

Security Now! #1075 - 04-21-26

Yes. Exactly.

This week on Security Now!

- A disgruntled developer discloses multiple Windows 0-days.
- Microsoft purchases its own bugs in massive campaign.
- VeraCrypt & Wireshark suddenly lost their dev accounts.
- A serious problem with re-captured 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.

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



Security News

Several Microsoft 0-Day Troubles

Last Thursday the 16th, BleepingComputer's headline was "New Microsoft Defender "RedSun" zero-day PoC grants SYSTEM privileges". That piece told the story of the disgruntled developer who had been publishing fully working proof-of-concept exploit code for his discoveries of privilege elevation vulnerabilities in Windows workstation and server. The following day, Friday the 17th, BleepingComputer followed that reporting with another piece titled: "Recently leaked Windows zero-days now exploited in attacks". They wrote:

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-Eclipse" 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.

Two of the vulnerabilities (dubbed BlueHammer and RedSun) are Microsoft Defender local privilege escalation (LPE) 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, since they had no official patches or updates to address them.

On Thursday, Huntress Labs security researchers reported seeing all three zero-day exploits deployed in the wild, 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 "hands-on-keyboard threat actor activity."

While Microsoft is tracking the BlueHammer vulnerability as CVE-2026-33825 and has patched it in the April 2026 security updates, the other two flaws remain unaddressed. As Bleeping-Computer previously reported, 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, even after applying the April Patch Tuesday patches.

The disgruntled researcher explained: "When Windows Defender realizes that a malicious file has a cloud tag, for whatever stupid and hilarious reason, the antivirus that's supposed to protect decides that it would be a good idea to just rewrite the file it found to its original location. The PoC abuses this behaviour to overwrite system files and gain administrative privilege."

When BleepingComputer contacted Microsoft earlier this week for more information on the disclosure issue reported by the anonymous researcher, a Microsoft spokesperson told BleepingComputer: "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."

Two days before that, on Wednesday, this person going by the moniker "Chaotic Eclipse" posted a diatribe over on blogspot. I think it's worth sharing since it gives some impression of who's

doing the grumbling. Dated Wednesday, April 15th, the blogspot post was titled "Public disclosure, a response for CVE-2026-33825 patch". It reads...

Here is the code, enjoy <https://github.com/Nightmare-Eclipse/RedSun> Now to address what some media articles wrote, first of all, I want to talk about MSRC official response regarding BlueHammer

"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."

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 are also aware that this one will be disclosed but again, they are ignorant.

Normally, 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 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 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.

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...

I will personally make sure that it gets funnier every single time Microsoft releases a patch.

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 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."* My presumption, without knowing anything specific, is that Microsoft almost certainly treated this researcher the same way they treat everyone else. But he believed that he deserved special treatment. 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 that he was counting on receiving a bug bounty payout that never came. So now Microsoft is to blame for his situation.

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 ~100% reliably to go from unprivileged user to SYSTEM against Windows 11 and Windows Server 2019+ 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.

Cldapi.dll provides support for the Windows Cloud API. In the next quote from Will, he refers to EICAR. E. I. C. A. R. is the abbreviation for the European Institute for Computer Antivirus Research. The file they produced, known as EICAR, is a popular pseudo-malware test file that can be used to deliberately freak out any good anti-virus tool without actually containing or doing anything malicious. In a follow-in Mastodon posting Will writes:

This exploit uses the "Cloud Files API", writes EICAR to a file using it, 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) to 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. Game over.

Our primary takeaway is that all fully-patched — as of last week's mega-patch Tuesday — Windows desktop and server are currently vulnerable to this exploit. 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 the odd file rewrite. That allows attacker-provided code to be run with full SYSTEM privilege. While it's not the end of the world, Huntress Labs is observing it under active use. So it would be nice if Microsoft were to issue a fix for this before April's patch Tuesday which is three and a half weeks away.

Microsoft has been buying their own bugs

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:

*Microsoft has awarded \$2.3 million to security researchers after receiving nearly 700 submissions during this year's Zero Day Quest hacking contest. Tom Gallagher, Vice President of Engineering at Microsoft Security Response Center (MSRC), 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**.*

Oh great ... and we've all been using that software. 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 chains, and cross-tenant access."

Last August, Microsoft announced that it would increase the prize pool at this year's Zero Day Quest hacking contest to \$5 million in bounty awards, which the company described as the "largest hacking event in history." The 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. After the hacking competition concluded, Microsoft announced it had paid \$1.6 million in rewards after receiving more than 600 vulnerability submissions.

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 requiring "an overhaul."

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. Learnings from the Zero Day Quest will be shared across Microsoft to help improve Cloud and AI security in alignment with SFI's core principles: securing by default, by design, and in operations."

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. In December, it also announced that security researchers would be paid for finding critical vulnerabilities in any of Microsoft's online services, even if a third party wrote the vulnerable code.

I think it's great that Microsoft's software will now, or soon be, more than 80 high-impact cloud and AI security vulnerabilities poorer. That's good for everyone. It does seem somewhat 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 by money. Uhhh... those were all bugs in your software, right? None of them should have been present to be found in the first place. But hey, Microsoft has way more cash on hand than it knows what to do with. So dangling increasing quantities of that cold hard cash in front of security researchers who will then be motivated to go bug hunting is definitely money well spent. Let's have more of that!

Microsoft quickly reinstates closed driver developer accounts

One of the recent bits of news that was bumped so that we'd have time to thoroughly examine Anthropic's Mythos was that Microsoft had, without apparent cause or reason, suddenly dropped a number of driver developer accounts. Products such as VeraCrypt and WireGuard incorporate kernel driver components in order to obtain the deep OS-level access they need. So this was a huge concern for many users of these products and I heard from many of our listeners who picked up on the news.

Well, as it turns out, it was probably just as well that I waited since last week only the news of the event and not the entire story was available. Today it is ... and the reason for Microsoft's suspension of these accounts, which turned out not to be a mistake but was entirely deliberate, should hit home with many of our listeners since it's an issue I've been talking about a lot recently.

Once again, BleepingComputer was on top of it under their heading "*Microsoft rolls out fast-track to reinstate Windows hardware dev accounts*", writing:

Microsoft has rolled out a fast-track process 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, 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 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 A. Donenfeld and others, who described being locked out while facing lengthy or unclear appeals processes.

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 partners about the requirement since October 2025. Microsoft requires identity verification for the Windows Hardware Program because it allows developers to sign and distribute kernel-level drivers, which run with high privileges and have been abused by threat actors in past attacks.

However, many developers 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. 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."

Under the new process, developers are told 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. The company also addressed issues developers reported with the support workflow. It advised partners to ensure they are signed in with the correct account when submitting tickets and to continue prompting Copilot to create a ticket if automated assistance fails. 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, so affected developers are advised to act quickly.

I would tend to believe the developers over Microsoft regarding the complete lack of attempts to inform them. As I noted earlier, Microsoft is no longer an entity that is able to care. It's just too big, and "caring" is a distraction. So someone, somewhere, decided that the best way to get developer attention would be to simply suspend all non-compliant accounts for non-compliance. This has the advantage 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 act together. As we know, this may involve getting an affiliated attorney or CPA to sign some attestation papers. But it's also true that running code in the Windows kernel is a privilege that no one wants bad guys to have. While Microsoft could have been way more gentle about it, this did get that job done.

A company with badly designed software goes out of business

This next piece of news beautifully exemplifies a problem we've seen before that's largely a consequence of the aging Internet and aspects of its design that were never very well thought through since its early designers could have never, and never did, foresee what their creation would become.

BleepingComputer's headline for this reporting was "*Signed software abused to deploy antivirus-killing scripts.*" And while that's factually true, it's more of the consequence of the problem than the problem itself. So let's start with what BleepingComputer's reported:

A digitally signed adware tool has deployed payloads running with SYSTEM privileges that disabled antivirus protections on tens of thousands of endpoints, some in the educational, utilities, government, and healthcare sectors. In a single day, researchers observed more than 23,500 infected hosts in 124 countries trying to connect to the operator's infrastructure, with hundreds of infected endpoints present in high-value networks.

Security researchers at managed security company Huntress discovered the campaign on March 22, when signed executables viewed as potentially unwanted programs (PUPs) triggered alerts in multiple managed environments. 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.

Huntress researchers say that the software was signed by a company called Dragon Boss Solutions LLC, involved in "search monetization research" activity and promoting various tools (e.g., Chromstera Browser, Chromnius, WorldWideWeb, Web Genius, Artificius Browser) labeled as browsers but detected as PUPs by multiple security solutions.

Beyond annoying users with ads and redirects, Huntress researchers say the browsers from Dragon Boss Solutions also feature an advanced update mechanism that deploys an antivirus killer. 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 and with no user interaction. It also installed the payloads with elevated privileges (SYSTEM), prevented users from disabling automatic updates, and checked frequently for new updates.

Okay. None of those things seem deliberately malicious. Having been harassed by false-positive A/V detections I at least understand the motivation behind creating exceptions for one's code. As we know, that's not the approach I take, but I get it. So mostly this seems like software written entirely with the convenience of its publisher rather than its user in mind. That's bad software, but that's also life. The reporting continues:

According to the researchers, the update process retrieves an MSI payload (Setup.msi) disguised as a GIF image, which is currently flagged as malicious on VirusTotal by only five security vendors.

Okay, that 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? They continue:

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.

The ClockRemoval.ps1 script also executes a routine when the system boots, at logon, and every 30 minutes, to make sure that AV products are no longer present on the system by stopping services, killing processes, deleting installation directories and registry entries, silently running vendors' uninstallers, and forcefully deleting files when uninstallers fail. 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 to 0.0.0.0).

Okay. 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 PUP — a potentially unwanted program. So these cretins have upped the ante by making their adware offerings even more obnoxious. Now 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 to sinkhole the connection from all infected hosts.

As such, 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, the researchers identified 324 infected hosts in high-value networks:

- *221 academic institutions in North America, Europe, and Asia*
- *41 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*
- *3 healthcare organizations (hospital systems and healthcare providers)*
- *networks of multiple Fortune 500 companies*

BleepingComputer 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 onto 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, and through an already established infrastructure.

Huntress recommends that system administrators look for WMI event subscriptions containing "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 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 here's the problem that the Internet's designers never considered: What happens when the progenitors of ill-begotten and very badly designed software, which continually reaches out to the Internet for updates, eventually (and probably inevitably) goes out of business? Their horrible software remains installed and alive and querying for updates. Then their various domains expire. Oops.

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 motherload — 221 academic institutions, 41 Operational Technology networks in energy, transport and critical infrastructure providers, 35 municipal governments, state agencies, and public utilities, 24 primary and secondary educational institutions, 3 healthcare organizations (hospital systems and healthcare providers) and the networks of multiple Fortune

500 companies. This abandoned software would literally have a ready-to-go backdoor into the networks of all of those 324 high-value entities.

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 for abandoned domains.

There is no accountability for the actions of the software while it's in use, nor after its parents have abandoned it. The entire "rent-a-domain-name" system the Internet has always used – which should be reminiscent of the AWS abandoned bucket problem – recycles abandoned domains. This behavior is assumed and encouraged, but it leaves us with serious potential for security problems.

Listener Feedback

A listener shared some musings over strategies for securing our open source repositories. It provided a perfect setup for an aspect of the future. Gene Hastings wrote:

A colleague and I often meet to talk about devops and related issues (you know, system and personal health). He is more dev, I am more ops. Both often cranky. 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.

I was already aware of version pinning, and then 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 then need to store a hash of the package and compare it on retrieval. 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 LetsEncrypt: A service for smaller developers that can do an automatic audit at a tolerable expense.

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.

The OWNERS of the agents are trustworthy!

Then there could be a future for us and the Internet.

*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 am 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. /Gene Hastings*

Gene is suggesting that a future solution might be a system in the form of LetsEncrypt where individual developers would need to have an AI-based agent audit their code for problems, unsuspected and 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 trust.

I think there's a more direct, cleaner and straightforward means of accomplishing the same thing: We simply 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. Code cannot leave the repository without first being checked by the AI agent.

And the process might not be autonomous. The repository's AI might have some questions for a package's author(s) that would need to be answered and negotiated before a new or updated package could be made widely available. And since the use of AI will certainly come at a non-zero cost for the foreseeable future, I'd imagine that there would be some form of rate-limiting on submissions for review and publishing. 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 submission limit imposed.

But I suspect that solves the problem. The vast majority of a repository's code is statically available. So an AI would only need to give it "the once over" once. And from then on those who pull it could rely upon its security more than ever before. And AI could also track submitter reputation. Although we know that reputation is not a perfect proxy for security, if the identities of valid long-time submitters cannot be spoofed, that would be another useful signal.

What I know is this: The year is currently 2026 when AI costs far more to run than it's able to generate. But the economics of AI will be radically different in the future, just as the economics of mass storage and computation have been utterly transformed over the past 50 years. Today, we're all walking around with globally connected pocket computers that would have boggled the minds of our grandparents. 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 grand kids as the Internet is by today's kids.

GP

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 haven't 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.

*My young one enjoyed the movie so much that they wanted to read the book, so we signed up to borrow it from the library. However, we were number 110 in the queue, so we opted for my old copy on Audible instead. Listening to the story again didn't 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 didn't ruin the book, but actually improved upon it. Good job to the production team on this one!
All the best, GP*

Leo?

Yes. Exactly.

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 of the data, I realized that leaving the answer to the question that title implied up in the air would be wrong, because there was no way Mythos was only marketing. I acknowledged that it was certainly also marketing, but it was also far far more than that.

I titled today's podcast "Yes. Exactly." because last Thursday, two days after our "*What Mythos Means*" that podcast was delivered, an incredibly significant group of industry veterans, who pretty much comprise a who's who of the cyber security industry, weighed-in with a formal emergency wakeup call for the entire cyber security 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: <https://labs.cloudsecurityalliance.org/mythos-ciso/>
The "AI Vulnerability Storm": Building a "Mythos-ready" Security Program

The paper enumerates its sixteen (16) primary contributing authors. Because this is important for appreciating the weight of the paper's stated concerns, I want to share them. They are:

1. Jen Easterly, CEO of the RSA Conference and Former Director of CISA
2. Bruce Schneier, renowned cryptographer, currently Chief of Security Architecture at Inrupt and Fellow and lecturer at the Harvard Kennedy School
3. Chris Inglis, The White House's former National Cyber Director
4. Phil Venables, Ballistic Ventures, former CISO, Google Cloud
5. Heather Adkins, current CISO of Google
6. Rob Joyce, the NSA's former Cybersecurity Director
7. Sounil Yu, the CTO of Knostic and former Chief Security Scientist for Bank of America
8. Katie Moussouris, the Founder and CEO of Luta Security
9. John N. Stewart, Talons Ventures and former CSTO for Cisco Systems
10. James Lyne, CEO of the SANS Security Institute
11. Dave Lewis, Global Advisory CISO for 1Password
12. Maxim Kovalsky, Managing Director of AI Security CoE for Consortium Networks
13. Jim Reavis & John Yeoh, CEO & CSO of the Cloud Security Alliance
14. Joshua Saxe, CTO and Co-founder at Security Superintelligence Labs, former AI and Llama Security Lead, Meta
15. Ramy Houssaini, CCSO, Cloudflare

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 won't read them since they are too numerous. But I've reproduced that page from the report in the show notes:

Reviewers

Many CISOs, and some other practitioners, assisted in reviewing and editing this document. These are the ones who would share their names publicly, in alphabetical order (by last name):

Mark Aklian, Founder & CISO, Silver Oak Cyber
Barry Anderson, Information Security Architecture Strategy and Engineering Manager, HESTA
David Aronchick, Co-Founder, Expanso / Kubeflow
Jake Bernardes, CISO, Anecdotes
Alan Berry, CISO, Centene Corporation
Bryson Bort, CEO, SCYTHE
Jeff Bryner, CISO, Independent
Michael Calderin, CISO
Daniele Catteddu, Chief Technology Officer, Cloud Security Alliance
Viswanath Chirravuri, Global Product Security Director, Thales
Mea Cliff, CISO, Cengage
Scott Clinton, Co-chair, OWASP Gen-AI Security Project
David B. Cross, CISO, Atlassian
Chris Cochran, Field CISr,O & VP AI Security, SANS Institute
Michael Colao, former Corporate CISO, AXA, Director, Island Cyber
Daniel Cuthbert, Associate Fellow, Cyber and Tech, RUSI
Julie Davila, VP Product Security, GitLab
Michael Douglas, Managing Partner InfoSec Innovations, and SANS Instructor, SANS Institute
Yoni Efrati, former CISO, Bank HaPoalim
Eliya Elon, EIR, Notable Capital
Sergej Epp, CISO, Sysdig
Chris Farris, Cloud Security Nerd, Securosis
Alex Foley, Data Security TISO, Wells Fargo
George Gerchow, CSO, Bedrock Data
Dan Glass, CISO, Delek US Holdings
Barry Greene, Co-Founder, Qubit Cyber
The grugq, Independent Security Researcher, Independent
Erik Hart, Global CISO, Cushman & Wakefield
Gary Hayslip, CISO in Residence, Halcyon, former CISO, SoftBank
Dustin Heywood, Senior Technical Staff Member, IBM
Ariel Herbert-Voss, CEO, RunSybil
Heather Hinton, CISO, Sitecore, Lecturer, Harvard Extension School
Dave Hoelzer, independent
Matt Holland, Director of Cyber and AI Security, Kainos
Igor Ignatov, Head of Security & Compliance, Cognichip Inc.
Waylon Janowiak, Head of Information Security & IT, Faire
Mike Johnson, CISO, Rivian
Avner Langut
Paul Lanzi, Principal Consultant, IDenovate
Rock Lambros, Director of AI Standards and Governance, Zenity, Founder, RockCyber
Josh Lemos, CISO, lululemon
Peter Liebert V, CISO, Salesloft
Ariel Litvin, former CISO, First Quality Enterprises
Bob Lord, former CISO, Yahoo, CSO, DNC
Myke Lyons, CISO, Cribl
Ciaran Martin, Head of Cyber Leaders Network, SANS Institute, Founder and former CEO, UK NCSC

Michael Machado, CISO & CDO, Hyland
Donald McFarlane, Principal Technical Advisor, Microsoft
Gal Malach, CTO and Co-Founder, Terra Security
Tomas Maldonado, CISO, NFL
Greg McCord, Founder and CISO, McCord Keystone Advisory
Ross McKerchar, CISO, Sophos
Greg Notch, CSO, Expel
Charles Nwatu, former Head of Security, Netflix
Helen Oakley, OWASP GenAI Security Project
Mark Orsi, CEO, Global Resilience Federation
Teju Oyewole, CISO, Brave Technology
David Quisenberry, CTO & CISO, Ferguson Wellman Capital Management
Rupa Parameswaran, VP, Security and IT, Ex-Handshake, Amplitude
Gavin Reid, CISO, Human Security
Gerardo Richarte, CISO, Satellogic
Joshua Scott, VP of Security and CISO, Hydrolix
Mark Seiden, Volunteer and Security Advisor, Internet Archive
James Shank, Director of Threat Operations, Expel
Samir Sherif, Global Field CISO, Fastly
Conor Sherman, CISO in Residence, Sysdig
Ed Skoudis, President, SANS Technology Institute
John Sotiropoulos, OWASP GenAI Agentic Security Initiative Co-Lead, Co-Founder, Deep Cyber
Sean Todd, CISO, trycoral.ai
Anna Sarnek, Senior Fellow, McCrary Institute for Cyber & Critical Infrastructure Security
Holger Spohn, CISO, Candriam
Rob van der Veer, OWASP AI Exchange
Mikael Vinding, CISO, AP Technology
Yabing Wang, CISO & CIO, Justworks
Mike Wilkes, Adjunct Professor, New York University
Jeff Williams, CISO, Sigma360
Steve Wilson, Chief AI Officer, Exabeam, Co-Chair OWASP GenAI Security Project
Jason Woloz, CISO, TransUnion

All the listed authors and reviewers represent only themselves, and not their employer(s).

Join the Cloud Security Alliance CISO community, email us at cisos@cloudsecurityalliance.org.

DISCLAIMER: These materials are provided for convenience only and may not be relied upon for any purpose. The contents of this document are not to be construed as legal, technology, or business advice; please consult your own attorney, technology, or business advisor for any such legal, technology, and business advice.

This document is released under the **Attribution-NonCommercial 4.0 International (CC BY-NC 4.0)** license.

Okay. So we've clearly established the provenance of this document. 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, which is, of course, why I immodestly titled today's podcast "*Yes. Exactly.*" This amazing group of experts even use many of the same phrases I used. Given the impossible-to-exaggerate significance of Mythos and the successor systems that are sure to follow, I think it's crucial for the listeners of this podcast to appreciate that it's not just me with a lone opinion. After bringing myself up to speed with the facts, it turns out I'm in the best company this industry has to offer.

The authors of the Executive Summary set it up as a sort of topical Q&A. They wrote:

What happened?

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.

How is this different from the status quo?

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.

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 processes, 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.

I think it should be clear from these Executive Summary 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 was 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 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 multi-factor 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 with its inherently open ports altogether and come up with an affirmatively more secure solution.

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.

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 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.

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. 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 MFA, 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 vulns, Identity and Access Management, and defense-in-depth/breadth all increase the difficulty for attackers. To lower latent risk, expanding these efforts while there is time is prudent.

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. 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.

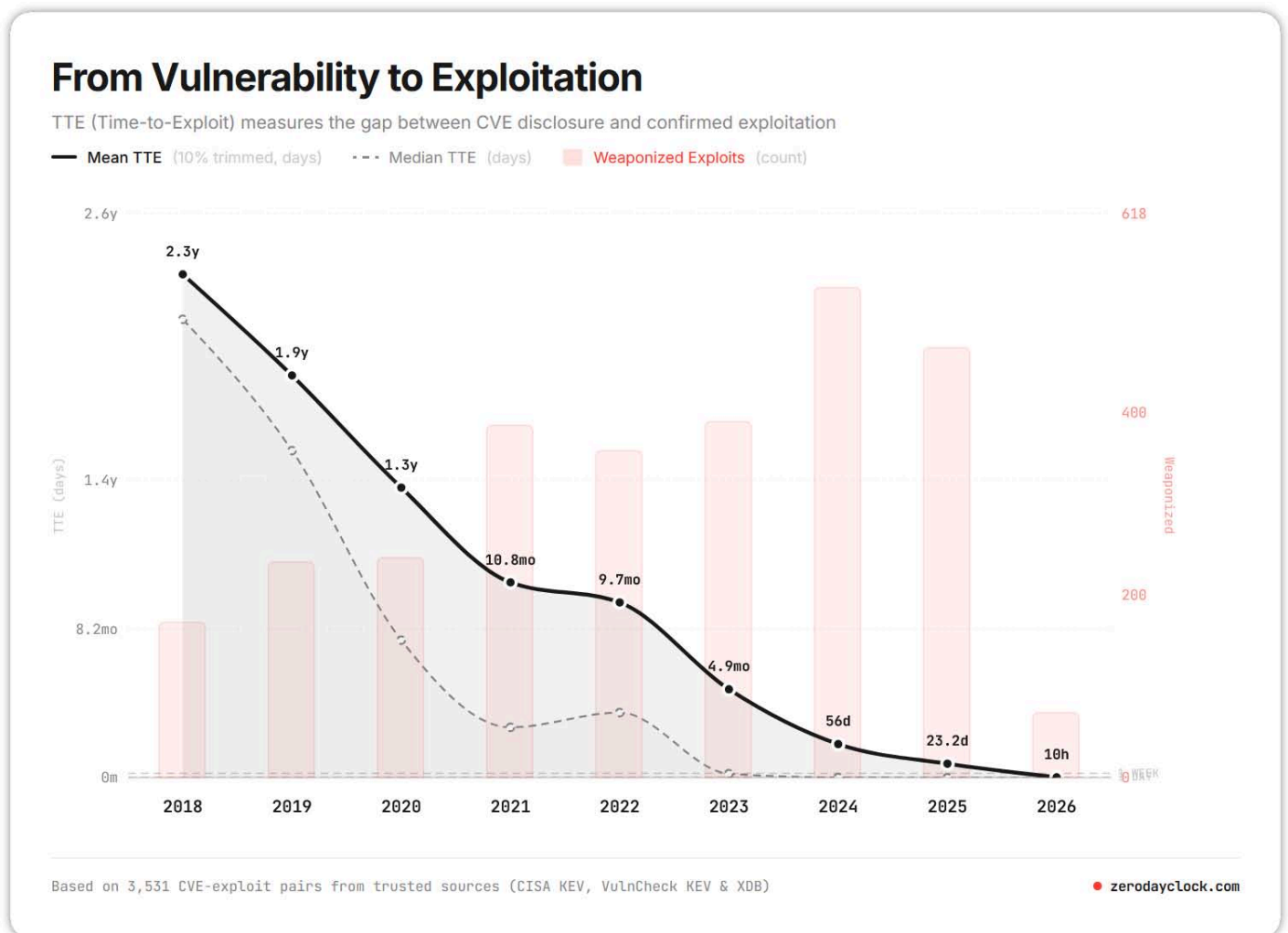
Evolve to a Mythos-ready Security Program.

Mythos 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.

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.



The bottom of page 17 of the show notes duplicates a breathtaking chart from the very cool and somewhat unnerving website: <https://zerodayclock.com/> The chart shows how the vulnerability vs exploit race has radically changed over just the past eight years.

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 average eight years ago. The next year, 2019, that exploitation gap had dropped from 2.3 to 1.9 years. In 2020 it was 1.3 years-to-exploit. 2021 averaged 10.8 months from CVE publication to exploitation. 2022 dropped that 10.8 months down to 9.7. 2023 was 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. So, from 2.3 years to 10 hours in eight years. You may wish to check out the <https://zerodayclock.com/> site for additional sobering truths.

Next, I'm going to share just the brief introduction that these cyber security industry expert authors wrote for this paper. They explain:

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. 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 new model, Mythos-class capabilities do represent a step-change, and will proliferate. 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. That 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.

And, finally, to place all of this into context, I want to share Appendix A of their paper which they title "Historical Precedent", writing:

This all began with the DARPA Cyber Grand Challenge, a landmark competition organized by DARPA in 2016 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.

By mid-2025, XBOW, an autonomous offensive security company, topped the HackerOne leaderboard. The DARPA AI Cyber Challenge (AIXCC) 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, Google) and Gadi Evron (CEO, Knostic) published a warning that attackers were racing toward a singularity moment, with autonomous vulnerability discovery and exploitation roughly six months away.

Wow. That's impressive. Their timing was exactly correct.

In February 2026 Anthropic, using Claude Opus 4.6, reported more than 500 high-severity vulnerabilities in open source software. AISLE found 12 OpenSSL zero-days, including a CVSS 9.8 vulnerability dating to 1998.

Linux kernel maintainers saw vulnerability reports climb from 2 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 reports.

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.

I 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. They end their historical timeline by telling us about the Zero Day Clock:

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. It is 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-day exploitation. The Zero Day Clock trend is a leading indicator of where attacker capability is heading, not a direct measure of current damage.

The AI-driven security research company AISLE responded to all of the Mythos buzz. Those are the guys we reported on in February who found the 15 vulnerabilities in OpenSSL, 12 of which drove a single OpenSSL update. This paper also just mentioned them. Anyway, they're grumbling somewhat that they were able to reproduce Anthropic's results without the mythical Mythos. They wrote: <https://aisle.com/blog/ai-cybersecurity-after-mythos-the-jagged-frontier>

We took the specific vulnerabilities Anthropic showcases in their announcement, isolated the relevant code, and ran them through small, cheap, open-weights 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 5.1B-active 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.

I don't doubt that AISLE did what they claimed, though there's much they didn't say. For example, even with isolated code, confirming any already known problem feels different from making brand new discoveries, which we know these guys have also done. We don't know how autonomous their system was. That was one of the main points that Anthropic has been making about Mythos. And it's only natural for AISLE, a commercial enterprise whose specific and narrow focus is to offer commercial vulnerability discovering services to enterprises, 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' outcomes for far less money. Once the dust has settled it's going to come down to who can do the most with the fewest resources.

However, everyone who's been paying attention knows that what the cyber-security industry most needs right now — this instant and without delay — is a swift kick in the pants. This Security Now! podcast informed its listeners of AISLE's AI-driven vulnerability discovery news in February. It's one of the reasons that Anthropic's claims for Mythos made sense. But AISLE didn't break through — Mythos has.

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 for all 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.

Let's turn this into another Y2K event that never happened, not because it wasn't real, but because it was, which everyone who needed to understand and then took action to prevent the apocalypse.

I have created two GRC shortcut links to this significant paper to make it even easier for our listeners to get it. You can either go to: <https://grc.sc/mythos> which will bounce you over to the PDF, or use this week's episode number: <https://grc.sc/1075> which will do the same thing.

