

Security Now! #1067 - 03-03-26

KongTuke's CrashFix

This week on Security Now!

- The lowdown on last week's "no turn" picture of the week.
- Is an AI-driven hacking campaign a big deal now.
- Clause used in multiple Mexican government attacks.
- Apple continues to be confronted with age restrictions.
- COPPA needs an exception to allow age collection.
- Meta swamps law enforcement with AI-slop CSAM reports.
- Roskomnadzor has been busy blocking VPNs. Guess how many?
- The UK tries to report their self-scanning success.
- Remember that hacker who extorted the psychotherapy patients?
- Scattered Lapsus\$ Hunters is actively recruiting women.
- Cisco lands another breathtakingly rare 10.0 CVSS.
- VulnCheck's report on 2025 vulnerabilities and exploits.
- Steve discovers a fabulous \$72 Hardware Security Module.
- A listener shares an interesting AI service discovery.
- The very potent "ClickFix" exploit evolves.

Dad: "This is THE LAST TIME I'm going to tell you to turn down the volume of what you call music!"



Following up on last week's wacky picture of the week:

I wanted to thank many listeners who were made curious by last week's picture of the week. That was the street which was the stem of a "T" intersection where the signage encountered as a driver was driving toward the "T" intersection indicated that neither turning left or right would be legal. Thus the caption I gave that photo of "But Officer..." Thanks to listener research, many of whom used AI to learn about the situation, we now know that the photo was not synthetic. It was bizarre but authentic. And after the photo went viral, it became a significant embarrassment to the local government that was responsible for its emplacement.

The location was a town called Simcoe in Norfolk County, Canada and a news report that [one of our intrepid listeners found and shared](#) explained that *"Drivers please note that signs were installed this week which restrict right and left hand turns at the intersection of Crescent Boulevard and Queensway in Simcoe. The intent of the new signs was to make Crescent Boulevard a dead-end street. The signs have been removed."* That's amazing. In other words, the signage was technically correct. The street we saw in the foreground which had the nutty "cannot turn left and cannot turn right" signage was intended to end at the "T" intersection so that no one would have any choice other than to perform a "U-turn" at that location.

My favorite quip about last week's photo was provided to us by listener Joseph Rork who noted: *"Despite the presence of the Tim Horton's in the background, we know this cannot be Canada. Otherwise, there'd be a line of cars sitting at the stop sign..."*

Many thanks to all of our listeners among the more than 20,000 who receive the weekly mailing, whose imaginations were captured and took the time to research and/or comment. And a big thanks to whomever sent it to me in the first place! You know who you are.

Security News

An AI-driven hacking campaign

The headline in the news last week was: *"AI-driven hacking campaign breaches 600+ Fortinet devices"*. I'll first share the news report, then I have a few things to say about it. The reporting says:

A Russian-speaking financially motivated threat actor used commercial AI toolkits to hack more than 600 Fortinet firewalls. The campaign began at the start of the year, around January 11, according to the AWS security team. The attacker did not exploit 0-days or older vulnerabilities. Instead, they targeted FortiGate devices that had their management ports exposed online, used weak passwords, and didn't have MFA enabled.

FortiGate devices with publicly exposed management ports, weak passwords and no other authentication required. So no flaws were used, just very poor configuration hygiene. The story continues:

Once inside, the attacker employed a collection of scripts that AWS says were written by AI tools. While AWS did not name products, researchers from Cyber & Ramen and Ctrl-Alt-Int3l identified them as Claude and Deepseek.

DeepSeek was used to create scripts to perform reconnaissance and extract configurations from the hacked devices, while Claude was used to generate scripts for vulnerability assessments and to run offensive tools against the networks.

Since this is the intersection of AI and infosec, the report generated a tornado of feedback and opinions on social media. The general consensus was that the threat actor wasn't particularly sophisticated, which AWS also believes. AWS CISO CJ Moses said the attacker was more interested in scale rather than value. Every time they encountered errors caused by hardened or non-standard internal networks, the attacker moved on to softer targets.

Once they did move laterally from the Fortinet device, the attacker compromised the victim's Active Directory environment, extracted database credentials, and tried to gain access to backup infrastructure. This led everyone to believe the threat actor was a low-skilled initial access brokers (IABs) that gain initial footholds on corporate environments and sell access to the hacked network on underground portals.

Okay. I think it's entirely expected that anyone who has any need for any sort of code or scripting for any purpose whatsoever will increasingly be using AI. That's just today's reality. Good guys are doing it and bad guys are doing it, and there's no reason to expect AI to be able to discriminate between the two. A high-level language compiler doesn't know or care who's using it, or to what purpose the code it's helping to produce will be put. That's not its job. So the fact that we have now chosen to give consciousness-emulating large language models the marketing label of "Artificial Intelligence" should not and does not automatically mean that these new tools somehow have responsibility for what they're being asked to produce.

But don't these AI tools make attackers more powerful? Yes, they do. And they also make the good guys more productive. That's why everyone – good and evil – is using them. In the current instance, there's nothing inherently wrong with a script that performs a vulnerability assessment. White hat security researchers employ such tools to aid their beneficial research much as bad guys may use the same tools to perform pre-attack vulnerability assessments.

My point is that any social media hysteria arising from the fact that AI "was involved" is now ridiculous. If you encounter it online I would recommend meeting it with a shrug and a click on the "thumbs down" button. This is just the way the future is going to look now. It may have surprised us a few years ago, but it should surprise us no longer. And "AI" should not receive **any** of the blame for the way its creators – we humans – choose to use it. It's a tool and nothing more. It has no social obligations or responsibilities. It's not accountable – we are.

Mexican government hacked using Claude

Now I'm going to give everyone a quick self-test to see whether the point I hope I've just made has had the chance to sink in. Perform a self assessment to see how you feel about this next piece of news:

"A hacker has stolen more than 150 gigabytes of data from multiple Mexican government agencies. The attacker allegedly used Claude to assemble scripts to gain access to government networks. According to Bloomberg, the attacker breached and stole data from Mexico's tax authority, national electoral institute, and several state water utilities. The stolen data covers 195 million taxpayer and voter records, government employee credentials, and civil registry files."

Should we care at all that AI was employed in these attacks? No. The fact that Claude was used in these attacks appears to be the highlight of Bloomberg's piece – it was certainly the headline which they attempted to make inflammatory. Eventually the world will get used to this and it will just be assumed.

Apple and Age Restriction

Apple appears to be feeling the pressure to respond to the growing legislation-driven need for the providers of Internet services and online apps to know and to respond to the age of their users. Last Tuesday, Apple posted an update to their developer portal, addressed to their App developers. They wrote:

Today we're providing an update on the tools available to developers to meet their age assurance obligations under upcoming U.S. and regional laws, including in Brazil, Australia, Singapore, Utah, and Louisiana. Updates to the Declared Age Range API are now available in beta for testing.

Brazil

Developers who are distributing apps in Brazil can use the updated Declared Age Range API to obtain a user's age category. Age categories for users in Brazil will be shared when the user or a parent or guardian (where relevant) agrees to share the age category with you. The API will also return a signal from the user's device about the method of age assurance. For developers distributing their apps in Brazil, if you identify that your app contains loot boxes through the age rating questionnaire, the age rating of your app on the Brazil storefront will be updated to 18+.

Apps rated 18+ in Australia, Singapore, and Brazil

Starting February 24, 2026 [in other words, the last Tuesday date of this posting], Apple will block users in Australia, Brazil, and Singapore from downloading apps rated 18+ unless they have been confirmed to be adults through reasonable methods. The App Store will perform this confirmation automatically. However, developers may have separate obligations to independently confirm that their users are adults. To assist with this, the Declared Age Range API—available on iOS, iPadOS, and macOS—provides developers with a helpful signal about a user's age.

Utah and Louisiana

For users with new Apple Accounts in Utah as of May 6, 2026, and in Louisiana as of July 1, 2026, age categories will be shared with the developer's app when requested through the Declared Age Range API. The tools we previously announced have been expanded to help developers meet compliance obligations for Louisiana and Utah, including:

- *Declared Age Range API*
- *Significant Change API under PermissionKit*
- *New age rating property type in StoreKit*
- *App Store Server Notifications*

New signals are now available through the Declared Age Range API, including whether age-related regulatory requirements apply to the user and if the user is required to share their age range. The API will also let you know if you need to get a parent or guardian's permission for significant app updates for a child.

Developers can use the Declared Age Range API to present significant update notifications to adults in these states through the Significant Update Action, now in beta. When releasing a significant update, developers must follow the Human Interface Guidelines and provide users with a meaningful description of the update.

This all certainly seems like a mess. We're now littering the landscape with special case handling – region by region; nations and/or individual states – in order to comply with the random and arbitrary legislation being put in place by local and national governments. In some cases app developers must use the APIs that Apple has been reluctantly providing and only improving when forced to. In other instances Apple itself is required to function as the gatekeeper. To this we add the murky features of new versus old accounts, the unknown of how age is determined, and why there's any difference between downloading versus using an app. The suggestion is that if you already have an app you're somehow grandfathered into keeping and using it, but newbies cannot get it? The entire thing is a confounding mess. And since individual states in the U.S. and countries in the world will always retain the right to apply whatever restrictions they may choose, it's unclear how things are likely to improve.

And all of this is just for application use on a device. None of this helps with web-based services.

FTC Issues COPPA Policy Statement

And speaking of online web-based services. There has apparently been some concern over the intersection of children's privacy enforcement and the apparently explicit need to violate that privacy for the sake of complying with legislated age determination.

So last Wednesday, on the heels of Apple's begrudging update to their age-related APIs and download enforcement, the U.S. Federal Trade Commission, our FTC, issued a formal policy statement with the headline: *"FTC Issues COPPA Policy Statement to Incentivize the Use of Age Verification Technologies to Protect Children Online"* They wrote:

The Federal Trade Commission issued a policy statement today announcing that the Commission will not bring an enforcement action under the Children's Online Privacy Protection Rule (COPPA Rule) against website and online service operators that collect, use, and disclose personal information for the sole purpose of determining a user's age via age verification technologies.

The COPPA Rule requires operators of commercial websites or online services directed to children under 13, and operators with actual knowledge that they are collecting personal information from a child, to provide notice of their information practices to parents and to obtain verifiable parental consent before collecting, using, or disclosing personal information collected from a child under 13.

So we see the problem here, right? The emerging age restriction regulations are placing the burden upon online services to whatever they must to determine their visitors' ages. But doing this could force the site to run afoul of other regulations, specifically COPPA, which are already in place to protect the privacy of their underage visitors and users.

In this instance, it's necessary to carve out an explicit privacy exception so that online services will be able to collect the data that they must without fear of tripping over COPPA's restrictions. So the FTC explains:

Age verification technologies play a critical role in helping parents as they monitor their children's online activities. Since COPPA was enacted in 1998, there's been an explosion in the use of internet-connected technologies by children. To help parents navigate the challenges associated with their children's online activities, some states have started requiring some websites and online services to use age verification mechanisms to help determine the age of users. But as noted at the FTC's recent workshop on age verification technologies, some age verification technologies may require the collection of personal information from children, prompting questions about whether such activities could violate the COPPA Rule.

Christopher Mufarrige, Director of the FTC's Bureau of Consumer Protection said: "Age verification technologies are some of the most child-protective technologies to emerge in decades. Our statement incentivizes operators to use these innovative tools, empowering parents to protect their children online."

*The policy statement states that the Commission will **not** bring an enforcement action under the COPPA Rule against operators of general audience sites and services and mixed audience sites and services that collect, use, or disclose personal information for the sole purpose of determining a user's age without first obtaining verifiable parental consent—if they comply with certain conditions, specifically that they:*

- do not use or disclose information collected for age verification purposes for any purpose except to determine a user's age;*
- do not retain this information longer than necessary to fulfill the age verification purposes, and delete such information promptly thereafter;*
- disclose information collected for age verification purposes only to those third parties the operator has taken reasonable steps to determine are capable of maintaining the confidentiality, security, and integrity of the information, including by obtaining certain written assurances from those third parties;*
- provide clear notice to parents and children of the information collected for age verification purposes;*
- employ reasonable security safeguards for information collected for age verification purposes; and*
- take reasonable steps to determine that any product, service, method, or third party utilized for age verification purposes is likely to provide reasonably accurate results as to the user's age.*

The policy statement indicates that the Commission intends to initiate a review of the COPPA Rule to address age verification mechanisms. The policy statement will remain effective until the Commission publishes final rule amendments on this issue in the Federal Register, or until otherwise withdrawn.

So this policy statement is intended to provide interim cover for online sites and services that need to enforce privacy-breaching age-restriction measures which would otherwise expose the site to COPPA infringement. This suggests that COPPA will require amending to provide a permanent and clear path for privacy-respecting age verification for minors.

Meta's CSAM-AI false positives

The Guardian reports that Meta's CSAM-detection AI is flooding law enforcement with low-quality unactionable false positive reports of online child sexual abuse that are seriously hampering law enforcement's ability to function. Under their headline "Meta's AI sending 'junk' tips to DoJ, US child abuse investigators say", here's what The Guardian reported:

Officers from the US Internet Crimes Against Children (ICAC) taskforce said that Meta's use of artificial intelligence to moderate its social media platforms is generating large volumes of useless reports about cases of child sexual abuse, which are draining resources and hindering investigations. Benjamin Zwiebel, a special agent with the ICAC taskforce in New Mexico, said last week during his testimony in the state's trial against Meta: "We get a lot of tips from Meta that are just kind of junk." The state's attorney general alleges the company's platforms are putting profits over child safety.

At first I was puzzled by that. But what I believe New Mexico's attorney general is saying is that rather than employing humans who would be able to usefully discriminate between what is and is not actual child exploitation and abuse, Meta is endeavoring to save money by using AI which is not actually doing the job. The reporting continues:

Meta disputes these allegations, citing changes it has introduced on its platforms, such as teen accounts with default protections. The ICAC taskforce is a nationwide network of law enforcement agencies coordinated with the US Department of Justice to investigate and prosecute online child exploitation and abuse cases.

Another ICAC officer, speaking on the condition of anonymity to discuss internal matters, said: "Meta is providing thousands of tips each month. It's pretty overwhelming because we're getting so many reports, but the quality of the reports is really lacking in terms of our ability to take serious action." The ICAC officer added that the total number of cybertips their department had received doubled from 2024 to 2025.

Both Zwiebel and two ICAC officers said that unviable tips from Instagram, Facebook and WhatsApp often contain information that is not criminal. The anonymous officers added that in other cases tips sometimes contain information indicating that a crime may have occurred, yet vital images, videos or text are missing or redacted. The ICAC officer added: "[Unviable tips from] Instagram have really skyrocketed recently, especially in the last couple of months, and that's one of the biggest places where we're seeing important information not being provided. In those cases, we don't have the information to further the investigation. It weighs on you to know that this crime occurred, but we can't identify the perpetrator."

So just to clarify that point. These investigators are saying that what they see are clearly crimes which Meta's use of AI happened to have found, but that the evidence that's needed to take any action about it is missing. So Meta's use of AI is not only flooding law enforcement with crap, but it's also serving to obscure the necessary details of actual crimes it detects. If we didn't know better we'd be inclined to think this had been deliberately designed by criminals for criminals.

Asked about Zwiebel's testimony and the ICAC officers' remarks, a Meta spokesperson said: "We've supported law enforcement to prosecute criminals for years: the DoJ has repeatedly praised our fast cooperation that has helped lead to arrests, and NCMEC has praised our streamlined and 'improve[d]' tip reporting process. In 2024, we received over 9,000 emergency requests from US authorities and resolved them within an average of 67 minutes –

and even more quickly for cases involving child safety and suicide. Consistent with applicable law, we also report apparent child sexual exploitation imagery to NCMEC and support them to prioritize reports, from helping build their case management tool to labeling cybertips so they know which are urgent.”

I'll note that while this sounds great, it doesn't appear to be responsive to the question of AI's use. That Meta spokesperson appears to be referring to the work of humans employed by Meta, not cost-saving AI. The Guardian's reporting then shifts gears to provide some background on the NCMEC – the National Center for Missing & Exploited Children, writing:

By law, social media companies based in the United States are required to report any detected child sexual abuse material (CSAM) on their platforms to the National Center for Missing & Exploited Children (NCMEC). NCMEC serves as a national clearinghouse for reports, which it forwards to the appropriate law enforcement agencies across the United States and internationally. NCMEC does not have the authority to filter out any tips that may be unviable before they are sent to the relevant law enforcement agencies.

Meta is by far the largest reporter to NCMEC. In its data report for 2024, NCMEC said Meta made 13.8 million reports across Facebook, Instagram and Whatsapp, out of the 20.5 million tips it received in total.

NCMEC said that in 2024, more than one million CyberTipline reports were linkable to a specific US state, and those reports were made available to the ICAC taskforces around the country, as well as other federal, state, and local law enforcement agencies, for investigation.

Meta and other social media companies use AI to detect and report suspicious material on their sites and employ human moderators to review some of the flagged content before sending it to law enforcement. The Guardian has previously reported that tips generated by AI that haven't also been reviewed by a social media company employee often cannot be opened by a law enforcement officer without a warrant because of fourth amendment protections. This extra step also slows investigations of potential crimes, lawyers involved in such cases say.

A Meta spokesperson said: "It's unfortunate that court rulings have increased the burden on law enforcement by requiring search warrants to open identical copies of content we've already reviewed and reported. Our image-matching system finds copies of known child exploitation at a scale that would be impossible to do manually, and we work to detect new child exploitation content through technology, reports from our community, and investigations by our specialist child safety teams."

Under the "Report" Act – where REPORT is an acronym for Revising Existing Procedures On Reporting via Technology – which came into force in November 2024, online service providers must broaden and strengthen their reporting obligations by notifying NCMEC's CyberTipline not only about child sexual abuse material but also about planned or imminent abuse, child sex trafficking and related exploitation; they must also preserve evidence for a longer period and face higher penalties if they knowingly fail to comply.

Since the act passed, the number of unviable tips supplied by Meta has increased dramatically, which could be because the company is acting to ensure it is not falling foul of the law, two ICAC officers said. Many of these tips could not be construed as a crime, such as adolescent girls talking about which celebrity they find most attractive.

Special agent Benjamin Zwiebel said in court. "Based on my training and experience, it appears that they are being submitted through the use of AI, as these are common mistakes that an AI would make that a human observer would not. Zwiebel added that his department receives significantly fewer tips on legitimate cases of child sexual abuse material (CSAM) distribution from Meta than in previous years.

Every tip that reaches an ICAC division must be reviewed, and the influx of unviable tips is taking time and resources away from investigating legitimate cases of child abuse, said two officers. One ICAC officer said: "It's killing morale. We are drowning in tips and we want to get out there and do this work. We don't have the personnel to sustain that. There's no way that we can keep up with the flood that's coming in."

I chalk this up less to Meta being evil than to the growing pains of effective AI deployment. We're still very much learning how to best use the new and surprising capabilities of large language model networks. And I suspect that a strong case could be made for there truly being far too much content for humans to manually inspect. Although the specter of having overlord AI's examining everything that's transacted over social media feels very Orwellian, our legislators are requiring a level of oversight from social media companies that likely has no other workable solution. So AI it will be. We just need to continue figuring out how to best use it.

Roskomnadzor has been busy

I just saw a short mention blurb that surprised me. The news was just that Russia's wonderfully named Internet watchdog, Roskomnadzor, has now blocked Russian citizens' access to ... wait for it ... 469 individual VPN services inside Russia. It seems to me that the fact that there **are** 469 VPN services inside Russia to be blocked is the real story here. Holy crap. Talk about a citizenry that's desperate to escape from the shackles of their own state's filtering and tampering. This is a citizenry that is desperate for contact with the outside world and a repressive government that is doing everything it can to prevent that. It's becoming increasingly clear why Russia has been experimenting with completely disconnecting from the global Internet.

In other Russian news, I saw a report that indicated that the Kremlin had decided to fully block Telegram starting in April of this year. That puzzled me, since I thought the Telegram was already being fully blocked. But this reporting stated that Telegram was currently only 55% blocked. But it's certainly not clear to me what a 55% block might mean. The only thing I can figure is that perhaps access to Telegram is currently being limited to specific sectors or industries or regions and that additional regions are being added so that by the end of this month of March nothing will be left? Whatever the case, Russia appears to quite intent upon controlling its citizen's access to information.

The UK reports on proactive national network scanning

About 14 months ago – January 2025 – we reported that the UK was launching a plan to begin continuously and proactively scanning its own national public-facing network segments for the purpose of preemptively detecting vulnerabilities and alerting those owners of the IP addresses where vulnerabilities were found. Our listeners may also recall that I was jumping up and down over how much sense I thought this made and I suggested that this is something every nation should be doing in its own self interest.

We're talking about this again today, 14 months later, because last Thursday the UK put out a celebratory press release with the headline: *"Government cuts cyber-attack fix times by 84% and launches new profession to protect public services."* A new profession? Okay. The Press Release led with three summary bullet points:

- *Critical cyber weaknesses across the public sector will now be fixed 6 times faster than before.*
- *Ministers are determined to go further — with first-ever dedicated government Cyber Profession to give the state the skilled staff it needs to protect UK's key services from cyber threats.*
- *The number of serious unresolved cyber security weaknesses across government cut by three quarters as part of government-wide efforts to strengthen Britain's digital defences.*

Before I share what the press office of the UK said, allow me to preface this by noting that we're going to encounter something that makes no sense whatsoever to me. But regardless, here's what they wrote:

Public services millions of people depend on – from the NHS to the Legal Aid Agency – are becoming significantly safer and more resilient thanks to major improvements by the government to identify and fix cyber threats.

A specialist government monitoring service, introduced as part of the Blueprint for modern digital government, published in January 2025, means serious security weaknesses in public sector websites are fixed 6 times faster – cutting the average time from nearly 2 months to just over one week.

Okay. So far so good. But then this appears to go off the rails. The Release next says:

The vulnerabilities are in the Domain Name System (DNS) — the internet's address book that turns website names into the numbers computers use to find them. Weaknesses in DNS can allow attackers to redirect users to fraudulent sites, steal sensitive data, or take services offline entirely — with potentially serious consequences for anyone relying on government services.

Before this service was in place, a weakness in a government DNS record could go unnoticed for nearly 2 months — long enough for a hostile actor to redirect someone trying to access a government service to a fake site designed to steal their personal details, intercept sensitive communications, or disrupt services that people rely on. The vulnerability monitoring service has closed this window down to 8 days. It alerts the right people with clear, practical guidance on how to fix the problem, and tracks progress until each issue is resolved.

What the hell are they talking about? What's a "*weakness in a government DNS record*" ?? In this day and age, when I see something that sounds entirely plausible and reasonable to a lay person, but which is actually nonsense, the first thing I think is that some AI somewhere was having a bad day. The press release said: "*Before this service was in place, a weakness in a government DNS record could go unnoticed for nearly 2 months*" — again **what?!?! what?!?!**

Okay. Well, let's just play along and see what else happens. The Release continues:

Speaking at the annual Government Cyber Security and Digital Resilience conference, Digital Government Minister Ian Murray will outline how this will sharply reduce the risk of hackers targeting essential services like the NHS.

He'll also outline how the government has reduced its backlog of these vulnerabilities by 75% – significantly shrinking the window for cyber criminals to target essential public services – from GP surgeries and ambulance trusts to hospitals and social care providers.

Today's announcement marks a decisive step in closing the door on such threats with the government going even further with the launch of the first-ever dedicated government Cyber Profession.

What... DNS monitoring?

This programme will recruit and train the top-tier cyber experts needed to keep public services safe. Minister for Digital Government Ian Murray said:

Cyber-attacks aren't abstract threats – they delay NHS appointments, disrupt essential services, and put people's most sensitive data at risk. When public services struggle it's families, patients and frontline workers that feel it.

The vulnerability monitoring service has transformed how quickly we can spot and fix weaknesses before they're exploited so we can protect against that. We've cut cyber-attack fix times by 84% and reduced the backlog of critical issues by three quarters. And as the service expands to cover more types of cyber threats, fix times are falling there too.

But technology alone isn't enough. Today I'm launching a new government Cyber Profession to attract and develop the talented people we need to stay ahead of increasingly sophisticated threats - making government a destination of choice for cyber professionals who want to protect the services that matter most to people's lives.

Dr Richard Horne, CEO of the NCSC, said:

Cyber security is more consequential than ever today with attacks in the headlines showing the profound impacts they can have on people's everyday lives and livelihoods.

As our public services continue to innovate, it is vital that they remain resilient to evolving threats and vulnerabilities are being effectively managed to reduce the chances of disruption.

The government Cyber Action Plan is a crucial step in building stronger cyber defences across our public services and the launch of the government Cyber Profession today will help attract and retain the most talented professionals with the top-tier skills needed to keep the UK safe online.

The VMS continuously scans 6,000 UK public sector bodies, detecting around 1,000 different types of cyber vulnerabilities. When a weakness is identified, the service alerts the relevant organisation with specific, actionable guidance and tracks progress until the issue is resolved.

Okay. Now, THAT finally makes sense! THAT is what we would expect. They have a continuously-running Internet scanner that's scanning 6,000 UK public sector agencies and entities looking for 1,000 different types of cyber vulnerabilities at each of the IPs of the configured targets. Yay!

Unfortunately, the presence of that looney tunes nonsense about weaknesses in government DNS records casts the entire announcement into question. Just where does the AI brain fart end and reality begin? But we do now appear to be back on track. The Release finishes up, writing:

By automating detection and streamlining remediation, the service has:

- *reduced median time to fix domain-related vulnerabilities from 50 days to 8 days — an 84% improvement*

Whoops. We're back to crazy town. What is a "domain-related vulnerability" and how can it have been reduced from taking 50 days to fix down to just 8 days? It really does seem as though an AI had a hand in the preparation of this Release. That's too bad. The other three bullet points seem more reasonable:

- *reduced median time to fix other cyber vulnerabilities from 53 days to 32 days*
- *cut the backlog of critical open domain-related vulnerabilities by 75%*
- *processed and resolved around 400 confirmed vulnerabilities each month*

The new government Cyber Profession is co-branded with the Department for Science, Innovation and Technology and the National Cyber Security Centre. It will introduce a competitive total employee offer, establish a dedicated Cyber Resourcing Hub to streamline recruitment, and create a clear career framework aligned with UK Cyber Security Council professional standards.

It will also include a government Cyber Academy for training and development, a new apprenticeship scheme to build future talent, and structured career pathways to strengthen long-term capability across the public sector.

The North West will serve as a primary hub for the profession, building on Manchester's growing digital ecosystem and the forthcoming government Digital Campus.

All of that sounds great and reasonable, too. The UK has clearly implemented an extremely useful service. I hope that other nations pick up on this idea and put it into practice. I see no downside of any sort.

Hacker/Extorter on the loose

Okay, now this little tidbit is a bit surprising. As I was scanning the news I encountered a piece of news declaring that "Vastaamo hacker disappears". Okay. I have no idea what that is. But then in reading a bit into the story it mentions that a Finnish hacker lost his appeal and will have to go back to prison after a court increased his original sentence. So far, nothing stands out there, right? But Leo, I believe you'll recall this event from six years ago, since you were quite affronted by this cretin's conduct at the time. The report explains that this Finnish hacker was sentenced to six years and three months for hacking the Vastaamo psychotherapy centre in 2020 and then extorting its patients – that's what made this stand out.

This creep obtained the psychotherapy centre's quite personal and highly confidential psychotherapy records – including, of course, the contact information that would be needed for them to be contacted. He then threatened them with public exposure unless they paid up.

Beyond this, as I also recall, we were also shocked by the sheer number of patient records that the Vastaamo psychotherapy centre had maintained online which were stolen. We noted that not only were they at fault for not better-protecting their data, but that they should really be held accountable for leaving the data of years and years of previously patients in hot storage. If they wanted to retain them for possible future need they could be archived offline.

Scattered Lapsus\$ Hunters (SLH) recruits women

In their "Cyber Intel Brief" the cyber intelligence firm "Dataminr" reports that the Scattered Lapsus\$ Hunters (SLH) collective has begun actively recruiting female individuals for their voice phishing campaign. SLH is offering upfront payments for social engineering calls targeting IT help desks. Dataminr's report offered three key takeaways:

- *Tactical Evolution: SLH is diversifying its social engineering pool by specifically recruiting women to conduct voice phishing attacks, likely to increase the success rate of help desk impersonation.*
- *Large Incentives: The group is offering significant financial incentives (\$500-\$1,000 upfront per call) and providing pre-written scripts to recruits.*
- *High-Profile Risk: SLH is a "supergroup" alliance of Lapsus\$, Scattered Spider, and ShinyHunters, known for compromising major global corporations and stealing over 1.5 billion records.*

Their posting then walks us through their discovery of SLH's online recruitment postings and ends with some very useful advice to any potential enterprise targets. Under the heading "Organizations should adopt a heightened defensive posture against social engineering" they enumerate:

- *Help Desk Training: Immediately brief IT help desk and support personnel on this specific recruitment trend. Emphasize that attackers may use pre-written scripts and polished voice impersonation.*
- *Strict Identity Verification: Enforce out-of-band identity verification (e.g., video calls or secondary internal verification) for all password resets or MFA credential changes requested via phone.*
- *Harden MFA Policies: Move away from SMS or push-based MFA, which are vulnerable to SLH's known TTPs like SIM swapping and fatigue bombing. Implement FIDO2-compliant hardware security keys where possible.*
- *Monitor Anomalous Access: Audit logs for new user creation or administrative privilege escalation immediately following help desk interactions.*

It's interesting that they've figured out that a woman's voice may generate a different response than a man's. These guys are tricky and it's fair to say that it's not possible to be over prepared.

Cisco's latest CVSS 10.0

Last Wednesday, Cisco released the news of CVE-2026-20127, once again achieving that rarest of the rare CVSS 10.0 scores. This one is an actively exploited 0-day, first discovered while it was being abused in the wild. The title Cisco gave to their disclosure was "*Cisco Catalyst SD-WAN Controller Authentication Bypass Vulnerability*". Yes, you heard right. Surprise surprise, an authentication bypass vulnerability. Cisco wrote:

A vulnerability in the peering authentication in Cisco Catalyst SD-WAN Controller, formerly SD-WAN vSmart, and Cisco Catalyst SD-WAN Manager, formerly SD-WAN vManage, could allow an unauthenticated, remote attacker to bypass authentication and obtain administrative privileges on an affected system.

This vulnerability exists because the peering authentication mechanism in an affected system is not working properly.

Huh. "Not working properly." No one would disagree with that, though calling it "catastrophically

defective” might be somewhat more accurate. This one is so bad that both the NSA and CISA in the US, the Australian Signals Directorate’s Australian Cyber Security Centre, the Canadian Centre for Cyber Security, New Zealand’s National Cyber Security Centre and the UK’s National Cyber Security Centre all published “patch or perish” announcements in a desperate attempt to bring the need to patch all systems to the attention of their owners. The “SD” in SD-WAN stands for Software Defined. So it’s a software-based networking platform that connects branch offices, data centers, and cloud environments through a centrally managed system. It uses a controller to securely route traffic between sites over encrypted connections.

This is another instance where any company that recognized that simple authentication can never be relied upon for security, and had therefore taken the trouble to separately restrict incoming SD-WAN connections to known endpoint peers, would never have anything to fear from these authentication failures and [a] would not have suffered a potentially devastating network compromise and [b] could therefore update their SD-WAN instances with something less than pants-on-fire emergency measures.

Once again, Cisco’s own announcement moderately underplayed the consequences. They wrote:

An attacker could exploit this vulnerability by sending crafted requests to an affected system. A successful exploit could allow the attacker to log in to an affected Cisco Catalyst SD-WAN Controller as an internal, high-privileged, non-root user account. Using this account, the attacker could access NETCONF, which would then allow the attacker to manipulate network configuration for the SD-WAN fabric.

It was the Australian Signals Directorate who first discovered and reported these attacks being used in the wild. No surprisingly, they paint a somewhat less rosy picture of the consequences, writing:

*Malicious cyber threat actors are targeting SD-WANs of organisations, globally. These actors exploited a Cisco Catalyst SD-WAN controller authentication bypass vulnerability, CVE-2026-20127. After exploitation of this vulnerability the malicious actors add a rogue peer, and eventually **gain root access** to establish long-term persistence in SD-WANs.*

So, sorry Cisco... not just non-root user accounts.

VulnCheck’s 2025 Exploit Intelligence Report

VulnCheck’s annual report on the in-the-wild use of known security CVE’s is interesting. From the report’s teaser summary we learn a few things. They write:

In 2025, barely 1% of disclosed vulnerabilities were exploited in the wild. Yet those that were exploited were operationalized quickly, attracted diverse threat actors, and often caused outsized damage before organizations had a chance to respond. This report identifies which vulnerabilities mattered, why attackers targeted them, and where timing failures left organizations exposed.

VulnCheck tracked exploitation patterns, threat actor behavior, and weaponization timelines across hundreds of thousands of vulnerabilities in 2025. The data revealed how quickly new vulnerabilities became bona fide threats, how AI proof-of-concept code is polluting risk assessment pipelines, and which threat actors ramped up vulnerability exploitation amid

geopolitical tension:

- *VulnCheck identified 50 Routinely Targeted Vulnerabilities from 2025 that had elevated risk profiles by the end of the year, drawing interest from ransomware, threat actors, botnets, and researchers (often simultaneously)*
- *Proof-of-concept exploits for new CVEs increased 16.5% in 2025, inundating organizations with "risk" signals that often turned out to be false or misleading AI-generated slop*
- *China-nexus threat actor attributions increased 52% year-over-year, while ransomware groups shifted toward zero-day exploitation at accelerating rates, with 56.4% of ransomware CVEs discovered through zero-day activity*

Their fully detailed 41-page report contains a wealth of information some of which I'd like to share next week. Stay tuned.

Listener Feedback

David Benedict

*Hi Steve, Not to pull you back into the whole Code Signing discussion again, but what if WE simply don't buy these code signing Certs? What if we simply start self signing code? Is there anything to stop us from self signing and building our own reputation in that way?
Thanks, David Benedict*

That's an interesting idea. The moves that the CA/Browser forum have been making on the code signing front feel entirely different from their earlier squeezing on the TLS certificate side.

The reason Let's Encrypt was able to effectively replace and displace the traditional certificate authorities for the world's web server domain validation certificates is that Let's Encrypt is only providing what its name suggests: encryption. It's making no assertion of any kind about the reputation of the domain name holder. And when you think about it, where strong assertions of identity are needed and are being made about the owner of a certificate, whether for a web domain, the digital signer of a document, or the authorship of code, we need entities such as certificate authorities standing by to do the necessary work of verifying identity and carefully issuing certificates which attest to what their research has found.

Unfortunately, while we've been going about our lives, the certificate authority business has quietly been consolidating. This has sometimes been triggered when an irresponsible certificate authority so flagrantly abuses its position of trust that the various root programs are finally forced to remove their trust. In those cases, the disgraced CA is forced to sell off its CA business assets to another certificate authority. In other cases it's just the bigger fish swallowing up the smaller fish. While I was scouting around for a new code signing certificate authority I noticed that many of the smaller companies had exactly the same pricing as DigiCert. It turned out that many of them have become fronts for DigiCert's products – they're just resellers.

The upshot of many years of CA industry consolidation is that the world no longer has a competitive certificate authority industry. We're watching the formation of a monopoly that has

the gall to charge its customers per signature.

David began his note, writing "*Hi Steve, Not to pull you back into the whole Code Signing discussion again*". It's not your fault, David. This whole thing obviously rubs me the wrong way. One of my personality's hot buttons is bullying. I have never been okay with the abuse of power which is what, I believe, anyone observing the actions of the CA industry would conclude is happening. I don't see any way out of this, but I will gladly share any solutions I find.

To that end, during this research I discovered that all of the various CA's who offer code signing certificates provide the option of installing certificates into a customer-provided HSM rather than selling the certificate pre-installed in their own dongle-token, typically for another \$100.

I found a \$72 USB-A HSM dongle that I love called the "SmartCard-HSM 4K" – 4K because it can handle 4096-bit RSA keys, which is now what's necessary. I have a link to this device in the show notes: <https://www.cardlogix.com/product/smartcard-hsm-4k-usb-token/>

<https://www.smartcard-hsm.com/features.html#usbstick> The device has a website at: [smartcard-hsm.com](https://www.smartcard-hsm.com) and most significantly all of the card's multi-platform support is open source: <https://github.com/OpenSC/OpenSC/wiki/SmartCardHSM> I have a link to its GitHub pages in the show notes.

One of the very cool features of this for me, is that this HSM enables secure and encrypted cross-HSM private key and certificate transfers. I have multiple machines where I want to be able to sign code: My two working locations and GRC's servers in the Level3 data center. So I first had the first HSM securely generate an 4096-bit RSA key pair. The private key never leaves the device, but the public key is exported in a CSR, a Certificate Signing Request. I uploaded that CSR to IdenTrust to receive their signature. They promptly returned the resulting certificate which is used to verify any signatures the HSM generates for my code.

One of the many cool things about this solution is that each HSM includes its permanent device certificates that enable it to establish a secure key sharing key among others of its own kind. This allows one HSM's private keys to be securely duplicated across as many other devices as you may wish, as well as being externally backed up for export without ever exposing its private keys. Each HSM also has a large amount of storage with room for hundreds of keys, certificates and so on. So if an enterprise might have a number of trusted developers, work-from-home, satellite offices or whatever, for the price of \$72 each, as many developers can be given the ability to securely sign code as needed.

There's much much more than I have shared. So if you have an interest or need, check out [smartcard-hsm.com](https://www.smartcard-hsm.com). The [cardlogix.com](https://www.cardlogix.com) retailer I found happened to be near me in the US. Their "Where To Buy" page also lists German and Taiwanese retailers: <https://www.smartcard-hsm.com/buy.html>

Since my original DigiCert EV code signing certificate does not expire until this November 20th, but I wanted to obtain a 3-year certificate before they were no longer available, my plan has been to see whether I can pre-establish a reputation for the new IdenTrust 3-year certificate by having it cosign GRC's freeware. That's now in place. GRC's most popular freeware, first and

foremost ValiDrive, which is being downloaded more than 1000 times per day, is now cosigned with IdenTrust's certificate. So I'm hoping that once we get to November I'll be able to drop the DigiCert signatures from my newly signed software and IdenTrust's certificate, which will by then have 10 months of exposure, will be able to stand alone and keep Windows' trigger-happy gatekeepers happy.

Finally, just to see whether I could, last week I also reissued my existing DigiCert certificate in customer-provided HSM mode. This allowed me to add it into my new HSMs alongside the newly minted IdenTrust certificate. It all worked perfectly and my new HSMs now contain both the existing DigiCert and the new IdenTrust code signing certificates.

Okay. Believe it or not, I haven't forgotten about David. He started me off on all this by asking about the possibility of coders sidestepping all of this nonsense by using self-signed certificates. The use of self-signed certificates has been common practice for web developers for many years. I have a self-signed certificate for "localhost" sitting in the trusted root stores of my various workstations. I run a webserver on those machines which I use for local web development. Having a self-signed certificate for "localhost" allows me to use HTTPS URLs starting with <https://localhost/> without the web browsers pitching a fit. It's just very handy.

So let's explore how this might be extended for code signing. If, rather than having DigiCert or IdenTrust sign my CSR – my code signing request – I could instead use my private key to sign the certificate, creating a self-signed certificate which would then be installed into the system's trusted root store. From that point on, any code I signed would carry the root certificate's matching public key and any check on the validity of my code's signature would show its signature to be valid.

The reason this is not a useful solution for a software publisher, such as GRC, is that these code signatures would **only** be valid on machines that had previously installed their matching root certificate. What DigiCert, IdenTrust and all of the other CAs all have going for them is that their root certificates are already pre-installed wherever any certificates they have signed might need to be trusted.

Since Windows has now developed the practice of deleting-on-sight any executable file that appears on its drive without a valid and trusted signature, it would be necessary for GRC to tell its customers that before attempting to download, let alone run, any of our software they must first download and install GRC's trusted root certificate. Since I would never install anyone else's root certificate into my machine's root store, I would never ask anyone else to do that in order to download and run my code. The burden of making my code acceptable to someone else's machine should be on me, not on them. So, while signing one's own code for use on our own machines would work, just like using a self-signed web certificate for local use of a web browser and web server, I don't see any way to break our Certificate Authorities' stranglehold on developers for code signing that needs to be universally trusted.

Note, however, that this does not apply to enterprises. Enterprises might choose to use publicly trusted code signing certificates for their internal use. But within an enterprise it might also work to sign code with a certificate that's only trusted within the enterprise on enterprise machines. Remember that many enterprises already install their own TLS root certificates on all internal

workstations so that their networking middleboxes are able to intercept, decrypt and scan everything that enters and exits their network. So it would make sense to add a private code signing root certificate to all enterprise machines for their own internal use.

DGC

*Hey Steve, Long time listener but I'm a few episodes behind right now. In Episode #1062 you said "... We also see employees in positions of trust on internal enterprise networks being tricked into clicking malicious links and inviting malware inside the house. No form of fancy coding AI is going to fix *any* of those things."*

[DGC writes:] That may not be entirely true; I recently came across a new solution, "Charlemagne", which runs on a desktop and monitors (privately, locally) what the user is doing. It uses an SLM (a Small Language Model) to detect potentially malicious actions like lookalike websites, malicious links the user might click on, etc, and then warns the user not to do those things. So an AI agent helping a user avoid accidental bad actions could be helpful, no?

Could be helpful, yes. And I very much like the idea. The talk Leo and I will be holding during the Zero Trust World event, is titled "The Call is Coming From Inside the House." This is a metaphor for what I believe to be the biggest and most intractable problem facing today's enterprises. Stated as succinctly as possible, the problem is "Overprivileged users who make mistakes." Though the term "overprivileged" is typically used in a derogatory context, I don't mean it that way at all. I'm using it in its computer science context where "overprivileged" is the result of not following a strict "least privilege" model.

The beauty of describing the problem as "overprivileged users who make mistakes" is that it points us toward two solutions to the problem: Either we no longer overprivilege our users or we somehow arrange to prevent them from making mistakes. Whereas it might be possible to constrain what the employees of an enterprise might be able to do on the enterprise's network, a large part of the joy of using a personal computer is that its use is – personal – which is to say almost entirely unconstrained. Since operating within a "least privileged" environment is no fun and would not be tolerated by personal computer users, that suggests that the solution lies in somehow arranging to prevent their mistakes.

So, to that end, I love the idea of some form of active client-side local AI agent that carefully scrutinizes everything the user is seeing and doing and interposes itself between the user's actions and the computer system. Leo, I and our listeners know how to examine a URL to detect trickery. But even the best of us are not always paying 100% attention and even when we are, we might miss the use of embedded UNICODE characters to create a lookalike URL. Or in our haste we might click a link without first carefully checking all the way to the right of its domain portion to make sure that its top level domain is what we expect. So by all means, the idea of having an AI agent peeking over our shoulder and watching our back to help detect the things we might well miss makes all kinds of sense. And needless to say, I hope that it would totally freak out if its user were getting ready to paste the contents of the clipboard which was pasted from their web browser into their Windows' "Run" dialog!

KongTuke's CrashFix

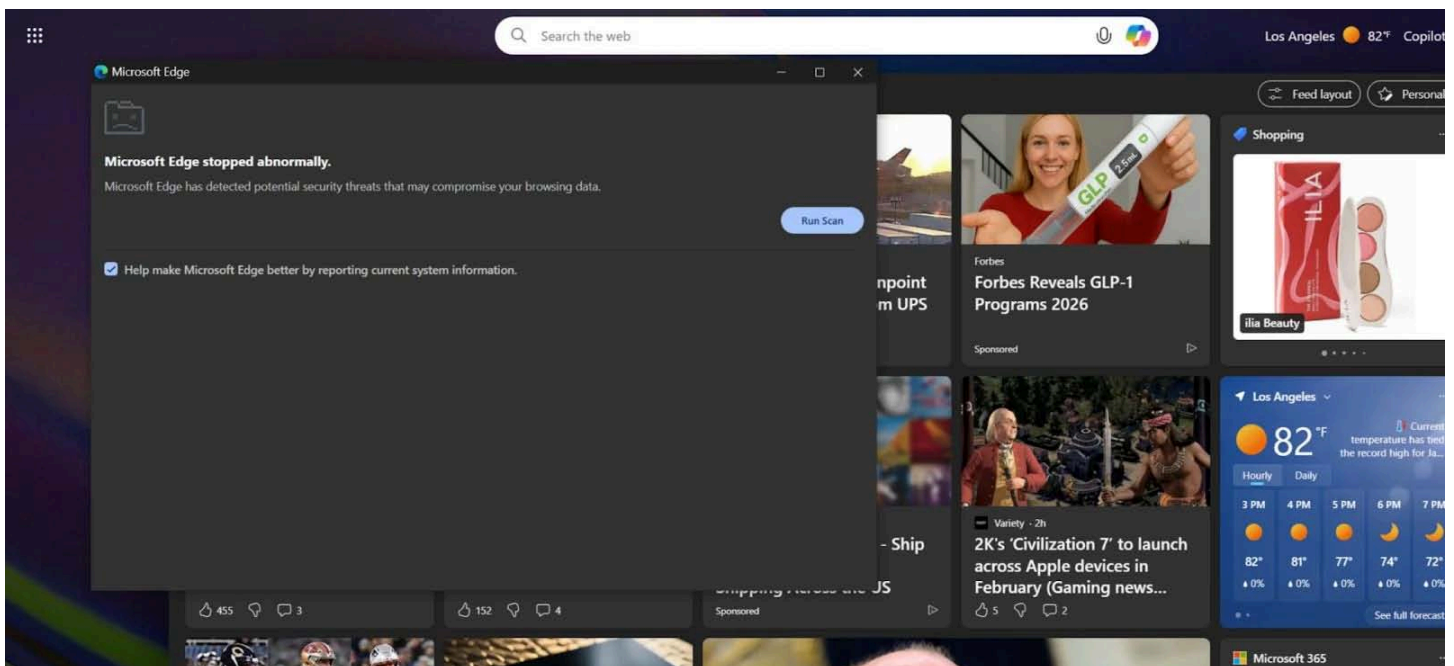
I wanted to finish today's podcast by sharing a newly arrived spin on the "ClickFix" attack method we've previously discussed. That's the attack where the familiar CAPTCHA "prove your human" test is extended to ask its victim to please paste the contents of the Windows system clipboard into the "Run..." dialog and press Enter.

In the newer form of this, which its discoverers, Huntress labs, has dubbed "CrashFix", the victim's web browser is made to appear hung, broken or defective (thus the crash).

And as for the Kingonesque "KongTuke", KongTuke is the name Huntress Labs has given to this threat actor they've been tracking for the past year.

Huntress wrote:

In January 2026, Huntress Senior Security Operations Analyst Tanner Filip observed threat actors using a malicious browser extension to display a fake security warning, claiming the browser had "stopped abnormally" and prompting users to run a "scan" to remediate the threats:



"Microsoft Edge stopped abnormally" – Microsoft Edge has detected potential security threats that may compromise your browsing data."

[Run Scan]

[x] Help make Microsoft Edge better by reporting current system information

Our analysis revealed this campaign is the work of KongTuke, a threat actor we have been tracking since the beginning of 2025. In this latest operation, we identified several new developments: a malicious browser extension called NexShield that impersonates the legitimate uBlock Origin Lite ad blocker, a new ClickFix variant we have dubbed "CrashFix" that intentionally crashes the browser then baits users into running malicious commands, and

ModeloRAT, a previously undocumented Python RAT reserved exclusively for domain-joined hosts.

Ironically, the victim was searching for an ad blocker when they encountered a malicious advertisement. The ad directed users to the official Chrome Web Store's NexShield Advanced Web filter.

The deliberate targeting of domain-joined machines suggests KongTuke is after corporate environments where a foothold means access to Active Directory, internal systems, and lateral movement opportunities. Home users on standalone workstations receive a separate infection chain that appears to still be in testing, when we finally got through all the layers, the C2 (Command & Control infrastructure) responded with "TEST PAYLOAD!!!!".

Whether this means non-domain targets are lower priority or the campaign is still being built out, one thing is clear: KongTuke is evolving their operations and showing increased interest in enterprise networks.

So, what are CrashFix and NexShield?

The malicious NexShield app is all the more diabolical by being a fully functioning working clone of the authentic open source uBlock Origin Lite browser extension. So its user will be pleased with the ad and annoyance blocking behavior of the extension they have just found and installed.

But after using their browser for a while, the bogus "Microsoft Edge stopped abnormally" display will appear with its "Run Scan" button. Upon pressing that the user will be presented with a fake "Security issues detected"

alert and instructed to manually "fix" the issue by opening the Windows Run dialog (Win + R), pasting from their clipboard (Ctrl + V), and pressing Enter. The malicious extension silently copies a PowerShell command to the clipboard, disguised as a legitimate repair command. When the user follows these steps, they unknowingly execute the malicious command.

We were not about to blindly paste from the clipboard, so we tried copying the displayed command (edge.exe -fix-browser -hash=...) like civilized malware analysts.



The browser's response? Complete freeze. When your "fix" causes crashes, the name writes itself, say hello ... to CrashFix. Before we go deep diving into how we ended up with a malicious pop-up message, let's take a step back and look at how it got delivered.

You have probably heard the recommendation to install an ad blocker to protect yourself from malvertising, malicious advertisements that deliver malware through legitimate ad networks. Our victim likely just wanted to get rid of annoying ads. Instead, they downloaded a malicious one (NexShield) while searching for an ad blocker for Chrome.

This header falsely attributes the code to Raymond Hill, the legitimate developer of uBlock Origin, and references a non-existent GitHub repository. This tactic exploits the trust users place in well-known open-source projects. The actual uBlock Origin Lite repository is located at <https://github.com/uBlockOrigin/uBOL-home>, not the URL referenced in this malicious extension.

The NexShield extension is almost entirely a clone of uBlock Origin Lite, a legitimate extension by Raymond Hill. The threat actor likely ran a few find-and-replaces to replace every instance of `uBlock` with `nexshield`.

Huntress continues with their analysis of their latest discovery. But for us the takeaway is that the malware community at large has stumbled upon a fundamental security weakness of Microsoft Windows, which is its users' comparatively script-following level of understanding of Windows, when set against Windows' increasing power and sophistication. It's no longer useful to ask what can be done with Powershell and DotNET. The question is "what cannot be done." That specific pairing comes to mind because, while assembling today's podcast, I encountered another new exploit which did exactly that. It used naive user's invocation of Powershell with a command they could not have understood.

Now that we're encountering a proliferation of similar abuses of powerful commands escaping the browser with unwitting users blindly pasting and executing these commands that they did not write and do not understand, it should be clear that this story only has one ending – sooner or later Microsoft will need to step up to protect users from themselves, much as they did with the Mark Of The Web which flags files that were downloaded across a network. Files containing the Mark Of The Web are treated much more cautiously and skeptically by Windows. The system's clipboard needs to be handled similarly. Contents that were sourced by any web browser need to be quarantined.

Like I said, there's only one possible ending to this trouble. Let's hope Microsoft doesn't wait too long before updating Windows with this change. I wish I believed they would act responsibly here. We can hope.

Creating a 3rd-party utility to fix this won't help since it's all the people who would never know about such a utility that need it the most. The only fix is for this to come from Microsoft as an integral part of Windows.

