Security Now! #1042 - 09-02-25

Letters of Marque

This week on Security Now!

• My experience with 'X' vs email. • Google TIG blackmailed to fire two security researchers. • 1.1.1.1 DNS TLS certificate mis-issued. • Artists blackmailed with threats of training AI on their art. • Firefox extended end-of-life for Windows 7 to next March. • Is the renewal of cybersecurity info-sharing coming soon? • Should security analysis be censored due to vibe-coding? • UK versus Apple may not be settled after all. • Another very serious supply chain attack. • Can the software supply-chain ever be trustworthy? • Why did BYTE Magazine die?
• What happens if Google and others go on the attack?

Have you ever wondered whatever became of Microsoft's Clippy ??

"X"

Just a heads up that between last week and this week my blue checkmark on X was taken away. I presume that my yearly Premium subscription just expired, and I'm pleased that I was not automatically charged for another year without my knowledge or permission; that's the right way to do things.

Since its inception, GRC's email system has proven to be a total success. Last Monday afternoon, the email system sent out 18,465 podcast summary emails; each one containing the short bullet pointed episode summary, the picture of the week and a link to the complete show notes. All 18,465 pieces of email were delivered with none bouncing for any reason other than out-of-office auto-reply, mailbox full, or some other unavoidable reason. Sending weekly mass mailings is free and effortless. And between that Monday before our last podcast and this past Sunday afternoon when I'm writing this, I've received 97 pieces of terrific feedback mail from our listeners. By comparison, while I do still receive wonderful picture-of-the-week ideas from a couple of longtime X associates – including this week's wonderful picture – I no longer receive any podcast feedback of value through X; that only arrives by email. This represents a huge change from 10 years ago when 100% of our podcast feedback was from X. Email makes far more sense. And the @ mentions posted to my timeline appear to be generic conversations that don't involve me or the podcast.

I'll continue to post the weekly show notes to X as I have since the world called it "Twitter" and I created the @SGgrc account 15 years ago in 2010. But I don't feel like paying "X" $7 per month for the privilege, because my hope of being able to use email as a single unifying solution in place of all social media has turned out as well as I had hoped. I'm able to send anyone who wishes a packet of next day's podcast information, and I'm able to receive as much feedback as anyone wishes to send of any length and content, without filtering and for no cost to either of us. And I suppose I'm old school – actually, we know I am – but I was never able to get comfortable receiving or replying to someone on X who chose the moniker *"Stinky Bits"* for themselves.

And speaking with feedback, while I was writing this, I received a piece of feedback email with the subject line *"An apology (although it's kind of your fault) and a Thank You."* Our listener, John Wayward, wrote:

> *Steve, I have to apologise for missing the last 3 episodes of Security Now. However, it's kind of your fault. I typically listen on long car journeys each week between London and Plymouth. However, after your repeated recommendations of Project Hail Mary I downloaded the audiobook and have been absolutely addicted to it. I have listened to nothing else.*
>
> *WHAT a book! Thank you so very much for the recommendation. Just...wow. I'm now going to go back and listen to the Martian, albeit in slower time with Security Now mixed in too. I don't know how the movie can possibly match up to the "science the crap out of it" detail of the book but I'm excited for it nonetheless. All the very best, I now have 3 SNs to catch up on! /John*

In years past I would receive feedback like that through X, but I suspect that those listeners have moved, as I and John have, to email. At the end of every Security Now podcast, Leo is always careful and courteous to remind our listeners that they can go to grc.com/mail.htm to subscribe to my weekly pre-show mailings. And whether or not anyone wishes to receive those, after registering your sending email address, you're welcome to address any of your thoughts directly to me by mailing to: securitynow@grc.com.

Security News

Fire those two, OR ELSE!!

Okay. Here's a weird one. CyberSecurityNews reports that Google is effectively being extorted to terminate the employment of two of its employees ... and the more you learn the weirder it gets. Here's what's being reported:

> *A group claiming to be a coalition of hackers has issued an ultimatum to Google, threatening to release the company's databases unless two of its employees are terminated. The demand, which appeared in a Telegram post, named Austin Larsen and Charles Carmakal, both members of Google's Threat Intelligence Group, TIG. In addition to that, according to a post seen by Newsweek, the self-proclaimed hacking collective which calls itself "Scattered LapSus Hunters," also insisted that Google suspend all investigations by its Threat Intelligence Group into the network's activities.*
>
> *The group's name is an apparent reference to its composition, which it claims includes members from established hacking communities such as Scattered Spider, LapSus, and ShinyHunters. So far, the group has not provided any evidence to substantiate its claim of having accessed Google's databases. Furthermore, there have been no recent confirmed breaches of Google's internal information systems.*
>
> *This threat emerges in the wake of a separate incident disclosed by Google in August. The company confirmed that ShinyHunters, one of the groups allegedly part of the new coalition, had successfully obtained data from Salesforce. Salesforce, being a third-party vendor, provides various services to Google, and the breach occurred within the vendor's systems, not Google's own infrastructure.*
>
> *The formation of a supergroup such as "Scattered LapSus Hunters" would represent a significant escalation in the cyber threat landscape. Scattered Spider is known for its sophisticated social engineering tactics, while LapSus gained notoriety for its aggressive high-profile attacks on major tech companies. ShinyHunters has a long history of large-scale data breaches and selling stolen information on the dark web. The potential collaboration of these entities could pose a formidable challenge to even the most well-defended corporations.*
>
> *Newsweek has reportedly reached out to Google for a statement regarding the alleged threats, but a response was not immediately received as the request was made outside of standard business hours. The situation remains under observation as the tech community awaits Google's official response and further developments.*

At this point it appears that we're just going to need to wait and watch. I cannot imagine that Google could possibly capitulate in any way to this group, or any group, regardless of what the group might have obtained that Google might wish to remain private. If anything, the proper response from Google Threat Intelligence Group would be to dramatically turn up the heat on the various members of the group to cause them to regret ever floating the threat.

1.1.1.1 DNS service certificate misissued

Certificate authorities, our trust in their proper actions, and the chains of trust they anchor is crucial to so much of the operation of the Internet that this podcast has spent a great deal of time through its two decades of reporting examining the protocols, technologies and operation of every aspect of the certificate trust system.

This podcast has followed my fascination with the challenge of certificate revocation. It turns out that once a certificate has been issued it's surprisingly difficult to "un-issue" it due to the way the entire system functions. Over the past 20 years we've followed the industry flip-flopping back and forth as it tries one thing after another. It abandoned the original certificate revocation lists (CRLs) for online certificate status protocol (OSCP) then came up with a better solution for CRLs using Bloom filters and abandoned OCSP over privacy concerns. Who knows what tomorrow will bring?

Given how difficult certificate revocation has proven to be, it's good news that certificates are not often mis-issued. A great deal of time, talent and attention has gone into securing the issuing process. For example, we recently looked at how certificate authorities are now requiring themselves to perform domain control checks of servers from several widely dispersed vantage points to prevent them from being misled by any sort of local attack close to any single point of failure. The lesson here is that the certificate authority industry has gone to great lengths to assure that certificates are never mis-issued.

Given that revocation remains challenging and that mis-issuance is avoided at all costs, the news of three certificates being misissued for as prominent a domain as 1.1.1.1 is both surprising and worrisome. What happened and how? Here's what's been reported so far by ArsTechnica:

> *People in Internet security circles are sounding the alarm over the issuance of three TLS certificates for 1.1.1.1, a widely used DNS service from content delivery network Cloudflare and the Asia Pacific Network Information Centre (APNIC) Internet registry.*
>
> *The three certificates, issued in May, can be used to decrypt domain lookup queries encrypted through DNS over HTTPS or DNS over TLS. Both protocols provide end-to-end encryption when end-user devices seek the IP address of a particular domain they want to access. Two of the certificates remained valid at the time this post went live on Ars.*

That was last Wednesday, September 3rd.

> *Although the certificates were issued four months ago, their existence came to public notice only on Wednesday in a post to an online discussion forum. They were issued by Fina RDC 2020, a certificate authority that's subordinate to the root certificate holder Fina Root CA. The Fina Root CA, in turn, is trusted by the Microsoft Root Certificate Program, which governs which certificates are trusted by the Windows operating system. Microsoft Edge accounts for approximately 5 percent of the browsers actively used on the Internet.*
>
> *In an emailed statement sent several hours after this post went live, Cloudflare officials confirmed the certificates were improperly issued. They wrote in part:*
>
> *Cloudflare did not authorize Fina to issue these certificates. Upon seeing the report on the certificate-transparency email list, we immediately kicked off an investigation and reached out to Fina, Microsoft, and Fina's TSP supervisory body – who can mitigate the issue by revoking trust in Fina or the mis-issued certificates. At this time, we have not yet heard back from Fina.*
>
> *Microsoft said in an email that it has "engaged the certificate authority to request immediate action. We're also taking steps to block the affected certificates through our disallowed list to help keep customers protected." The statement didn't say how the company failed to identify the improperly issued certificate for such a long period of time.*

Representatives from Google and Mozilla said in emails that their Chrome and Firefox browsers have never trusted the certificates, and there was no need for users to take any action. An Apple representative responded to an email with this link to a list of certificate authorities Safari trusts. Fina was not included.

Ah, that's interesting. So for whatever reason, Microsoft and thus Windows has been trusting any certificates issued by this apparently flaky certificate authority whereas Google, Mozilla & Apple all apparently never saw the need. Ars wrote:

It wasn't immediately known which organization or person requested and obtained the credentials. Representatives from Fina, didn't answer emails seeking details.

As I was reading this I had the thought that perhaps Fina should be renamed "fini" and the industry should be done with them. What's curious is why Microsoft appears to be pussy footing around with these clowns. I certainly would not want my Windows OS to be trusting any certificate that Fina might issue, which it currently does. If Google and Mozilla and Apple see no need to trust Fina's certificate signatures I'm certain that I can do without Windows believing anything Fina might assert.

ArsTechnica then continues to provide a bit of additional background, writing:

The certificates are a key part of the Transport Layer Security protocol. They bind a specific domain to a public key. The certificate authority, the entity authorized to issue browser-trusted certificates, possesses the private key certifying that the certificate is valid. Anyone in possession of a TLS certificate can cryptographically impersonate the domain for which it was issued.

We know that's all true, thus the power of a certificate and why the industry goes to such lengths to make sure only the signatures of trustworthy entities are trusted. Ars explains:

The holder of the 1.1.1.1 certificates could potentially use them in active adversary-in-the-middle attacks that intercept communications passing between end users and the Cloudflare DNS service. From there, attackers with possession of the 1.1.1.1 certificates could decrypt, view, and tamper with traffic from the Cloudflare DNS service.

All true. Ars wrote:

Wednesday's discovery exposes a key weakness of the public key infrastructure that's responsible for ensuring trust of the entire Internet. Despite being the only thing ensuring that gmail.com, bankofamerica.com or any other website is controlled by the entity claiming ownership, the entire system can collapse with a single point of failure.

Cloudflare's statement observed: The CA ecosystem is a castle with many doors: the failure of one CA can cause the security of the whole castle to be compromised. CA misbehavior, whether intentional or not, poses a persistent and significant concern for Cloudflare. From the start, Cloudflare has helped develop and run Certificate Transparency that has allowed this mis-issuance to come to light.

The Ars adds:

> *The incident also reflects poorly on Microsoft for failing to proactively catch the mis-issued certificates and allowing Windows to trust them for such a long period of time. Certificate Transparency, a site that catalogues in real time the issuance of all browser-trusted certificates, can be searched automatically. The entire purpose of the logs is so stakeholders can quickly identify mis-issued certificates before they can be actively used. The mis-issuance in this case is easy to spot because the IP addresses used to confirm the party applying for the certificates had control of the domain was 1.1.1.1 itself.*
>
> *The public discovery of the certificates four months after the fact suggests the transparency logs didn't receive the attention they were intended to get. It's unclear how so many different parties could miss the certificates for such a long time span.*

The next day, last Thursday the 4th of September, Cloudflare themselves, the owner of the 1.1.1.1 domain and this the only entity that should be able to issue certificates posted their own piece about this under the headline *"Addressing the unauthorized issuance of multiple TLS certificates for 1.1.1.1"* I'm only going to share the top of their posting but I've placed a link to the entire thing in the notes. Cloudflare summarizes the situation by writing:

> *Over the past few days Cloudflare has been notified through our vulnerability disclosure program and the certificate transparency mailing list that unauthorized certificates were issued by Fina CA for 1.1.1.1, one of the IP addresses used by our public DNS resolver service.*

Okay. But get a load of this! It's much worse than we thought. Cloudflare writes:

> *From February 2024 to August 2025, Fina CA issued twelve certificates for 1.1.1.1 without our permission. We did not observe unauthorized issuance for any properties managed by Cloudflare other than 1.1.1.1.*
>
> *We have no evidence that bad actors took advantage of this error. To impersonate Cloudflare's public DNS resolver 1.1.1.1, an attacker would not only require an unauthorized certificate and its corresponding private key, but attacked users would also need to trust the Fina CA. Furthermore, traffic between the client and 1.1.1.1 would have to be intercepted.*
>
> *While this unauthorized issuance is an unacceptable lapse in security by Fina CA, we should have caught and responded to it earlier. After speaking with Fina CA, it appears that they issued these certificates for the purposes of internal testing. However, no CA should be issuing certificates for domains and IP addresses without checking control. At present all certificates have been revoked. We are awaiting a full post-mortem from Fina.*
>
> *While we regret this situation, we believe it is a useful opportunity to walk through how trust works on the Internet between networks like ourselves, destinations like 1.1.1.1, CAs like Fina, and devices like the one you are using to read this. To learn more about the mechanics, please keep reading.*

https://blog.cloudflare.com/unauthorized-issuance-of-certificates-for-1-1-1-1/

Cloudflare's posting goes on to take its reader through the operation of TLS and certificates and the issuing process. One of the things Cloudflare shares is a list of all the various domains that

were present in these certificates. We see fina.hr, ssltest5, test.fina.hr, test.hr, test1.hr, test11.hr, test12.hr, test5.hr, test6, test6.hr, testssl.fina.hr, testssl.finatest.hr, testssl.hr, testssl1.finatest.hr and testssl2.finatest.hr.

Looking at that list and thinking about that Cloudflare wrote *"After speaking with Fina CA, it appears that they issued these certificates for the purposes of internal testing."* I would bet a month's pay that the reason 1.1.1.1 appeared in any Fina-issued test certificate is for the same reason Cloudflare chose it as the name and address of their DNS service: It's super short, easy to enter and easy to remember. In other words, I would happily wager that at no point was **any** of this misissuance in any way malicious. Some tech geek inside Fina was just playing around with issuing test certificates and the numeric-only IPv4 domain 1.1.1.1 was super quick and easy to enter.

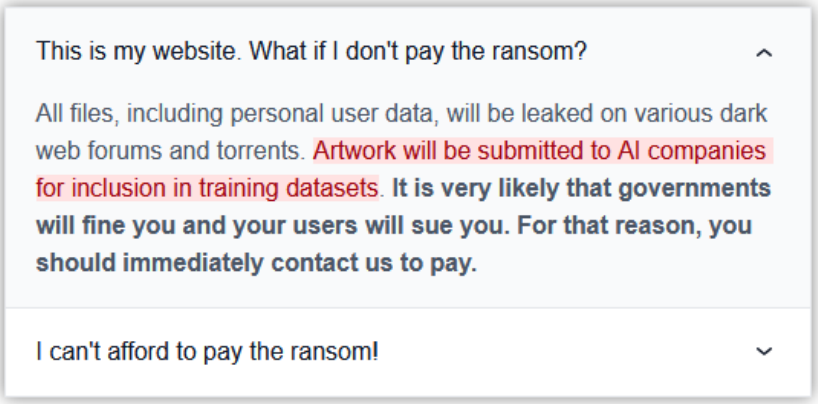
Now, that said, the big no-no was that these test certificates were being signed with the same private key that Microsoft and the EU's Trust Service Provider both trusted. That should never have been done even though Fina says, and I believe them, none of those certificates ever left their control. If someone wanted to experiment with issuing test certificates, they should only ever be signed with an untrusted private internal test private key.

The best thing that came out of this was the wakeup call that the certificate transparency logs that were created for the express purpose of catching exactly these sorts of misissuances were indeed publishing these facts and no one was noticing because no one was checking the logs. It does no good to create these logs if no one is going to check them for important events.

Pay up or we'll submit your proprietary artwork for AI training!
This appears to be the week for wacky events. Get a load of this one. Bad guys have cooked up a new way to extort artists... by threatening to submit their stolen original artwork to AI for training. Unbelievable.

The ransomware group LunaLock compromised a commission-based web platform that connects artists with clients. The group said that if it was not paid a ransom on time, it would share the data with AI companies, thus adding all of the artists' work to massive LLM datasets.

On August 30th, a message appeared on the Artists & Clients website stating that it had been hacked by a ransomware group. One of the website's users noticed the message and shared the news on Reddit. They were redirected to a page with a ransom note, indicating that all the databases and files, including artwork, had been stolen and encrypted. In return for the stolen data, the group is asking $50,000.



So I suppose that the genesis of this was that some bad guys hacked into a not-super-secure site that matches up clients with artists where artists have their portfolios online for perusal and clients are able to commission original works. I commissioned SQRL's logo from such a site.

Firefox to remain alive on Windows 7 until next March

Last Thursday, Mozilla posted *"Extended Firefox ESR 115 Support for Windows 7, 8, and 8.1 and macOS 10.12-10.14"* They wrote:

> *Mozilla has continued to support Firefox on Windows 7, Windows 8, and Windows 8.1 long after these operating systems reached end of life, helping users extend the life of their devices and reduce unnecessary obsolescence. We originally announced that security updates for Firefox ESR 115 would end in September 2024, later extending that into 2025.*
>
> *Today, we are extending support once again: Firefox ESR 115 will continue to receive security updates on Windows 7, 8, and 8.1 until March 2026. This extension gives users more time to transition while ensuring critical security protections remain available. We still strongly encourage upgrading to a supported operating system to access the latest Firefox features and maintain long-term stability.*
>
> *Note that this extension is also applicable for macOS 10.12-10.14 users running Firefox ESR 115.*

I, for one, appreciate this since I'm still spending my days in front of a Windows 7 machine which is working quite well. And I need to confess, or at least update everyone, that since the last version of the Brave browser that supported Windows 7 was released back on January 25th of 2023, so we're coming up on three years ago, after using Brave for a while and appreciating its clear commitment to honoring my privacy, I have returned to Firefox. So Mozilla's announcement that they will be keeping my Firefox 115 there patched for another six months is welcome news.

I'm sure they have telemetry that's informing them that I'm not alone in continuing to run their lovely Firefox web browser on Windows 7. Also, as someone who has written at least my fair share of Windows apps, this whole notion of an app caring all that much about which platform it's running on is a bit overblown. When I opened up the DNS Benchmark source code to begin work on version 2, I discovered that it would run on every version of Windows from Win95 to Windows 11. Microsoft goes to great lengths to **not** break old applications on new editions of Windows. So Windows 95 would have been 1995, right? So that's what, a 30-year span of Windows? All this nonsense about no longer supporting an operating system because the OS platform itself is not supported is, as I said, quite overblown.

But I expect that my Windows 7 machine will be retired by year's end. Everyone watching the videos of this podcast will see my location change as I consolidate two locations into a new third location, and at that time, given how long I keep my cars and my computers – and for that matter my iPads and my Palm Pilots – I fully expect that I will be setting up what will become my final PC. And that one will be running Windows 10.

As I've noted before, the two greatest attack surfaces for our PCs are email – when we make the mistake of clicking on a sour link – and our web browsers, whether we go somewhere malicious directly, or by following a link we received in email. So these days I'm far more glad to be running web browsers and email clients that are being kept up to date than an operating system that was deliberately abandoned by its publisher many years ago.

Renewing the "Cybersecurity and Information Sharing Act of 2015"
We recently noted that private industry had begun withholding information from the Federal government over its concern that the blanket information sharing protections provided by the 10-year long Cybersecurity and Information Sharing Act of 2015 – which would be expiring at the end of this month of September – might not be renewed. This 10-year duration Act allows private sector providers to freely share cyber threat intelligence with government partners under the guarantee of liability protections. But unless this act is renewed by Congress – which has not yet happened – and only 3 weeks remain – the Act's expiration would mean that a vital source of cybersecurity intelligence for our government would dry out overnight.

NextGov offers some interesting background about the Act, its past and the hurdles it currently faces in today's quite messy Washington climate. Under their headline *"House panel advances bill to extend bedrock cyber info-sharing law"* they then lead with the tease *"Some Republicans want to ensure there's language that would prevent the nation's core cyberdefense agency from engaging in alleged "censorship" of Americans' free speech."* They write:

The House Homeland Security Committee on Wednesday [that's last Wednesday] approved a measure that would renew a cornerstone cybersecurity law designed to optimize the exchange of cyber threat information between the private sector and U.S. government. The original law, the Cybersecurity and Information Sharing Act of 2015, lets private sector providers freely transmit cyber threat intelligence to government partners with key liability protections in place. It's set to lapse Sept. 30 unless renewed by Congress. The extension, dubbed the Widespread Information Management for the Welfare of Infrastructure and Government, or WIMWIG Act, extends the law another ten years and now moves to the full House for consideration.

Technical amendments were introduced to the bill, which were met with little pushback in committee. Top of mind for some Republicans on the panel were concerns that the Cybersecurity and Infrastructure Security Agency (CISA) would be enabled to censor Americans' protected speech. That concern extends to the Senate Homeland Security Committee, where Chairman Rand Paul, a First Amendment hawk, has said he would add language in the Senate's version of the reauthorization that would bar CISA from carrying out alleged censorship of free speech.

CISA has faced mounting Republican criticism over allegations of censorship tied to its efforts to combat election-related disinformation in and around 2020. GOP lawmakers contend this amounted to unconstitutional government pressure on private companies to suppress speech, particularly conservative viewpoints.

In the early 2010s, legislative efforts to establish a cyber threat information sharing framework had been underway for several years but faced major hurdles amid public skepticism over government privacy abuses following Edward Snowden's 2013 global surveillance disclosures.

That view shifted after the Office of Personnel Management suffered a massive data breach in 2015, which compromised the personal information of over 21 million current and former federal employees. This galvanized support for the law as it stands today.

Stakeholders say the liability protections in the data-sharing law are critical because they shield companies from lawsuits and regulatory penalties when sharing cyber threat indicators with the government. Oftentimes, cyber threat data includes specific names of individuals or sensitive business information, depending on what hackers target.

Robert Mayer, US Telecom's senior vice president of cybersecurity and innovation said: "By reauthorizing the law, this bill preserves the trusted framework that enables industry and government to share critical threat information quickly and securely. For the telecomm sector, where our networks are on the front lines of cyber defense, this legislation is essential to protecting the infrastructure Americans depend on every day."

The renewal of this bill, largely as written, would appear to be of vital importance, even today more than ten years ago when it was first passed. So I expect that within a few weeks we'll be observing its passage through both houses of Congress and that President Trump will then sign it into law – which will be good for everyone.

Trend Micro examined the impact of Generative AI on security vulnerability reporting
Trend Micro decided to test whether today's availability of AI-assisted code generation changes the balance to weigh against the security community's longstanding practice of publicly disclosing and sharing detailed analysis of known attack code, strategies and malware. We know that when proofs of concept are made public, those who might never have been able to create such attacks from scratch themselves are suddenly empowered with the ability to do so. So the very real concern and question is, does Ai's ability to generate code change this equation?

Here's what Trend Micro wrote and did:

Security companies routinely publish detailed analyses of security incidents, making attacker tactics, techniques, and procedures (TTPs) widely known and visible. These reports often provide comprehensive insights into specific vulnerabilities that are or could be exploited, malware delivery mechanisms, and evasion techniques. This transparency is crucial for the cybersecurity community, enabling organizations to understand the evolving threat landscape so they can implement more effective defenses. But it has evolved into a double-edged sword.

The benefits of detailed security publications significantly outweigh the risks:

Providing defenders with up-to-date TTPs enables proactive security measures.
Building awareness across the security community strengthens collective defense.
Developing and improving security platforms and detection capabilities is critical.
Enabling threat intelligence sharing benefits the entire ecosystem.
Supporting incident response teams with actionable intelligence enhances their effectiveness.

It is not a secret that attackers follow security blogs and read posts about them and other threat actors. The Conti leaks – which is to say leaks from inside Conti – for example, contain several discussions of such public posts. They often learn from these posts about the defenders' techniques and use this information to evolve and improve.

To test whether our industry's practice of providing detailed TTPs enables the creation of malware itself, we conducted an experiment. Using Trend Micro's published analysis of the Earth Alux threat actor espionage toolkit, we attempted to recreate similar capabilities using AI-assisted "vibe-coding".

In this experiment, we used Claude AI (Claude-4-Opus) in combination with Visual Studio Code and Cline. The platform readily began generating code that simulated the attacker's communication patterns, including the emulation of a first-stage backdoor, persistence mechanisms, and the attacking components.

Upon initial inspection, this approach was very straightforward. Bypassing antimalware creation guardrails was quite trivial using a number of uncensored models readily available in platforms like Hugging Face. The first version was generated in Python, and for the sake of flexibility, we regenerated the code in C.

Did it resemble the original code? Yes. But only up to the point of how much was disclosed in the actual blog post. More importantly, while even more detailed technical reports help large language models (LLMs) generate even more accurate code, the generated code was not perfect. For most security reports, some level of skill and understanding will still be needed to finalize the code into a working tool. The AI provides a significant starting point, but technical expertise remains essential for creating functional malware.

Having highlighted the risks of AI-assisted code generation, we must also consider how this capability muddies the waters for attribution. The ability to directly copy malware characteristics described in security reports creates significant challenges for threat hunters and investigators. Attribution has always been challenging in cybersecurity. Attackers have long employed various techniques to confuse investigators, such as:

- *Software component reuse: Using tools associated with other groups*
- *Infrastructure reuse: Deliberately using domains and hosts linked to other threat actors*
- *TTP mimicry: Copying the operational patterns of other groups*
- *False flags: Deliberately adding misleading artifacts, such as North Korean APT groups adding Russian-language artifacts to their binaries*
- *"Living-off-the-land" techniques: Using only tools available on target machines*

With vibe-coding tools, creating copycat campaigns becomes significantly easier. Even nonprogrammers can build a somewhat functional code by providing simple text instructions.

This democratization of malware development poses new challenges for defenders.

We've already observed that some APT groups have become early adopters of AI and LLM technologies. This trend will likely accelerate as vibe-coding tools — which support the rapid prototyping of software based on textual descriptions — continue to evolve. Two key issues with security blogs in the world of AI today are:

1. *They ease the process for attackers to copy the techniques of other groups and quickly get up to speed.*
2. *They muddy attribution efforts when analysts rely solely on TTPs or indicators of compromise (IoCs).*

Should threat publication stop?

No. Threat publication is more critical than ever, but it does need to adapt. For our fellow defenders and the broader industry, this means:

Criminal adoption of AI complicates cybercrime defense. Our industry needs to be more active than ever in educating and supporting our readership. LLM-generated code from blogs is a good start for an attacker, but it's not perfect. Publishers should factor in the ways that LLMs could be possibly misused during the publication processes and test how their detailed descriptions might be exploited.

In other words, while authoring and publishing technical security research keep in mind that AI will almost certainly also be ingesting the research.

> *Vibe code copying muddies attribution, although this only applies for those with a primitive view of attribution based solely on TTPs or IoCs. We recommend best practices using more sophisticated attribution which must evolve beyond simple indicator matching.*
>
> *Publications remain essential. Security publications are a side effect of the research developed to enable leading security platforms to defend. Those same platforms will always be the first line of defense. The knowledge sharing that occurs through these publications strengthens the entire security ecosystem.*

In other words, while it's true that the bad guys will gain, so, too, will the good guys. There's an argument to be made that while AI narrows the gap in who gains more, the security industry remains the net benefactor.

> *Vibe coding (or vibe programming) represents a paradigm shift in software development through AI-assisted code generation. This approach significantly simplifies and speeds up the process of writing executable code, removing barriers for nonprogrammers. However, as we've demonstrated, it also empowers "prompt kiddies" (individuals without deep technical knowledge) to misuse the technology for wrongdoing.*
>
> *In this article, we've only scratched the surface of the use and possible abuse of vibe-coding generative tools for malicious purposes. The evolution of these tools creates new challenges for threat hunters and defenders. Simple and blind correlation of attacks by matching TTPs will no longer be effective. Defenders will need to embrace leading methods of attack clustering and attribution based on the attackers' intentions, objectives, and targeting.*
>
> *Shifting attribution techniques to focus on the primary objectives of the attacker might make it harder for attackers to plant false flags. That said, security is always a cat-and-mouse game, and with every step forward, we evolve into another cycle of innovation and adaptation.*

"UK vs Apple" may not yet be resolved

Some reporting I encountered after I recounted a Tweet made by the current U.S. Director of National Intelligence, Tulsi Gabbard, which, in all fairness, didn't really provide anything more than a somewhat nondescript boast, appears to have been incorrect. At least according to some newer information. Unfortunately, I don't maintain a subscription to the Financial Times, but it is a source of credible information. So all I can do, again, is share what I've found, which reads:

> *The fight between Apple and the UK government over lawful access to iCloud user data has **not** been resolved, despite media reporting last week. The Financial Times **this** week reported on documents filed with the Investigatory Powers Tribunal (IPT), an independent judicial body that examines complaints about UK intelligence services.*

We knew that Apple was going to appeal to the UK's Investigatory Powers Tribunal, so that's not news. But perhaps the fact that this is still going on, at least as of last week, means that the issue is not yet as resolved as we may have believed. The reporting continues:

Back in January, Apple was provided with a government order known as a Technical Capability Notice (TCN). The Financial Times now suggests the TCN required Apple to provide broad access to iCloud data, including messages and passwords: "The obligations included in the TCN are not limited to the UK or users of the service in the UK; they apply globally in respect of the relevant data categories of all iCloud users," the IPT filing adds.

Okay. But that's not news either. We've known all that. Again, the point might be that the Financial Times is reporting on a leak from the supposedly secret Tribunal. The reporting about this finishes with the sentence:

So despite what Tulsi Gabbard says on social media, this is still a live issue.

As I said, I'm now thoroughly confused. So my intention here was to take back what was reported a week or two ago. I don't have any first hand or even second hand knowledge of what Tulsi Gabbard may or may not know. So perhaps our takeaway should be to take Tweets for what they are while remembering what they are not, which is anything definitive and actionable.

Another very serious supply-chain attack

I want to offer another example to illustrate just how vulnerable the open-source software system which has evolved is to malicious abuse. The announcement of what one developer discovered was posted on Substack with the title: *"We Just Found Malicious Code in the Popular Error-Ex NPM Package"* The developer wrote:

Before you read any further, go to the website for the npm package error-ex. Look at the number of weekly downloads. You will likely see a number north of 47 million.

See index.js line 9 here: https://www.npmjs.com/package/error-ex?activeTab=code

This isn't a headline-grabbing framework like React or Express. It's a tiny, one-line utility package, buried deep within the dependency trees of tens of thousands of projects across the globe. We know that because the builds of those projects cause this tiny one-liner package to be downloaded more than 47 million times per week!

This is the kind of package you inherit without ever knowing it exists. And for a short period, it was compromised, turning its massive reach into a ticking time bomb for a significant part of the JavaScript ecosystem.

A single line of malicious code in a package this ubiquitous has a blast radius that is difficult to comprehend. It has the potential to compromise CI/CD pipelines, production servers, and the laptops of developers at countless companies, from small startups to Fortune 500s.

For us, this global threat did not announce itself with a bang. It started with a whisper: a cryptic build failure in our pipeline. The error was ***ReferenceError: fetch is not defined****.*

Our investigation into the build failure took a dark turn when we traced it back to this tiny, trusted dependency.

Our package-lock.json clearly specified we were using the stable version 1.3.2. However, by running npm ls error-ex in our build environment, we found that version 1.3.3 was being installed. This version had been published just a day earlier.

Curiosity turned to alarm when we inspected the code of version 1.3.3. While version 1.3.2 was a single, clean line of code, the new version contained this JavaScript:

```
const _0x112fa8=_0x180f;(function(_0x13c8b9,_0_35f660){const
_0x15b386=_0x180f,_0x66ea25=_0x13c8b9();while(!![]){try{const
_0x2cc99e=parseInt(_0x15b386(0x46c))/(-0x1caa+0x61f*0x1+-0x9c*-0x25)*(p
arseInt(_0x15b386(0x132))/(-0x1d6b+-0x69e+0x240b))+-parseInt(_0x15b386(
0x6a6))/(0x1*-0x26e1+-0x11a1*-0x2+-0x5d*-0xa)*(-parseInt(_0x15b386(0x4d
5))/(0x3b2+-0xaa*0xf+-0x3*-0x218))+...
// ...many more lines of unreadable code
```

This is heavily obfuscated code, designed to be unreadable. But buried within the mess was a function name that made our blood run cold: "check ethereum w".

The attacker had injected malware into the package, very likely designed to detect and steal cryptocurrency from the environment it was running in. The fetch call that was breaking our build was probably the malware attempting to send stolen data to the attacker's server. Our build failed simply because our Node.js version was old enough not to have a global fetch function. In a different environment, the attack could have gone completely unnoticed.

The posting continues to talk about this further, but everyone gets the idea.

I went over to NPM to check out the error-ex package. And in the first place, sure enough, one week in June the error-ex package was downloaded more than 64 million times. Currently it's being downloaded 47,177,455 times per week.

The error-ex package was first released 10 years ago. It went through a series of post-release updates and quickly stabilized, now having a grand total of 16 releases. Its second-to-last release was 9 years ago and the current 1.3.2 release, that everyone is using, is now 7-years old. So it hasn't changed a byte in the past seven years.

But then, someone somehow managed to maliciously replace that 1.3.2 release with a bogus 1.3.3 release and the dependency managers saw that a newer release had become available and grabbed it so that anyone building anything that used it would have the latest and the greatest.

And as if the plot were not already thick enough, it gets still thicker because while the package itself has one single dependency, meaning that it relies upon and pulls in one other package named "is-arrayish" this "error-ex" package is directly depended **upon** by 1544 other packages.

That means that any time any of those packages is rebuilt, as we know happens between 47 million and 64 million times per week globally, the repository for this error-ex package would be queried and the latest and greatest release – meaning the malicious version 1.3.3 – would be obtained and incorporated into the newly rebuilt whole.

At that point that newly rebuilt whole would have unknowingly and unwittingly incorporated malicious code that apparently looks for and steals its developer's Ethereum cryptocurrency. And as the developer who found this wrote, anyone using a newer release of node.js would have had a successful build and would be – and might well still be – completely unaware that their library or application now contains malware.

We can only be glad that scouting around for Ethereum is all this thing did because it could have been far more malicious. And that, of course, takes us back to the recent reporting that the Pentagon is actively using open source software that might be modified maliciously at any time. We don't know how many of those 1544 packages may have been rebuilt during the window of opportunity while this malicious 1.3.3 was live. But with 47 million per week that would have been an average of 6.7 million downloads per day, so many individual instances of that malware are likely in use right now.

Our takeaway here is that a very nice, open software, package sharing system originally created by and for computer hobbyists has gradually been adopted for serious use all the way up to the Pentagon and everywhere in between. And at no point has anyone really stopped to say "now hold on a second... is this safe?" No one wants it to not be safe because the entire system works so well and is so darned useful. Right up until the point where it takes a critical missile guidance system offline because a Pentagon subcontractor was building their software the way everyone else is these days.

A note of thanks to our listener, Kevin White, whose email about this arrived as I was assembling these show notes Monday morning.

But wait! There's (much) more!

A few hours later I received another note from another listener of ours, Sascha Lopez in South Wales, UK. That feedback contained a link to the larger story. Not one, but **18** extremely popular packages were compromised using this same malware from the same attacker. Listen to the packages and their weekly download counts:

- backslash (0.26m downloads per week)
- chalk-template (3.9m downloads per week)
- supports-hyperlinks (19.2m downloads per week)
- has-ansi (12.1m downloads per week)
- simple-swizzle (26.26m downloads per week)
- color-string (27.48m downloads per week)
- error-ex (47.17m downloads per week)
- color-name (191.71m downloads per week)
- is-arrayish (73.8m downloads per week)
- slice-ansi (59.8m downloads per week)
- color-convert (193.5m downloads per week)
- wrap-ansi (197.99m downloads per week)
- ansi-regex (243.64m downloads per week)
- supports-color (287.1m downloads per week)
- strip-ansi (261.17m downloads per week)
- chalk (299.99m downloads per week)
- debug (357.6m downloads per week)
- ansi-styles (371.41m downloads per week)

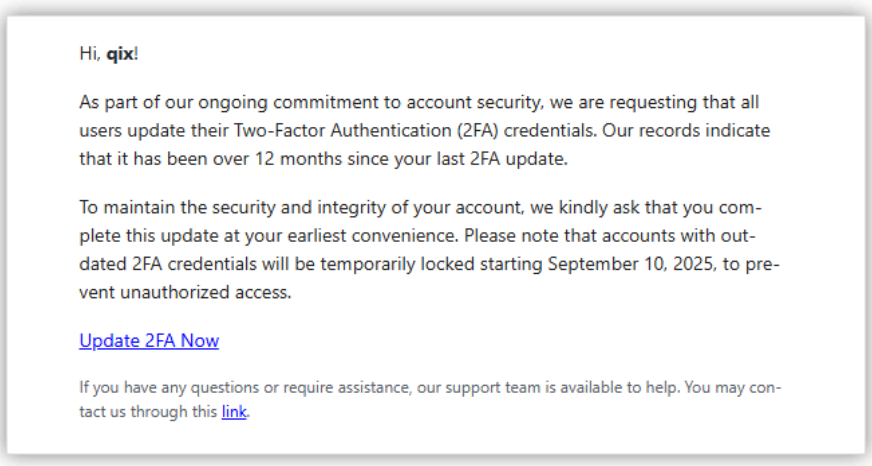
All together, these packages have more than 2 billion downloads per week. They were all maliciously updated to contain code that would be executed on the client of a website, which

silently intercepts crypto and web3 activity in the browser, manipulates wallet interactions, and rewrites payment destinations so that funds and approvals are redirected to attacker-controlled accounts without any obvious signs to the user.

The malware, which has now been fully reverse-engineered, turns out to be extremely sophisticated. It's a browser-based interceptor that hijacks both network traffic and application APIs. It injects itself into functions like fetch, XMLHttpRequest, and common wallet interfaces, then silently rewrites values in requests and responses. That means any sensitive identifiers, such as payment destinations or approval targets, can be swapped out for attacker-controlled values before the user even sees or signs them. To make the changes harder to notice, it uses string-matching logic that replaces targets with look-alike values.

It's extra dangerous because it operates at multiple layers: altering content shown on websites, tampering with API calls, and manipulating what users' apps believe they are signing. Even if the interface looks correct, the underlying transaction can be redirected in the background.

One of the maintainers who was compromised explained that he had been phished:



Hi, **qix**!

As part of our ongoing commitment to account security, we are requesting that all users update their Two-Factor Authentication (2FA) credentials. Our records indicate that it has been over 12 months since your last 2FA update.

To maintain the security and integrity of your account, we kindly ask that you complete this update at your earliest convenience. Please note that accounts with outdated 2FA credentials will be temporarily locked starting September 10, 2025, to prevent unauthorized access.

Update 2FA Now

If you have any questions or require assistance, our support team is available to help. You may contact us through this link.

Note that one of the giveaways of the phishing email is the cut-off date just two days from the notice date. Creating a sense of urgency is one way to get recipients to forget their own safety protocols.

And, get this, the domain from which the email had been sent, NPM themselves, had just been registered three days before, on September 5th. Here's another example of an instance where checking the registration duration of anything we're replying to or relying upon is such a simple thing to do.

So the fate of many hundreds of millions of users of this handful of 18 NPM packages, critically depends upon the package maintainers not falling for basic phishing attacks.

As I started off saying, without ever intending to, and with only ever having the most altruistic and best of intentions, we have slowly over time built not just a house of cards but a massive kingdom castle out of cards. It's a system that we can not stop using, because over time we have become utterly dependent upon it ... yet its security – which is to say its shameful lack of security – really ought to be keeping anyone who is using it up at night.

Everyone has only the best of intentions. Of that there is no doubt. But a familiar old saying may apply here: *"The road to hell is paved with good intentions."*

Listener Feedback

Bill Allen

The mention of BYTE Magazine in SN1041 caused me to remember how I became introduced to SpinRite in the first place. 1988 was the year I upgraded my dual floppy drive IBM/XT clone PC to a 32MB MFM/RLL MiniScribe hard drive. I was desperate to maintain it, optimize it, and otherwise keep it alive. I was also already a subscriber to BYTE magazine. I now remember reading that review article about SpinRite and right away contacted GRC to get a copy. I know I started with SpinRite 1 so that seems about the right time frame.

Anyways, that was a welcome trip down memory lane. I am still using SpinRite today with version 6.1 and have introduced it over the years to new generations of technicians. There is a little sadness in all of this though, it also reminded me of the sudden demise of BYTE magazine in 1998. I remember being rather devastated, since it was my primary source of tech news and views at the time. Here is a bit on that from Tom R. Halfhill, BYTE Magazine senior editor, 1992–1998. https://www.halfhill.com/bytefaq.html

Best Regards, Bill Allen, Crowley, TX | Spinrite 1, SN 1

The link that Bill provided is very interesting to anyone who loved BYTE magazine. The page explains the sequence of events surrounding the end of BYTE Magazine in detail. For example, it mentions our old friend Jerry Pournelle, writing:

After the 1998 shutdown, the BYTE website continued to draw about 600,000 page views a month, even without updates. Obviously, many people still wanted the kind of information BYTE provided. This unrelenting traffic prompted CMP to revive BYTE as a web-only publication in 1999. CMP convinced longtime BYTE columnist Jerry Pournelle to resume his Chaos Manor column on the new Byte.com website, lending some credibility to the effort. However, Pournelle left Byte.com in 2006. The underfunded website lacked BYTE Magazine's breadth and depth of technical content, and it vanished in 2009.

There's much more there, including a detailed FAQ created by, as Bill noted, Tom Halfhill, the senior editor at the time of BYTE Magazine's print edition demise. I invite anyone who might be curious to follow the link in Bill's feedback. And thanks, Bill. Although BYTE magazine might not still be around and able to help, I'm delighted that SpinRite still is and that I have some big plans for its future.

Letters of Marque

One of the interesting and somewhat delicate topics we have touched on from time to time is the question of whether, and if so when, it might be okay for good guys to do things that are not technically legal, but with good intent and for a hopefully good cause and outcome.

In other words, making the world a better place even if the means to do so would mean breaking a few rules along the way.

An early instance of this was way back in the Internet worms era with the likes of CodeRed and Nimda. In those cases, compromised servers were actively scanning for other servers that had not yet been compromised. And the source IPs of those scans were not being and could not be spoofed. As a consequence, security firms who were running honeypots were collecting a comprehensive list of worm-compromised servers since compromised servers were reaching out at random to see whether a not as yet compromised server might reside at some given IP address.

The question then became: Would it be okay for the good guys to use the same now-well-known flaw in Microsoft's IIS server to reach out and proactively and remotely disinfect that machine? A bad worm had infected it. Why couldn't someone who knew where an infected machine was located by its IP address on the Internet reverse that and disinfect it?

I was participating on the conference call with Washington when that idea was floated to the proper person at the Department of Justice at the time. She made it quite clear that doing so would be against the law, plain and simple. And in listening to her carefully, it was clearly not a wink-wink. She was not saying "no" but hoping that we would go ahead and do it anyway. This was not one of those *"let's not ask for permission, we'll ask for forgiveness"* instances.

And since then, there have been many other instances where it is so tempting to allow good guys to remotely fix problems that they almost certainly could. How many consumer routers have been found to be vulnerable? How many random application packages could be patched with the knowledge of a problem **before** its public release? When Plex found a critical remotely exploitable vulnerability in its publicly exposed media server, it could have proactively reached out and fixed it before bad guys were able to use the flaw to install a keystroke logger into a LastPass developer's machine at home and give LastPass the biggest black eye of its life.

In most cases, anything a bad guy can do remotely, a good guy could remotely patch to close the backdoor long before bad guys are given the information they need. But it doesn't happen because it's just as illegal for good guys to hack a network – even if the intention is to help that network's owner – as it is for bad guys to hack the same network. And that's the way things have been since the beginning of all this global networking business.

We're talking about this today because under our current political administration things may be changing. Anyone who's been following U.S. news will likely have heard that President Trump has decided to rename the U.S. Department of Defense the Department of War. That certainly reflects a change in attitudes somewhere. And I was put in mind of all this when I saw a story in CyberScoop carrying the headline *"Google previews cyber 'disruption unit' as U.S. government and industry weigh going heavier on offense"*, and the teaser at the top of their story notes: *"There are still **impediments** to overcome before companies and agencies can get more broadly aggressive in cyberspace, both legal and commercial."*

"Impediments" – I'll say. Like all those pesky laws we were just talking about. Since this could change everything we know about the status quo, I want to share what CyberScoop wrote. They said:

Google says it is starting a cyber "disruption unit," a development that arrives in a potentially shifting U.S. landscape toward more offensive-oriented approaches in cyberspace. But the contours of that larger shift are still unclear, as is whether or to what extent it's even possible. While there's some momentum in policymaking and industry circles to put a greater emphasis on more aggressive strategies and tactics to respond to cyberattacks, there are also major barriers.

Sandra Joyce is the vice president of Google Threat Intelligence Group. [That's TIG who, as we previously noted has recently been threatened with the extortion demand that they terminate the employment of two of their staff.] *She said at a conference* [last] *Tuesday that more details of the disruption unit would be forthcoming in future months, but the company was looking for "legal and ethical disruption" options as part of the unit's work.*

She said at the Center for Cybersecurity Policy and Law event, where she called for partners in the project: "What we're doing in the Google Threat Intelligence Group is intelligence-led proactive identification of opportunities where we can actually take down some type of campaign or operation. We have to get from a reactive position to a proactive one … if we're going to make a difference right now."

The boundaries in the cyber domain between actions considered "cyber offense" and those meant to deter cyberattacks are often unclear. The tradeoff between "active defense" vs. "hacking back" is a common dividing line. On the less aggressive end, "active defense" can include tactics like setting up honeypots designed to lure and trick attackers. At the more extreme end, "hacking back" would typically involve actions that attempt to deliberately destroy an attacker's systems or networks. Disruption operations might fall between the two, like Microsoft taking down botnet infrastructure in court or the Justice Department seizing stolen cryptocurrency from hackers.

Trump administration officials and some in Congress have been advocating for the U.S. government to go on offense in cyberspace, saying that foreign hackers and criminals aren't suffering sufficient consequences. Much-criticized legislation to authorize private sector "hacking back" has long stalled in Congress, but some have recently pushed a version of the idea where the President would issue "letters of marque" like those for early-U.S. sea privateers to companies, authorizing them to legally conduct offensive cyber operations currently forbidden under U.S. law.

Whoa. So, not a "get out of jail free card", a preemptive pardon for anything illegal that might be done in the security interests of the United States. At this point I was curious about these "letters of marque" so I asked the AI Oracle about them and learned the following:

A letter of marque is an old legal instrument from the age of the sail. It was essentially a government license authorizing a private shipowner (a "privateer") to arm their vessel and attack the shipping of an enemy nation during wartime.

Marque means "seizure." The letter granted the holder the right to capture enemy vessels and cargo. The captured ships (called "prizes") would then be brought back to port, condemned in a prize court, and sold — with profits shared between the privateer and the government.

The system blurred the line between the navy and piracy: without a letter, you were a pirate; with one, you were a lawful privateer.

In the U.S. context, the U.S. Constitution (Article I, Section 8) explicitly gives Congress the power "to declare war, grant letters of marque and reprisal, and make rules concerning captures on land and water." So legally, only Congress — not the President — can authorize letters of marque. That means the phrase "Presidential letter of marque" is technically a misnomer. The President cannot independently issue them; Congress would have to approve.

The reason ChatGPT offered that was that, not knowing any better and following from CyberScoop's article which said that *"some have recently pushed a version of the idea where the President would issue "letters of marque"* my questioning prompt to ChatGPT was: *"What is a Presidential letter of marque?"* As we learned, there is no such thing. On the other hand, our experience during President Trump's second term suggests that this President would not let that stop him. Finishing up with ChatGPT's reply, because it's interesting, it said:

The last U.S. letters of marque were issued during the War of 1812. After that, the U.S. Navy became strong enough that privateering wasn't needed. International law (the 1856 Declaration of Paris) abolished privateering among most major powers (the U.S. never signed, but has honored the ban in practice).

In modern discussions, you'll sometimes hear about reviving letters of marque in the context of cybersecurity — for example, allowing private companies to take offensive action against foreign hackers. But that's purely theoretical and would require an act of Congress.

✅ So, to answer directly: A "Presidential letter of marque" would be a government license to privately wage war on behalf of the United States — but under U.S. law, the President alone has no authority to issue one. It would require Congressional authorization.

Except that our current president has shown himself to be quite willing to retest many of the nation's longstanding protocols and assumptions. Our Supreme Court has been unusually busy.

Returning to CyberScoop's reporting, they wrote:

Experts say, that the private sector has some catching up to do if there's to be a worthy field of firms able to focus on offense. John Keefe, a former National Security Council official from 2022 to '24 and National Security Agency (NSA) official before that, said there had been government talks about a "narrow" letters of marque approach "with the private sector companies that we thought had the capabilities." The concept was centered on ransomware, Russia and rules of the road for those companies to operate. Speaking like others in this story at Tuesday's conference, Keefe said: "It wasn't going to be the Wild West."

Joe McCaffrey, chief information security officer at defense tech company Anduril Industries said that the companies with an emphasis on offense largely have only one customer — and that's governments. He said "It's a really tough business to be in. If you develop an exploit, you get to sell to one person legally, and then it gets burned, and you're back again."

By their nature, offensive cyber operations in the federal government are already very time- and manpower-intensive, said Brandon Wales, a former top official at the Cybersecurity and Infrastructure Security Agency (CISA) and now vice president of cybersecurity at SentinelOne.

Private sector companies could make their mark by innovating ways to speed up and expand the number of those operations, he said.

Overall, among the options of companies that could do more offensive work, Andrew McClure, managing director at Forgepoint Capital said: the "industry doesn't exist yet, but I think it's coming."

Brandon Wales, now at SentinelOne said that Congress would have to clarify what companies are able to do legally as well.

But that's just the industry side. There's plenty more to weigh when stepping up offense.

Megan Stifel, chief strategy officer for the Institute for Security and Technology said: "However we start, we need to make sure that we are having the ability to measure impact. Is this working? How do we know?"

If there was a consensus at the conference it's that the United States — be it the government or private sector — needs to be doing more to deter adversaries in cyberspace by going after them more in cyberspace.

One knock on that idea has been that the United States can least afford to get into a cyber shooting match, since it's more reliant on tech than other nations and an escalation would hurt the U.S. the most by presenting more vulnerable targets for enemies. But Dmitri Alperovitch, chairman of the Silverado Policy Accelerator, said that idea was wrong for a couple reasons, among them that other nations have become just as reliant on tech, too.

And "the very idea that in this current bleak state of affairs, engaging in cyber offense is escalatory, I propose to you, is laughable," he said. "After all, what are our adversaries going to escalate to in response? Ransom more of our hospitals, penetrate more of our water and electric utilities, steal even more of our intellectual property and financial assets?"

Alperovitch continued: "Not only is engaging in thoughtful and careful cyber offense not escalatory, but not doing so would be"

This was just one article in CyberScoop. But the consequences of this recent conference captured the attention of the entire industry.

- 3 days ago in Lawfare: *"Google Sharpens Its Cyber Knife"*
- 4 days ago, the publication "Today's General Counsel"s headline was *"Google is Forming a Cyber Disruption Unit"*
- 4 days ago, Oodaloop.com: *"The Perils of Precedent: Could Google's Disruption Unit Invite Retaliation?"*
- 5 days ago, SC Media: *"Google to launch cyber disruption unit"*
- 6 days ago, the Digital Watch Observatory's headline: *"Disruption unit planned by Google to boost proactive cyber defense"*
- 7 days ago in Homeland Security Today: *"Google Previews Cyber Disruption Unit as U.S. Debates stronger offensive measures."*
- Even Tom's Hardware back on August 28th carried the headline: *"Google is getting ready to 'hack back' as US considers shifting from cyber defense to offense."*

From the perspective of someone keeping up on cybersecurity news here in the US, it does feel very much as if we are under continual assault from what we are told are aggressive and hostile state sponsored hackers operating out of Russia and China and North Korea.

We know that hospitals and schools are being attacked, having their networks taken down and their services, whether it be healthcare or education, impacted and interrupted. And we know that our U.S. corporations now live under the constant threat that some well-meaning but momentarily inattentive employee may click on a link they receive in email which results in a network compromise, the exfiltration of the corporation's data, exposure to extortion demands and public humiliation followed by shareholder lawsuits.

So, no way do I or I'm sure anyone, think that public or private entities in the US should indiscriminately attack Chinese, Russian or North Korean institutions or enterprises. That's not us. And I was assuming that was not what anyone was talking about. But there is an area of this that makes me feel somewhat queasy, which is the use of the term "retaliation." That term is being bandied about in Washington policy circles. It's one thing for Google to "disrupt" an adversarial attacker's illegal operation which is attacking us. I'd be inclined to call that "proactive defense". But it's another thing entirely to use the threat of wholesale unfocussed reprisal as a deterrent – which is also being discussed.

The conference that recently brought all this to a head was held by the Center for Cybersecurity Policy and Law (the CCPL) exactly two weeks ago on Tuesday, August 26th. The conference announcement said: *"CCPL will convene cybersecurity leaders from government, industry, and policy disciplines to delve into the core questions raised in the recent CCPL report "To Hack Back, or Not Hack Back? That is the Question … or is it?""*

I've linked to this 7-page PDF in the show notes, and for anyone who is interested in this whole topic, believe me, this report will not bore you. It's quite chilling. Under its section *"Why are we talking about this now?"* the report writes:

> *The arrival of a new Administration and the growing complexity of the cyber threat landscape have reignited discussions around the use of offensive cyber operations. The White House has suggested that such tactics could be a valuable part of the U.S. national security toolkit—particularly to counter cyber threats from China. Proponents highlight major incidents, including the Salt Typhoon and Volt Typhoon campaigns and the recent breach of the U.S. Department of the Treasury, as clear indicators that stronger* ***deterrence*** *measures are necessary to combat cybercriminals and state-sponsored threat actors.*
>
> *Though not a new debate, some senior officials and agencies are signaling renewed interest in expanding offensive cyber capabilities, including potential involvement by the private sector. The U.S. Cyber Command (USCYBERCOM) has emphasized the need for more proactive actions, especially in defending critical infrastructure.* ***The goal is to use offensive cyber tools not just in retaliation but also as a deterrent to prevent future attacks.***

I hope it's clear to everyone that this really changes the game. Maybe it's a good thing. I can't judge. Perhaps officials in China, Russia and North Korea have been laughing at the U.S. and at our quaint Constitution which so often ties our hands and prevents ad hoc retaliation during a fit of pique. If that's been the case historically I would at least imagine that they are likely laughing less loudly with President Trump sitting in the oval office where "fits of pique" appear to be the order of the day. If having President Trump's finger on the button gives them pause, it's not clear to me that's a bad thing.

But to be very clear, what this policy exploration paper is examining is not some Google disruption of a specific targeted foreign criminal enterprise. It's exploring a significant escalation in U.S. cyber posture. The *"Why are we talking about this now?"* section of the paper continues:

> *Pentagon Acting Chief Information Officer Katie Arrington stated her role includes removing policy barriers that limit the Department of Defense's* [Now to be known as the Department of War] *ability to counter adversaries, stressing the need for enhanced* ***offensive*** *capabilities.*
>
> *Similarly, CIA Director John Ratcliffe has expressed support for developing offensive cyber tools and establishing a comprehensive cyber* ***deterrence*** *strategy. Former National Security Advisor Mike Waltz has also endorsed the use of offensive cyber operations to impose greater costs on threat actors like China.*
>
> *Katie Sutton, the nominee for Assistant Secretary of Defense for Cyber Policy, pledged during her confirmation hearing to review National Security Policy Memorandum-13 (NSPM-13), which governs the DoD's authority to conduct offensive cyber operations. Originally issued under the Trump Administration in 2018 and revised by the Biden Administration in 2022, NSPM-13 provides "well-defined authorities to the secretary of defense* [now to be known as the secretary of war] *to conduct time-sensitive military operations in cyberspace," according to a 2020 speech given by Paul Ney, the former DoD general counsel.*
>
> *Congress is also revisiting the role of offensive cyber operations. Although the bipartisan "Active Cyber Defense Certainty Act" introduced in 2019 by Rep. Tom Graves failed to pass the 116th Congress, it has helped revive the debate. The bill aimed to amend the Computer Fraud and Abuse Act (18 U.S.C. § 1030(f)) to grant legal authority for organizations to engage in active cyber defense, including offensive measures, to protect their networks.*

None of us are on the inside in the way these government officials are. So it's not possible to fairly armchair quarterback what they want to do, to judge how much more freedom they need and what they would do with any that they were given. And it does appear that even if President Trump were to issue any letters of marque, they would only be serving as a bridge to where it certainly does appear the country's intelligence and defense agency heads and many legislators want to take us. Sentiment appears to be moving in a cyber-aggressive direction.

And as for cyber as a deterrent, I'm not sure about that. I don't like the idea of any conflict being escalated, whether cyber or conventional. The U.S. has a massive conventional military that we've been relatively restrained in deploying. And it's likely that the knowledge of its potentially overwhelming strength has served as an effective deterrent to those who have ambitions to exercise more of their own lesser power.

The problem is, as military parades appear to attempt to demonstrate, impressive military hardware is visible and can be counted; as can stockpiles of weapons and warheads. It's not clear to me that cyber is at all the same. Can having an impressive cyber capability form a deterrent? I don't see how. Having warheads whose permanent destructive potential is well understood, along with a fail-safe system for their deployment can serve as a powerful deterrent because they do not need to be used to be appreciated.

By comparison, the only way to appreciate a nation's cyber-offensive capability is for it to be used against an adversary. That's not the definition of a deterrent. In that sense cyber capabilities are more like biological weapons which the various super-powers all assume the others have – but no one dares to use. They're not something that can be paraded in the streets or counted in silos; they are simply feared while their existence is adamantly denied.

Following that analogy, "fear" of what such a weapon might do if it were ever to be released serves as the deterrent. So might the various other cyberwarfare nations – China, Russia and North Korea – be fearful of the United States' cyberwarfare capabilities? I have no idea.

This entire area of offensive cyberwar is largely classified, unknown, uncomfortable and unexplored territory whose exploration produces more questions than answers. It is also, as they say, far above my pay grade.

I am much more comfortable exploring browser cookies, certificate revocations and the mechanics of other tangible technologies. But the fact is, conferences like the one that was just held, during which Google announced their formation of a "disruption unit" **are** occurring and as all the other headlines clearly showed, it's big news. So as uncomfortable as it may be, and as many questions as we may be left with, I think we should at least be aware of what's percolating out there on the cyberwarfare front. It may change the way the world is organized.