probably - remember I said that last week's show notes was revision 3. In the first revision I wrote: "This could be the most significant podcast in the history of Security Now!."

Leo: Woohoo.

Steve: Now, I did some deep breathing, and in one of the follow-on revisions I removed that sentence. But, you know, my point was we're talking about potentially - well, then, of course, then my working name was "Mythos: Marketing or Mayhem" because this could be a big deal. So anyway, I just want to - we've got an amazing document that I'm going to share. And the other thing that is useful is I've heard from some listeners who are having a hard time convincing upper management that they need to respond because of course any response is going to be expensive; right? I mean, it's going to require expenditures of talent, equipment, upgrade, upheaval, whatever.

Leo: Well, that's why the flaws are still there in the first place.

Steve: Exactly. And of course, and then there's the issue of the new things that haven't yet been found. So one of the things that this document offers, and in fact this is also the first time I've had two shortcuts in one podcast for the same file because you can get to this two ways: GRC.sc/mythos or episode number, GRC.sc/1075.

Leo: Nice.

Steve: Because I want there to be no possible reason that our listeners can't get the PDF and send it up to the C-suite folks because it is written for them. There are takeaways and bullet points and priority lists, and this is what you have to do because a tsunami is very likely coming.

Leo: Wow.

Steve: And in fact, I realized there's a - I'm already giving this away. I've got this whole thing in my head. There's very much of a Y2K aspect to this; right?

Leo: Oh.

Steve: Think about it; you know? Everyone who said after we went into the year 2000, oh, look, that was nothing. Nothing happened. Well, folks, there was a reason nothing happened. It's because everybody who needed to actually took it seriously and prevented something from happening. So anyway, we're going to have, I think, a great follow-up today to last week's - last week was just my opinion. Today we've got everybody's opinion. But we're going to talk about a disgruntled developer who has been disclosing multiple Windows zero-days because he's upset with Microsoft. Microsoft purchasing its own bugs in a massive campaign. The story behind something that's a couple weeks ago. Many of our listeners wrote to me. I didn't know what to say about it.

I could have talked about it last week, but actually I bumped it because last week's podcast was full, about how VeraCrypt and Wireshark and some other projects suddenly lost their dev accounts at Microsoft. They were like, what happened? Like they were unable to do revisions of their software for some reason. We now have the whole story there. So it's kind of good I waited a week because we were going to talk about it anyway. We've got a serious problem of recaptured domain names, which is reminiscent of the bucket reuse that we talked about with AWS a couple weeks back. Also a listener feedback-inspired exploration of how exactly AI might help to secure, might best help to secure open source repositories. A listener wrote to say, hey, I never heard you and Leo talk about your opinions of "Project Hail Mary."

Leo: Oh, yeah.

Steve: Could you say a few words? So we will. And then we're going to end with what cybersecurity professionals across the industry tell us about what Mythos means.

Leo: Oh, that'll be interesting.

Steve: So, and of course, again, the title of today's podcast is "Yes. Exactly."

Leo: That'll give you some idea of what's to come. Awesome. We also have a lovely Picture of the Week, which I haven't seen, but I know it's lovely because it always is. They're always fun.

Steve: I think this one, this one is a little bit of a hoot. So, yeah.

Leo: A bit of a hoot coming up.

Steve: A bit of a hoot.

Leo: All right. I am prepared to show all the Picture of the Week.

Steve: Okay. So I gave this picture the caption, "Hyphen usage is uncommon, but there are times when there's no substitute."

Leo: All right. This is going to be another punctuation issue?

Steve: Hyphen usage is uncommon, but there are times when there's no substitute.

Leo: Okay. Okay. Somebody took this quite literally. Do you want to describe it?

Steve: Yes. Yes. So we have a sign, you can sort of see a keyboard on a shelf above, and maybe a monitor. There's some sort of a - and there's a power strip in the back. There's some sort of some sensitive electronics and, you know, PC stuff.

Leo: We had signs like this all over our studio because I would always bring my coffee in and spill it.

Steve: Yes. And the sign says "No drinks back here unless then have a screw on top." And then it says "Thank Management." Now, they clearly meant a screw-on top, not a screw on top. So what we have them is a Styrofoam cup with one of those little plastic Styrofoam lids, and a long, about an inch and a half wood screw sitting on top of the cup. So it has a screw on top. Which of course...

Leo: All you need. It satisfies all requirements.

Steve: Yes. And had it said screw-on, then it would have been clear that you didn't want a screw on the top of the cup, you wanted a screw-on top...

Leo: Again, punctuation...

Steve: ...container. Yes.

Leo: ...can be very important.

Steve: Yes, not many people use hyphens. But I like the hyphen.

Leo: I like hyphens.

Steve: I do, too. Not quite clear when you need them, but I just say, hey, err in the direction of hyphenating because what the hell. Okay. So...

Leo: So we mentioned last time that Patch Tuesday last week was the second biggest patch Tuesday of all time. I mean, big numbers.

Steve: Yes. And we don't know yet whether that's Mythos-related, but we know that Microsoft is one of the companies that was named that Anthropic have given access to Mythos. And you have to wonder because, you know, I've heard people talk about, oh, you know, like somebody was saying on some show, oh, you know, what do they think? There's no way that it's not going to escape. It's not going to get out. It's like, they're not, you know, they're giving them access to the model online. They're not, like, giving them Mythos-to-go.

Leo: No. Here's your Mythos. We'll wrap it up for you. Would you like to wear it out, or...

Steve: Yes. Please make sure it's not like one of those iPhones that gets left at a bar when you walk, you know, an unreleased iPhone.

Leo: No, yeah.

Steve: So anyway, there's no problem. But the fact that it's cloud-based means that Microsoft would need to trust their competitor because of course they're all OpenAI, whereas this is coming from Anthropic. They would need to trust Anthropic with their source code uploads into Anthropic's cloud in order to have Mythos rummaging around in Microsoft's source. So there's that. But then on the other hand, everybody is going to have to trust Anthropic in that fashion because it's their cloud. There's no local Mythos yet, as far as I know.

Anyway, last Thursday the 16th, BleepingComputer's headline was "New Microsoft Defender 'RedSun' [that's the name that's been given to it] zero-day PoC grants SYSTEM privileges." Which, as we know, elevation of privilege is almost as important as remote code execution. Because oftentimes the remote code that you're executing is in the context of a user login, where the whole OS has got security wrapped around the user to keep the user from misbehaving. So you need to first get into the user account, but then you need to get out of the user account into the system account, you know, into root. So again, you know, elevation of privilege is a big deal.

So BleepingComputer's piece told the story of the disgruntled developer who had - and I'll share some of what this guy wrote because we'll get a - it sounds like he himself is a little more than disgruntled. He's a little sketchy. But anyway, this guy has been publishing - this is like not even the first - fully working proof-of-concept exploit code for his discoveries of privilege elevation vulnerabilities in both workstation and server. From like 2019 on, it's like Server 2019 on have been vulnerable to this.

So the following day, on the 17th - so on the 16th BleepingComputer's headline was "New Microsoft Windows Defender zero-day proof-of-concept grants system privileges." The next day they followed up that reporting with another piece titled: "Recently leaked Windows zero-days now exploited in attacks." In other words, this guy put the proof-of-concepts up on GitHub, and the next day bad guys had found them and were exploiting them to hurt Windows users. So, not good.

BleepingComputer said: "Threat actors are exploiting three recently disclosed Windows security vulnerabilities in attacks to gain SYSTEM or elevated administrator permissions. Since the start of the month, a security researcher known as 'Chaotic Eclipse' or 'Nightmare'" - oh, there's a hyphen - "'Nightmare-Eclipse'" - so he's hyphenating - "has published proof-of-concept exploit code for all three security issues in protest to how Microsoft's Security Response Center (MSRC) handled the disclosure process." And we're not getting much visibility into what that means, exactly.

Bleeping said: "Two of the vulnerabilities (dubbed BlueHammer and RedSun) are Microsoft Defender local privilege escalation flaws, while the third (known as UnDefend) can be exploited as a standard user to block Microsoft Defender definition updates. At the time of the leak, the security flaws these exploits targeted were considered zero-days by Microsoft's definition" - which, remember, is a little different than the industry's - "since they had no official patches or updates to address them." Normally zero-day is about surprise. In this case, it's about response, essentially, to something that hasn't yet been patched or updated.

They said: "On Thursday" - and this is of last week - "Huntress Labs security researchers reported seeing all three zero-day exploits deployed in the wild" - meaning in use to hurt people - "with the BlueHammer vulnerability being exploited since April 10. They also spotted UnDefend and RedSun exploits on a Windows device that was breached using a compromised SSLVPN user, in attacks showing evidence of what they're calling 'hands-on-keyboard threat actor activity.'" Meaning not just automated scan stuff, but an attacker logged in through an SSL VPN hitting keystrokes in order to explore and exploit the vulnerable connection.

They said: "While Microsoft is tracking the BlueHammer vulnerability as CVE-2026-33825 and has patched it in April 2026 security updates" - so it got fixed last Tuesday, which was Patch Tuesday this month - "attackers can use the RedSun exploit to gain SYSTEM privileges on Windows 10, Windows 11, and Windows Server 2019 and later systems when Windows Defender is enabled" - that actually uses Defender in order to leverage its attack, so weirdly disabling Defender disables the zero day - "even after applying the April Patch Tuesday updates." So this is a vulnerability post-Patch Tuesday of this month. So we don't know when it's going to get fixed. Maybe an emergency out-of-cycle update, who knows.

The disgruntled researcher explained. So this is the researcher saying: "When Windows Defender realizes that a malicious file has a cloud tag" - meaning you know, like the, remember the Mark of the Web which tags are able to get saying we're going to treat this differently because you downloaded this off the Internet. Says: "When a malicious file has a cloud tag" - this is the disgruntled researcher writing - "for whatever stupid and hilarious reason, the antivirus that's supposed to protect decides it would be a good idea to just rewrite the file it found to its original location. The proof-of-concept abuses this behavior to overwrite system files and gain administrative privilege." And we'll actually get a little more detail about that in a second.

So they wrote: "When BleepingComputer contacted Microsoft earlier this week for more information on the disclosure reported by the anonymous researcher, a Microsoft spokesperson told BleepingComputer" - and of course this is going to be as helpful as they generally are: "Microsoft has a customer commitment to investigate reported security issues and update impacted devices to protect customers as soon as possible. We also support coordinated vulnerability disclosure, a widely adopted industry practice that helps ensure issues are carefully investigated and addressed before public disclosure, supporting both customer protection and the security research community." So thank you for that, Microsoft.

So two days before that, on Wednesday, this person - this is me talking now - this person going by the moniker "Chaotic Eclipse" posted his own diatribe over on Blogspot. And I think it's worth sharing since it gives us some impression of who's doing the disgruntling. Dated Wednesday, April 15th, the Blogspot post was titled "Public disclosure, a response for CVE-2026-33825 patch." So it reads, posted by the guy: "Here is the code. Enjoy." And he's got a GitHub link, github.com/Nightmare-Eclipse/RedSun.

So he said - so it looks like Nightmare Eclipse is the guy's name, right, and RedSun is the exploit. "Now to address what some media articles wrote, first of all, I want to talk about MSRC official response regarding BlueHammer." That's his previous release of a zero-day exploit proof-of-concept code. He said: "Microsoft has a customer commitment to investigate reported security issues and update impacted devices" - oh. So this is him quoting what BleepingComputer just showed us they had said, for what it's worth.

So he said: "This is a very generic response, almost as if they don't care, and they don't. Why? Because MSRC was fully aware of this public disclosure, a case was filed but was dismissed by them, and they're also aware that this one will be disclosed. But again, they are ignorant." This is, again, Mr. Disgruntled.

"Normally," he writes, "I would go through the process of begging them to fix a bug. But to summarize, I was told personally by them that they will ruin my life, and they did. And I'm not sure if I was the only person who had this horrible experience, or a few people did. But I think most would just eat it and cut their losses. But for me, they took away everything. They mopped the floor with me and pulled every childish game they could. It was soo bad" - two o's on soo bad - "at some point I was wondering if I was dealing with a massive corporation, or someone who is just having fun seeing me suffer. But it seems to be a collective decision."

Leo: Wow.

Steve: I know. "And one other thing, they do everything but support the research community, I won't disclose details, but they sabotage people a lot. I mean, just look at the past. Microsoft is the only major company who had a track of multiple vulnerabilities being publicly disclosed just because the researchers were soo upset by how MSRC treated them. Unfortunately, the folks who have the capacity to stop those disclosures not only don't care, but also seem to push harder for even worse exploits to be released. I didn't want to be evil, but they are actively poking me to start releasing RCEs which I will be doing at some point." Dot dot dot. And he finishes: "I will personally make sure that it gets funnier every single time Microsoft releases a patch."

Okay. So we've talked before about vulnerability discoverers feeling that their brilliance is not being sufficiently recognized or rewarded. In an earlier posting on March 26th, last month, this person wrote: "I never wanted to reopen a blog and a new GitHub account to drop code. But someone violated our agreement and left me homeless with nothing. They knew this will happen, and they still stabbed me in the back anyways. This is their decision, not mine." Okay. So my presumption, without knowing anything specific, is that Microsoft almost certainly treated this researcher the same way they treat everyone else. But he believed, he or she believed that they deserved special treatment.

You know, we've certainly shared horror stories in the past about the way some researchers have been treated. But Microsoft is not evil. It's full of good people, but a great many good people. So the result is that it's a big lumbering machine that doesn't "care" about anything, but only because "caring" is not what big lumbering machines are optimized to do. This researcher appears to have adopted an "I didn't want to, but you made me do it" rationale for his actions. Reading between the lines, my guess is maybe

he was counting on receiving a big bug bounty payout that he desperately needed, which never came. Sort of sounds like he may have released a proof-of-concept before Microsoft went through the formal disclosure process and so blew his opportunity because he pre-released. And so now he's complaining about that. So now of course he's blaming Microsoft for this.

It's unfortunate that this person is having trouble with life. I looked at the details of the proof-of-concept that he designed, and it's a slick bit of work. The well-known security researcher, Will Dormann, from whom BleepingComputer often seeks confirmation of complex issues, posted about this new RedSun exploit over on Mastodon.

Will wrote: "From the same author as BlueHammer we now have RedSun. This works around 100% reliably to go from unprivileged user to SYSTEM against Windows 11 and Windows Server 2019 and beyond with April 2026 updates, as well as Windows 10, as long as you have Windows Defender enabled. Any system that has cldapi.dll should be affected." Okay. So cldapi sounds like the cloud, the Windows cloud API, and is.

In the next quote from Will, he refers to EICAR, E-I-C-A-R. That's the abbreviation for the European Institute for Computer Antivirus Research. The file they produced, which itself is known as EICAR, is a popular pseudo-malware test file that can be used to deliberately freak out any good AV tool without actually itself containing or doing anything malicious. It's just used as a test file to see if AV detects it.

So in a follow-on Mastodon posting, Will writes, to explain what this thing does, he says: "This exploit uses the 'Cloud Files API,' writes EICAR to a file using it" - meaning using the Cloud Files API - "uses an oplock to win a volume shadow copy race, and uses a directory junction/reparse point to redirect the file rewrite with new contents into C:\Windows\system32\ TieringEngineService.exe. At this point, the Cloud Files Infrastructure runs the attacker-planted TieringEngineService.exe (which is the RedSun.exe exploit itself) as SYSTEM." And he writes: "Game over." In other words, this is the proof-of-concept that the Disgruntled Developer engineered, which as I said is some slick work. I mean, that's kind of tricky to figure that out.

Anyway, our primary takeaway here is that all fully patched, as of last week's mega-Patch Tuesday - as you said, Leo, second biggest ever, but apparently not quite big enough - Windows desktop and server are currently vulnerable to this exploit, which is now being actively used in the wild. This is not the end of the world, since something bad must first get onto a machine so that it's able to trick Windows Defender into performing that odd file rewrite dance. And that allows attacker-provided code to be run with full SYSTEM privilege. But the attacker has to first get in there and provide the code. So as I said, while it's not the end of the world, Huntress Labs is observing it under active use. So it would be nice...

Leo: Wow.

Steve: Yeah, if Microsoft were to fix the issue for this before May's Patch Tuesday, which is still a full 3.5 weeks away. I mean, this is a bad problem; and, you know, Microsoft didn't get there in time, and they probably should get this updated.

Leo: In the meantime, I think it's hilarious, turn off Windows Defender.

Steve: Yes. Actually, that is the only - that's the mitigation.

Leo: That's hysterical.

Steve: Is turn off Defender because Defender's being used, this weird behavior that Defender has, and I'm sure Microsoft knows why they're doing this. But, you know, it ends up you can leverage that in order to get yourself attacked. Yeah. I wouldn't turn off Windows Defender. I don't know what I would do. Maybe the 0patch guys. I didn't think to look over at 0patch.com. Maybe they have a quickie patch. You know, they offer for free patches which Microsoft - patches for vulnerabilities that are known, but which Microsoft has not yet provided fixes. And they might do that. So that might be an opportunity, if they address this.

Leo: I'm checking it right now just to see. I'll let you know.

Steve: Cool. Meanwhile, Microsoft has been buying up their own bugs. And while we're on the topic of Microsoft and bugs, BleepingComputer also reported that Microsoft has been breaking records for bug bounty payouts. Before I take note of the irony inherent in this, I'll share what BleepingComputer wrote. They said: "Microsoft has awarded \$2.3 - get this - \$2.3 million to security researchers after receiving nearly 700 submissions during this year's Zero Day Quest" - which is the name of it, Zero Day Quest, you know, ZDQ - "hacking contest. Tom Gallagher, Vice President of Engineering at Microsoft Security Response Center" - that MSRC we were just talking about - "said that over 80 of the flaws found during the live event at Microsoft's Redmond campus were high-impact cloud and AI security vulnerabilities."

So that's just great. We've all been using that software, which has 80 high-impact cloud and security vulnerabilities. And actually Leo, that may more account for the April Patch Tuesday than Mythos stuff.

Leo: You don't really need Mythos. There's plenty to go around.

Steve: Well, just pay. Because \$2.3 million, I mean, what we've seen is this bug bounty concept paying - you need to motivate researchers who have, you know, only so many hours in the day to go running around chasing after Microsoft vulnerabilities, although doesn't seem that there's any scarcity of those.

So BleepingComputer continues: "Gallagher said: 'During the 2026 live hacking event, Microsoft partnered with the global security research community, representing more than 20 countries and a wide range of professional backgrounds, from high school students to college professors. Researchers conducted all testing within authorized environments in accordance with Microsoft's Rules of Engagement, demonstrating potential impact without accessing customer data or other tenant systems. Within these constraints, researchers identified critical paths involving credential exposure, SSRF (server-side request forgery) chains, and cross-tenant access.'" So lots of cloud, lots of AI.

BleepingComputer said: "Last August, Microsoft announced that it would increase the prize pool at this year's Zero Day Quest hacking contest to \$5 million in bounties, which the company described as the 'largest hacking event in history.' In 2025, Zero Day Quest also generated significant participation from the security community, following Microsoft's offer of \$4 million in rewards for vulnerabilities in cloud and AI products and platforms." So they bumped it from last year's 4 million to this year's 5 million. After the hacking competition concluded, Microsoft announced it had paid \$1.6 million in rewards after receiving more than 600 vulnerability submissions. So last year they offered 4 million,

they paid out 1.6 after 600 vulnerabilities. This year they offered 5 million and paid out 2.3 million. So lots more actual problems found.

"The Zero Day Quest contest is part of Microsoft's Secure Future Initiative (SFI), a cybersecurity engineering effort launched in November 2023, following a scathing report from the Cyber Safety Review Board of the U.S. Department of Homeland Security that found the company's security culture 'inadequate' and required 'an overhaul.'" And of course we talked about this at the time. I mean, it was just really raked him over, you know, charging excess money for logging security events. So they could just turn it on. But it's like, oh, no, let's make some more money from having so many bugs that people have to log in order to track them. Right.

Anyway, Bleeping said: "Last August, Gallagher said: 'As part of our Secure Future Initiative (SFI), we will transparently share critical vulnerabilities through the CVE program, even if no customer action is required.'" And here I hate this word, "Learnings from the Zero Day Quest will be shared across Microsoft" - they're sharing their learnings.

Leo: They love that word. I don't know why, but...

Steve: Oh, god, you know, I mean, yeah. The things we learned.

Leo: It's corporate talk, I think.

Steve: The things we learned is, well, the way I learned, my learnings, "to help improve Cloud and AI security in alignment with SFI's core principles: securing by default, by design, and in operations." Apparently, however, not in code.

Finally: "Earlier last August, Microsoft announced it had paid a record \$17 million to 344 security researchers across 59 countries through its bug bounty program between July 2024 and June 2025." You know, there's a mixed blessing, right, of bragging about how many millions of dollars you have paid to 344 security researchers who have found really bad problems with your software. On one hand, okay, I think it's great that Microsoft's software will now, or will soon be, you know, that many bugs fewer in cloud and AI security vulnerabilities. That's, of course, good for everyone. But as I said, it seems a little ironic to have Microsoft gleefully bragging about how many hundreds of bugs researchers were just able to find throughout their products when they were sufficiently motivated to do so.

So anyway, as we know, none of those bugs should have been there in the first place to be found. But hey, Microsoft has way more cash on hand than it knows what to do with. So dangling increasing quantities of cold hard cash in front of security researchers, who will then be motivated to go bug hunting, that's definitely money well spent. Let's have more of it because Microsoft can certainly afford to pay.

We're going to talk about the mysterious disappearing developer accounts, Leo, after you tell our listeners about one of our sponsors who's still here with us. They've not mysteriously disappeared.

Leo: No. Unlike those developer accounts. I just, I like the, what was it, hands-on keyboard? That's another good one. I never.

Steve: The hands-on keyboard attacker, yeah.

Leo: I love it. Is typing.

Steve: You can see, like, oh, Boris is hunting and pecking, yes.

Leo: Wow. Wow. Redcon5 in our Discord says the term "learnings" was not in common use in 19th and 20th centuries, although the countable noun sense "learning," as in things learned, dates to Middle English, and the plural "learnings" to early modern English. Note that early use of learnings often have the sense or connotation "teachings." Yeah, I've heard "teachings" before.

Steve: Yeah.

Leo: As was the case of learned generally has found occasional use for centuries, including by Shakespeare. In parallel construction...

Steve: So I guess if you're subject to teachings, what you come away with is learnings.

Leo: Learnings. I don't like it either. I agree with you 100%. Feels like corporate speak.

Steve: It feels like you've - what are you smoking on the peninsula out there?

Leo: Steve.

Steve: I think those little - remember when there were all those fake sweepstakes sites?

Leo: That's another way, yeah. Because we know that that happened on Facebook. That was the Cambridge Analytica scandal. They were making quizzes, like which Star Wars figure am I? And just by simply virtue of taking those quizzes, not only did they get all your Facebook information, they got all your friends' Facebook information. So even if you said, oh, I'm not going to do that, if your friends do it, they're revealing it. That hole's been plugged, but there's always more because they're like cockroaches. You can't get rid of them. All right. Anyway. On we go. Let's talk about the missing Microsofties.

Steve: Okay. So one of the recent bits of news that was, as I said, bumped so that we'd have time to thoroughly examine Anthropic's Mythos last week was that Microsoft had, without apparent cause or reason, suddenly dropped a number of driver developer accounts. Products such as VeraCrypt and WireGuard, I mean, like the well-known WireGuard, like, next-generation VPN. They both, they inherently incorporate kernel driver components in order to obtain the deep OS level access they need in order to operate. So this was a huge concern for many users of these products, and I heard from many of our listeners who picked up on this news.

Okay. As it turns out, it was, as I said before, it was just as well that I waited, since last week the news, the only news that we had was that the accounts had been dropped. We didn't know why. Today we know. The reason for Microsoft's suspension of these accounts turned out not to be a mistake, but was entirely deliberate, which is also a process or a reason for some concern. And I know probably it's going to hit home for many of our listeners since it's an issue that I've been talking about recently during that whole process of updating my code-signing certificate, what I had to go through. Which, you know, like getting my CPA to sign an attestation letter that yes, he'd just laid eyes on me, and I was real, and he was putting his license on the line in order to vouch for me. It's like, yikes. None of that had ever been needed before.

So once again, BleepingComputer was on top of this, under their heading "Microsoft rolls out fast-track to reinstate Windows hardware dev accounts." Kind of an oops. Anyway, BleepingComputer explained, writing: "Microsoft has rolled out a fast-track response to help developers regain access to accounts recently suspended from its Windows Hardware Program, following widespread complaints that they were locked out without warning. Last week, the company suspended Windows Hardware Developer accounts used to publish Windows drivers and updates for widely used tools like WireGuard, VeraCrypt, MemTest86 [super popular], and Windscribe. The suspensions prevented developers from releasing new Windows builds and security patches, raising concerns about potential delays in responding to vulnerabilities.

"VeraCrypt developer Mounir Idrassi stated that his account - so this is VeraCrypt, the successor to TrueCrypt, as we know, that was taken over by someone in France, this Idrassi guy, he said his account had been terminated without warning, and that he was unable to reach a human support representative, leaving him unable to publish Windows updates. Similar experiences were reported by WireGuard maintainer Jason Donenfeld and others, who described being locked out without facing any - or without access or facing lengthy or unclear appeals processes." You know, there was the machine, the Microsoft machine was just sort of ignoring him.

"After many developers took to X to report the suspensions, Microsoft Vice President Scott Hanselman said the accounts were suspended for failing to complete identity verification in the Windows Hardware Program, and that the company had been emailing these people, which they call partners, about the requirement since October of 2025." Right? So six months of we're trying to reach you, and you're not replying.

"Microsoft requires identity verification for the Windows Hardware Program because it allows developers to sign and distribute" - actually it doesn't anymore, but okay. Remember Microsoft is doing the signing now - "and distribute kernel-level drivers." It does allow developers to develop kernel level drivers under the program, "which run," writes BleepingComputer, "with high privileges" - well, yeah, in the kernel, like it could do anything - "and have been abused by threat actors in past attacks.

"However," they write, "many developers claimed," and there are so many it's probably true, "claimed they had not received any prior notification, including emails, before they were suspended. While Hanselman and others at Microsoft have been working to reinstate accounts, Microsoft yesterday introduced a temporary process to fast-track reinstatement for suspended accounts. An update to Microsoft's advisory adds: 'We've heard your feedback.'" Uh-huh. "'We know that some partners whose accounts were suspended following Account Verification are experiencing challenges regaining access to the Hardware Dev Center (HDC). Protecting the security of the Windows ecosystem remains our highest priority, and we are adding a temporary process to accelerate the reinstatement experience for partners who are able to resolve outstanding compliance requirements.'" Wow.

"Under the new process, developers are told," they wrote, "to open a support case through the Hardware Program as the fastest way to reinstate accounts. Requests must include a clear business justification explaining how access to the Hardware Dev Center will be used. Microsoft says that once reinstated, all outstanding compliance requirements must still be resolved before full access is restored." So this is sort of an interim, you know, we're on a provisional trial basis, we're going to give you your account back that you previously had access to for years, until we decided nope, no more. Suddenly we don't know who you are. But now you're going to have to tell us who you are and prove it, of course.

So Microsoft said it "advised partners to ensure they're signed in with the correct account when submitting tickets and to continue prompting Copilot to create a ticket if automated assistance fails." Whatever that means. "For those unable to submit requests through standard channels, Microsoft provided an alternative support contact to help initiate the process. Microsoft has not said how long this accelerated process will remain in effect, that is, what this grace period is going to be, so affected developers are advised to act quickly."

Okay. So I would tend to believe the developers over Microsoft regarding this complete lack of attempts to inform them. As I noted earlier, Microsoft is no longer an entity that is actually able to care. Caring is not something that it does. It's just too big, and caring is a distraction. So someone, somewhere, doubtless decided that the best way to get developer attention or just to remove dead accounts would be to simply suspend all currently non-compliant accounts for non-compliance. This has the advantage, as I said, of weeding out any older accounts that no one really cares about that much, since they won't be immediately inconvenienced by their inability to access Microsoft's developer portal.

And conversely, those who are inconvenienced will be highly motivated to get their identity proving act together. As we know, this may involve getting an affiliated attorney or CPA to sign some attestation papers. It's what I went through. Basically it's the same process in order to - you need to have, you know, identity verification that is bulletproof. And of course that's true because running code in the Windows kernel is a privilege that no one, none of us want bad guys to have. So we do want Microsoft to make that as bulletproof as possible.

You know, while Microsoft could have been way more gentle about this, this did get the job done. So, you know, that's what that was all about. The reason Microsoft suddenly suspended a bunch of developer accounts, many who were immediately inconvenienced because they were using them actively, now, you know, basically it's like, okay, we're going to give them back to you for a while, but you need to get your identity made compliant. So that will happen.

Okay, now, this next piece of news beautifully exemplifies a problem we've seen before that's I think largely a consequence of the aging Internet and aspects, critical aspects of its design that were never very well thought through since, and in defense of its designers, they could have never, and never did, foresee what their creation, which we call the Internet, would become. They, I mean, I'm in awe of the original design of these protocols that have stood the test of time. There are some aspects that haven't.

BleepingComputer's headline for this reporting was "Signed software abused to deploy antivirus-killing scripts." Not a great headline. While that's factually true, it's more of the consequence of the problem than the problem itself. Okay. But let's start with what BleepingComputer reported.

They said: "A digitally signed, meaning there's a real company behind it, so it's digitally signed, and pretty much these days anything has to be, a digitally signed adware tool, so

not malware, not, you know, not evil, but just unwanted, a digitally signed adware tool has deployed payloads running with SYSTEM privileges that disabled antivirus protections on tens of thousands of endpoints, meaning host computers where it was installed, some in the educational, utilities, government, and healthcare sectors. In a single day, researchers observed more than 23,500 infected hosts across 124 countries trying to connect to the operator's infrastructure, with hundreds of infected endpoints being present in high-value networks."

Okay. So they're saying 23,500 PCs have this adware tool. Hundreds of them are in, like, really important networks. And they're all reaching out, trying to phone home to this operator's infrastructure.

Bleeping wrote: "Security researchers at managed security company Huntress discovered the campaign on March 22nd, when signed executables viewed as potentially unwanted programs" - love that, PUPs, Potentially Unwanted Programs. Do you really want this? Which is what that original OptOut tool that I wrote for that old Radiate or Aureate adware was. Anyway, they said: "Potentially Unwanted Programs triggered alerts in multiple managed environments." So Huntress is in the environment management business, and they saw these things doing this. They wrote: "PUPs, or adware, are regarded more as a nuisance than malicious, as their role is typically to generate revenue for the developer by showing advertisement pop-ups, banners, or through browser redirects." You know, they'll infect browser URLs to bounce through some other redirect before they go to that site that you actually intend.

"Huntress researchers say that the software was signed by a company called Dragon Boss Solutions LLC [sounds kind of Chinese] involved in 'search monetization research' [whatever that is] activity and promoting various tools, for example, the Chromstera Browser" - oh, Leo, don't leave home without the Chromstera Browser - "Chromnius [whatever that is], the WorldWideWeb [oh, that's catchy], Web Genius, and the Artificius Browser." I don't know if I want the Artificius browser. Anyway, "all labeled as browsers but detected as PUPs by multiple security solutions."

Leo: Yeah, yeah.

Steve: So they're recognized as are you sure you want this? "Beyond annoying users with ads and redirects, Huntress researchers say the browsers from Dragon Boss Solutions also feature an advanced update mechanism that deploys" - and get this - "an antivirus killer." In other words, they found out that there's things that don't like them, so let's kill that because we don't want to be unliked. "Huntress researchers discovered that the operation relied on the update mechanism from the commercial Advanced Installer authoring tool to deploy MSI and PowerShell payloads.

"Analyzing the configuration file for the update process revealed several flags that made the operation completely silent, no user interaction required." You don't want to bother users with those pesky permission dialogs. "It also installed the payloads with elevated system privileges, prevented users from disabling automatic updates, and checked frequently for new updates." So basically badly behaving malware. I would agree potentially unwanted. Probably definitely unwanted. That would be DUP, Definitely Unwanted Programs.

Leo: It's not a PUP, it's a DUP.

Steve: That's right. Okay. So none of those things seem deliberately malicious; right? Having been harassed by false-positive AV detections, I can at least understand their motivation behind creating exceptions for one's code. As we know, that's not the approach I take, like killing off AV that bothers you. Mostly this seems like software written entirely with the convenience of its publisher, rather than its user, in mind. That's bad software, no doubt about it. But that's also life.

So the reporting continues, saying: "According to the researchers, the update process retrieves an MSI payload (Setup.msi) disguised" - this is weird - "disguised as a GIF image, which is currently flagged as malicious on VirusTotal," but only by five out of 69 or 70 security vendors. So not many false positives, or positive positives. Anyway, it does seem a little sketchy. Why would any software publisher who thinks of themselves as legitimate retrieve a Windows Setup.msi file disguised as a GIF image? What?

They continue, writing: "The MSI payload includes several legitimate DLLs that Advanced Installer uses for specific tasks, such as executing PowerShell scripts, looking for specific software on the system, or other custom actions defined in a separate file named '!_StringData' that includes instructions for the installer.

"Huntress says that before deploying the main payload, the MSI installer conducts reconnaissance by checking the admin status, detecting virtual machines, verifying Internet connectivity, and querying the registry for installed antivirus (AV) products from Malwarebytes, Kaspersky, McAfee, and ESET. The security products are disabled using a PowerShell script named ClockRemoval.ps1, which is placed in two locations. The researchers say that installers for the Opera, Chrome, Firefox, and Edge browsers are also targeted, likely to avoid potential interference with the adware's browser hijacking." Yeah, you want to turn off anything that might get in the way.

"The ClockRemoval.ps1 script also executes a routine when the system boots, and at logon, and every 30 minutes, to make sure that antivirus products are no longer present on the system by stopping services, killing processes, deleting installation directories" - wow, you know, really wiping them - "and registry entries, silently running vendors' uninstallers [love that] and forcefully deleting files when uninstallers fail to successfully uninstall. It also ensures that the security products cannot be reinstalled or updated by blocking the vendor's domains through modifying the hosts file and null-routing them, redirecting them to 0.0.0.0." Wow. So again, not technically malicious, but you just don't want that.

So what's clearly going on here is that the publishers of this mal-behaving crapware have previously experienced well-deserved run-ins with a handful of alert anti-crapware utilities that want to warn their users that this is a potentially unwanted program, in spades. So these cretins have upped the ante by making their adware offerings even more obnoxious in the things they do to get anything that doesn't like them off the system and keep them off. Like you can't even contact those AV companies any longer because your browser will not resolve the domain because the hosts file has been edited in order to null route them. Wow.

Okay. So here's something curious and interesting: "During the analysis, Huntress found that the operator did not register the main update domain (chromsterabrowser[.]com) or the fallback one (worldwidewebframework3[.]com) used in the campaign, presenting them with the opportunity" - presenting them, Huntress - "with the opportunity to sinkhole the connection from all infected hosts." In other words, domain got abandoned. Huntress saw nobody had registered it, so they did.

"As such," writes Bleeping Computer, "they registered the main update domain and watched 'tens of thousands of compromised endpoints reach out looking for instructions that, in the wrong hands, could have been anything.'" Based on the source IP addresses

of the endpoints" - these PCs that have this crap on them - "the researchers identified 324 infected hosts residing in high-value networks." Remember, that's 324 out of 23,500.

So there's 23,500 PCs overall, 324 in high-value networks. Specifically, "221 academic institutions in North America, Europe, and Asia; 41 OT, as we're calling them now, Operational Technology networks in the energy and transport sectors, and at critical infrastructure providers; 35 municipal governments, state agencies, and public utilities; 24 primary and secondary educational institutions; three healthcare organizations (hospital systems and healthcare providers); and the networks of multiple Fortune 500 companies." So if a bad guy had registered that domain before Huntress did, they'd have access into all of those networks. BleepingComputer wrote that they tried to reach out to Dragon Boss Solutions, but could not find contact information as their site is no longer operational.

"Huntress warns that, while the malicious tool currently uses an AV killer, the mechanism to introduce far more dangerous payloads into infected systems is in place, and could be leveraged at any time to escalate the attacks. Additionally, since the main update domain was not registered, anyone could claim it and push arbitrary payloads to thousands of already infected machines with no security solutions protecting them, by design, and through an already established infrastructure.

"Huntress recommends that system admins look for WMI event subscriptions containing [the string] 'MbRemoval' or 'MbSetup,' scheduled tasks referencing 'WMIload' or 'ClockRemoval,' and processes signed by Dragon Boss Solutions LLC. Additionally, review the hosts file for entries blocking AV vendor domains and check Microsoft Defender exclusions for suspicious paths such as 'DGoogle,' 'EMicrosoft,' or 'DDapps.'"

Okay. So this particular incident is not the end of the world. But as I noted at the start, it's another perfect example of something the Internet was never designed to handle. Some random company may itself not be explicitly evil, but might have sloppy, uncaring, and abusive coders who install software that does things to its hosting PCs that would raise serious concerns from anyone who understood what was going on. But as we know, the phrase "from anyone who understood what was going on" is almost never going to include the end-user who decided, "Hey, you know, I'll bet that Chromstera Browser would be a lot better than Chrome. So I'm using Chromstera instead of Chrome." It's like, okay.

Leo: Can't wait.

Steve: Yeah. So here's the problem that the Internet's designers never considered: What happens when the progenitors of ill-begotten and very badly designed software, and not necessarily even that, like any software which is now using an infrastructure to phone home to check for updates, and then has the power to automatically download them and put them in place, what happens when that software, which continually reaches out to the Internet for updates, eventually, and if it's a fly-by-night company, probably inevitably goes out of business? Their horrible software remains installed and alive and querying for updates. I know that all of us have stuff on our PCs that we installed some time ago and then stopped using, but probably haven't taken the time to remove because it's not bothering us. But then their various domains also expire. Oops. Now anybody could reregister them.

Fortunately, in this instance, Huntress are the good guys who re-registered those expired domains for the sake of their research. But if bad guys were to do this, they would have stumbled upon the mother lode: 221 academic institutions; 41 Operational Technology

networks and infrastructure providers; 35 municipal governments, state agencies, and public utilities; 24 school systems; three healthcare organizations; and the networks of multiple Fortune 500 companies. They could get into all of them. Ransomware, anyone? This abandoned software would literally have a ready-to-go backdoor into the networks of all of those 324 high-value targets.

And here's the concern to think about: This cannot be an isolated event. This particular discovery was Huntress showing that they're awake and alert and doing their managed security thing. That's great. But similar events are doubtless happening across the Internet. Companies are abandoning their previous failed software offerings which included technology to phone home. Then home is abandoned, too. Note that it's one thing when some random website's domain is abandoned. But it's an entirely different matter when automation that's been silently installed into user machines is making those queries. This creates a ready-made backdoor into every one of the networks that's reaching out to abandoned domains.

You know, we're in a world where there is no accountability for the actions of the software while it's in use; right? I mean, people can download this crapware, and it does that to their machines. Horrible things, installing scheduled tasks, stripping AV out, running the AV uninstallers. And if that doesn't work, removing their registry entries and manually deleting the software from their machines, blackholing their domains by putting 0.0.0.0 in the hosts file. And the user doesn't know. They said, yeah, I really want the Chromstera browser. Sounds great. And this happens. There's no accountability in our current environment. Companies can do whatever they want, including this kind of crap.

You know, basically we're in a world where we have a rent-a-domain-name system; right? We rent a domain name, and as long as we're willing to pay for it, we get to keep it. But when we decide we don't want to rent that domain name any longer, after it expires, it's up for grabs. Just like the AWS abandoned bucket problem was, where bad guys could grab abandoned buckets that still had activity on them. So unfortunately, this reregistering a domain is assumed and encouraged. But it leaves us with some serious potential for security problems. It's not something our forefathers on the Internet thought about because they could have never imagined that the 'Net would become what it has. But this problem of recycling domains, it creates a whole new world of security problems.

Leo: What an interesting story, yeah. Chromstera. I can't wait to get it. By the way, I just saw this news cross the wire. Mozilla is saying now that it used Mythos on Firefox, and that it found 271 bugs, which they patched in their current version 150. So this is the first that we've seen of an actual admission that Mythos was used, and by an independent third party.

Steve: Yup.

Leo: 271 bugs in shipping software.

Steve: Yeah.

Leo: It's been tested and tested and tested.

Steve: Oh my god, pounded on. And we know it is the largest attack surface on anyone's computer is the web browser.

Leo: Yeah, yeah.

Steve: Very cool.

Leo: One of the things Mozilla said is our belief is that tools have changed dramatically, and there were categories of bugs that you could find with human analysis, you couldn't find with automated analysis. Which means that threat actors, you know, had an advantage. If they were willing to spend the time and energy, we couldn't keep up.

Steve: And now it is finding them with automated analysis. Wow.

Leo: Yeah. "Every piece of software" - this is, by the way, this is Holley, Bobby Holley, Firefox's CTO - "every piece of software is going to have to make this transition because every piece of software has a lot of bugs buried underneath the surface that are now discoverable. This is a transitory moment that is difficult and requires coordinated focus and a lot of grit to get through, but I think this is a finite moment, even as the models become more advanced." He said, "Yes, we are flooded now with things we have to fix. But at least we know about them."

Steve: Yeah. And when AI is in the pre-delivery pipeline, we are not going to be there again. So as I said, we are going to have - it's transient mayhem potentially. It's Y2K. And Y2K is a perfect model.

Leo: I think this confirms what you just said, exactly. Yup. It is Y2K. It's hair on fire. But for a limited time only.

Steve: Yes. And now, and you know that going forward, with the Mozilla team having seen this, they will vet anything they do now through AI like Hyperlint in order to catch...

Leo: Exactly.

Steve: ...any of the problems before they ship.

Leo: That's what Holley is saying, basically, is this is now incumbent on everybody.

Steve: It's the new model.

Leo: It's the new model. It's the future. But this is a real confirmation that Mythos, it wasn't merely marketing hype, that there is something going on. If you can find 271 bugs in a highly tested current version of...

Steve: Please tell Jeff and Paris. I'm so annoyed with them. Like, is it really? Yes. Read something.

Leo: Yes. It is now. Well, we didn't, I mean, to be fair, we weren't sure. You know, because...

Steve: I was.

Leo: Yeah. You said that last week.

Steve: I read the [crosstalk].

Leo: Yeah, yeah. And I mentioned that to them. But now we have absolute confirmation. This is the real deal. They've been using automated tools before. This is not - this is a special category.

Steve: And I will, when we get to our main topic, I will - the guys at AISLE, remember A-I-S-L-E, they're the guys that found all the problems in OpenSSL. And so they have a little bit of pushback against Anthropic, which I'll share.

Leo: Okay.

Steve: To round this out.

Leo: Okay.

Steve: But anyway, as I said, this podcast is titled "Yes. Exactly."

Leo: Exactly. And you called it. You were absolutely right. And we've got some feedback. What? You're muted. Hello?

Steve: Thank you. Okay. So, feedback. A listener shared some musings over strategies for securing open source repositories, and it provided a perfect setup for looking at this aspect of the future. So his name is Gene Hastings, who listens to us. I'm familiar with the name. He's sent email in the past.

He wrote: "A colleague and I often meet to talk about DevOps and related issues, you know, system and personal health. He's more dev, I'm more ops. Both often cranky." One of our listeners. "In any event, we were talking about the nightmare that is having a project's dependence on libraries all over the 'Net, and what steps might be taken to provide some degree of defense." He said: "I was already aware of version pinning, and there was the recent news about a compromised package where the infection modified it without changing the version. I recalled long after our conversation that one would need

to store a hash of the package and compare it on retrieval." Right, because a modification would get detected, the hashes would match.

He said: "Little protection against a compromised new version or a first-time use, but some nonetheless. There is also the concern as to the trustworthiness of the package's own dependencies. All this led me to reflect that what may need to happen next is to have each package and its components, not only signed by the author, but also by an independent auditor. Obviously, this does not scale physically or financially. So the next step is to have a trusted Agentic Auditor that does not charge a fortune for each signing. Such automation will be necessary soon.

"This led me to a further thought. Imagine a new project philosophically akin to Let's Encrypt, a service for smaller developers who can do an automatic audit at a tolerable expense," he says, "if all of the following are true: The agents, like Mythos and descendants, are competent. The agents are efficient. The agents are trustworthy. The agents are not priced out of reach, with some flavor for everyone. And the OWNERS of the agents are trustworthy!" He has an exclamation point on that one. He said: "Then there could be a future for us and the Internet." Apparently otherwise forget about it. It's all over.

He said: "As an aside, I am an AI skeptic. I do not trust that which cannot be explained. Getting back to operations, if I don't have a half decent idea what a system and its configuration is doing, I am very reluctant to put my name on it. I am willing to trust people who are able to understand the systems to assure me that I can be fairly reassured. At the moment, such people are hard to find amid the tsunami of hype. I'm not as concerned about the quality of the technologies as I am about the people pushing them. I wouldn't trust simple driving directions from the likes of Sam Altman, Mark Zuckerberg, or Jeff Bezos. I do not trust their motives or plans. Signed, Gene Hastings."

Leo: Love it. Thank you.

Steve: So as he said, "often cranky." Cranky Gene. He's suggesting a future solution which might be a system in the form of Let's Encrypt where individual developers would need to have an AI-based agent audit their code for problems, and unsuspected or unwanted behaviors, and would then sign the library all for a low cost. The trouble with this is that then we need some authority to manage the trust in these AI agent signatures and, on the trusting end, some sort of new root store that users of these signed libraries could use to lookup and verify the trust. In other words, a whole bunch of new stuff.

I think there's a more direct, cleaner, and straightforward means of accomplishing the same thing. We simply move to a world, very much like what we were just talking about, Leo, with Mozilla. We move to a world where anything that a public code repository offers for broad public consumption first passes the scrutiny of an AI agent. An AI will be guarding the exits, essentially. Code cannot leave the repository without first being checked by the AI agent. And the process might not be autonomous. You know, the repository's AI might have some questions for a package's authors that would need to be answered and negotiated before a new or updated package could be made widely available.

And since the use of an AI will certainly come at a non-zero cost for the foreseeable future, at least, I mean, there'll probably always be some cost because this is always going to be some compute, I'd imagine that there would be some form of rate-limiting on new submissions being made available publicly for review and publishing. You know, non-professional authors who are in the habit of constantly revising their code to make

an endless series of incremental improvements might have a release delay or some sort of submission limit imposed. But the idea being, in the same way that Mozilla will be running their Firefox code from now on through AI, the solution is for repositories to do the same thing, to clean anything that is being released to the public before it gets out there.

And I suspect that solves this problem. You know, the vast majority of a repository's code is mostly static; right? So an AI would only need to give it the once-over one time. And from then on those who pull it could rely upon its security more than they ever have been able to before. And most code only changes incrementally. So an AI could retain the context that it developed during that original once-over, and then bring itself back up to speed and only look at the changes, all the deltas to the code, in order to minimize the recurring cost of continuing to review code which incrementally changes over time. So I think the whole system can be made practical.

And so what I know is this: The year is currently 2026 - see, I got it right this time, it's not 2024, 2026 - when AI costs today far more to run than it's able to generate in revenue. I am sure that the economics of AI will be radically different in the future, just as the economics, for example, of mass storage and computation have been utterly transformed over the past 50 years.

Leo: Yeah. There's a rich history of this. This has always happened, yeah.

Steve: Today we're all walking around with globally connected pocket computers that would have boggled the minds of our grandparents.

Leo: Yeah, yeah. Our parents. Forget the grandparents, yeah.

Steve: Yes, yes. It should be clear to everyone that AI, which continues to boggle our minds today, will be just as accepted and taken for granted by our grandkids as the Internet is by today's kids. So, I mean, kids growing up today, they've always had the Internet. That's just like, yeah, they don't know life without it. We're still sort of like, wow, remember those days? Remember books?

Leo: I used to - remember CDS?

Steve: That's right.

Leo: DVDs?

Steve: That's right.

Leo: Records?

Steve: And finally, GP, our listener, says: "Dear Steve and Leo. Given April's security-related news, I can see how thoughts on the 'Project Hail Mary' movie might have been pushed to the wayside. I'm wondering what you gentlemen thought of the film and its

treatment of the source material? I felt the movie struck a nice balance; it did justice to the book while allowing those who have not read it to enjoy the story without being overwhelmed by a flood of science, which could have easily turned it into a five-part miniseries."

Leo: Oh, yeah. There's a lot of science in the book, yeah.

Steve: Yeah, there is. Well, and that's why we love Andy Weir's writing.

Leo: Right.

Steve: So he said: "My young one enjoyed the movie so much that they wanted to read the book."

Leo: Oh, good.

Steve: Yeah.

Leo: That's good.

Steve: Yes. So we signed up to borrow it from the library. However, we were number 110 in the queue of the public library to borrow the book.

Leo: Wow. Yeah. It's on the best-seller list again, I think, yeah.

Steve: Yeah. "So we opted for my old copy on Audible instead." Although I salute him for reading because I'm still a - I think reading is primal. But anyway, he said: "Listening to the story again did not diminish the movie; it only enhanced the experience for both me and my little one. It's like getting the 'inside story,' if you get my drift. This is one of the few times in recent history where a movie did not ruin the book, but actually improved upon it. Good job to the production team on this one. All the best. GP." So Leo?

Leo: Yeah, I'd agree 100%. In fact, I'm relistening to the book, which I started right after the movie. The other thing we did do, though, is we also rewatched "The Martian" because Lisa and I had a little inside bet. Because after the "Project Hail Mary," I said, oh, that was as good as "The Martian." She said no, it wasn't. I mean, she said it was really good, but it's not as good as "The Martian." And I said, oh. And then we watched "The Martian," and I have to agree with her. "The Martian" was remarkably good.

Steve: Yeah.

Leo: I think that's partly because Ridley Scott directed it. I think the directors of "Project Hail Mary," who chiefly are famous for "The LEGO Movie," maybe have a

little bit more of a kiddy sensibility. I can see how it appealed to his little ones. But yes, you know, Grace, Ryland Grace is, like, there's a lot of times he goes "Oh ho ho ho," you know, things that kids would like.

Steve: Yeah.

Leo: But it's a little over the top. That bugged me a little. I do feel it was very true to the book. The book has infinitely more detail.

Steve: Yes.

Leo: Because you had to cut all that stuff out. I'd forgotten how much science there is in the book.

Steve: Yeah.

Leo: And so there's stuff that I thought, oh, boy, they left that part out of the movie. But I had forgotten it.

Steve: And for example, I loved the details of breeding astrophage from the book.

Leo: Yes. Yes.

Steve: It was so good.

Leo: Right.

Steve: And we just got a little suggestion of it in the movie.

Leo: Almost all the science is suggested, you know, in the movie, yeah.

Steve: Yeah, yeah.

Leo: Yeah. They focus on the drama, the interpersonal relationships. And the science gets a second best.

Steve: So my theory, because, you know, I reread the book when we knew there was going to be a movie because I read it originally when Andy wrote it. And I thought about this question of movie versus book a lot. Of course famously I've complained here that "Jurassic Park," when I was watching the movie, I was incensed because so much was left out that was in...

Leo: Right.

Steve: I mean, some arguably really important stuff. On the other hand, look, "Jurassic Park" was a phenomenon as a movie. So who can say that, like - what I've decided is it's really not fair to compare.

Leo: Yes.

Steve: They are two - what they have in common is a similar plot. So they have the concept and the plot. But you really are addressing two different audiences. A book reader or Audible listener is a different audience than somebody who wants to go to a movie in two hours and be entertained.

Leo: They're different media.

Steve: Yes.

Leo: Absolutely. And you have to be native to the medium. Otherwise it just isn't going to work. And I understand that. But I do agree with you that this, which is unusual, this movie makes you want to read the book, which is really great. And you don't feel disappointed in either direction. Which is very unusual. I almost always feel disappointed by science fiction books not living up to the movie.

Steve: Yeah.

Leo: In this case, no, I think for both "The Martian" and "Project Hail Mary," the movies are great. They really do a good job, yeah. So we're in agreement.

Steve: Yeah. Okay. Let's take a break, and then we're going to plow into what the experts say and what - you just shared a perfect example from Mozilla, what they found when they ran Mythos against their Firefox codebase.

Leo: Yeah, yeah, very interesting. I do have to point out, Redcon5 asked in the Discord chat, our Club TWiT chat, how many of the 271 bugs were severe or were, you know, and they actually didn't talk about severity. So I don't know. They might have been smaller bugs. We don't know. So that's the next question. But I guess a bug is a bug is a bug. I mean...

Steve: And we know how often bugs can be elevated...

Leo: Right.

Steve: Yes, into something more severe.

Leo: Right, right. Fifteen of the Firefox CVEs were low, 18 were moderate, 13 were high, at least of those. At least that's according to JokinBokin. So, on the YouTube.

Steve: And I realize that the proper response to the guy in the Club is listen to what the Mozilla guy is saying.

Leo: Yeah.

Steve: He is saying...

Leo: You've got to do this.

Steve: ...this is significant.

Leo: Yeah.

Steve: You know, this wasn't dust that was found, you know. They were like, whoa. So...

Leo: Yeah. And that number is huge, 271 is mindboggling. But if 13 were high, this is from version 149 to 150. This is huge.

Steve: Yeah.

Leo: Anyway, let's talk about it.

Steve: And just one package. We're talking about, I mean, all - think of all the software in the industry. Think of how minutely Firefox has been curated and developed over time, how much scrutiny it's received. And even so, AI found what people could not.

Leo: Yup.

Steve: Now, imagine that typical software that's just thrown together and out the door.

Leo: Think about Windows, how many hundreds of millions of lines of code.

Steve: Well, and how many bugs they know about and, like, eh. Remember, didn't they ship famously 7 with like 10,000 or 20,000, like, known bugs? Like, what? How does it even get off the ground?

Leo: Oh, my gosh. Yeah, all right.

Steve: It's a revolution. Okay. So.

Leo: Yes, exactly, Steve.

Steve: Yeah. As I noted several times last week, my original working title for last week's podcast was "Mythos: Marketing or Mayhem?" But once I'd assembled and examined all the data, I realized that leaving the question, or the answer to the question, that that title implied up in the air would be wrong because there's no way, really, after looking at the facts and just with no bias, there's no way that Mythos was only marketing. We had evidence of it. So, you know, I acknowledged also that it was certainly also marketing, but it was also far more than only that. And I think that's where people get confused is they just mistrust people's motives to such a degree these days...

Leo: That's true. They don't trust anybody.

Steve: That it's like, oh. But again, it could be both. And it was. It happens that Anthropic used this for marketing. But I'm going to make the point at the end of the podcast, thank god, because it broke out. That's the difference. And this breakout is what we're talking about today. I titled today's podcast "Yes. Exactly." Because last Thursday, two days after, as I said at the top of the show two days after our "What Mythos Means" podcast was delivered, an incredibly significant group of industry veterans who pretty much comprise a who's who of the cybersecurity industry all weighed in with a formal emergency wakeup call for the entire cybersecurity world. The organizer and publisher was a group calling themselves the cloud security alliance, and I have a link to the most recent version of their 23-page paper in the show notes. They titled it "The AI Vulnerability Storm: Building a Mythos-Ready Security Program."

So the paper enumerates its sixteen (16) primary contributing authors. Because this is important for appreciating the weight of the paper's stated concerns, I'm going to share them briefly. They are: Jen Easterly, CEO of the RSA Conference and Former Director of CISA. Bruce Schneier, who we all know, renowned cryptographer, current Chief of Security Architecture at Inrupt and Fellow and lecturer at the Harvard Kennedy School. Chris Inglis, the White House's former National Cyber Director. Phil Venables, Ballistic Ventures. He is formerly the CISO of Google Cloud. Heather Adkins, current CISO of Google. Rob Joyce, the NSA's former Cybersecurity Director.

Sounil Yu, the CTO of Knostic and former Chief Security Scientist for Bank of America. Katie Moussouris, the Founder and CEO of Luta Security. John N. Stewart, Talons Ventures and former CSTO for Cisco. James Lyne, CEO of the SANS Security Institute. Dave Lewis, Global Advisory CISO for 1Password. Maxim Kovalsky, Managing Director of AI Security CoE for Consortium Networks. Jim Reavis and John Yeoh, who are the CEO and CSO respectively of that Cloud Security Alliance. Joshua Saxe, CTO and Co-founder at Security Superintelligence Labs, former AI and Llama Security Lead at Meta. And finally Ramy Houssaini, CCSO for none other than Cloudflare.

So as I said, the who's who. In addition to those primary contributing authors, the paper's content was also reviewed by a list of CISOs that pretty much includes everyone else. I'm not going to read them since there are too many of them. But I've reproduced that page from the report in the show notes. So you can just see it. I mean, it is, like, if there's anybody who I didn't just read, a former head of security for Netflix. CISO for Brave Technology. You know, I mean, Global Field CISO for Fastly. You know, your eye just drops on any of them. I mean, so, you know, everybody basically understood what Mythos meant.

Okay. So we've clearly established the provenance of this document. So I want to first share the Executive Summary overview, then the key takeaways for CISOs, followed by their brief summary of why Mythos is so important. Much of this will sound exactly like I did last week, two days before this was published, which is, of course, why I immodestly titled today's podcast "Yes. Exactly." This amazing group of experts even use some of the same phrases that I used. Given the impossible-to-exaggerate significance of Mythos and the successor systems that are sure to follow, and not only from Anthropic, but I get it. As I said last week, they're just first. But they were the one that broke through. And breaking through is what really needed for our industry to get the wakeup call it needs. So I think it's crucial for the listeners of this podcast to appreciate that it's not just me with a lone opinion here.

Okay. So the authors of the Executive Summary set it up as a sort of topical Q&A. They wrote: "What happened?" Answered: "AI, as demonstrated by Anthropic's Mythos" - so again, noted that even they didn't fall to their knees in front of Mythos. They're saying: "AI as demonstrated by Anthropic's Mythos has significantly increased the likelihood of attackers discovering new vulnerabilities, creating new exploits, and using them in complex automated attacks at scale. While AI also increases the speed of patch development and reduces defects in new software, defenders still face a heavier relative burden due to the inherent limitations of patching. Attackers gain asymmetric benefits."

And that's what I referred to last week when I was talking about the existing installed base of software that hasn't had the opportunity to be screened through AI. It's already deployed. It's in devices and appliances, and many of it has been forgotten. But not by the attackers who want to use it to get in.

So they ask the question: "How is this different from the status quo?" And answered: "In the near term, security organizations will likely be overwhelmed by the need to apply patches and respond to AI-discovered vulnerabilities, exploits, and autonomous attacks.

"What to do now to deal with the current risk spike? Adjust risk calculations and re-orient security program resources for increasing volume of patches, decreasing time to patch, and more persistent and complex attacks. Focus on the basics and harden your environment further. Segmentation, egress filtering, multifactor authentication, and defense-in-depth/breadth all increase the difficulty for attackers."

"What do we believe will happen next? The storm of vulnerability disclosures from Project Glasswing is the first of many large waves of AI-discovered vulnerabilities that may occur in rapid sequence. The capabilities seen in Mythos will quickly become more widely available, dramatically increasing the number and frequency of complex, novel attacks organizations will face."

And finally: "What else should start now to be ready for the next waves? Prioritize robust dependency management to reduce vulnerabilities in third-party and open-source components. Enforce automated security assessments consistently in your development process, including using LLM-powered agents to find vulnerabilities before attackers do. Introduce AI agents to the cyber workforce across the board, enabling defenders to match attackers' speed and begin closing the gap. Re-evaluate your risk tolerance for operational downtime caused by vulnerability remediation, to account for shorter adversary timelines. Update governance for more efficient vendor onboarding and increase headcount to facilitate a faster cycle deployment of new AI-based defenses. As an industry we need to strengthen our coalitions, cooperation, and coordination."

Okay. So I think it should be clear from these Executive Summary bullet points that the cyber security industry's posture on Mythos is that there is less than no time to waste. This is not the time to adopt a wait-and-see posture and to be reactive to events. By the time a reaction is indicated, it will be too late. Despite these clear alarms being rung by

many security professionals who have no profit stake in any of this being true, inertia being what it is, many organizations will nevertheless wait to see if anything really happens.

For what it's worth, I did not wait. Although GRC's border security has always been as strong as I have been able to make it, as I've mentioned before, I did have two deliberately exposed SSH servers listening for connections from any U.S. domestic IP. Foreign IPs have always been hard blocked. I'm referring to them now in the past tense because, after Mythos, they're already shut down. I've used those SSH links to allow me to deal with the rare IP changes in my two Cox Cable connections. An SSH session allows me to update the firewall filters that block all other connections from anywhere other than my two remote work locations. Even though those SSH servers are both using the strongest multifactor identity authentication available, that might not matter if some bypass vulnerability is found.

I don't need those SSH servers as much as I need security. So I'm going to take a wait-and-see approach in the opposite direction. Rather than waiting to see whether a problem is found and then hoping I get the news quickly enough, I'm going to assume that someone using Mythos might discover something unforeseen in the SSH server software I'm using. So I'm going to wait and see about that before I feel safe to poke my head out again. And in fact I may drop SSH completely, with its inherently open ports, altogether and come up with an affirmatively more secure solution. Leo, like you were talking about using Tailscale in order to get into your inside because Tailscale is able to do NAT penetration, in which case you don't have to have any open ports.

Leo: Yeah. That's what I use, and it's great. I love it.

Steve: Yup. So this wonderful call-to-action paper next offers some key takeaways for CISOs. Here's what the paper's authors recommend CISOs to consider: "Use LLM-based vulnerability discovery and remediation capabilities." They said: "Unlike defensive AI technologies, LLM-based vulnerability discovery capabilities are already mature and can be used to your advantage. Start immediately by asking an agent for a security review of any code, and build toward a VulnOps capability. Update your risk metrics. With the shifting landscape, many of your metrics and risk assessments may be outdated and could affect business reporting. Consider how to update these, and communicate the challenge with stakeholders.

"Accelerate your team by the use of coding agents." And you were just talking about this on MacBreak Weekly, how some group at Apple...

Leo: The Siri Group is being sent to learn how to vibe code, almost 200 of them.

Steve: Yeah.

Leo: Because I guess they weren't - they didn't...

Steve: Take it seriously.

Leo: Take it seriously, yes.

Steve: They didn't realize what the benefits were. So these guys are saying to CISOs: "Accelerate your team by the use of coding agents. While defensive AI technologies are lagging behind offensive ones, agents can already accelerate human action across the board, from incident response to GRC. Encourage and require your team to use these agents to accelerate their capabilities. Triage and test patches, red team your environment, automate audit data collection, and accelerate security operations overall.

"Prepare to respond to more incidents. Run tabletop exercises for multiple simultaneous high-severity incidents occurring within the same week, and have playbooks in place for high-level critical incidents." I mean, these guys are literally predicting a storm is coming. "Examine how to automate remediation capabilities to the degree possible. Verify and enable mitigating controls such as segmentation, egress filtering, Zero Trust architectures, phishing-resistant multifactor authentication, and secrets rotation to limit impact when exploitation occurs. The supply chain will be affected.

"Increase focus on the basics. The basics remain valid and can be prioritized for risks that cannot otherwise be mitigated. Segmentation, patching known vulnerabilities, Identity and Access Management, and defense-in-depth and breadth all increase the difficulty for attackers. To lower latent risk, expanding these efforts while there is time is prudent." In other words, do it now before it's too late. They said: "We cannot outwork machine-speed threats. Re-prioritize, automate, and prepare for burnout.

"The cadence and volume of vulnerability disclosures will exceed anything we have experienced before." I mean, they're literally saying, understand, everybody, bad guys, China and Russia and North Korea, they're going to get this capability, and they are going to come at us hard. They wrote: "The cadence and volume of vulnerability disclosures will exceed anything we have experienced before. Consider how you manage current priorities, and request additional headcount and budget for reserve capacity to avoid exhausting available resources, or potentially burning out existing staff. This, in parallel with adoption of coding agents, re-prioritization, putting more automation in place, and helping your team through career uncertainties and upskilling challenges." Yikes.

"Evolve to a Mythos-ready Security Program. Mythos," they wrote, "is likely one of many changes coming to cybersecurity risk. If not already underway, seriously consider incorporating Mythos and its implications into your strategy.

"Build Collective Defense Now. Attackers already operate as syndicates, crowdsourcing, sharing tools, and moving as a collective. Engage now with sector coordinating groups, ISACs, CERTs, and standards bodies to share threat intelligence, coordinate response, and produce sector-specific guidance for this moment. Defenders must do the same and leverage our coordinating groups, especially when considering organizations that fall below the Cyber Poverty Line, as introduced by Wendy Nather."

So just to pause, a little over three years ago, back in 2023, Cisco's CISO, Wendy Nather, articulated a concept she termed "The Cyber Poverty Line." It was the point below which an organization cannot afford to invest in the minimum required security to remain safe on the Internet. So, like, you do need to invest in security.

The bottom of page 17 of the show notes duplicates a breathtaking chart from the very cool and somewhat unnerving website, zerodayclock.com. The chart shows how the vulnerability versus exploit race has radically changed over just the past eight years. At the bottom of page 17, a beautiful chart.

Eight years ago, in 2018, the average TTE - Time To Exploit - was 2.3 years. In other words, just eight years ago, on average there was a 2.3-year gap between the public

disclosure of a security vulnerability in a CVE and its confirmed use in an attack exploit. 2.3 years.

Leo: Wow, we had a lot of time back in the day.

Steve: We did.

Leo: Not anymore.

Steve: Look at this chart, Leo, at the bottom of page 17.

Leo: How many days now do we have in a zero-day?

Steve: Well, watch how this happens. The next year, in 2019, that exploitation gap had dropped from 2.3 to 1.9 years. In 2020, a year later, it was down to 1.3 years-to-exploit. 2021 averaged 10.8 months from CVE publication to exploitation. A year later, 2022, dropped that 10.8 months down to 9.7. The next year, 2023, was down to 4.9 months. 2024, just two years ago, we were down to 56 days. Last year, 23.2 days. And, shockingly, so far this year we are seeing exploits appear an average of 10 hours after their CVE vulnerabilities have been published.

Leo: That's AI; right? I mean, that's got to be AI.

Steve: That is, and we've been talking about this on the podcast, mostly theoretically because it was obvious it was going to happen, bad guys are sitting waiting for new vulnerabilities to be published. And they instantly jump on them.

Leo: Ten hours.

Steve: Ten hours. And so, I mean, there is just no time. As the writers of this paper said, humans cannot outperform machine-driven attacks. It can't, it won't, it doesn't happen. So think about that. Eight years, Leo, gone from 2.3 years to 10 hours. So everybody should check out the zerodayclock.com. It's got this chart and a bunch of others, where these sorts of stats are being maintained. And it is breathtaking.

Okay. So next, I'm going to share just the brief introduction that these cyber security industry expert authors wrote for the paper. But Leo, let's first take our final break.

Leo: Good thinking. Thank you for remembering, Steve. All right. On we go. Oh, Steve, I think you're muted again.

Steve: Did it again. I didn't want you to hear me typing, so...

Leo: No problem.

Steve: Okay. Thank you. Okay. So the brief introduction that these cybersecurity industry expert authors wrote for the paper. They explain, well, I'm going to point our listeners to recommend that they point their bosses, anybody who doesn't understand this, the paper was written for the C-suite guys to understand. And that's why it's got the who's who behind it.

So they wrote: "Many of our assumptions about the capabilities of AI in vulnerability research, exploitation, and autonomous attacks may be outdated. Throughout 2025 and into 2026, we've seen continuous examples of increasing capabilities, both in research and in actual in-the-wild attacks. AI-driven vulnerability discovery and exploitation has been accelerating for over a year.

"Anthropic's Claude Mythos Preview represents a step change in that trajectory, autonomously finding thousands of critical vulnerabilities across every major operating system and browser, generating working exploits without human guidance, and empowering autonomous attack orchestration, all at a speed and scale that outpaces any prior capability.

"The asymmetry this creates is structural. AI lowers the cost and skill floor for discovering and exploiting vulnerabilities faster than organizations can patch them." This is what I was talking about last week when I said, you know, now script kiddies can be expert attackers and exploiters because you just ask AI for some attacks. "The window between discovery and weaponization has collapsed to hours. Attackers gain disproportionate benefit; and current patch cycles, response processes, and risk metrics were not built for this environment.

"While many of these capabilities pre-date this model, Mythos-class capabilities do represent a step-change, and will proliferate." Meaning Anthropic is only first. They're not the last. "The organizations that respond well will be those that build the muscle now - the processes, the tooling, and a culture willing to adopt AI as a core part of how security gets done. The adaptability will help determine who meets the next wave on their own terms.

"This moment requires reprioritizing resources, reviewing risk levels and controls, and leveraging AI where feasible. At the time of this writing, most AI defensive controls and approaches are not yet mature. That said, AI attacker technology may be used for defense purposes, and coding agents will help."

Okay. And to finally place all this into context, I want to share Appendix A of their paper which they titled "Historical Precedent," meaning where we came from. Because this will sort of help everybody to put this in context.

They said: "This all began with the DARPA Cyber Grand Challenge, a landmark competition organized by DARPA in 2016, so a decade ago, that demonstrated the potential of fully automated cybersecurity systems. Teams developed autonomous platforms capable of identifying, exploiting, and patching software vulnerabilities in real time, without human intervention. The challenge highlighted a shift toward machine-speed cyber defense, showing how automation and artificial intelligence could significantly enhance vulnerability management and incident response, while also raising important questions about trust, control, and the future role of human operators in cybersecurity." Meaning humans are going to be obsolete.

"By mid-2025, XBOW, an autonomous offensive security company, topped the HackerOne leaderboard. The DARPA AI Cyber Challenge found 54 vulnerabilities in four hours of compute. Google's Big Sleep discovered real zero-days in open source. Anthropic was used to automate full attack chains from reconnaissance through

exfiltration. And open source tools such as raptor proved autonomous vulnerability research is available to anyone able to use an agent.

"In September 2025, Heather Adkins (CISO for Google) and Gadi Evron (CEO of Knostic) published a warning" - okay, September 2025. They published a warning "that attackers were racing toward a singularity moment, with autonomous vulnerability discovery and exploitation roughly six months away." Wow. Well, that's impressive. Their timing was exactly correct. That was six months ago.

"In February 2026, Anthropic, using Claude Opus 4.6, reported more than 500 high-severity vulnerabilities in open source software. AISLE (A-I-S-L-E) found 12 OpenSSL zero-days, including a CVSS 9.8 vulnerability dating back to 1998. Linux kernel maintainers saw vulnerability reports climb from two to 10 per week, largely hallucinated at first, but that changed rapidly. The volume has held steady, but the reports are now all verified as real bugs.

"The curl project, which originally discontinued its bug bounty program because it was drowning in hallucinated vulnerability reports ('AI slop'), last week echoed the observation from the Linux team, reporting an increasing number of AI-supported high-quality security incidents. Sysdig documented an AI-based attack that reached admin-level access in eight minutes. This week, Gambit released a report on the AI-led compromise of Mexican government infrastructure, originally reported in February."

And actually, I saw that and skipped over reporting that due to show length. But briefly, an attacker used a combination of both ChatGPT and Claude to attack, rapidly penetrate, inventory and exfiltrate a much larger amount of data from the Mexican government than would have ever been possible without the aid of AI automation. They used an AI-automated-based attack. So they end their historical timeline by telling us about the Zero Day Clock, writing: "In March, Sergej Epp and others introduced the Zero Day Clock, visually demonstrating the disappearing time-to-exploit development, demonstrating the drastic fall in time to exploitation to less than a day in 2026." Yeah, 10 hours.

"It's worth noting that the historical collapse in time-to-exploit has not yet produced a proportional increase in the impact of exploitation. Many of the most consequential incidents of recent years involved credential abuse, social engineering, or supply chain compromise rather than zero-days. The Zero Day Clock trend is a leading indicator of where attacker capability is heading, not a direct measure of current damage." So it's predicting what's going to happen shortly.

Okay. So then the AI-driven security research company AISLE (A-I-S-L-E), remember that we talked about them at the time, they found the problems in OpenSSL. They responded, a little disgruntled themselves, understandably, to all of the Mythos buzz. And so it was in February that we reported on them finding those 15 vulnerabilities in OpenSSL, 12 of which entirely composed a major update to OpenSSL. And as we know, this paper briefly mentioned them in passing. They're grumbling somewhat, saying that they were able to reproduce Anthropic's results themselves without the mythical Mythos. They wrote, and I have a link to their report in the show notes.

They said: "We took" - this is AISLE. "We took the specific vulnerabilities Anthropic showcases in their announcement, isolated the relevant code, and ran them through small, cheap, open-weight models. Those models recovered much of the same analysis. Eight out of eight models detected Mythos's flagship FreeBSD exploit, including one with only 3.6 billion active parameters costing \$0.11 per million tokens. A 5.1 billion active token open model recovered the core chain of the 27-year-old OpenBSD bug. And on a basic security reasoning task, small open models outperformed most frontier models from every major lab. The capability rankings reshuffled completely across tasks. There is no stable best model across cybersecurity tasks. The capability frontier is jagged."

In other words, they're just saying hold on here, you know, we've got our own small cheap models that we are able to deploy that do the same thing as Mythos.

And I don't doubt that AISLE did what they claimed, although there's much they didn't say. For example, even with isolated code, confirming an already known problem feels different from making brand new discoveries, although I know in theory there should be no difference. We also don't know how autonomous their system was. That was one of the main points that Anthropic has been making about Mythos is that you just ask it pretty please to attack somebody, and it's able to. And it's only natural for AISLE, a commercial enterprise whose specific and narrow focus is to offer commercial vulnerability discovering services to enterprise, to be somewhat miffed over all the breathless industry and media coverage Mythos has generated.

They should be celebrating their own systems if they're able to meaningfully compete with Mythos's outcome for far less money. As I said, once the dust has settled, it's going to all come down to who can do the most with the fewest resources. So if AISLE's got some bunch of tricks up their sleeve that allows them to offer these services much less expensively, far more economically, then I say that's great. Bravo.

However, everyone who's been paying attention knows that what the cybersecurity industry most needs right now, this instant, without delay, is a swift kick in the pants. This Security Now! podcast informed its listeners of AISLE's AI-driven vulnerability discovery news back in February. It's one of the reasons that Anthropic's claims for Mythos made so much sense to us; right? Because, like, we saw this coming. This made sense. But AISLE did not break through in February. Mythos did.

Even if Mythos were hype, which none of these experts who should know believe it to be, it should be abundantly clear - even looking at AISLE's results with OpenSSL from February - that the next stage of AI-driven rapid vulnerability discovery and exploitation is here now, and that as all of these experts also agreed, we're not ready for it. So I'm all for the hype this industry is able to muster, if it will help to instill some much-needed fear and action from an industry which appears to have become far too comfortable with the status quo.

You know, as I said at the top and a couple times, let's turn this. Let's have another Y2K event that never happens, not because it isn't real, but because it is; and everyone who needed to, understood and then took action to prevent the apocalypse from ever happening when everything rolled over to the year 2000.

Anyway, as I said, I've created two GRC shortcut links to this very significant paper to make it even easier for our listeners to get to it. You can either go to grc.sc/mythos, which everyone should be able to remember, and that'll just bounce you over to the PDF; or this week's episode number, grc.sc/1075. That'll do the same thing.

Leo: Nice.

Steve: And I think it is very clear that, I mean, I get it that when we're talking about increasing headcount and reshuffling priorities and all this, I mean, these are expensive things to ask for a problem that hasn't yet manifested. The problem is, by the time it does, it could be too late. It's like, wait.

Leo: And a lot more expensive.

Steve: Yes. It's like waiting to see what, you know, like if the elevators stop running on January 1st of the year 2000. Like I'd rather not get stuck in an elevator, thank you very much.

Leo: Okay. Now, here's the question. Models are going to continue to get better. I think there's no doubt about that. There was some question a year ago maybe that maybe we'd hit a plateau, and models weren't getting better fast. I think we all see that that's not the case.

Steve: And we're learning how to use this new thing.

Leo: Yeah.

Steve: Like the notion of parallel agents and...

Leo: Yeah, we're getting better at it, yeah.

Steve: And the collection of different capabilities that are brought in.

Leo: Right.

Steve: So, yeah, so we're learning basically how to ask.

Leo: So presumably the 271 bugs we found in Firefox this time, next time we might find more. We might find more again as models get better. We might find more again. Is there a point where software just becomes perfect, and there are no more bugs?

Steve: I think there is.

Leo: Yeah.

Steve: I think there is. Software is math. Math is 100% predictable.

Leo: Right.

Steve: You know, there's no random number generator like there is in AI. There's no random number generator in our software. You know, it is deterministic. And I hold that something that doesn't get lost in the details, basically humans have created software that is too complex for them to hold in their head.

Leo: Yes. Right.

Steve: That's what's happened is we don't understand our own creation. But AI can be scaled to be able to understand - and I use "understand" in air quotes. I know it's not conscious. It's not actually understanding.

Leo: Right, right.

Steve: But to weave through all the combinatorial ins and outs. And who was it, there was another person who I just saw, I think it might have been in an email feedback, somebody else - oh, it was one of our listeners. I'll share it next week. It was one of our listeners who has been maintaining a package that is exposed to the Internet, and it involves SQL, and he was curious, so he aimed Claude Code at the software that he wrote, and it found a vulnerability that astonished him. And he said it wasn't super critical because it only, blah blah blah blah, I know there were lots of ways that you had to have it. But he was amazed by what it found in his own code. And so he stood there thinking, my god, is this true? And then he actually, oh, he didn't want to upset it by asking it for an exploit, so he wrote the exploit himself.

Leo: Actually, Glenn Fleishman on Sunday was reporting something very similar. He's had web-facing or Internet-facing, a tool running for I think he said like 20 years, a long time, just a little thing that he runs, some sort of book search or something. And he said he fired Claude Code at it. It found bugs that had been running for 20 years no one's seen, he hadn't seen, that he was able to fix. I mean, I think that - I think you nailed it, which is that it's gotten impossible for us as human beings to make perfect software. But this is a machine. It is tireless. It doesn't make the same kinds of mistakes.

Steve: And use chess, fall back to chess again. You know, super chess grand masters are able to look at a board and see things in it I can't even begin to describe.

Leo: Right.

Steve: They were able to hold their own for a long time. No more.

Leo: No. Not even close.

Steve: That's gone.

Leo: It's no longer even close.

Steve: And that suggests that now we have computers that are able to look at this same thing. Again, there's no - no one's rolling dice in chess. It is a deterministic board game. And they can take us down now. We are, you know, we're - okay. You're a little younger than I am. You're in your late sixties. I'm now 71.

Leo: No, I'm almost 70. I'll be 70 in November. I'm not so far behind you, Steve.

Steve: Okay. So we're still going to be here in another 10 years.

Leo: I hope so. God willing.

Steve: And the world is going to be different.

Leo: I know, and I love that. I thought I was going to miss the apocalypse.

Steve: It's going to happen so fast. What's so cool is there's so much money behind software development that there will be a huge push to make this happen.

Leo: Oh, yeah. And the other thing that's encouraging is these tools are getting more efficient, which means they're taking less hardware to do more. Which means not only will they improve, but they will be more accessible.

Steve: Yes.

Leo: They will be less expensive.

Steve: Yes. And I'm convinced cloud crap is going to go away.

Leo: I hope so, yeah.

Steve: We're going to have local running models because we'll have little, you know, AI boxes in our homes that we talk to, and they're able to do what we want.

Leo: Yup. Yup. That's what I'm working on right now. That's exactly it. I'm trying to make - because I'm so dissatisfied with Alexa and Siri and all these other assistants, I'm trying to make an assistant that works the same way but is local and knows me.

Steve: Nice.

Leo: And has memory and all of that stuff. And I'm getting closer than I ever thought I would. And I think, you know, by the time I'm 80...

Steve: And here's the key, Leo. You are having fun.

Leo: Oh, and it's the best game ever.

Steve: You are having fun.

Leo: It's like coding. It's similar, I mean, I still love coding. But coding is more like hand-building furniture.

Steve: Leo, it is modern coding.

Leo: It is.

Steve: This is what coding is going to become.

Leo: Yeah.

Steve: People are going to be removed from the code-generating loop. And we will be directing AI to write our code.

Leo: Yes. Yes.

Steve: And this is me, who is still coding in assembly language. I'm saying it's over, folks. People are going to be taken out.

Leo: Wow. And, you know what, maybe that's the right thing because...

Steve: We had our time, yeah.

Leo: And computers are going to do a better job of this. This is their native tongue. You know? Steve Gibson does such a good job. I'm so glad we have you to rely on. And let's hope we get to keep doing this for many, many more years.

Copyright (c) 2014 by Steve Gibson and Leo Laporte. SOME RIGHTS RESERVED

This work is licensed for the good of the Internet Community under the Creative Commons License v2.5. See the following Web page for details:
<http://creativecommons.org/licenses/by-nc-sa/2.5/>