Security Now! #1007 - 01-07-25

AI Training & Inference

This week on Security Now!

The consequences of Internet content restriction. The measured risks of 3rd-party browser extensions. The consequences of SonicWall's unpatched 9.8 firewall severity. The incredible number of still-unencrypted email servers. Salt Typhoon finally evicted from three telecom carriers. HIPAA gets a long-needed cybersecurity upgrade. The EU standardizes on USB-C for power charging. What? Believe it or not, a CATCHA you solve by playing DOOM. And once we've caught up with all of that: What I learned from three weeks of study of AI.

Security Now!'s 1st Caption Contest

Before we launch into this first podcast of 2025 I want to take a moment to assure everyone that this podcast, named "Security Now!", is not morphing into "AI Now!" I'm quite conscious of the fact that through the end of 2024 and yes, today for this first podcast of 2025, we have and will spend time looking at what's been quietly simmering in the backrooms of university and commercial labs for years. This podcast has, from time to time, veered rather far afield, touching on topics of health, science fiction, the Voyager spacecraft and even homemade portable sound guns. What underpins all of these diversions is the underlying science and technology that makes them go. And in this most recent case of my focus upon and fascination with AI, all of the feedback I've received from our listeners has suggested that this is a topic of interest that is deeply shared.

Over the holidays, during the three weeks we've been apart, I focused upon bringing myself up to speed about what's been going on. I've come away with an understanding of the big picture and I have a number of observations I'm excited to share. So we'll get to that. But I also suspect that this will be it for a while. Eventually, I suspect that fallout from AI research will bear directly upon the security of our software and on the elimination of software flaws. But I suspect we're probably still a year or two away from those more highly focused applications of this very promising and still quite young technology. Given the high level of interest our audience has shown in this, I won't be shying away from it, but neither will we be spending an undue amount of time on it going forward. Essentially, I've satisfied my curiosity and I hope to satisfy yours.

Security News

Unrestricted Internet Content

Questions surrounding restrictions on access to Internet content are controversial and nuanced. They factor in the individual's age and location, the nature of the content and the prevailing government. If ten different people are asked about restrictions on access to Internet content ten different answers will be returned. And where questions of access to Internet content by children arise, even parents and guardians will disagree. But I do know from conversations with many parents of young children – many of whom take time from their lives every week to listen to this podcast – that managing what their kids are exposed to on the Internet is a source of significant concern. The first thing many of our listeners do when setting up a new network at home is choose to use a DNS filtering provider that offers a "family oriented plan" to filter and remove access to the Internet's more unseemly websites.

One place where nearly everyone agrees is that "age appropriateness" is a thing. There is content on the Internet that requires some maturity and perspective to understand correctly. Back in the days before the Internet, a world many of us remember well, our rough age could be determined just by glancing at us. So if, at the tender age of 10 or 11 we were to try to get into a bar or a strip club, those who stood to lose their license to operate the facility would go to great lengths to prevent our entrance. Everyone is familiar with the concept of a "fake ID." The only reason for needing to fake an ID is to enable its holder to do something that the law forbids at their true age.

What's different today is that we have the Internet, and no one knows how old anyone is in cyberspace. Although there can be benefits to this, it's also subject to abuse. And this represents a profound change from the physical world many of us grew up in. Having been born in 1955, I was 34 years old when, in 1989, Tim Berners-Lee came up with the idea for the world wide web. That means there was never a time when a website might ask me to verify that I was at least 18 years old, when that wasn't true – I was nearly twice that age by then. But there's no doubt that with gossip, curiosity and peer pressure being what it is, plenty of today's children who are far short of their 18th birthday are clicking those "You betcha I'm 18!" buttons. It's not my intention

to moralize, and I'm not doing that here. If today's Internet existed when **I** was 14, I have no doubt that I would have been pressing those age verification buttons to see what was behind them ... once I had first bounced my connection through a handful of Tor nodes.

I suspect few parents would disagree that where "age appropriateness" is concerned, a world of difference separates access to the sort of hardcore pornography that's readily available on the Internet from viewing TikTok cat videos. The difference is so stark that the Internet's premiere adult-content website already blocks its access across much of the U.S. Southern states and it just went dark across Florida last Wednesday, taking preemptive action as the Sunshine State's latest legislation went into effect.

But what about those cat videos? I chose this as our first topic of 2025 because, as we start into this new year, more and more states are enacting and have enacted Internet age restriction legislation aimed at the far more benign gray area of modern social media. And much of this new legislation just went into effect last Wednesday, January 1st. While there are increasingly well understood pros and cons to all of this, it seems very clear that the world is beginning to take up questions of the regulation of who is able to access which of the Internet's services, and how.

As this legislation is surveyed, the thing that stands out most is just how random and ad hoc it all appears to be. To give a sense of this, here's a timeline starting with the summer before last:

- July 1st, 2023 - Connecticut SB 3: Requires social media platforms to obtain parental consent before allowing minors to open accounts. Jumping forward a year, we have:

- July 1st, 2024 - Louisiana Act 456: Requires social media platforms to impose limitations and restrictions on certain accounts, implement age verification for account holders, and obtain parental consent.

- September 1st, 2024 (4 months ago) - Texas HB 18: requires digital service providers such as social media platforms to get consent from a parent or guardian before entering into an agreement with minors younger than 18, including to create an account.

- October 1st, 2024 - Maryland Kids Code: requires social media platforms to set default high privacy settings for users under 16, ban the collection of children's data for personalised content, ensure age-appropriate design, implement age verification, and obtain parental consent for younger users.

  Utah HB 464 & SB 194: The Social Media Regulation Act requires parental consent for minors to create social media accounts and mandates age verification by social media companies. It also restricts social media use between 10:30 PM and 6:30 AM for users under 18 without parental consent.

- 1st January 2025 - Tennessee HB 1891: Requires social media companies to verify the age of users attempting to create or maintain accounts. It mandates that platforms obtain parental consent for minors under 18 and enforces stricter privacy and safety measures for these users. The law aims to protect minors from potential online harms by ensuring that social media companies comply with these new regulations.

- 1st July 2025 – Florida HB 3: Requires social media platforms to verify users' ages, obtain parental consent for users under 18, protect minors' personal data, and limit their exposure to harmful content.

Georgia SB 351: Known as the "Protecting Georgia's Children on Social Media Act of 2024," requires social media platforms to implement age verification processes for users, mandates parental consent for minors to create accounts, and restricts social media use in schools.

Minnesota MN HF3488: Sets rules for compensating minors who contribute to online content creation. It requires content creators to keep records and set aside earnings for minors, and it allows for legal action against violators. Also mandates the removal of content featuring minors upon request.

The U.S. Congress also has some legislation that's been floating around since 2023 known as the *"Protecting Kids on Social Media Act"*. Its future is unclear, and I have no idea what position the incoming administration and next congress will adopt on such measures. On one hand there's the politically popular promise of "protecting the children" whereas the flip side is the U.S. Constitution's first amendment guarantee of freedom of speech.

A well-known website featuring adult content greets its visitors with this statement: *"Did you know that your government wants **you** to give your driver's license before you can access [this site]? As crazy as it sounds, it's true. You'll be required to prove you are 18 years or older such as by uploading your government ID for every adult content website you'd like to access. We don't want minors accessing our site and think preventing that from happening is a good thing. But putting everybody's privacy at risk won't achieve that."* It's unclear what would prevent anyone from uploading a photo of someone else's ID, or just synthesizing one from scratch to upload. But the larger point here is to note that there are consequences to this move from the real world to the cyber world and that the unfettered anonymity and freedom we've enjoyed through the first 24 years of the 21st century Internet, may soon be challenged.

It may be that none of this will come to pass. Or that at least if it does, it won't be until its consequences have received significant legal and constitutional scrutiny. In reaction to Florida's new laws, last October the Computer & Communications Industry Association and NetChoice, whose members include the likes of Google and Meta, filed a federal lawsuit challenging the constitutionality of the various restrictions being imposed by the new Florida law. The lawsuit's text stated: *"In a nation that values the First Amendment, the preferred response is to let parents decide what speech and mediums their minor children may access — including by utilizing the many available tools to monitor their activities on the internet."* This feels as though it's headed to the Supreme Court because U.S. legislators are going to need to have some clarification about what they can and cannot require of social media and other companies.

What seems clear today is that these long simmering issues are beginning to come to a boil and that the parents and guardians of minors may soon be put in the loop and at the very least be given the controls they need to allow their households to abide by the prevailing laws of their locality.

But how can this also be done while preserving the privacy of the individual? As I started out saying *"no one knows how old anyone is in cyberspace."* That also applies to you and me, too. No one looking at me today in the physical world would mistake me for a minor. But when any of us connect to any website, there's no indication of any kind how long we've been breathing this planet's air. That's been a freedom we've all enjoyed. So we need to consider what it means for that to change, since that's what we're talking about here. No one would argue that our children need to be protected from harm – even while we're going to need to work out a exact- enough definition of harm to be actionable – but as that notice on the premiere adult content website noted, the ultimate consequence of that may be us needing to affirmatively show that we're **not** minors who are in need of State-mandated protection. How do we do that without sacrificing a great deal of the privacy we currently enjoy?

The Cyberhaven Chrome Extension Compromise

I want to share the following news to highlight the very real threat users of increasingly popular web browser extensions face, which is a compromise of the extension which is then downloaded to or updated by a user's browser. Several times in the past we've talked about the threat of an extension's author abandoning an extension or selling his installed base to an unscrupulous 3rd party.

The other clear and present danger is a deliberate attack on and compromise of an extension's publisher for the purpose of turning an extension malicious. This is what recently happened to the security firm "Cyberhaven" and at least 35 other known Chrome browser extensions that are known to have been compromised as part of a concerted effort. So what happened? Two days after this past Christmas, on December 27th, Cyberhaven posted under their headline *"Cyberhaven's Chrome extension security incident and what we're doing about it"*. They wrote:

Our team has confirmed a malicious cyberattack that occurred on Christmas Eve, affecting Cyberhaven's Chrome extension. Public reports suggest this attack was part of a wider campaign to target Chrome extension developers across a wide range of companies. We want to share the full details of the incident and steps we're taking to protect our customers and mitigate any damage. I'm proud of how quickly our team reacted, with virtually everyone in the company interrupting their holiday plans to serve our customers, and acting with the transparency that is core to our company values.

On December 24, a phishing attack compromised a Cyberhaven employee's access to the Google Chrome Web Store. The attacker used this access to publish a malicious version of our Chrome extension (version 24.10.4). Our security team detected this compromise at 11:54 PM UTC on December 25 and removed the malicious package within 60 minutes.

- *Version 24.10.4 of our Chrome extension was affected.*
- *The malicious code was active between 1:32 AM UTC on December 25 and 2:50 AM UTC on December 26. (So for a little over 25 hours.)*
- *Chrome-based browsers that auto-updated during this period were impacted.*
- *Our investigation has confirmed that no other Cyberhaven systems, including our CI/CD processes and code signing keys, were compromised*
- *For browsers running the compromised extension during this period, the malicious code could have exfiltrated cookies and authenticated sessions for certain targeted websites.*
- *While the investigation is ongoing, our initial findings show the attacker was targeting logins to specific social media advertising and AI platforms.*

Our Response:

- *We notified affected customers December 26 at 10:09 AM UTC*
- *We also notified all other customers not impacted*
- *The compromised extension has been removed from the Chrome Web Store*
- *A secure version (24.10.5) has been published and automatically deployed*
- *We have engaged an external incident response firm for third-party forensic analysis*
- *We are actively cooperating with federal law enforcement*
- *We have implemented additional security measures to prevent similar incidents*

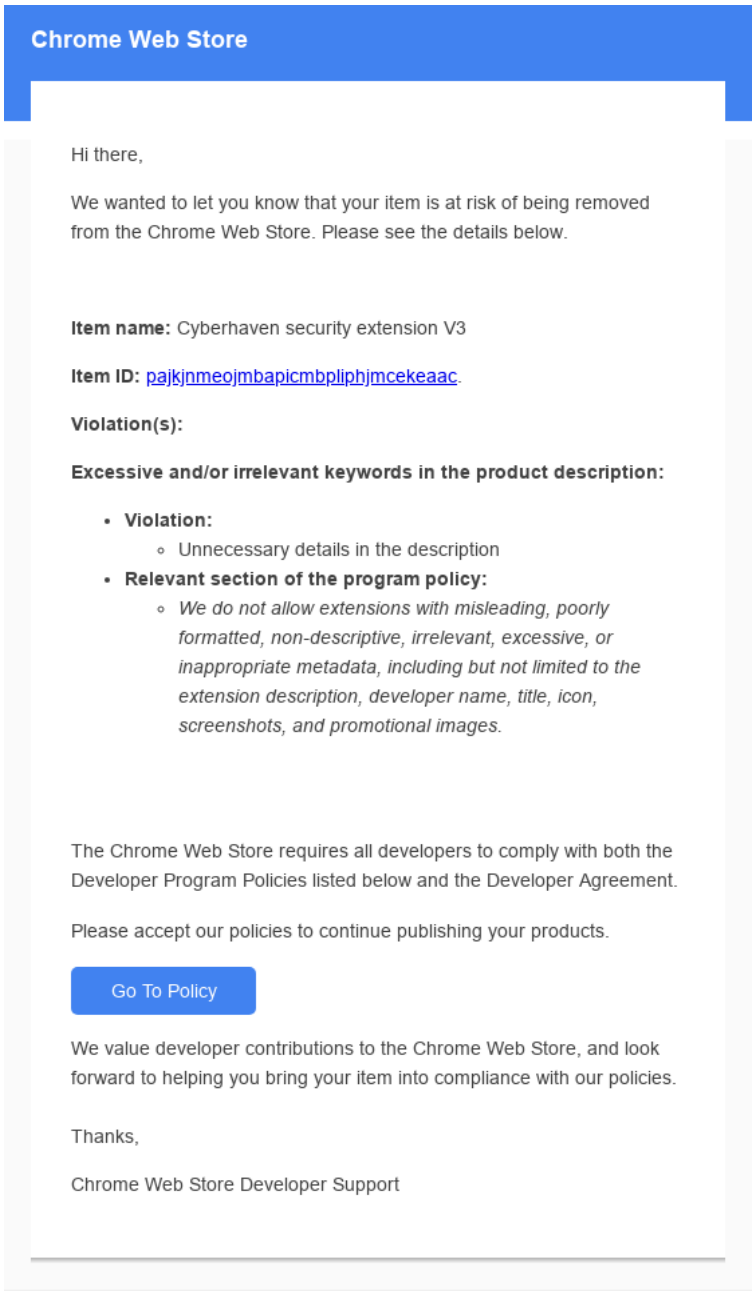
For customers running version 24.10.4 of our Chrome extension during the affected period (December 24-26, 2024), we strongly recommend:

- *Confirm if you have any browsers running the Cyberhaven Chrome extension version 24.10.4 and force an update to version 24.10.5 (currently available in the Chrome Web Store) or newer.*
- *Rotate Facebook personal and business account passwords for accounts on impacted machines.*
- *Review all logs to verify no outbound connections to the attacker's domain or other malicious activity.*

It's good to see that this security firm acted appropriately in every way. They responded immediately. They determined the original attack vector – how the bad guys penetrated their perimeter security – an employee fell victim to a crafted phishing attack. They replaced their compromised extension quickly, verified that this was the extent of the penetration, and notified the public without delay. They fessed up to the mistake and made no attempt to downplay it ... and they did all this on Christmas Day.

It's likely no coincidence that the phishing email attack was launched on December 24th, the day before a span of holiday that was doubtless intended to maximize the period of time the extension's malicious modification would go undetected.

I'd have to say that this particular phishing attack might have caught any developer unaware. The show notes has a snapshot of the perfectly formatted HTML notification that was received:



Chrome Web Store

Hi there,

We wanted to let you know that your item is at risk of being removed from the Chrome Web Store. Please see the details below.

Item name: Cyberhaven security extension V3

Item ID: pajkjnmeojmbapicmbpliphjmcekeaac.

Violation(s):

Excessive and/or irrelevant keywords in the product description:

- **Violation:**
  - Unnecessary details in the description
- **Relevant section of the program policy:**
  - *We do not allow extensions with misleading, poorly formatted, non-descriptive, irrelevant, excessive, or inappropriate metadata, including but not limited to the extension description, developer name, title, icon, screenshots, and promotional images.*

The Chrome Web Store requires all developers to comply with both the Developer Program Policies listed below and the Developer Agreement.

Please accept our policies to continue publishing your products.

Go To Policy

We value developer contributions to the Chrome Web Store, and look forward to helping you bring your item into compliance with our policies.

Thanks,

Chrome Web Store Developer Support

Once the employee clicked on the email, they were taken to the standard Google authorization flow for adding a malicious OAUTH Google application called *"Privacy Policy Extension"*. The trick here was that the employee was actually authorizing the installation of a malicious application. But by naming it *"Privacy Policy Extension"* what was really happening was obscured.

The authorization page was hosted on Google.com and was part of the standard authorization flow for granting access to third-party Google applications. The employee followed the standard flow and inadvertently authorized this malicious third-party application. The employee had Google Advanced Protection enabled and had MFA covering his account. The employee's Google credentials were not compromised.

The attacker gained requisite permissions via the malicious application ("Privacy Policy

Extension") and uploaded a malicious Chrome extension to the Chrome Web Store. After the customary Chrome Web Store Security review process, the malicious extension was approved for publication.

This malicious extension (24.10.4) was based on a clean previous version of the official Cyberhaven Chrome extension. The attacker made a copy of the clean extension and added malicious code to create a new malicious extension. This extension was uploaded to the Chrome webstore and replaced the clean official Cyberhaven Chrome extension. The malicious Chrome extension was now available and distributed to a portion of our customer base.

The Cyberhaven guys reverse-engineered the malicious modification to their extension to identify its behavior. In a subsequent posting they wrote:

> *In our analysis of compromised machines, the extension was targeting Facebook.com users. If the user was logged into Facebook.com and navigated to the Facebook.com website, the extension would execute the malicious code path. Here is what the malicious flow would execute:*
>
> - *Get the user's Facebook access token*
> - *Get the Facebook user's ID*
> - *Get the user's account information (if available) via Facebook API*
> - *Get the user's business accounts (via Facebook API)*
> - *Retrieve user's ad account information (via Facebook API)*
> - *Package all of this information, along with Facebook cookies and the user agent string, and send it to the C&C (Command & Control) server*
>
> *After successfully sending all the data to the C&C server, the Facebook user ID is saved to browser storage. That user ID is then used in mouse click events to help attackers with 2FA on their side if that was needed.*

So the web browser extension attackers were interested in attacking the accounts of any Facebook users whose Chrome browsers might update to the malicious extension before it was detected and removed from the Chrome Web Store. Obtaining a user's Facebook access token cookie allows for full and immediate impersonation of that user.

Another security site, Secure Annex, provided a broader perspective into attackers behind this campaign. By pivoting from the known-malicious Cyberhaven extension, indications of compromise were obtained, which is how we know that many more Chrome web extension developers have also fallen victim to these phishing attacks. The earliest known instance of one of this group's many attacks was last May. So they've been active since at least then. I think it's important for everyone to have some sense for the scope of this. So here's 19 of the compromised Chrome web extensions: **VPNCity** with 10k users, **Parrot Talks** with 40K users, **Uvoice** with 40K users, **Internxt VPN** with 10K users, **Bookmark Favicon Changer** with 40K, **Castorus** with 50K, **Wayin AI** with 40K, **Search Copilot AI Assistant for Chrome** with 20K, **VidHelper Video Downloader** with 20K, **AI Assistant, ChatGPT and Gemini for Chrome** with 4K users. **Vidnoz Flex - Video recorder & Video share** with 6K, **TinaMind, The GPT-4o-power AI Assistant!** with 40K users, **Bard AI chat** with 100K users, **Reader Mode** with 300K, **Primus (previously PADO)** with 40K, **GPT 4 Summary with OpenAI** with 10K, **GraphQL Network Inspector** with 80K, **YesCaptcha assistant** with 200K and **Proxy SwitchyOmega** with 10K.

Every one of those Chrome web extensions was compromised last year, exposing as many as 1,060,000 users of Chrome to malicious browser-side code. The good news here, if there is any, is that the attackers appeared to be focused solely upon Facebook users. But that was this time, and they were certainly willing to go out of their way to compromise those accounts.

It wasn't long ago that we were talking about the move from Chrome's v2 extension manifest to the significantly more limited v3, and how, as a consequence, uBlock Origin won't ever be offered in its full v2 strength version. I'm certain that the Chromium team understands how much value the 3rd-party browser extension ecosystem brings to their Chrome browser. But given this attack campaign as just one example, it's not difficult to see why they would be anxious to curtail the damage that aberrant extensions could do.

And note that none of this is ever about an extension's user doing anything wrong. It was the extension's developers whose account was accessed and abused. So this is another form of supply-chain attack. As users of Chrome, the one thing we can do is practice good browser extension hygiene. That means keeping the set of extensions we're loading and using to a minimum and removing any "dead wood" that might needlessly expose us through that extension's compromise. Every additional extension presents another potential breach entrypoint.

SonicWall vulnerability patching

Back in August, SonicWall, a manufacturer of popular Network Security Appliances (NSAs) revealed a vulnerability in their SSL VPN Firewall product that they rated with a severity of 9.3. However, NIST officially gave it a 9.8 and shortly afterward CISA formally warned of the serious potential for its exploitation. They called it an "SonicOS Improper Access Control Vulnerability" and noted that it was "potentially" being successfully attacked in the wild.

Among the reporting on this, I particularly liked the write up by the security intelligence firm "Field Efect". They wrote:

> *While it's unclear what SonicWall means by 'potentially' exploited, Field Effect can confirm that we have seen an increased targeting of SonicWall firewalls since CVE-2024-40766 was announced on August 23.*
>
> *However, further investigation is required to determine if threat actors are specifically targeting CVE-2024-40766 or other, older, unpatched vulnerabilities. Traditionally, when vendors disclose critical vulnerabilities in edge devices, it draws the attention of threat actors toward the devices in general and that could be what we have observed in relation to the SonicWall firewalls.*

I really appreciate how rational and sober these guys are being. There's no breathless hyperbole here. They finished by noting:

> *SonicWall firewalls are very popular among critical infrastructure industries and corporate environments and are thus frequently targeted by threat actors looking to obtain initial access into networks of interest. According to the Shadow Server Foundation, approximately **400,000** SonicWalls are deployed worldwide, representing a significant potential attack surface for threat actors who possess SonicWall exploits.*

So that was back in August where and when we have an estimated 400,000 Internet-facing SonicWalls with a known remote authentication vulnerability. Where are we today? Two days after Christmas, on December 27th, a Japanese security researcher posted his own update on the state of play with SonicWall devices today.

> *In August 2024, the SonicWall NSA vulnerability CVE-2024-40766 was disclosed. I have found strong indications that the ransomware groups Akira and Fog are still exploiting this vulnerability for unauthorized access. Through my ongoing investigations, I found that, as of December 23, 2024, the number of companies suspected to have been compromised by these two groups via this vulnerability had exceeded 100. In this article, I will share the details of this investigation and highlight the current situation in which at least 48,933 devices* ***remain*** *vulnerable to CVE-2024-40766.*
>
> *Since the vulnerability was disclosed, I have been investigating whether the organizations listed on various ransomware groups' leak sites own SonicWall Network Security Appliance devices. Focusing on the 218 organizations identified as victims of Akira and Fog, I found that over 100 (approximately 46%) were running SonicWall. Considering that the SonicWall NSA ownership rate among organizations victimized by other ransomware groups (excluding Akira and Fog) remains at around 5% or less, this figure is remarkably high.*

In other words, whereas the general rate of overall SonicWall presence among companies who have been breached and listed by ransomware groups other than Akira and Fog is only 5%, the fact that around 46% of the organizations listed by those two ransomware groups are currently exposing a SonicWall device to the Internet strongly suggests that those two groups have successfully designed an exploit for the vulnerability and are working their way through the inventory of still-exploitable and unpatched SonicWall device owners.

This Japanese researcher wrote:

> *I developed a proprietary method to evaluate patch status by examining the HTML structure of SonicWall devices to assess mitigation efforts for CVE-2024-40766. For SonicWall NSA devices with SNMP exposed, it is possible to obtain accurate model and version information. By comparing the results of my custom method with SNMP data from around 5,000 devices, I've confirmed the accuracy of this detection approach.*

He then posted a chart showing the lackluster patch status of these devices. The United States has more than half of the globally deployed SonicWall devices with 238,678 out of an identified 390,474 total. Sadly, of the identified 48,933 currently known vulnerable SonicWall devices, 29,107 are detected as still being vulnerable four months after their publisher’s and CISA’s warning of a 9.8 CVSS exploitable vulnerability.

Something needs to change, and is it any surprise that ransomware continues to be a scourge across the Internet? On the one hand, any company being victimized with their proprietary data exfiltrated and then held for ransom is a crime. But we all know that Internet security can never be one-and-done install and forget. The connection of an internal corporate network to the global public network is incredibly empowering... but with it comes the responsibility of managing the security of that interconnect. To ever take that for granted is to risk everything the organization holds dear.

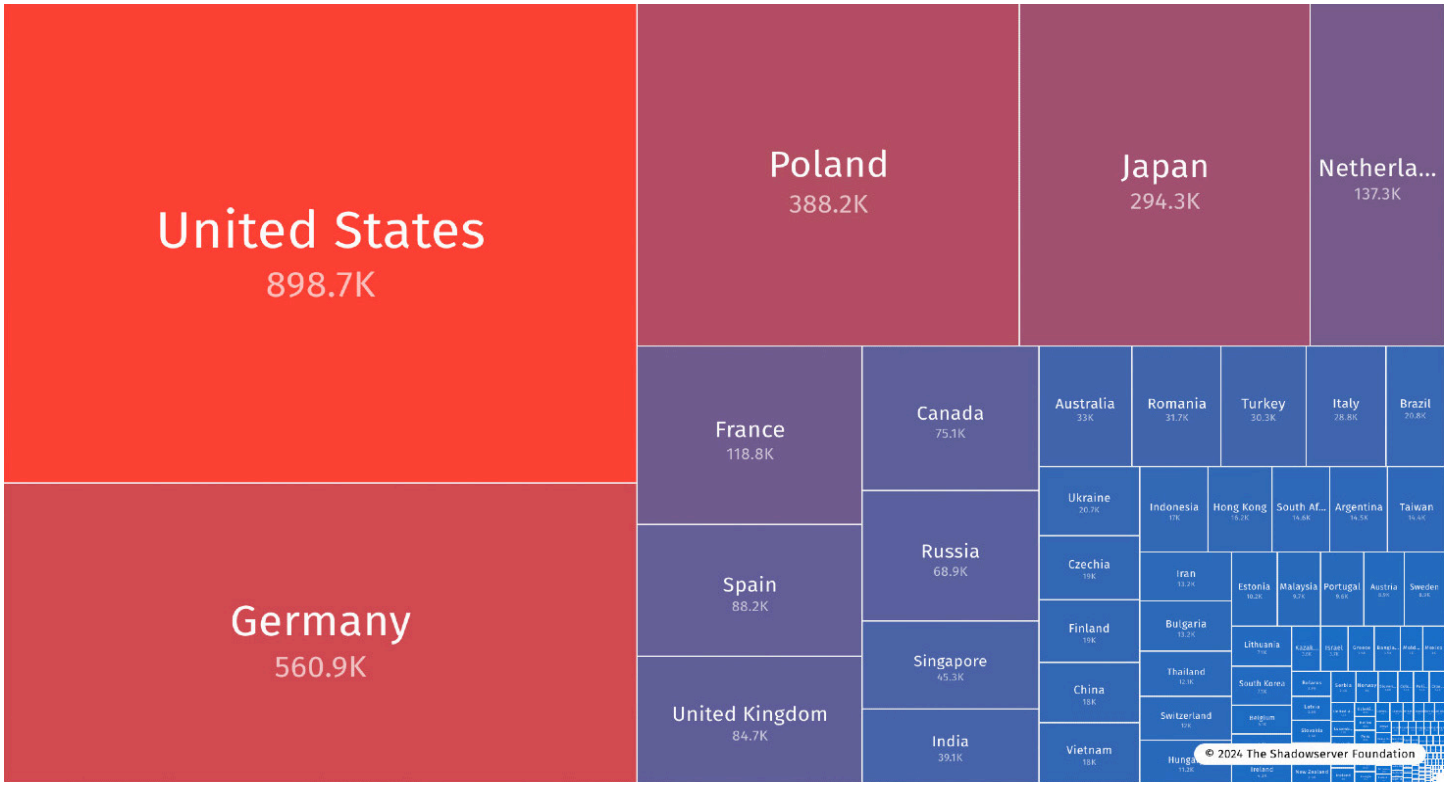
Shadowserver Foundation & eMail Encryption
Speaking of the Shadowserver Foundation, on New Years Eve morning, they posted to their Bluesky.social account:

> *We've started notifying owners of hosts running POP3/IMAP services without TLS enabled, meaning usernames/passwords are not encrypted when transmitted. We see around 3.3 Million such cases with POP3 & a similar amount with IMAP (most overlap). It's time to retire those services!*

This is something we don't talk about often, but it bears reminding everyone. Like the rest of the entire original Internet – meaning the WEB, FTP, DNS and all the rest, electronic mail exchanged over the SMTP, POP and IMAP protocols was not originally encrypted. It was sent over simple unencrypted TCP connections in ASCII plaintext – thus making it all completely readable by anyone tapping into any location, whether near to any sender or receiver – such as by any ISP or wireless hotspot operator – or over the public Internet wherever traffic is moving past.

With inertia being the prevailing force that it obviously is on the Internet, the Shadowserver Foundation reminds us that a sizable portion of email servers have never bothered to move to encryption. You know, no one has ever made them encrypt. Unlike the web with HTTPS where encryption became mandatory, email security has largely fallen through the cracks, even while it has arguably become more important than ever as we depend upon it as our identity authentication of last resort. That means that all of the email these 3.3 million server's send and receive has remained the same unencrypted plaintext that it was 35 years ago. Right now. Today. Those emailed "Oops! I forgot my password" recovery links. The "We just sent you a super-secret 6-digit one time code to authenticate yourself because it's so important" emails.

And lest we imagine that these 3.3 million email servers must be scattered among backwater countries no one has ever heard of and can't spell, the Shadowserver Foundation thoughtfully provided a heatmap showing just where these utterly security-negligent machines are located:



Guess which country leads the pack? Yup. None other than the good old U.S. of A. Within our proud borders lie some 898,700 completely unencrypting email servers. Those nearly 899 thousand email servers are right now, today, this very moment, exchanging email for people who probably have no idea that everything they are sending and receiving is in the clear and readable by anyone who might even be the least bit curious. Takes very little effort. And we know that none of these are people at home since residential ISPs long ago blocked SMTP's port 25 due to rampant spam abuses. So these are organizations of some size who probably think it's super-spiffy to save some money by running their own email while apparently never stopping to think or care that all of the email they are transacting is readable by anyone. I said there were a total of 3.3 million and we've accounted for the U.S. taking the top slot at nearly 899 thousand instances. So there are others. Germany takes the second spot at 560,900 unencrypted email servers. Poland takes 3rd place at 388,200, followed by Japan at 294,300 and then The Netherlands at 137,300.

Having seen these numbers it would be very interesting to know what is going on here. Who are these 899 thousand entities in the U.S. who probably run encrypted web servers with up to date TLS certificates because the world insists upon it, but perhaps never bothered to think about email.

Email servers, like web servers, connect to each other using the TCP protocol. So just like web servers, it's very possible for email servers to add a layer of authentication and encryption by negotiating TLS certificates with each other. This allows them to each verify the other's identity and to agree upon a shared secret key to use for encrypting and decrypting each other's traffic.

The $64,000 question is: How is this ever going to be made to change? Because we know that the phrase "being made to change" is the only way it will ever happen. Web browsers, thanks to the tightly coordinated efforts of the CA/Browser forum, were able to force the entire web to move to encrypted connections by rightfully scaring anyone using a browser that was unable to establish an encrypted connection to a remote web server. At first it was a frightening experience; today one really needs to work at it. Since most Internet users really have no idea what's going on with any of this, and since they pretty much exist on the Internet in a constant state of fear, anything that seems the least bit sketchy will cause them to bolt immediately. Thus it didn't take long for all web servers to obtain TLS certificates.

As we know, this transition to HTTPS everywhere was tremendously aided by the creation of Let's Encrypt and the ACME protocol which automated the issuance and installation of free web server domain validation TLS certificates. Unfortunately, nothing like Let's Encrypt exists for email servers. The ACME protocol is able to verify a server's control over a domain through the presence of a transient signature file located in the .well-known root directory of a web server, or by querying a TXT record with that domain's DNS. But there is no similar direct support for email servers despite there being clear demand for it evidenced within Let's Encrypt feedback forums. There are, however, various workarounds.

All of GRC's email transactions are, of course, encrypted. At the moment, once every year, after I've updated all of GRC's web servers with a new certificate from DigiCert, I need to manually reformulate the certificate from binary to ASCII base64 encoding and install it into GRC's beloved hMailserver. That's a manual process which I don't mind performing one per year. But as, and if, certificates continue their apparently inexorable reduction in lifetime, any sort of manual process will obviously become increasingly problematic. Since I have multiple Windows and UNIX servers that need to be kept synchronized with wildcard domains, this entirely pointless reduction in certificate lifetime will eventually force me to roll my own solution to keep everything running without my intervention.

I've received a great deal of feedback from our listeners who have chimed in with their own issues surrounding shortening certificate lifetimes – and the headaches this is creating for them – for their non-web services. Certificates are used not only for web and hopefully for email, but for many other purposes which are all being ignored. It appears that the CA/Browser forum is being somewhat myopic in their apparent belief that the entire world is the web.

I haven't looked deeply enough into this mess to determine whether it might be possible to delineate the use of short-life certificates for web services where automation is convenient and supported while allowing non-web server TLS certificates to remain reasonably multi-year. Alternatively, since we know that web browsers are able to – and have said they would be – eventually independently rejecting any certificate having an out-of-spec total lifetime, everything could be left as it is with web browsers being the sole enforcers of short-lift web certificates.

But I've wandered well off course now. My point is, without some means of enforcing the use of TLS certificates for email, history shows us that nothing will ever move these recalcitrant email servers to encryption. If they don't see any problem today why would they ever? The only obvious mechanism for forcing this change would be for those web servers that do support encryption to refuse to accept any insecure email connections. Out of fear of missing anyone's important email, I've historically configured GRC's email server to accept unencrypted email over port 25 while offering to dynamically upgrade the connection to full security with STARTTLS, which is an SMTP command that allows cooperating email servers to add encryption over a traditionally unencrypted port. Now I'm thinking that perhaps it's time to end that practice, since port 25 has largely become the domain of spammers.

But for those 3.3 million unencrypted email servers in the world, nearly 899 thousand of which are in the United States, before they're going to be able to move to encryption they're going to need some means of obtaining reasonably priced and reasonably maintained TLS certificates. That doesn't exist today for small independent servers.

Salt Typhoon Evicted

Following up on the news that the Chinese-backed advanced persistent threat group known as "Saly Typhoon" had infiltrated all telecom providers, three US providers – AT&T, Verizon, and Lumen – all say they've now evicted Salt Typhoon from their networks. After this widespread and frighteningly successful hacking campaign came to light CISA suggested that we should not be relying upon the security of telecom carriers and should, instead, add the strong encryption provided by 3rd-party apps such as Signal. In the aftermath of these attacks, remaining with CISA's recommendation would seem prudent when holding any sensitive conversations.

HIPAA gets a cybersecurity update

On December 27 the U.S. Department of Health and Human Services (HHS) issued a Notice of Proposed Rulemaking (NPRM) to modify HIPAA — the aging Health Insurance Portability and Accountability Act of 1996. The HIPAA regulations will be getting a bunch of new, welcome and needed cybersecurity rules including the mandatory use of encryption, multifactor authentication, network segmentation, vulnerability scanning, and more.

Miscellany

USB-C becomes EU standard

Under the label of true miscellany, I wanted to mention in passing that the EU, apparently having nothing more pressing to legislate at the moment, has taken the time to establish USB-C as the official common standard for charging electronic devices throughout the European Union. There's actually an official document bearing the headline: "One common charging solution for all" In part, the official EU legislation reads:

> *The Commission promotes solutions that favour technological innovation in electronic device charging while avoiding market fragmentation. The voluntary approach did not meet consumer, European Parliament or Commission expectations, so we put forward a legislative approach. The common charger will improve consumers' convenience, reduce the environmental footprint associated with the production and disposal of chargers, while maintaining innovation.*

In other words, the market didn't settle into any sane and rational standard by itself, so we're going to impose some legislation here.

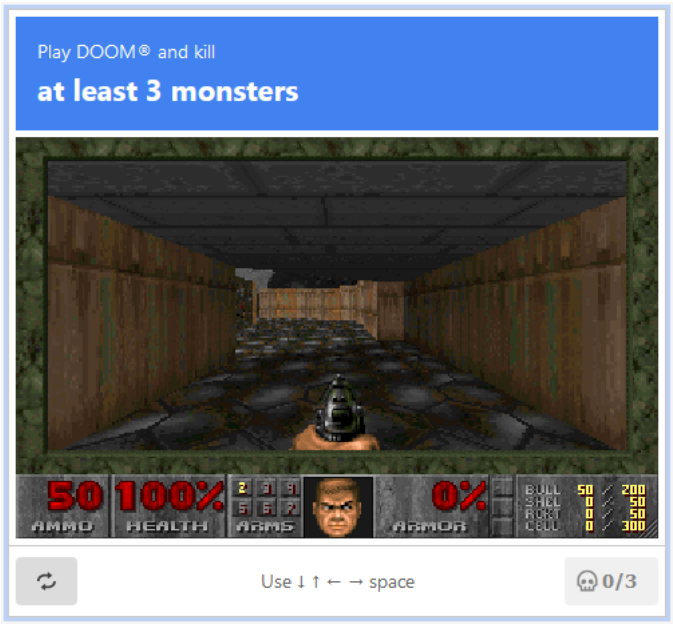
> *The 'common charging' requirements will apply to all handheld mobile phones, tablets, digital cameras, headphones, headsets, portable speakers, handheld videogame consoles, e-readers, earbuds, keyboards, mice, and portable navigation systems as of 28 December 2024. These requirements will also apply to laptops as of 28 April 2026. Such transition periods will give industry sufficient time to adapt before the entry into application. The main elements are as follows:*
>
> *A harmonised charging port for electronic devices – USB-C will be the common port. This will allow consumers to charge their devices with any USB-C charger, regardless of the device brand.*
>
> *Harmonised fast charging technology – Harmonisation will help prevent different producers from unjustifiably limiting charging speed and will help to ensure that charging speed is the same when using any compatible charger for a device.*
>
> *Unbundling the sale of a charger from the sale of the electronic device - Consumers will be able to purchase a new electronic device without a new charger. This will limit the number of chargers on the market or left unused. Reducing production and disposal of new chargers is estimated to reduce the amount of electronic waste by 980 tonnes yearly.*

Wow. 980 tons worth of chargers eliminated. No more drawers full of unneeded, unwanted unused and forgotten chargers. So before long those in the EU will be spared the experience of opening the box and thinking: *"Oh shoot, not another damn charger!"*

They did note that since the wireless magnetic induction charging market is, so far, behaving itself and is not showing undue fragmentation, they did not feel the need to impose any order there. But that market, too, might need some harmonization if things start going all wild and wooly! So they're keeping a watchful eye on it.

The "DOOM" Captcha

Yes, that's right. Since nobody likes CAPTCHAs, an enterprising software engineer has created a DOOM-based CAPTCHA system where you have to kill at least 3 bad guys in the DOOM video game to proceed to a website:



https://doom-captcha.vercel.app/

Its author wrote: https://grc.sc/doom

> *A CAPTCHA that lets you play DOOM® to prove you're human (for educational and entertainment purposes) The project works by leveraging Emscripten to compile a minimal port of Doom to WebAssembly and enable intercommunication between the C-based game runloop (g_game.c) and the JavaScript-based CAPTCHA UI. Some extensions were made to the game to introduce relevant events needed for its usage in the context of a CAPTCHA. Started out with a minimal, SDL-based port of Doom that can be efficiently compiled to WebAssembly then tweaked the build to make it compatible with the shareware version of wad (doom1.wad) for legal use.*

So it actually is a working CAPTCHA which supports the official JavaScript CAPTCHA API. I'm no video gamer, so was promptly killed right off the bat while I was working out the arrow keys and spacebar for movement and firing. But it's also not difficult to kill three baddies since I was able to pull that off on my second try. Since I imagined that some of our listeners might enjoy seeing what this is all about, I gave it an easily remembered URL of: grc.sc/doom

AI Training & Inference

As I said at the top of today's podcast, Security Now! Will not be evolving into the "AI Today" podcast. But that said, aside from the fact that the recent truly astonishing advances in AI are going to directly impact everyone's lives outside of the security sphere, I'm also very certain that we're going to be seeing AI's impact upon the security of our software and operating systems – and we may not need to wait for long. So over the course of the next few years the topic of AI will be reemerging in security.

Our listeners have been following my journey through this topic, which has not been in a straight line. More than anything else, I endeavor to be an honest researcher. An honest researcher will readily revise their entire belief system, as required, when presented with new facts and information. Clutching to obsolete dogma simply because it's familiar and comfortable is not the way of science. It was because I was puzzled and confused by what I was experiencing firsthand that I went searching for that information. I believe I have found it, I understand it – at least as much as is possible without actually implementing it. And I have been changed by it.

Three weeks ago I said that I might have something to say about this before we met again today and that, if so, I would probably enjoy sharing that with this audience with a special email over the holidays. That possibility induced more than 1100 of our listeners to take the time to subscribe to GRC's Security Now! mailing list. So for that reason alone, due to that declaration of interest, I felt that I had to say something. Today, I have much more to say on the topic than I did nine days ago, last Monday, December 30th. But let's start with what those 15,060 subscribers received from me last week, then I'll expand a bit on what I think are the most important points and what I've continued to learn since. I wrote:

When I first set about writing this email, my plan was to share what I had learned during the first half of our 3-week hiatus from the podcast. But it quickly grew long (even longer than this) because I've learned quite a lot about what's going on with AI. Since I suspect no one wants to read a podcast-length piece of email which I would largely need to repeat for the podcast anyway, I'm going to distill this into an historical narrative to summarize a few key points and milestones. Then I'm going to point everyone to a 22-minute YouTube video that should serve to raise everyone's eyebrows.

So here it is:

- *Everything that's going on is about neural networks. This has become so obvious to those in the business that they no longer talk about it. It would be like making a point of saying that today's computers run on electricity. (Duh!)*

- *AI computation can be divided into "pre-training" and "test-time" (also called "inference-time"). Pretraining is the monumental task of putting information into a massive and initially untrained neural network. Information is "put into" the network by comparing the network's output against the expected or correct output, then back- propagating tweaks to the neural network's vast quantity of parameters to move the network's latest output more toward the correct output. A modern neural network like GPT-3, which is already obsolete, had 175* ***billion*** *parameters interlinking its neurons, each of which requires tweaking. This is done over and over and over (many millions of times) across a massive body of "knowledge" to gradually train the network to generate the proper output for any input.*

- *Counterintuitive though it may be, the result of this training is a neural network that actually contains the knowledge that was used to train it; it is a true knowledge representation. If that's difficult to swallow, consider human DNA as an analogy. DNA contains all of the knowledge that's required to build a person. The fact that DNA is not itself intelligent or sentient doesn't mean that it's not jam-packed with knowledge.*

- *In fact, the advances that have **most** recently been made, which we'll get to in a bit, are dramatic improvements in the technology for extracting that stored knowledge from the network. That's why I titled today's podcast, "AI Training and Inference".*

- *The implementation of neural networks is surprisingly simple, requiring only a lot of standard multiplication and addition, pipelined with massive parallelism. This is exactly what GPUs were designed to do. They were originally designed to perform the many simple 3D calculations needed for modern gaming, then they were employed to solve hash problems to mine cryptocurrency. But they now lie at the heart of all neural network AI.*

- *Even when powered by massive arrays of the fastest GPUs rented from cloud providers, this "pretraining" approach was becoming prohibitively expensive and time consuming. But seven years ago, in 2017, a team of eight Google AI researchers published a truly ground-breaking paper titled **"Attention is all you need."** The title was inspired by the famous Beatles song "Love is all you need" and the paper introduced the technology they named "Transformers" (because one of the researchers liked the sound of the word). The best way to think of "Transformer" technology is that it allows massive neural networks to be trained much more efficiently "in parallel." This insightful paper also introduced the idea that not all of the training tokens – the long string of data being fed into a model during one training iteration – needed to be considered with equal strength because they were not all equally important. More "Attention" could be given to some than others. These breakthroughs resulted in a massive overall improvement in training speed which, in turn, allowed vastly larger networks to be created and trained in reasonable time.*

  Thus, it became practical and possible to train much larger neural networks ... which is what gave birth to LLM's – Large Language Models.

- *The "GPT" of ChatGPT stands for Generative Pre-trained Transformer.*

- *But over time, once again, researchers began running into new limitations. They wanted even bigger networks because bigger networks provided more accurate results. But the bigger the network, the slower and more time consuming – and thus costly – was its training. It would have been theoretically possible to keep pushing that upward, but a better solution was discovered: Post-training computation.*

- *Traditional training of massive LLM's was very expensive. The breakthrough "Transformer" tech that made LLM-scale neural networks feasible for the first time was now being taken for granted. But at least the training was a one-time investment. After that, a query of the network could be made almost instantly and, therefore, for almost no money. But the trouble was that even with the largest practical networks the results could be unreliable – known as hallucinations. Aside from just being annoying, any neural network that was going to hallucinate and just "make stuff up" could never be relied upon to build "chains of inference" where its outputs could be used as new inputs to explore consequences when seeking solutions to problems. Being able to reliably feed back a network's output into its inputs would begin to look a lot like thinking – and thus inference for true problem solving.*

- *Then, a few years ago, researchers began to better appreciate what could be done if a neural network's answer was not needed instantly. They began exploring what could be accomplished **post-training** if, when making a query, some time and computation – and thus money – could be spent working with the pre-trained network. This is known as "test-time computation" and it the key to the next level breakthrough.*

- *By making a great many queries of the pre-trained network and comparing multiple results, researchers discovered that the overall reliability could be improved so much that it **would** become possible to create reliable inference chains for true problem solving. Using the jargon of the industry, this is often referred to as Chains of Thought (CoT). Inference chains would allow for problem solving behavior by extracting the stored knowledge that had been trained into these networks, and the pre-trained model could also be used for the correction of its own errors.*

- *I should note that the reason asking the same question multiple times results in multiple different answers is that researchers long ago discovered that introducing just a bit of "random factor" – which is called "the temperature" – into neural networks resulted in superior performance. (And, yes... if this all sounds suspiciously like VooDoo, you're not wrong – but it works anyway.)*

- ***OpenAI's recently released o1 model** is the first of these more expensive test-time inference-chain AI's to be made widely available. It offers a truly astonishing improvement over the previous ChatGPT 4o models. Since o1 is expensive for OpenAI to offer on a per-query basis, subscribers are limited to 7 full queries per day. But the o1-mini model, which is faster and still much better, but not as good, can be used without limit.*

- *But wait! – there's more! The big news is that during their celebration of the holidays, OpenAI revealed that they have an o3 model that blows away their brand new o1 model. It's not yet available, but it's coming. What IS available are the results of its benchmarks and that's why I believe you need to make time to watch this YouTube video: https://youtu.be/YAgIh4aFawU (https://grc.sc/1007)*

- ***Is it AGI?** OpenAI is saying "not quite", but there's little question that they're closing in on it. As you'll see in that video, the performance of OpenAI's latest o3 model when pitted against independent evaluation benchmarks designed specifically to measure the general reasoning strength of AIs – when confronted by problems that were absolutely never part of the AI's training set – demonstrate reasoning abilities superior to most humans.*

- *Even if it were AGI, that doesn't mean it's taking over. The "AGI" designation is only meant to indicate that over a wide range of cognitive problem solving tasks an AI can outperform a knowledgeable person. Computers can already beat the best Chess, Go and Poker players. I think it's very clear that today's AIs are not far from being superior to humans at general problem solving. That doesn't make them Frankenstein's Monster to be feared; it only makes AI a new and exceedingly useful tool.*

Many years ago I grabbed the domain "clevermonkies.com" just because I thought it was fun. It occurs to me that it takes very clever monkies, indeed, to create something even more clever than themselves. All the evidence I've seen indicates that we're on the cusp of doing just that.

Okay. So that, with a bit of editing to improve it, is what many of our listeners received from me over the holidays.

If you take nothing else away from this discussion of AI today, **here** is the one point I want to firmly plant into everyone's mind: ***Nothing that was true about this field of research yesterday will remain true tomorrow.*** Nothing. This entire field of AI research is the fastest moving target I have ever experienced in my nearly 70 years of life.

There are a number of consequences to this fact. For one, no book about AI that was written a year ago or six months ago – or even last month – will be usefully up to date about what's happening now. Books written in the past can definitely be useful for describing the history of AI, and as a snapshot of a point in time. But even their predictions will prove to have been wildly wrong.

The guys at OpenAI who are working on this and ought to know, believed two years ago that at least another decade – another 10 years – would be needed to achieve what they announced last month and are getting ready to unveil. They thought it would take ten years, it took two.

One of the factors in facilitating this astonishing speed of development is that it turned out that much of what was needed was scale, and a weird side effect of cloud-side computing is that it's massively scalable. If you can pay to rent it, you can use it. So investor dollars were pumped into the training of ever more complex models and they kept seeing surprising improvements in performance.

Leo's original appraisal of Large Language Models as fancy spelling correctors was an accurate and useful from-the-hip summary of OpenAI's ChatGPT-3 model. And that's their take on it, too. ChatGPT-3 produced grammatically correct language, but it only coincidentally and occasionally produced anything highly meaningful. If it was left to keep talking it would soon get lost and wander off course to produce grammatically correct nonsense.

Even so, back then, highly creative people who operate on the cutting edge, like MacBreak Weekly's own Alex Lindsay, were using the GPT-3 model as a source of new ideas and inspiration. As I wrote this I was reminded of how popular formal "Brainstorming" once was where sometimes random ideas were tossed out without filtering – and that was the entire point: To say something as a means of inspiring some new perspective. So even ChatGPT-3 was useful for the nonsense that it sometimes produced.

As a consequence of everything I've learned over the past three weeks, and of the events which have transpired since, our previous podcast title **"The Wizard of Oz"** no longer seems to fit and I'm a bit embarrassed by what I wrote because it no longer reflects reality. As I said earlier, an honest researcher may need to discard previous belief systems when confronted with new information and facts. Never has that been more true than it is here. I'm needing to continuously update my ***own*** internal model.

There is an unfortunate downside emerging, however. Unfortunate, but I suppose, inevitable.

With startling speed, AI has moved from a curio in the corner of university and corporate R&D labs into big business. That meant that the suits in their neckties with their non-disclosure agreements descended upon the labs of the once freely and fruitfully collaborating academia-oriented researchers and dropped the cone of silence over all their ongoing work.

In the Distinguished Lecture Series at the Paul Allen School, one of OpenAI's leading researchers, Noam Brown, gave a lecture titled *"Parables on the Power of Planning in AI: From Poker to Diplomacy."* (I have a YouTube link to Noam's excellent talk at the end of the show notes.) During his lecture you could so clearly see Noam's unbridled enthusiasm and love of his subject, and also his disappointment when he was forced to stop himself short to prevent sharing some detail of his work that was now deemed to be proprietary and no longer his to share.

We only have Google's breakthrough Transformer and Attention technology – which was the sole enabler of the subsequent LLM revolution – because seven years ago, back in 2017 when things were still moving somewhat slowly, Google AI researchers were freely publishing their work as the academic curiosity it was at the time. They were working on improving Google's inter-language translation capabilities and this inspiration emerged unbidden from a chance meeting of eight Googler's from various parts of the organization. Would such a breakthrough be published in today's climate? That seems unlikely.

And now OpenAI is seeming less open than it once was. We know that ChatGPT-3 used a neural network containing an astonishing 175 billion neuron-interlinking parameters. We know that because OpenAI freely told us. But we have no similar information about any of their succeeding models. The sizes of the various ChatGPT-4 models, not to mention o1 and o3 have become closely held secrets – as has details of their operation.

Fortunately, a massive amount of detail – all detail needed for recreating much of what we see today from the corporate side – had previously been shared in the public domain and research continues with new vigor and doubtless with new funding within academia. And remember that it wasn't so long ago that Apple was getting patents on Andy Hertzfeld's clever stepwise circle drawing algorithm for bitmaps. Very little of anything that's really useful remains secret forever and it seems clear that before long we're going to have AI everywhere.

I would love to spend more time talking about the way neural networks function in detail because there are some very cool aspects of that, too. But that's not the purpose of this podcast and perhaps I'll find another opportunity for that in the future. There are also already tons of videos on YouTube talking about all of this for anyone who's interested, and YouTube's recommendation engine appears to be quite excellent.

I do need to point out a series of astonishingly well-conceived and produced instructional videos on this topic by a guy named Grant Sanderson. Grant's website is 3blue1brown.com and Grant's short bio says:

> *These videos, and the animation engine behind them, began as side projects as I was wrapping up my time studying math and computer science at Stanford. After graduating, I worked for Khan Academy producing videos, articles and exercises, primarily focussed on multivariable calculus. Since the end of 2016, my primary focus has been on 3blue1brown and its associated projects.*
>
> *In those years, I've also had the pleasure of contributing to a number of different outlets for math exposition, including spending a semester lecturing for an MIT course on computational thinking, contributing a Netflix documentary about infinity, writing for Quanta, and collaborating with many other educational YouTube channels.*

Grant produced a coherent series of eight videos, all available commercial-free on YouTube, which take its viewer from the basics of how neural networks operate all the way through where

and how they're able to store knowledge, how and what transformers transform, and how "Attention" is managed.

https://www.3blue1brown.com/topics/neural-networks

I recommend these without reservation to anyone who's interested in understanding more of the inner workings of the comparatively ancient technology of neural networks.

This old technology has recently been given new life thanks to the scalability of cloud-based computing and the presence of GPUs which are able to perform massive amounts of simple computational operations. So long as we have sufficient processing power it appears that the world is facing a true breakthrough thanks to the scale of compute and training we've been able to throw at the problem.

Though what we have today works and is working, it's also incredibly inefficient. It works only due to the massive scale we've managed to throw at neural network technology which is, itself, an extremely flexible but inefficient technology. It's possible to train a neural network that has just a handful of neurons to perform a simple binary adder function. But the same thing can be done far more efficiently with a couple of logical NAND gates. The thing that makes the handful of neurons potentially more interesting is that the same network could be trained to perform other simple functions. But the fundamental problem remains that any simple function that a neural network could be trained to do could be reduced to a far more efficient couple of NAND gates.

So here's what I think will eventually emerge someday. And I have no idea whatsoever when that might be. My hunch is that just as with the handful of neurons that can be trained to perform simple logic functions, we're going to eventually discover that there is a far simpler way to solve the same AI implementation problems much more efficiently than we're currently solving them with massive scale inefficient neural networks. I have no idea what that might be. But the intriguing thing is that cognitive science researchers now have a crude sort of brain that does manage to store a useful amount of knowledge and is able to use that knowledge to solve novel problems and I suspect, before long, to invent truly new things. People are already beginning to ask how exactly it does this ... because, believe it or not, that remains a mystery.

What is no mystery is what transpires here every Tuesday as it will next Tuesday and for many more Tuesdays to come.

Parables on the Power of Planning in AI: From Poker to Diplomacy: Noam Brown (OpenAI)
https://www.youtube.com/watch?v=eaAonE58sLU