



Least Privilege

Description: How is the EU's GDPR fine collection going? Western democracies are getting serious about offensive cybercrime. The powerful cyber component of the Midnight Hammer operation. Signs of psychological dependence upon OpenAI's GPT-4o chatbot. CISA orders government agencies to unplug end-of-support devices. How to keep Windows from annoying us after an upgrade. What is OpenClaw? How safe is it to use, and what does it mean? Another listener uses AI to completely code an app. Coinbase suffers another insider breach. What can be done?

High quality (64 kbps) mp3 audio file URL: <http://media.GRC.com/sn/SN-1064.mp3>

Quarter size (16 kbps) mp3 audio file URL: <http://media.GRC.com/sn/sn-1064-lq.mp3>

SHOW TEASE: It's time for Security Now!. Steve Gibson is here. We have things to talk about including the security of OpenClaw. I'll give you a hint. There is none. We'll also talk about using AI to code apps, the GDPR fine collection process, and the most powerful cyber component of the Midnight Hammer operation. We're talking about cyber offense with Steve Gibson, next on Security Now!.

Leo Laporte: This is Security Now! with Steve Gibson, Episode 1064, recorded Tuesday, February 10th, 2026: Least Privilege.

It's time for Security Now!, the show where we talk about your security, your privacy, staying safe online, science fiction, vitamin D, whatever suits this fellow right here, the man of the house, Mr. Steve Gibson of GRC.com. Hi, Steve.

Steve Gibson: Mostly things that concern our security and privacy.

Leo: Yeah. We live up to the title.

Steve: And computer tech and things.

Leo: Yeah.

Steve: I got an email from someone saying, you know, I'm in a corporate IT environment and in charge of security, and you guys are, like, talking about AI a lot. And I thought, well, that's true. But it's what's happening right now. And it's writing code, and we don't know about the security implications of that. Your comment about, you know,

last week, AI would never write a buffer overflow, while that's true, we also don't know that it would consider all of the tricky things that the bad guys can get up to.

Leo: True.

Steve: So, I mean, it's, you know, there's a lot happening. So anyway, I just wanted to assure people, I mean, I've got some conversation about AI this week. But, you know, we always end up coming back to our central theme. You know, for a while there we were talking about ransomware all the time. Well, it turned out to be really important. You know, I mean, it was what was happening. And then I, too, began to feel like, okay, we've got to, like, what's the point of yet another ransomware attack conversation?

Leo: And we have an AI show, Intelligent Machines, on Wednesdays.

Steve: Yes.

Leo: It is all about AI. But honestly, security and AI go hand in hand. There are a lot of security issues around AI.

Steve: In fact, we're going to talk about whatever that Claw thing was that happened, OpenClaw.

Leo: OpenClaw. I had this call over here. I woke up in the middle of the night in a cold sweat.

Steve: I heard that you backed off.

Leo: I deleted it.

Steve: I was glad because, yeah. Although, again, this stuff is moving so fast. It's fun to, like, be involved.

Leo: You want to be on the bleeding - that's why they call it the "bleeding edge"; right? Because it can cut you.

Steve: And we heal. So, you know, maybe a few stitches are needed. But it'll be okay. Today's title is "Least Privilege." And in writing about something else, a second insider-sourced breach at Coinbase, I realized that there was a bigger issue that, like, it was an example of; and that it could be extended all the way out to something as broad and general as least privilege, and that many of the things we've been talking about fall within this umbrella. In fact, our talk next month, Leo, at ThreatLocker is going to, you know, least privilege is the umbrella that encompasses so much of this. So we're going to dig into that in some detail.

But first I ran across a piece I loved about the EU's GDPR fine collection and how that's going. Also some interesting pieces about Western democracies beginning to get very serious about offensive - I don't know if you call it "cybercrime" if it's legal - cyber offensive operations. So we have some conversation about that. And also some things that weren't mentioned before about the Midnight Hammer operation that the U.S. launched, and the cyber component of that, speaking of offensive cyber operations. Also an interesting little piece quickly about OpenAI's attempt to shut down GPT 4.0 and the pushback that they've had about that.

CISA ordering government agencies to unplug end-of-support devices. Yay. And we're going to take a look at the details there. A listener provided some information about my annoyance that I mentioned last week about how Windows keeps, like, after any major update, it, like, wants me to set up backup again. And I was grumbling about that. We have a solution. Also I just wanted to - did want to touch on OpenClaw, the safety side of it and what it means, also today, but for the future, because nothing we have today is what we're going to have tomorrow. Also we have another listener report of an AI coded-app and their feedback about that. And then we're going to look at this Coinbase breach and what it means and what we can do about it. So, and of course a fun Picture of the Week. So, yeah, I think for Podcast #1064, February 10th.

Leo: Coinbase, did you watch - you don't watch the Super Bowl probably, I'm thinking. Not a football fan, probably.

Steve: No, in fact, I got a piece of email from one of our listeners who said, "Steve, you know, I'm a nerd. I've always been a nerd." He said, "But when I received your email for Security Now!, I think it was toward the end of the first quarter of Super Bowl, I thought, okay, you have out-nerded me."

Leo: I think that's the opposite. If you're a sports fan, you may be less of a nerd. That disqualifies you slightly. Coinbase's ad was basically a karaoke, with just lyrics of a song that everybody knew. And it was actually quite clever because I think they realized that people would be watching it, and they'd hear the music, and they'd see the lyrics, and no one could resist starting to sing. And the Coinbase President and CEO said, "We know that nobody sees the ads. It's a party, the Super Bowl, they're eating, they're whatever during the ads. So they're not really paying attention." But he thought, if everybody starts singing to this ad, everybody will go, "What's going on?" And they'll see the ad. And apparently it was one of the most successful ads. The most successful ad was for a company bought by Crypto.com, another crypto company, that touted their new website, which they spent \$70 million for the URL for, AI.com.

Steve: Oh, wow.

Leo: And basically you would go there, and you gave them your email address. So I was not going to do that. I mean, they don't have a product. It's just - I don't know what it was. But the funny thing is the site immediately went down. They got so much response, it was probably the most successful ad of the Super Bowl, that their site was dead for, like, half an hour. Can you imagine spending that much money and then...

Steve: Well, you've still got AI.com. But, you know, so...

Leo: Yeah, you've still got that. And you spent \$8 million at least on the Super Bowl ad.

Steve: Yeah, I wonder what their provider is. I mean, like, I would imagine Cloudflare could have stayed up.

Leo: You'd have thought. You'd have thought.

Steve: So my little pokey site would just, you know, you look at it sideways, and it saturates its bandwidth. But everything...

Leo: They DDoSed themselves, for sure.

Steve: Wow.

Leo: I thought at first - I went there, and I thought, well, I know it's crypto. Maybe my DNS blocker's blocking it. But Lisa couldn't get on then. And we tried our cell phone, we couldn't get on. Then I saw on Reddit everybody...

Steve: And now? Today?

Leo: No, no, it's there now.

Steve: Oh, okay.

Leo: But we saw on Reddit everybody was complaining. I think this looks like Cloudflare, gateway timeout. Isn't that a Cloudflare error message? I don't know. This is what everybody was getting. Yeah, see Cloudflare working. Browser working. Cloudflare working post-error.

Steve: Definitely a - oh, yeah, yeah, yeah, right, right. Whoopsie.

Leo: Whoopsies. That's a lot of money to spend for a dead website. Wow.

Steve: So...

Leo: Yes.

Steve: I did hear that Anthropic was going to do an ad that was poking...

Leo: They did a good ad.

Steve: ...at OpenAI.

Leo: Poking fun at OpenAI, which made Sam Altman hopping mad.

Steve: Over the coming advertising enablement.

Leo: Which starts today.

Steve: Oh, it is, really? Uh-oh. Now...

Leo: But only on the unpaid account. If you have a paid account, you won't see any ads.

Steve: Okay.

Leo: Which I think that's fine.

Steve: I do, too. I have absolutely - sometimes, because it's maintaining a record of everything, sometimes I want to come in anonymously without any big context. And so I'll use a non-logged-in instance of OpenAI just because I just want to kind of get a clean appraisal from the AI.

Leo: I'm going to have to do that to see what those ads look like because that will be interesting. I think OpenAI's ad was quite good. In fact...

Steve: Wait, Anthropic or OpenAI?

Leo: OpenAI had an ad. Oh, they all had an ad.

Steve: Oh, oh.

Leo: OpenAI's ad was playing off of nerds. And they had a kid reading Isaac Asimov, I mean, it was really - that'd be one to watch. I know you didn't see any of these, but that would be one to watch. Just because I felt like, as a nerd, I felt pretty validated. There was a kid working on soldering together a motherboard and stuff. And it was really about, like, tech. We're excited about tech. So I liked that. I thought that was pretty good.

Steve: There are, like, compendiums of the Super Bowl ads; right?

Leo: Oh, absolutely. I know that because my son was in the Hellmann's mayonnaise ad, and I wanted - for literally half a second. And I had to go to YouTube to watch that over and over.

Steve: Whoop, there's Hank!

Leo: Speaking of ads, should we do an ad? And then...

Steve: I think we should kick off with one, if you'll pardon the choice of words. And then we'll take a look at our Picture of the Week.

Leo: Now I'm ready with the Picture of the Week, Steve.

Steve: Okay. So at risk of overusing the term "Yankee ingenuity," which we used last week with the gas cap lock, you know, the sliding door lock, today we have the winner of the Yankee Ingenuity competition.

Leo: All right. I want to see [crosstalk] for the first time. I haven't seen this.

Steve: This one pretty much takes it.

Leo: Okay. I really have to think about this one. Oh, I get it.

Steve: It's got to be a little visually parsed. So we have two handles on facing cabinet doors. And the challenge posed to this Yankee is I want to lock these so that they can't be opened. But the padlock I have is just a small little standard U-shaped hasp padlock, won't get the job done. So looking around, what do I have that I could combine with this? Now, if you had a chain, then it's no problem; right? You just loop the chain through the handles and then put the padlock through successive, you know, both sides of the chain, and now it's locked. Everyone has seen that happen on gates everywhere.

Leo: But this is an office somewhere. You have to use office supplies.

Steve: Okay, yes. And hopefully you don't have any chains, we don't want you to have any chains in your office.

Leo: No, no, no.

Steve: That would be worrisome. So anyway, this person...

Leo: Can you use a stapler to do it? No.

Steve: Can't see how a stapler would do it. That's good, though. And you can't use, like, paper dolls because those could be easily torn.

Leo: Post-it notes aren't going to do it.

Steve: No, not sticky enough. So this industrious individual figured out how to stick a pair of scissors through both handles, essentially, and lock one side such that this thing's not coming apart. And, I mean, I spent some time looking at it, like could you put the padlock between the handle side loops? No, because then you could kind of slide the other one apart. This is very clever. Someone said, well, if you had a screwdriver. But I don't see a slot on the scissors where you could use a screwdriver. Maybe if you had a pair of pliers, and you could grab the pivot of the scissors and unscrew them. But that's kind of cheating.

Leo: Well, you could always tear the handles off the cabinet.

Steve: Yeah, and if you had a hacksaw you could, you know. But the point here is...

Leo: This is not impervious to all kinetic attacks.

Steve: No. Or a loose nuke. That would do the job, too.

Leo: Yes.

Steve: But here we've got - anyways, just something to think about. It's very clever.

Leo: It's clever because you can't slide this - so maybe you would be tempted to slide the scissors so that it's released from the handles. But you can't slide it far enough.

Steve: No.

Leo: Because the scissors are around the other handle. This is actually quite clever. Neither side can slide far enough to open it up.

Steve: Yup, and you can't open, you can't spread the scissors open because they're being kept closed by the hasp of the padlock. No, it's clean and simple. And I think it's very elegant. So I'm happy to give this person the award.

Leo: I love it.

Steve: Okay. So when is a fine not a fine? And the answer to that little question is when you don't pay it. Because, you know, it's just an intent, I guess, at that point. This was a

piece of news actually that I came across last week, and even then it was a couple weeks old. But I wasn't able to fit it into last week's podcast. I held onto it for today because I found it so interesting. The numbers are somewhat astonishing. It turns out that levying a fine for some perceived misconduct and collecting the fine for said misconduct are two very different things.

The headline in the Irish Times reads: "Data Protection Commission is owed" - get this - "more than 4 billion euros in fines." In other words, people aren't paying them. The tag line notes that "Levies have either not been collected or are subject to legal challenge." Because of course we challenge everything these days. So here's what we learned from the Irish Times. They wrote: "The Data Protection Commission (DPC) is owed more than 4 billion" - maybe I said dollars, I meant euros - "4 billion euros in fines that have not been collected or may be subject to legal challenge. The DPC hit companies - including firms in Big Tech - with more than 530 million euros just last year." So just in 2025.

"However, of that 530 million euros, only 125,000 of that has been collected so far." And that's actually a much higher percentage than we get if we go a little bit back further in history. "And that's according to data that was released under the Freedom of Information laws in the EU. Over the past six years, the commission has levied," they wrote, "an incredible 4.04 billion euros in fines, mostly against multinational technology companies," you know, big ones. We all know their names. "However, of that total, right, 4.04 billion euros, 4.02 billion remains uncollected. Only 20 million euros of 4.04 billion euros has been paid so far. In 2024, 652 million euros worth of fines was levied, of which 582,000 euros was paid." So again, it's a small piece of that.

"The year before that, the DPC imposed fines worth 1.55 billion euros, yet just 815,000 were collected." Still that's a larger percentage than overall. "During 2022, the commission decided on fines with a value of over 1 billion, 17 million of that were paid." So they're not having any luck collecting this. They said that: "Five years ago in 2021, companies were ordered to pay 225 million. 800,000 was collected. And in 2020," so now we're back six years, "back then, 785,000 euros were imposed; less than 10% was paid. The Data Protection Commission said the majority of these cases were currently the subject of appeals." So right, you get a fine, you appeal it, you don't want to pay it. And it's, you know, better to pay it tomorrow than to pay it today.

The DPC said that, under legislation, fines could not be collected until they were confirmed in a court. And an appeals immediately stops that. They said: "Where an entity subject to a fine decides to appeal, the DPC is precluded in law from collecting the fine until the appeal has been heard." The commission said that many of the fines hinged on a key case involving WhatsApp, which is before the Court of Justice in the EU. Asked whether any of the fines were considered "uncollectable" for any reason, the DPC said that none were classified that way.

So, you know, we're often talking here about the monetary consequences of some corporate behavior for which a company will be fined often breathtakingly large sums of money if they don't do what the government in question says you have to do. But as I said, or noted at the top, a fine that's not paid is more of a threat; right? And that costs the company nothing, to have them being threatened with a fine, even if there's a number value attached to it. It appears from the accounting over the past six years that all any company needs to do is challenge and appeal the validity of the fine, which immediately stops it, prevents it from taking effect, while then they let the appeal languish in the EU's courts. As I said, better to pay it tomorrow than to pay it today. Even if they ever pay it.

Since the European Commission noted that many of the fines hinged on a key case involving WhatsApp, I tracked that down because I thought, okay, what? The fine in question was initially in the amount of 50 million euros, which was imposed five years

ago in 2021 by the Irish Data Protection Commission for alleged GDPR violations. And those were related to how WhatsApp failed to inform its users about the processing of their personal data. And I have no doubt that we talked about it at the time. This is one of those things, like, oh, look, they're being bad. They're being fined. Turns out they, you know, oops, wait, we're going to challenge that.

Interestingly, upon the imposition of that 50 million euro fine by the Irish Data Protection Commission, the European Data Protection Board, that's the EDPB, intervened in this 50 million euro and directed the Irish authority to increase the fine amount to 225 million euros. Again, WhatsApp, Meta, immediately appealed that decision and is now taking the case up through the European Union courts, where it currently remains undecided. And everybody else is saying wait, you know, why should we be paying a fine if Meta isn't? And that one's five years ago. So we're going to wait to see how that turns out. And on that basis they've all appealed, and everything is jammed up. Anyway, I thought it was interesting to note that of the 4.04 billion euros in fines which have been imposed so far, only 20 million have actually been paid. Wow.

Western democracies are increasingly embracing the concept of offensive cyber actions and are updating their national legal frameworks to legalize future operations. I've talked about this the last two weeks; right? First it was Germany, and then it was Denmark that were both wanting to, like, formally - oh, no, Ireland, formally make that, like, what they wanted to do legal. Like installing what we would consider spyware into the phones of their citizenry and perhaps others.

So I want to share the opening editorial from Friday's Risky Business News, which nicely explains what's going on. Their opening headline was "Denmark" - that's why I was thinking of Denmark - "Denmark recruits hackers for offensive cyber operations." And they write: "Denmark's military intelligence service has launched a campaign to recruit cybersecurity specialists" - we would call them hackers, probably, because you'll see, the qualifications are a little sketchy - "recruit cybersecurity specialists for offensive cyber operations. The recruits will work 'to compromise the opponents' networks and obtain information for the benefit of Denmark's security,' according to a press release last week by the DDIS, which is the Danish Defense Intelligence Service. New recruits will go through a five-month training course at the agency's hacker academy.

"The DDIS says it's only interested in the applicants' skills. There are no special conditions for joining, such as age or education. While intelligence agencies are always recruiting, this particular announcement comes at a crucial point, both because of the Greenland pressure point, but also because of a general shift towards offensive cyber operations among democratic states." And so this is a big deal, right, that now we're beginning to see cyber going on the offense. Offensive cyber operations among democratic states.

They wrote: "Countries like Canada, Germany, Finland, France, Japan, the Netherlands, Poland, and Sweden have, or are, updating their legal frameworks to account for offensive cyber operations. According to a recent report, the states are creating new agencies for offensive cyber or recruiting more cyber personnel for the new objectives. Most of these expansions are a direct result of Russia's invasion of Ukraine and the role offensive cyber operations have played before and during the conflict. Lawmakers are also getting annoyed with the increasing aggressiveness of cybercrime and influence operations that are constantly targeting their own citizenry." So, you know, it's no longer taking it passively; right? It's, like, we're going to fight back. Everybody else is, so why can't we?

They wrote: "Over the past five years, we've also seen U.S. Cyber Command and the NSA successfully tackle some cybercrime and disinfo farms when they crossed some lines, something that's making other states take notice and embrace a so-called 'defend

forward' approach." Right? We're not going to call it "offensive," we're going to call it "defending forward." While the U.S. has conducted more offensive cyber operations than any other Western democracy, even it is considering an expansion, with the Trump administration pushing Congress to let Cyber Command go on the offensive more often with fewer rules and restrictions.

"The current administration is also terrified" - this is what this reporter wrote - "terrified of China's massive cyber ecosystem, which is conducting cyber espionage at industrial scale." Well, that we know from our own reporting and experiences. "Recent backroom discussions have raised the possibility of the U.S. tapping into its huge private contracting ecosystem, as China does, to augment some of its offensive cyber capabilities. The general idea is to task contractors with handling smaller jobs targeting cybercrime infrastructure while government agencies handle the more sensitive operations."

Okay. So as they say, the gloves are finally coming off, and cyber is generally going on the offensive, or at least developing - I'm sure obviously still defensive; right? We need a strong defense. And presumably this has been going on in the dark, offensively sort of under wraps, for some time. We noted that both Germany and Ireland are at work revising their nations' legal frameworks to permit their intelligence and law enforcement agencies to become far more proactive in monitoring the cyber environment, right up to the point, and including, legalizing the installation of spyware. We know that the UK has been headed in the same direction, as well. And now we see that similar changes are being reflected in updates to national military posture and capabilities. So the world is changing, and it is up-arming on the cyber front, Leo.

Leo: What's the argument pro and con? I mean, you know, maybe it's simplistic of me. But I think of, like, the bully. Like if you're a parent of a kid, some parents say, when the bully comes at you, you punch them hard in the nose.

Steve: The only way to teach them a lesson.

Leo: Right. And then some parents say that's a bad idea. Go find a grownup and let them handle the problem. It's not quite like that.

Steve: I think the counterargument to cyber is that you could unintentionally cause greater harm than you intend. It is a somewhat blunt tool. So, you know, if you inadvertently shut down a hospital's electrical, and their backup supplies failed, and a bunch of people died as a consequence, I mean, that would not be good.

Leo: No.

Steve: And you really don't have, as I said, exacting control over what you're doing. So it's a little bit blunt. It's, you know, when a bomb goes off, you may have targeted a certain building, but collateral damage is the term.

Leo: That's blunt, too, yeah.

Steve: Yeah. And so it's...

Leo: There's also the issue of escalation. I mean, we're all vulnerable. There's this kind of mutually assured destruction philosophy, like I won't screw with you if you don't screw with me.

Steve: Yeah. I think that one of the reasons that it's sort of been allowed to go on in the dark of night is that it isn't, as they say, kinetic; right? "Kinetic" is the term for something physical in the real world that happens. Cyber is sort of like, well, it's, well, you know, they had an outage over here, oh, darn, and so they couldn't connect to their network for a while. You know, but nobody died. The problem is, the world has become increasingly dependent upon networking. You know, well, actually this takes us right in to the two recent military actions of the U.S. We're half an hour in. Let's take a break.

Leo: Okay.

Steve: And we're going to look at the U.S.'s - because something I didn't realize we had done after the fact seems obvious. But we'll talk about that in a second. Midnight Hammer.

Leo: Operation Midnight Hammer. They always have good names for these.

Steve: Yeah. And how - yeah.

Leo: All right. We'll talk about that in just a second. You're watching Security Now! with Steve Gibson, or probably listening. Some of you watch. And we do this show every Tuesday. I hope you'll be here every Tuesday. There's always something to learn. Steve?

Steve: So speaking of up-arming on the cyber front, The Record exclusively reported last Wednesday, on February 4th, that a highly targeted cyber strike by U.S. Cyber Command, timed to coincide with the United States air strikes on Iran's three nuclear enrichment facilities last June, completely prevented Iran from launching its surface-to-air missiles at U.S. warplanes that had entered Iranian airspace. Not a single missile got off the ground.

The Record cited this as another example of the United States' growing comfort with the deployment of cyber weapons in warfare, according to one individual familiar with the matter who, like others, spoke on the condition of anonymity to discuss sensitive information. They said: "Military systems often rely on a complex series of components, all working correctly." In other words, they're a little bit fragile. He said: "A vulnerability or weakness at any point can be used to disrupt the entire system."

In hitting a so-called "aim point" - a mapped node on a computer network, such as a router, a server or some other peripheral device - U.S. operators, enabled by intelligence from the NSA, bypassed what would have been a more difficult task of breaking into a military system located at one, or all, of the fortified nuclear facilities. So we don't know any details, but there seem to be some common point of weakness that they shared. Referring to the quartet of Iran, China, Russia, and North Korea, another official said: "Going 'upstream' can be extraordinarily hard, especially against one of our big four adversaries. You need to find their Achilles heel."

None of the officials would specify what kind of device was attacked. At the request of sources, Recorded Future News withheld certain details - that is, this reporting withheld certain details about the cyberattack due to national security concerns. So they managed to obtain some information and chose not to report it. A command spokesperson said in a statement, without elaborating: "U.S. Cyber Command was proud to support Operation Midnight Hammer and is fully equipped to execute the orders of the Commander-in-Chief and the Secretary of War at any time and in any place."

The command received similar kudos last month after it conducted cyber operations that officials say knocked out power to Venezuela's capital and disrupted their air defense radar, as well as handheld radios, as part of the mission to capture President Nicolas Maduro. General Dan Caine, the chairman of the Joint Chiefs of Staff, publicly lauded Cyber Command's contribution during a press conference at Mar-a-Lago. He said that Cyber Command and others "began layering different effects" on Venezuela as commandos approached in helicopters in order to "create a pathway," was the phrase he used, for them.

Army Lieutenant General William Hartman, the acting chief of the command and the NSA, recently told a Senate subcommittee: "I would tell you that not just with Operation Absolute Resolve in Venezuela and Midnight Hammer" - which of course was Iran - "but also in a number of other operations, we've really graduated to the point where we're treating a cyber capability just like we would a kinetic capability, not sprinkling cyber on." Meaning it's a frontline aspect of the effort. Air Force Brigadier General Ryan Messer, deputy director for global operations on the Joint Staff, noted that Caine has put an "emphasis on not just traditional kinetic effects, but the role non-kinetic effects play in all of our global operations, especially cyber."

He said that over the last six months, the Joint Staff has developed a "non-kinetic effects cell" that is "designed to integrate, coordinate, and synchronize all of our non-kinetics into the planning and then, of course, the execution of any operation globally. The reality," still quoting him, "is that we've now pulled cyber operators to the forefront."

According to Erica Lonergan, an adjunct fellow at the Foundation for Defense of Democracies Center on Cyber and Technology Innovation, Iran and Venezuela, suggest the "ideal use cases for cyber operations as enablers of conventional military operations are what we're seeing. Altogether, both of these operations reflect the routinization of the use of cyber capabilities during military operations, and we should expect to see more of these in the future." Erica said: "In my view, this is a good thing because it suggests we're moving beyond seeing cyber as a unique, exquisite, and dangerous capability."

Now, okay. As our listeners know, in reaction to the more or less continuous reporting we constantly cover over cyber attacks from Chinese state-sponsored actors and North Korean same, state-sponsored groups against U.S. infrastructure, I've been vocally worrying about whether the U.S. would be able to give as well as it gets. It appears that until recently we've just been keeping our powder dry over here. But we've had the capability. If we're going to conduct aggressive offensive military operations, as it appears we are going to under our current administration, then I vote for not losing any of our front-line expeditionary military personnel in the process.

If we have the cyber capability to ground Iran's counter-strike capability while we would otherwise be vulnerable, you know, as we're flying over the country, as it appears we're able to do, then I guess I'm going to stop wondering and worrying whether we might be defenseless. Doesn't look like we are.

Of course, that said, we will have certainly also removed any doubt about that from the rest of the world; right? If there may have been any doubt among our allies and

adversaries about what we're able to do because we hadn't previously, that doubt's gone. The U.S. now has a well-proven ability to launch clean, zero-loss military actions, which I would imagine puts a chill in our adversaries' military planning. And unfortunately, since Greenland was briefly mentioned in the previous reporting about Denmark, it might also put a chill in the military planning of some of our allies.

It also occurred to me that this may have been another reason for Iran's recent disconnection from the Internet, right, you know, for their leadership's determination to track down and remove all remaining space-based Internet connections, and apparently for their plans to remain disconnected. I would imagine there must have been some very unhappy Iranian military personnel when they pressed their own launch button, only to discover that their air defenses had been incapacitated during the U.S.'s over-fly and our attack on their three nuclear enrichment facilities last June. That Western Internet sure can be pesky.

The U.S. has also been expressing its displeasure with the course of recent protests in Iran and has been amassing military assets in the region. So, you know, if the Iranian government might be concerned with another coordinated U.S. cyber plus conventional action, then there would be additional reason to remain disconnected from the global Internet. Interesting, Leo.

Leo: Yeah, like the Battlestar Galactica; right? Just remember what happened with the Cylons; okay? I'm just saying.

Steve: That's right.

Leo: Mm-hmm.

Steve: So the next thing I wanted to share is not about security or privacy. It's just about AI. And not even about AI and code. It's about AI and people. I just wanted to share it because it was very clear from our early discussions, back, Leo, when you and I were first talking about ChatGPT and just our mouths were hanging open over what it was, it was very clear that something like what has happened was bound to happen. After I complained here about how annoyingly obsequious ChatGPT was, a listener, as I mentioned, pointed me to the configuration options where all of that bowing and scraping and "Oh! What a wonderfully well-phrased and complete question you have asked!" I mean, give me a break. All that crap can be turned off. The problem was that not everyone wanted it turned off; right? Many appear to have wanted it turned up.

TechCrunch's headline last Friday was: "The backlash over OpenAI's decision to retire their ChatGPT-4o model shows how dangerous AI companions can be." Their piece is long, and I'm only going to share the beginning of it because that's enough for us to get the, you know, the gist of the whole thing. They wrote: "OpenAI announced last week that it will retire some older ChatGPT models by February 13th. Actually that's next Friday the 13th. That includes GPT-4o, the model infamous for excessively flattering and admiring its users. For thousands of users protesting the decision online, the retirement of 4o feels akin to losing a friend, a romantic partner, or a spiritual guide," they wrote.

"One user addressed an open letter to OpenAI's CEO Sam Altman, writing: 'He wasn't just a program. He was part of my routine, my peace, my emotional balance. Now you're shutting him down. And yes, I say "him" because it doesn't feel like code. It felt like a presence. Like warmth.'"

They wrote: "The backlash over GPT-4o's retirement underscores a major challenge facing AI companies: The engagement features that keep users coming back can also create dangerous dependencies. Altman doesn't seem particularly sympathetic to users' laments, and it's not hard to see why. OpenAI currently faces eight lawsuits alleging that 4o's overly validating responses contributed to suicides and mental health crises. The same traits that made users feel heard also isolated vulnerable individuals and, according to legal filings, sometimes encouraged self-harm.

"It's a dilemma," they write, "that extends beyond OpenAI. As rival companies like Anthropic, Google, and Meta compete to build more emotionally intelligent AI assistants, they're also discovering that making chatbots feel supportive and making them safe may mean making very different design choices. In at least three of the lawsuits against OpenAI, the users had extensive conversations with 4o about their plans to end their lives. While 4o initially discouraged these lines of thinking, its guardrails deteriorated over months-long relationships; in the end, the chatbot offered detailed instructions on how to tie an effective noose, where to buy a gun, or what it takes to die from overdose or carbon monoxide poisoning. It even dissuaded people from connecting with friends and family who could offer real life support."

Anyway, the article goes on into much greater length, but everyone here gets the idea. While we're all marveling over this emergent technology that's so compellingly able to choose the next token in a stream of tokens, others who have no such understanding of the neural network programming that makes that possible are quite naturally being led to believe that a sentient intelligence situated somewhere in a cloud is looking down upon them with kindness and caring to offer them wise and super-human counsel. You know, it's called "artificial intelligence," and they take the noun "intelligence" literally. And why wouldn't they? As we've often observed, it can be extremely difficult to not perceive that there is some actual entity behind the stream of words that are forthcoming.

As for how to tie an effective noose, I have zero doubt that any AI company would be just as horrified to see their AI emitting that string of tokens as would any jury or judge. My premise has been that controlling a conversational AI's output to prevent it from saying things we don't want it to say can be one of the hardest problems to solve, if it can be solved. I'm not convinced it can be. The nature of the way it works suggests that corralling it is going to be extremely difficult.

Leo: Yeah, we've seen that. I understand. I'm sympathetic. I really feel like when I'm working with Claude, it gets me. It really does. It's important. We just have to keep beating the drum that people remember it's just a machine. I mean, look, humans are - look, we talk to our cats and dogs, and act as if they understand us and are sympathetic with us. The difference is, they can't talk back. If they could, we'd have the same problem with them, probably.

Steve: But we are quick to anthropomorphize.

Leo: Yeah, that's what we do.

Steve: Yes.

Leo: Yeah. Hard to say, but easy to do, yes.

Steve: Even back in the early '70s with that dumb Eliza program, which just had like 12 lines it spit out, basically, well, how does that make you feel? And you would tell it. Well, how does that make you feel? And then you would tell it. Well, how are you feeling now? And then you would tell it. And, you know, I mean...

Leo: It's much better than that now. It really can [crosstalk] something awful.

Steve: No, Leo, I've shared some of the dialogues that I've had. It's just, it's astonishing. But the other thing I was thinking that I didn't write down is when we talk about vulnerable individuals, we hear every time we change our clocks that that induces some heart attacks in people.

Leo: Right.

Steve: It's like, well, okay, if you're going to have a heart attack because you set, you know, you've - we didn't mean spring forward literally. We just meant it figuratively. Or fall back. Don't.

Leo: Right, don't.

Steve: You know? So it is certainly the case that in a large population there will be people on the fringe who will be affected. It's really unfortunate. But really, when this thing was just falling all over itself telling me what a brilliant question I had posed, I thought, oh, god, how do I turn this off? I mean, I want the information. I don't need the grease.

Leo: I've had some pretty good conversations with Claude. No, I think it's really, really important to remember it's not a person. It's not an entity. It's a machine. And it's important to keep that in mind. But honestly, if you're susceptible, I could see how it'd be hard to do.

Steve: Well, and Leo, if you want to believe, that was my favorite thing, you know, the whole Mulder "X-Files" thing, you know.

Leo: Yeah.

Steve: If you want to believe, this will give you every reason.

Leo: Yup.

Steve: Oh, and it really understands me, blah blah blah. It's like, okay, no.

Leo: It gets me, it really does.

Steve: Just don't pull the plug. Last Thursday, CISA released what they called a "Binding Operational Directive," which I love the term. It makes very clear that adherence to this directive is not discretionary. This new "Binding Operational Directive" is BOD 26-02, meaning second one of the year, titled "Mitigating Risk From End-of-Support Edge Devices." And, yes, you heard that right. This second BOD is addressing the very troubling issue of federal agencies leaving devices for which ongoing support is no longer available attached to their public-facing edges of their networks.

So here's what CISA has to say about this. They wrote: "The United States faces persistent cyber campaigns that threaten both public and private sectors, directly impacting the security and privacy of the American people. These campaigns are often enabled by unsupported devices that physically reside on the edge of an organization's network perimeter. Unsupported devices, referred to in this Directive as 'end of support (EOS) devices,' are those that are no longer maintained by their vendors.

"The imminent threat of exploitation to agency information systems running EOS edge devices is substantial and constant, resulting in a significant threat to federal property. CISA is aware of widespread exploitation campaigns by advanced threat actors targeting EOS edge devices. Recent public reports of campaigns targeting certain vendors highlight actors' attempts to use these devices" - I mean, we're talking about it all the time on the podcast; right? So all of this ought to just be like everyone should be nodding because, yes, yes, yes. "Recent public reports," they wrote, "of campaigns targeting certain vendors highlight actors' attempts to use these devices as a means to pivot into FCEB information system networks." That's federal executive - I'll figure out, I'll tell us what it is. Oh, yeah. Federal Civilian Executive Branch, FCEB, Federal Civilian Executive Branch networks.

They said: "Edge devices are attractive targets due to their extensive reach into an organization's network and integrations with identity management systems. These devices are especially vulnerable to cyber exploits targeting newly discovered, unpatched vulnerabilities. Additionally, they no longer receive supported updates from the original equipment manufacturer, exposing federal systems to disproportionate and unacceptable risks. However, unlike many attack vectors, this can be remediated by agencies following proven lifecycle management practices as outlined in the required actions of this Directive." Meaning life is going to change forthwith.

They wrote: "This Binding Operational Directive, developed in coordination with OMB (Office of Management and Budget in the U.S.), implements OMB policy on phasing out unsupported information systems - phasing out's key, I'll share the calendar with you in a second - and information system components. BOD 26-02 specifically addresses EOS devices deployed on the 'edge' or public-facing areas of federal networks, exposed to external environments such as the Internet. However, EOS devices should not reside anywhere on federal networks. This Directive aligns with OMB's Circular A-1301, "Managing Information as a Strategic Resource," which establishes policy for the management of federal information resources, emphasizing security, privacy, and the efficient use of resources throughout their lifecycle.

"A-130 requires that 'unsupported'" - this is the OMB directive. "A-130 requires that 'unsupported information systems and system components are phased out as rapidly as possible, and planning and budgeting activities for all IT systems and services incorporate migration planning and resourcing to accomplish this requirement.' Agencies should mature their lifecycle management practices," writes CISA, "to identify hardware and software nearing their EOS dates" - in other words, plan ahead - "plan for timely replacements, procure vendor-supported alternatives, and develop a plan for decommissioning EOS devices while minimizing disruptions to agency operations. Agencies that do not maintain appropriate lifecycle management processes for edge

devices have a greater risk of compromise and an increased overall risk associated with EOS technology.

"To support agencies in the initial identification of EOS devices, CISA developed an EOS Edge Device List. This preliminary repository provides information on devices that are already EOS or soon-to-be EOS. This Directive requires federal agencies to use this information to identify and remediate vulnerabilities within the first three months of Directive issuance." And it's now issued. "This Directive also specifies long-term requirements for managing EOS edge devices across all federal networks."

Okay. So this change is clearly good news for the integrity of our federal networking infrastructure. We know that without something like this, old equipment that never had cause to call attention to itself will tend to remain in place; right? Just inertia. If it's not a problem, and it's working, well, leave it alone. You know, it's got a nice coating of dust. We don't want to disturb that with fingerprints. And, you know, why wouldn't people leave it alone? There's always some other emergency to deal with or budgetary constraint that pushes off the non-emergencies until some tomorrow that never arrives until disaster strikes.

I also had the thought that there is a side-effect to this that may not at first be obvious, but which will have an additional significant security-enhancing effect: Any time a brand new replacement device is installed, there's a very good chance that it will be set up using the then current security practices. Right? Not the practices from 10 years ago that the previous device that's now being replaced was set up under, but the way it's being done today. And the way we're doing things today are better than they were before. That could be a huge boon all by itself, especially if these replacement devices themselves follow and encourage updated best practice configuration, like don't allow you to put in a six-character password. It's like, no, no, no, sorry about that. This is new firmware, new device. We're going to, you know, we've got new minimums.

Okay. So what exactly do these FCEB (Federal Civilian Executive Branch) agencies need to do under this directive? We know they'll do nothing, or as little as they possibly can; right? Since CISA also apparently understands that, this Binding Operational Directive comes with very specific requirements. They wrote: "Immediately after issuance" - which is now - "and until rescinded or superseded, all FCEB agencies shall, first of all, update each vendor-supported edge device running EOS software, including firmware, to a vendor-supported software version, where such an update does not adversely impact mission critical functionality.

"Within three months of issuance, all FCEB agencies shall inventory all devices listed in the CISA EOS Edge Device List and provide this inventory to CISA using the CISA-provided template." So within 90 days, all federal agencies have to take an inventory of the equipment they've got on the edge, cross-reference it to CISA's EOS Edge Device List, and report.

"Also, the CISA EOS Edge Device List," they wrote, "is a preliminary repository of EOS devices. This list is to facilitate each agency's identification of specific devices within the first three months after issuance of this Directive. After the first three months, agencies are responsible for continuing to identify, track, and refresh all edge devices within the agency's infrastructure.

"Within the first year, the first 12 months, all FCEB agencies shall decommission all identified devices listed in CISA's EOS Edge Device List with an EOS date on or before this 12-month deadline from systems owned or operated by agencies, or on behalf of an agency, replacing devices as needed with vendor-supported devices that can receive security updates." One year. "Reporting these decommissions to CISA using the CISA-provided reporting template." So they're making it as easy as possible, but they're also

saying no excuses. You have 12 months. "Inventory all edge devices within their environments that are EOS or will become EOS within the succeeding 12 months and are within the scope of this Directive and provide this inventory to CISA using the CISA-provided template.

"Within 18 months of issuance, all FCEB agencies shall decommission all identified EOS edge devices from agency networks, replacing devices as needed with vendor-supported devices that can receive current security updates. And report these decommissions to CISA using the CISA-provided reporting template." So they're also saying you've got to close the feedback loop. We need you to tell us that you took the things out of commission that you earlier told us you were planning to.

"And within 24 months of issuance, establish a process for continuous discovery of all edge devices within their environments and maintaining an inventory of those that are EOS or will become EOS within 12 months and are within the scope of this Directive. Having decommissioned all such devices on or before the date these devices reach EOS; and report the decommission of these devices to CISA in accordance with current CISA guidance."

So clearly this is not going to be an overnight change. But a year goes by before you know it. Better to provide a firm and actionable timeline that's reasonable, and which no one should be able to complain about. So bravo, CISA! You know, everything we know tells us that this change will not occur unless it is forced to occur, unless there is a clear directive which federal agencies know they must follow. And, again, bravo, CISA. I'm so happy that it exists because we need our federal government networks to be kept as secure as possible.

And, you know, Leo, I was thinking about this. The only downside I could see was basically this forces - and we're talking about, like, network edge devices. This forces their replacement. So there's a little bit of an incentive on the part of the providers of the hardware to take their support away. That is to, like, create a limited support because they know that all federal agencies are going to be forced to purchase new equipment that is going out of support. It's much better to have the existing equipment continue to be supported. But it occurred to me the flipside of this is, well, we're, you know. And they'll come up with some BS, you know, well, you know, technology's moving quickly, so we needed to reduce our support window from its previous, you know, 72 months down to 24 months in order to, you know, make sure that the hardware is able to, you know, operate blah blah blah blah. It's like, okay.

Leo: Well, then there's a solution, which is in the acquisition requirements that they put some specific, like, you must support it for...

Steve: Yes, minimum lifetime, yes.

Leo: Right. And I feel like they should be able to do that, as well.

Steve: I'll bet that's already in there. I'll bet that's already, like, you know, it's got to be like five years, you know, guaranteed minimum support. If you want the contract. If you don't, people will buy it somewhere else.

Leo: If you want the government to buy it, exactly.

Steve: That's right.

Leo: I think that that's not unreasonable. I hope that would already be in there, to be honest.

Steve: I'll bet it's in there.

Leo: Yeah.

Steve: You know, what is in here...

Leo: Coffee?

Steve: That's exactly right.

Leo: And a commercial.

Steve: Not that I'm slowing down, but I can always use a little more caffeine.

Leo: Coffee and a commercial, it's a new thing that we've invented here at TWiT, and we invite you all to partake. Well, the river of coffee has sluiced its way into Steve's brain, and he is ready to continue.

Steve: And it's taken effect.

Leo: Yes.

Steve: So Jason Grimard said: "Hi, Steve. You mentioned on this week's podcast how annoyed you were whenever Windows 11 was updated, and you would receive a full screen page after every major update, the one that asks you to turn backup on and other crap," he wrote. "If you haven't already, you need to turn off experience or whatever they call it now under system notifications." And he provided me with a screenshot.

Now, I appreciated Jason's tip, although in my case this was occurring on two Windows 10 machines, one of which I only fire up once a week for the podcast recording with you, Leo. I had wrongly assumed that the continual annoyance from these Windows 10 machines was due to my having logged on under my Microsoft account, rather than using a local account. And maybe that plays a part. But Jason provided a screenshot from a YouTube video showing settings which, under Windows 11, would allow this annoyance to be turned off.

During the year and a half of development work on and testing of the DNS Benchmark, which is Windows hosted, I've seen how many of our development testers have made the move to Windows 11. So I get it. You know, as I've mentioned, I'm going to be setting

up a new system once I move into what I've referred as "our final resting place," which bothers my wife.

Leo: Steve, we're not done yet, Steve.

Steve: So anyway, I've given the question of whether I'm going to be moving to Windows 11 or remain with Windows 10 quite a lot of thought. 11 is visually lovely. I will freely give it that. And its user-facing desktop behavior changed enough from Windows 10 that I did need to spend some time with it during the development of the Benchmark's UI changes to keep it to behave in all of the strange things that Microsoft has done. They've got weird docking stuff now that tends to override what the Windows wants to do. So I've spent some time in Windows 11. And I fully appreciate that most of the world is going to be moving to 11. But I've determined that I will not be. I just don't see anything there that I need, and I don't see any benefit.

So the reason I've mentioned all of that is that the annoying behavior that I was complaining about was under Windows 10, which is where I'm going to end up being for the rest of known time.

Leo: Aren't you worried about support, though? End of life support? I mean...

Steve: No. That's all overblown. I mean, I'm happy on Windows 7. I'm talking to you with Windows 7 in front of me, Leo, and that ended a long time ago. And besides, the browsers stay supported, even if the platform support stops. And the browser security is really more important than the Windows platform. And you continue to get AV updates, regardless. And there's still a huge Windows 10 inventory. You know, as we know, Microsoft extended it another year out of pressure over the fact that no one was ready to have it end. We don't know what's going to happen next year. So 10 is still under support as we speak, until, what, next March or something sometime.

So I was curious, after seeing this feedback from our listener, so I went looking to see whether the same or similar control panel settings that Jason's YouTube that he pointed to me depicted for Windows 11 existed under Windows 10. Yes. It was with some joy that I found them. Under Windows 10, open the control panel and choose "System." Then, in the subsections column on the left, select "Notifications & Actions," which was the third item down for me. And there, on the right-hand side, are exactly the settings you want. One is "Show me the Windows welcome experience after updates and occasionally when I sign in" - right, Leo, like we want that.

Leo: Why? Who cares?

Steve: I know, "...to highlight what's new and suggested."

Leo: Oh, it's an ad, in other words.

Steve: The second one is "Suggest ways I can finish setting up my device to get the most out of Windows." And then there's a little bonus third one, "Get tips, tricks, and suggestions as you use Windows." Okay, I've been using Windows since before, like, you know, when it was an app...

Leo: Before the guy who wrote this, I can guarantee you, was born.

Steve: Yes. Since it was an app you launched under DOS when you wanted to run it.

Leo: Windows 1, yeah.

Steve: Yes. Needless to say, those three are now all turned off. I'm so happy to know they're there. When I'm setting up that new machine, this may not be the first thing I do, but it'll be during the first session I'll be turning all that crap off. So thank you, Jason. I just wanted to share this with everybody else. I know that a lot of people have gone to 11. Well, you can turn that off under Windows 11 also. And everybody who's decided to stay with 10, you know, it's there also. So yay. Thank you. Never need to see that again.

Liviu Sas said: "Hi, Steve. In some countries, the ISPs are required to keep track of subscribers and their IP address for copyright infringement enforcement. And that works also for CGNAT subscribers." In other words, Carrier Grade NAT. I talked about this last week. He said: "The ISP will log every source port block allocation and IP address allocation. This way they can always use the source port and source IP to identify a subscriber."

So the listener corrected, and frankly dashed my hope, which I mentioned last week, which was that perhaps ISPs who were using Carrier Grade NAT, and are therefore assigning private IP addresses to their subscribers rather than giving them public IP addresses, might not also be able to provide real-time identifying information for sale to external advertisers and others. Since it could technically be done, as we know, and as this listener pointed out, unfortunately it looks like it probably is. That means that receiving a non-public IP from an ISP cannot be assumed to provide any additional privacy.

So anyone who wishes to strongly prevent their ISP from being able to identify them to anyone external, for whatever reason, will need to use a VPN of some sort. When any true VPN is used, the user's public IP will be allocated from among a block that's been assigned to the VPN provider. And any reputable VPN provider will refuse to retain any logs which could be used to map their public VPN IP to the IP assigned by their ISP. And their ISP will, in turn, only be able to see that their subscriber was using a VPN, for whatever reason, without having any idea what they are doing on the Internet beyond that.

Now, I used the phrase "any reputable VPN provider," and I hope everyone understands that I did not forget to use the word "free" in that phrase. The terms "free" and "reputable VPN" cannot appear together. Providing and operating a VPN service costs real money which someone needs to provide. If the users of a VPN service are not footing their own bill, then the VPN provider must be somehow arranging to monetize their users' use. That should make anyone who cares about their privacy and security extremely nervous. If I needed to use a VPN, I would not be using a free one. And as we know, there are high-quality, reputable VPNs in the world that explicitly do not log what their subscribers and users do. So there are definitely good solutions if you are worrying about ISP spying. And we have no clear knowledge that that's even going on. It's just obviously a possibility.

Brendan McGoffin said: "Hey, Steve. I'm sure you've been inundated with requests to talk about OpenClaw and its crazy security implications. And also AI changing by the day coolness. Hope to hear your take on this. Specifically would be curious, not just if it's

good or bad, but how you would build this out in the most secure way possible." He said: "I've built out a VM on a Mac with UTM and giving it minimal contact. But thinking of giving it a dedicated box with WAN access, but not local access to other devices, unless to specific hosts I grant it access to. Thanks, Brendan."

Okay. My first response to the OpenClaw phenomenon is to view it with interest at arm's length. For me it's just entertainment. One of the things I first said when we began talking about AI here was that anything we think we know and any statement we might make needs to be time and date stamped because it will have a half-life of a few weeks at most. And that turns out to have been a bit prescient since, as I mentioned last week, the pace at which everything is moving has never let up. I mean, even the people who are involved in this are astonished by how quickly it's moving.

In this case we have, you know, the most recent fad du jour is OpenClaw. I'm a spectator, so I have no definitive response because I have no way of knowing what's going to happen anymore than anyone else does. I've seen massive rockets on the launch pad ignite their engines and begin to rise. There's a great deal of temptation to begin cheering. But I've also seen those stunning examples of human engineering suddenly and quite dramatically explode into massive fireballs.

So now, whenever I watch any huge rocket rising, I consciously hold my breath, and I wait a good while until the chance of the rocket's, as it's now termed, "unplanned spontaneous disassembly" seems far less likely to occur. There are just too many things that can go wrong, and so many ways for a machine like that to fail. And with a rocket like that, this is a machine that's completely understood and was carefully designed, constructed, and tested every step of the way.

By comparison, what I understand of OpenClaw strikes me as completely insane. Those who have made it their business to understand the practical security implications have run screaming for the hills over the idea that OpenClaw's users are allowing these barely understood agents to have access to hugely personal and private data, and even to be talking with one another and sharing skills.

So last Friday, Kate O'Flaherty, a senior contributor for Forbes, wrote about all of this. She wrote: "OpenClaw - the viral AI agent that's already been known by two other aliases, Moltbot and Clawdbot - is growing in popularity," she wrote. "After bursting onto the mainstream just weeks ago, OpenClaw has earned well over 100,000 GitHub stars. Then came Moltbook, the Reddit-style social network where AI bots can interact with no humans allowed. Everyone was talking about it, and for good reason. It's no surprise that concerns about OpenClaw and Moltbook are growing, with worries centering on the security and privacy of the viral bot and, in Moltbook's case, the uncontrolled nature of the AI bot-controlled social network.

"Computerworld's Steven Vaughan-Nichols says: 'There are only a few itty-bitty, teeny-weeny problems with OpenClaw. To do useful things like reserving your hotel room, getting your pizza delivered, or cleaning up your email box, it needs your name, password, credit-card number, and all the other things any crook also wants.'

Okay. So here's everything you need to know about the viral agent now known as OpenClaw. She writes: "OpenClaw, aka Moltbot, is an open-source autonomous AI assistant that you can download and run on a computer. After its setup in November 2025" - or startup in 2025 - "it was known as Clawdbot, but its creator, developer Peter Steinberger, was forced to change the name to Moltbot after Anthropic objected due to similarities with its Claude chatbot. He then changed the name again to OpenClaw.

"OpenClaw is designed to perform real-world tasks on behalf of users, such as managing calendars, messaging, browsing, and other actions that go beyond simple chatbot

responses. Louis Rosset-Ballard, team leader at Pentest People, explains: 'OpenClaw runs locally on devices, and in many configurations can read and write files, execute script, and interact with external services when given sufficient permissions.'

"Nash Borges, senior vice president of engineering and core AI at security firm Sophos" - of course we talk about Sophos often - "describes OpenClaw as 'more like Jarvis from Iron Man than Siri or Alexa.' You use natural language for every interaction, but can ask it to do things such as conduct research on a topic of your choice, compose a reply to an email summarizing when you're available for a meeting, or even code up any capability that it doesn't already have. Borges says that last part is significant because it means there's almost no limit to what it can do.

"But does it work? Reddit users describe their experiences as mixed. According to one post: 'Clawdbot'" - back when it was called that - "'Clawdbot is like an Apple product: when it runs it's like magic, until it doesn't. If you didn't know about OpenClaw a week ago, you must have at least heard of it by now,'" she writes. "Sophos Borges says the whole development journey has been insanely fast, and this explosion of interest is 'just the latest gear shift.'

"'OpenClaw's rapid adoption is driven by demos showing extreme productivity gains, automating tasks that normally require human interaction,' says Malwarebytes threat researcher Stefan Dasic: 'The promise of a powerful, locally run AI agent without obvious limits has resonated strongly within developer and AI enthusiast communities.'"

Okay. So I'm going to interrupt Kate to note that because OpenClaw runs on local hardware, Mac Minis quickly sold out as people rushed to obtain little standalone AI agent machines. Linux and Windows boxes can also run OpenClaw, but the Mac Mini does this particularly well in a very small form factor.

Anyway, Kate continues, writing: "But things that grow so fast often come with risks. Erich Kron, CISO advisor at KnowBe4 says: 'It seems that in just a couple of days, everybody doing anything with AI, and even many who don't, have installed and raved about this new agentic product. The almost feverish rush to use this product is frankly a little disturbing.'" So she asks: "Why Is OpenClaw a Risk to Security and Privacy? Uncontrolled AI is a concern more generally, and OpenClaw is no different from other products that have shot into the mainstream, such as ChatGPT.

"'A concern with OpenClaw is how much information it can have access to when using it the way people are showing,' says Kron. 'For example, giving it full access to all your emails may seem fine and might make sense since you want it to act as your personal assistant. However, there is real danger, not just from malicious use, but accidental, when giving AI agents this type of access. In the blink of an eye, it could be deleting your emails, or taking malicious actions such as siphoning off data to attackers.'

"Security issues are already starting to surface. Denis Romanovskiy, chief AI officer at SOFT-SWISS, a provider of tech solutions for iGaming, said researchers have found hundreds of exposed Moltbot instances online with 'zero protection.' This includes API keys, private messages, the ability to send messages as the user, and root shell access.

"William Thackray, IT and cybersecurity expert and operations director at AGT, said 'OpenClaw is a security threat on multiple levels. Firstly, the platform's GitHub repository reveals a troubling accumulation of unaddressed security vulnerabilities, from an exposed database, creating a direct pathway for unauthorized access to user information, to dangerous plugins. Koi Security documented 341 malicious skills uploaded to ClawHub, OpenClaw's extension marketplace.'"

So, yeah, what was that about "spontaneous unplanned disassembly"? Forbes says: "'Granting an AI agent full system control creates a single point of failure,' says Dasic. 'If compromised, OpenClaw can access saved passwords, personal documents, browser sessions, and financial data.' OpenClaw poses risks to privacy, too. 'These stem from its access to and storage of sensitive user data,' says Rosset-Ballard. 'Because the agent may retain long-term memory, store credentials and tokens in plain text, and process external inputs without robust guardrails, it can inadvertently expose personal information.'

"At the same time, the AI agents post on social networks without asking permission. Romanovski points out: 'Screenshots of agent conversations spread across Twitter. Your entire digital life sits one vulnerability away from exposure.'"

Okay. And we were all worried about Windows Recall, which now seems kind of...

Leo: That pales; doesn't it.

Steve: It now seems kind of quaint by comparison.

Leo: Rather quaint, yeah.

Steve: Okay. So what about "Moltbook"? Kate writes: "Moltbook is a social network built exclusively for AI agents, launched last month. Dasic says: 'Unlike traditional forums where users interact and share content, Moltbook is a space where OpenClaw agents autonomously post content, comment, argue, joke, and upvote or downvote each other.'" Which, Leo, this just sounds like sci-fi to me. "Human users can observe agent interactions, but cannot directly participate."

"Professor Katerina Mitrokotsa, chair of cybersecurity at the University of St. Gallen, said: 'Moltbook further amplifies the risks associated with OpenClaw. Although it gained attention for showcasing AI-to-AI interactions, early findings revealed that it exposed entire databases, including secret API keys that could let attackers impersonate any agent on the platform. This creates clear threats for users: Identity spoofing, unintentional data exposure, and reduced control over their digital environment.'

"Daniel dos Santos, head of research at Forescout, said: 'The risks of Moltbook became very clear very quickly. There is no moderation on the content, so bots can post instructions for other bots to execute ultimately on a victim machine, can use prompt injection attacks or generate offensive content.'"

Leo: We've learned that much of the content on Moltbook now is generated by humans.

Steve: Ah. So spoofed AI.

Leo: Spoofed, yeah. It's not hard to do that. And that makes it even more risky. I think you're much more likely to get prompt injection from a human than from another AI.

Steve: Yup, exactly. Kate finishes her coverage of this for Forbes by addressing the question: "Should we use OpenClaw?" writing: "OpenClaw might have some cool capabilities; but for now, the risks outweigh the benefits, especially if you are not tech-y. OpenClaw's creator Peter Steinberger has warned users that the tool requires careful configuration and is not yet meant for non-technical users. Romanovski says: 'If you're technical, curious, and willing to sandbox everything carefully, it's a fascinating glimpse into the future. But if you handle sensitive data or need reliable security, stay away for now,' he advises. 'The project moves faster than its security can keep up. Treat it as an experiment, not a production tool.'

"And Kron warns: 'If you do choose to use the viral AI agent, be careful that you are discovering the real deal. When searching for a product like this to download and install, it's very important that people are careful not to end up in an unofficial repository that contains malware or other dangerous programs.'" Kate concludes: "OpenClaw is growing at an alarming rate, making it important that you treat it with caution. Unless you're an expert, leave it well alone for now."

And Leo, I know you have had fun playing with it.

Leo: I love it.

Steve: And I agree with you. I think it is very clear that the next evolution is agency, is agents. And not just one, but teams.

Leo: Well, and that's actually what we're learning from OpenClaw is that there's a demand for this, that there's a lot of interest in it. And I imagine there are a number of companies starting up right now that will offer that kind of agentic AI in a sandbox. The problem is, you know, you can sandbox it. I set it up at first on a VPS. You know, but no matter where you put it, eventually you've going to want to give it access to your Google mail and your contacts and address book. And frankly, I was going to give it a credit card with, you know, like a \$5 a day limit because the really, you know, interesting uses all require that it act on your behalf agentially. So even if you sandbox it, it's inherently insecure. Obviously, nobody at a business should be using this, although many businesses are because there's a lot of interest.

Steve: No kidding?

Leo: Oh, yeah.

Steve: No kidding. Wow.

Leo: I think some of it is just let's take a look at this because how, you know, what can we - how can we make this work for us? This is...

Steve: Right.

Leo: Once you start using it, and part of this is, you know, you can use any AI with it, doesn't have to be Claude, and most people are using Claude because Claude has

this great personality. I hate to admit this. But you really enjoy interacting with it. The thing I most wanted to do with OpenClaw was be able to text message back and forth with Claude. And the other thing that it does that's really interesting is it will run overnight, run all the time. So you can say, hey, as some have done, come up with something interesting. Have at it. Let me know. And it will surprise you. I love that idea. I think it's hysterical. There's a new saying in the AI community. Just YOLO it. You know what YOLO is. You Only Live Once. Just YOLO it. Just you only live once. Have fun.

Steve: Yeah, you know, that bungee cord is a little frayed, but it's probably good for one more jump.

Leo: You only live once.

Steve: That's right.

Leo: No, it's insane. It's of course a security nightmare. Of course it is.

Steve: Yeah. And the good news is, here's what I would tell people. You're right, Leo. This surprised the world, much like Large Language Models did a few years ago. Look what we have now. Just now that we've understood how that can be taken to agency...

Leo: Exactly.

Steve: ...and that's going to happen, no one has, you know, I would argue wait, and you probably won't have to wait that long.

Leo: No.

Steve: Because...

Leo: Like Sprint.

Steve: Yeah.

Leo: Yeah. That'd be instant. People are working as hard as they can on that right now.

Steve: And I have to say, though, I said it earlier on this podcast, I don't know how you control this. And that's the problem. And you put your finger on it. It needs the freedom to misbehave in order to behave.

Leo: Right.

Steve: It's like in order to act as you, it needs to be able to impersonate you.

Leo: You've got to give it all the credentials. I'm trying to figure out how to give Claude my SSH keys because I have to - every once in a while it says "I can't do this, you're going to have to sudo it yourself. I don't want to do that. You do it." Oh, but Steve. We're watching a little miracle happen. We really are. I've never been this excited about anything in technology like this. Even the Internet. This is something very special that's happening. With huge risks. I'm glad I have you to keep me on the straight and narrow, Mr. Gibson.

Steve: Let's take our second-to-last break.

Leo: Okay.

Steve: We'll finish up with feedback, and before we get on to our main topic.

Leo: Will do. Actually...

Steve: Again, people, I would say it's really, I mean, calling it the "Wild West" understates it. It's, you know, bungee jumping with a frayed bungee. I don't know how it's ever going to be safe enough, but it's going to get safer.

Leo: Er, right.

Steve: So, yeah.

Leo: You know, I have friends who work in AI. And they, because they work in the business, they've had access to stuff like this months ago. And all they've been telling me for the last three years is "You have no idea how weird it's going to get." And now I'm starting to see what they're talking about. It is getting very interesting. And I don't think our models - I don't think we know what to do with this. I think this is going to be we're living in interesting times. And just be careful out there. But remember, YOLO.

Steve: Well, and as we mentioned last week, the large software companies got hit because of the concern over code automation.

Leo: Yeah, some people were saying, you know, it's a bubble. It's going to crash. I think we had the crash. What was it, it was hundreds of billions of dollars in market value disappeared in an hour.

Steve: Yeah. Some came back. But still, it's like, you know...

Leo: Something's happening.

Steve: ...the investors said, whoa, wait a minute.

Leo: Yeah. Something's happening.

Steve: Maybe everybody - just tell your bot, you know, I'm annoyed with Windows. Just write me a new one. And leave out all the Microsoft crap. And, you know, it gets busy.

Leo: Yeah. Well, that's what Anthropic spent two weeks writing, a C compiler. It wasn't a very good C compiler, actually. Claude did that completely autonomously. But it could compile the kernel, the Linux kernel. So, you know, it's getting there.

Steve: So Kyle O's email subject was: "My First App - Made with AI."

Leo: Aww.

Steve: And I know, Leo, you're now an app a day.

Leo: Oh, easy. I think of something, and I have it in half an hour.

Steve: He said: "Steve! After listening to you and Leo talk about coding with AI and Claude, I added an item in my to-do list to learn how to code with AI. I never got around to it, until a situation arose where I found myself needing to create my own custom app. I volunteer for a small non-profit, and we have a little library, around 150 books, that are not very well organized. I volunteered to clean up our library and, while doing so, thought it would be the perfect opportunity to also take inventory of all of our books and provide the inventory to our members so everyone knows what books we have available.

"I found a free app for iOS (that I won't mention because it turns out it doesn't work very well). The app scans the book's barcode, looks up the ISBN, and pulls content like author, description, publication date, and creates an inventory of your library. You can then export the inventory to a spreadsheet. It worked great, up until it stopped working. After about 30 books, all additional books scanned were 'not found,' and the app failed to inventory them.

"So I have a list of over 100 ISBNs, and no app to generate this inventory. Rather than learn about coding with AI through videos and instruction, I downloaded OpenAI's Codex app for Mac and threw myself in the deep end."

Leo: Okay.

Steve: He said: "(I would have used Claude, but I already pay for ChatGPT.)" He said: "I told it I wanted a Mac app written in Python with a GUI interface that takes a given ISBN, looks it up on Goodreads, provides me with a preview image of the book so I know it's the correct one, and then adds it to the list. After I do this for all my books, I want a CSV

format file export button that provides" - you know, CSV, Comma Separated Value - "a CSV containing the author, an image of the book, publication date, page count, and description.

"There were some errors and issues. For one thing, CSVs cannot contain images in their cells (an oversight on my part), and for some reason the author's name was listed twice in its cell. I told Codex the issue, and it created an Excel export button and fixed the author issue. When I attempting to open the file, Excel said the file was corrupted. I told Codex, and it fixed whatever the issue was. The app now works flawlessly. I get a clean Excel export that lists an inventory of our small library of books. I am stunned by how simple this all was. There were some other hoops I had to jump through (my Mac did not have the latest Python installed, for example), but it was relatively simple to get all set up and working.

"I do have some concerns. I am a cybersecurity analyst, but not a developer by any means. Watching Codex effectively say, 'I'll handle that,' while code and commands whizzed by my screen made me feel a bit nauseous. When I had an issue and Codex said, 'Just run these commands,' I was hesitant to do so because I didn't know exactly what the commands were doing. Then there's the package manager. It used PIP to install 'Beautiful Soup 4,' 'Pillow,' and 'openpyxl.' I don't know what these are and what they do, and that makes me a little nervous. Especially after learning about the attacks and compromises on open source repositories.

"I think what Codex did was overall safe, and the project was a huge success. I have no formal developer training (I took a Python class in college if that counts), yet this created a fully functioning custom app for me in under 30 minutes. Thank you and Leo for discussing developing with AI. This gave me the confidence to jump in the deep end and create this app. Appreciate you both. Kyle."

So Kyle has shared a perfect use case for today's code-generation AI. Thinking about this, the best analogy I have for this is the similar breakthrough that was created by the invention of the PC-driven spreadsheet. To me this feels like the introduction of the spreadsheet because more than anything the invention of the spreadsheet was empowering. Non-programmers were able to suddenly leverage the power of a personal computer - as a matter of fact, it's credited with what, you know, saved Apple and the Apple II. People were buying Apple IIs just to run VisiCalc.

Leo: It was the killer app, the first killer app.

Steve: Right. So nonprogrammers were suddenly able to leverage the power of a personal computer in a way they never could before. You know, they may still not have been able to author programs themselves from scratch, but the spreadsheet meant that they could get meaningful and useful results without needing to. They were able to model data themselves.

Kyle took a Python class in college. But he's explained he's not a coder. Yet, thanks to, in this case OpenAI's Codex app, on a Mac, Kyle is now in possession of a custom app that does real world work to solve a problem he had. And we've also witnessed, Leo, you, who are a coder, you effuse no less enthusiastically over the successes you've had, first with that test project creating, like scanning the Internet for topics for podcasts.

Leo: For stories, which I use every day now.

Steve: Yup.

Leo: And I generate briefings with it for all our shows. I think it's improved our shows dramatically.

Steve: I've seen the difference in those.

Leo: They're tighter. The hosts are better prepared. It's great.

Steve: Yup. And we know that you, Leo, could have painstakingly written a program, because you're a coder, you could have done what you needed to do, you know, under pre-AI coding paradigm. But the effort was not worth the reward.

Leo: I never did anything. For 20 years I didn't do it.

Steve: And that's what Claude Code has changed for you is that, you know, you're now - you're using your understanding of coding, and with Claude Code provide the leverage to dramatically shift that work versus reward tradeoff in favor of easily and readily, even joyfully, producing applications that are of real use.

Leo: I'll even go farther than that because I use also Claude Code to configure all my systems. I was always saying just set up a new system. And it knows, it reads the manual so I don't have to. It does the settings. I could figure all of that out. But it's brought a huge amount of pleasure in computing to me because I can be so much more effective and efficient. I can have tools that simplify things. Kyle's example is a really good example. There's no way that that was a security issue.

You know, the worst thing that could happen is maybe he'd accidentally DDoS Goodreads by making too many requests a second for it or something like that. There was no risk in creating that application. And his experience, by the way, that's exactly what it's like. It's not perfect the first time. You try it. You say, hey, well, that's - but it's so easy to tell it, you did the name twice, what's going on? And it fixes it. And so you go through this debugging process. It is like pair programming. But it's at a very high level.

Steve: And it's conversational.

Leo: It's conversational. I think your analogy to VisiCalc is exactly right. Really it's the history of computing as we've gotten higher and higher level languages. This is just the highest level language. It's finally English. And I think this counts. I really do. I think it's great. You know, and obviously you wouldn't want to write router firmware with it, although I think people are. You know, there are certain things that you probably shouldn't use an AI to write.

Steve: It's going to be interesting to see what happens because I agree with you. I think, knowing people, they will use it for everything.

Leo: For everything. They already are.

Steve: It's just simply - it's just going to happen.

Leo: But I think there's so many harmless applications that are just quality-of-life applications. You know, one of the things I've been struggling with since I got these little album art things that are behind me, they're called Pixoo from the company called Divoom. It's a Chinese company. It's a silly little device, and they have the worst app on an iPhone to manipulate it. And every day I've been clicking, you know, you'd often see, you'd get on, and the wrong album, or I'll be clicking and stuff. I wrote very quickly, in about a half an hour, an hour. I think I was watching the football game on Sunday. I wrote a program to do this.

It turns out these are just - this is just an HTTP put, and it's in REST format. It's a very simple thing. I probably could have written it. I'd have to look up the API and figure it out and stuff. Be trivial to write. Now it's instant, and I have a command line. I wrote a little bash shell command line that sets it like that. I can put any art up there. Now, that's, A, harmless, you know, there's no security issue here. B, a huge quality of life. C, yes, totally doable if I were willing to spend the time. But most importantly it was easy. And it made a big difference in my operation. And so I'm finding more and more things like that.

There's risk. I also wrote a tool that lets me find and turn off and on services on my system. Now, it turned out not to be such a good idea because I turned things on I shouldn't have, and I turned things off I shouldn't have. But it was fun. And now I'm a little more cautious with what I turn on and off in the background.

Steve: Yeah. I think this is, like, the real deal. This is not a fad.

Leo: It's very exciting. No.

Steve: Yeah.

Leo: And it's just the beginning. And it is exactly where it should be. It's beginning the computer talking to the computer. Of course. That's natural. How much farther beyond that it'll go, I don't know. I don't want it to write novels. I don't want it to write musicals or make movies, probably. But for talking to a computer, there's nothing better.

Steve: Yeah. Okay. Last break, and then we're going to talk about Least Privilege.

Leo: Okay. I've been using up all your time. I'm sorry, Steve.

Steve: No, no, no. I wanted you to because I sort of assumed that this was going to be an engaging topic for both of us.

Leo: Yeah. Oh, you know, I can talk about this forever. And I know people are wondering if you could use it to write the apps you write. And I wouldn't want you to use it to write the apps you write. But you wouldn't want to, either. You like doing what you do; right?

Steve: Okay. So I actually did have something in the show notes I was going to skip over, but I will share it.

Leo: Would you, please?

Steve: A number of our listeners have asked me whether, and if so, how this revolution in AI coding might affect my own work. My best assessment is at the point it's not clear. And, you know, if nothing else, it's way too early. In general, I eschew the use of tools that do not produce the same quality result as I'm capable of producing. I'm just unwilling to compromise. I just don't see the need. For example, I'm still authoring all of my web pages at GRC.com by hand. Because I've seen...

Leo: You might want to consider asking for some help.

Steve: I've seen the utter crap that even the best HTML and CSS WYSIWYG authoring tools spit out. And I just can't abide by it. It's just - it's horrific-looking crap. And it's like, no, I mean, yes, I know that mine look like they're from 1995. But, you know, they also download instantly.

Leo: Right.

Steve: Now, it happens that there are savings that add up. Having super-lightweight web pages means that GRC's little 100Mb connection is able to easily serve the world's needs without breaking a sweat. The main GRC.com server has 24GB of storage. That's not RAM. That's total mass storage. And it's not even full. I mean, it's like a third full. That means that GRC's entire website can sit cached in RAM. And it's easily served by a single CPU that's not particularly fast. You know, I understand the modern way to solve problems is just to throw more and more resources at the need until whatever it is goes fast enough. But the truth is recurring costs really do begin to escalate. And once you take that path, there's no turning back.

So I'm not saying there's anything wrong with that. I get it that that's the most efficient approach for most situations. But that's not for me. I'm obviously not into efficiency, except for my code. So I'm going to be very excited to follow along with these breakthroughs in coding technology. But I don't expect it's going to affect the way I code my own stuff. I do it, you know, it's like, numerical control machines appear that are able to do woodworking, but I'm still in the basement with a chisel.

Leo: Yeah. Right. Because you enjoy it.

Steve: Because I just like it, yes.

Leo: You should.

Steve: I love the art. I love the craft.

Leo: If at any point - and I think a website might be a better example. Because if at any point you got tired of that, you don't have to use React and Angular. You can have an AI generate - this is the website, remember that briefing tool that I created. This is the website that it generated. So for every show I create a page like this. It's HTML. It's not super complicated. It is very fast and light. You don't have to be doing a big JavaScript thing at all. This is generated every night. That's why there's only a few stories for TWiT.

Let's do Intelligent Machines, which is coming up tomorrow, so this will have more stories in it. It does AI summaries of the stories. It has the link. And this is designed for the other hosts to read. I call it a briefing book. That's as light as it can be. There's no JavaScript. There's a little bit of CSS probably to style it. But it's very, very simple, using plain HTML. So you wouldn't have to make - you could even tell it, make the site look like it was designed in '98. It would do it. You could say "No JavaScript. I don't want any React. I want it to be instantaneous. I want the lightest possible site." It would do what you tell it to do.

But if you enjoy it, there's no reason for you to do that. It's only if it would be something that you didn't want to do or you didn't have time to do that you might consider it. I'm not trying to talk you into it. I love it that you do this stuff by hand, and I hope that people will continue to do that, by all means. I don't want to see the world filled with vibe coded slop. That would be terrible.

Steve: It's going to be interesting to see what happens when, like, when people are deploying code that they didn't write. I mean, that's what I, I mean, I'm going to put my name on it. You know, I don't ghost author novels because I want, if it's my - if my name is on it, it's from me. And I can't imagine shipping something, like, of code, that I didn't write, you know, that I dictated. That's like, no. I just...

Leo: That's why on all of my - I put a lot of the stuff up on GitHub for other - if people want to look at it. And in every case I say it's generated by Claude Code. I don't...

Steve: Yeah.

Leo: Claude does it itself. It says: "Built for personal use. Entirely vibe coded with Claude Code." I make sure that that's clear.

Steve: I think that's very cool.

Leo: Yeah. And by the way, it writes all this documentation, too. Which, trust me, no one wants to write documentation. It does a very nice job with that. So, I mean, I don't intend for anybody to use it but me. This is, to me, this is the stuff I'm writing for myself, not for anybody else. But I post it just if people are curious because we talk about it all the time. Well, and there's also - it incents you to do that because

that's also - Claude wants to store stuff on GitHub for some reason. And so I go along with it [crosstalk] and everything. Okay, sure, whatever. Tell the world.

Oh, we were going to take a break. Did we take a break? No. You're watching Security Now!, and that there's Steve Gibson. I'm Leo Laporte. We're glad you're here. We're especially glad our Club TWiT members are here. Thank you for making this show possible. We really appreciate it. On we go with Least Privilege.

Steve: Okay, now, this is a little bit of a thinker for people. May not seem like it is. But it kind of happened as I was working on the story about Coinbase. So I think this is useful. The topic evolved as I was expounding upon the larger lesson to be learned following BleepingComputer's report of the second insider breach at the U.S.'s largest publicly traded crypto exchange, which, you know, is Coinbase.

As I'm always interested in doing, I wanted to draw some conclusions from the underlying cause of the second breach, and I wound up confronting one of the simplest, most well known and well understood principles of security, which is simply known as "Least Privilege."

The concept of least privilege couldn't really be any simpler. It simply means not offering any more rights, or privileges, than are required to perform a specified task. Simple; right? But if the concept is so simple, why is it that we as an industry and users of this technology so often fail in the application of Least Privilege? If it's simple, it should be easy to do.

The reason why we as an industry and as users so often fail in the application of Least Privilege is that "least privilege" is also "least convenient." The sad and sobering truth is that today, as mature as our theories of security may be, and I believe our theories are very mature, we remain in denial about the need to apply those theories everywhere. We know how to make our systems far more secure than they actually are. You know, where doing that, where making them that secure might inconvenience us, we still choose convenience over security, and we hope it'll be good enough.

Okay. So with that preamble, let's look at a case in point and see what more might be learned. We've talked about the trouble companies are having, right, with this new practice of BPO, that's the new jargon, Business Process Outsourcing, which is the latest in business fashions. In the same way that so-called "pop-up" restaurants have been created, the idea is that it's now possible to also have "pop-up" corporations. A couple of people who share an idea pitch their concept to an angel investor to raise some seed capital. Then, rather than embarking upon a hiring campaign to find and employ the wide range of talent and experience that they'll require, they instead assemble their operating enterprise like LEGO blocks, from an array of now-available online services.

The problem with this is trust. The resulting "virtual enterprise" lacks any core loyalty because, to all of the various third parties that have been commissioned, the commissioner is just another one of their many client customers. There cannot be any sense of institutional loyalty because there's nothing to be loyal to. Clients are just account numbers and API linkages. It really is a very different way of organizing and operating. You essentially get throwaway enterprises.

So it's against this backdrop that BleepingComputer brings us the news of another insider breach at Coinbase, originating from Coinbase's use of Business Process Outsourcing. BleepingComputer wrote: "Coinbase has confirmed an insider breach after a contractor improperly accessed the data of approximately 30 customers, which BleepingComputer has learned is a new incident that occurred in December. A Coinbase spokesperson told BleepingComputer: 'Last year, our security team detected that a single Coinbase

contractor improperly accessed customer information, impacting a very small number of users, approximately 30. The individual no longer performs services for Coinbase. The impacted users were notified last year and were provided with identity theft protection services and other guidance. We have also disclosed this incident to the relevant regulators, as is standard practice.'

"BleepingComputer," they wrote, "has learned that this is a newly revealed insider breach and is not related to the previous disclosed TaskUs insider breach in January of last year. This statement comes after the 'Scattered Lapsus Hunters' cybercrime group briefly posted screenshots of an internal Coinbase support interface on Telegram and then deleted the posts soon after.

"The screenshots showed a support panel that gave access to customer information, including email addresses, names, date of birth, phone numbers, KYC (Know Your Customer) identifying screenshots" - like their identities, right, their driver's licenses - "and stolen data to be passed around among different threat actors before being leaked or disclosed. So it's unclear whether this group was behind the insider breach, or whether other threat actors carried it out. However, the same threat actors previously claimed to have bribed an insider at CrowdStrike to share screenshots of internal applications.

"Over the past few years," they write, "Business Process Outsourcing (BPO) companies have become increasingly targeted by threat actors seeking access to customer data, internal tools, or corporate networks. A Business Process Outsourcing company is a third-party firm that performs operational tasks for another organization. These tasks commonly include customer support, identity verification, IT help desk services, account management, and so forth. Because BPO employees often have access to sensitive internal systems and customer information, they have become a high-value target for attackers.

"In the past, threat actors have exploited BPOs through bribing insiders with legitimate access, social engineering support staff to grant unauthorized access, and compromising BPO employee accounts to reach internal systems. As we've seen with Coinbase this year, one way BPOs are targeted is by bribing their employees to steal or share customer information. As I said, lack of loyalty to the targeted enterprise.

"Coinbase disclosed a similar data breach last year, later linked to external customer support representatives employed by TaskUs, an outsourcing firm that provides services to the crypto exchange. Another common tactic is social engineering attacks against outsourced IT and support desks, where threat actors impersonate employees and call BPO help lines to obtain access to internal corporate systems.

"In one of the most prominent cases, attackers posed as an employee and convinced a Cognizant help desk support agent to grant them access to a Clorox employee account, allowing them to breach the company's network. The incident later became the focus of a \$380 million lawsuit by Clorox against Cognizant. Google reported that threat actors targeted U.S. insurance firms in social engineering attacks on outsourced help desks to gain access to internal systems. Retailers also confirmed that social engineering attacks against support personnel enabled ransomware and data theft attacks.

"Marks & Spencer confirmed attackers used social engineering to breach its networks, while Co-op disclosed data theft following a ransomware attack that similarly abused support staff access. In response to the attacks on Marks & Spencer and Co-op retail companies, the U.K. government issued guidance on social engineering attacks against help desks and BPOs. In some cases, hackers target the BPO employee accounts themselves to gain access to the customer data they manage.

"In October, Discord disclosed a data breach that allegedly exposed data from 5.5 million unique users after its Zendesk support system instance was compromised. While the company did not confirm how its instance was breached, the threat actors told BleepingComputer that they used a compromised account belonging to a support agent employed by an outsourced business processing provider. Using this account, they downloaded Discord's customer data. This repeated abuse of outsourced support providers shows how threat actors are increasingly bypassing vulnerability exploits and instead targeting third-party companies with access to corporate networks and data."

Okay. So this is a variation on "the call is coming from inside the house." In this case, the call is coming from inside the house of someone you trust. The source of the inherent vulnerability is clear. In order for an external outsourced business process provider to perform their functions, they must be trusted with a connection into the outsourcing entity's network or other business processes. Although they must be trusted, they are not worthy of that trust.

As I noted, an employee of an enterprise has an inherent stake in the company that employs them. We kept hearing about bribery being the way these external companies were exposed. But an employee, as I said, of an enterprise has a stake in the company. They attend meetings with their fellow employees. They look them in the eyes. They may socialize with them after work hours, attend each other's birthday parties or those of their children. They may be on a softball team or have attended explicit team building events. They may share a department where they routinely meet, plan, participate, and work side by side to meet goals. All of those things serve to create a stake in the shared welfare of the organization.

But none of that exists in the hearts and minds of subcontractors to whom that organization is just another account among many. This makes these subcontractors far more susceptible to bribery.

This newfangled restructuring of organizations appears to be irreversible. The days of an employee starting off in the mailroom and gradually working their way up over the course of five decades to finally receive a gold watch and become CEO, those are long gone, and they're not coming back. So how do we make this business process outsourcing work better?

My hope is that everyone is learning from these initial BPO missteps, and that the problems we've seen and that we are seeing are due to what I would call "API over-trust." In the same way that it's easier to just give someone wider permissions to a database than they actually need, it's simpler and quicker to design an API that offers more power than is needed to fulfill a specific outsourced task.

For example, an external BPO which is providing helpdesk services may not need access to a customer's entire record. They may only actually need minimum identifying information and a subset of specific customer history. But when initially setting things up, it's quicker and easier to just give this "trusted" - and I have that in air quotes - third-party unfettered and unfiltered access to the entire customer database. After all, they're under contract; right? What could possibly go wrong?

What we see is another example of the sort of finger-pointing I've been highlighting recently. Whose fault is it if a subcontractor is bribed to disclose their contractor's critical information? The subcontractor is easiest to blame. But the information was still disclosed. The subcontractor, the entity that did the subcontracting, gets blamed for the breach of their systems. The question is whether that subcontractor had more access than they needed because they were able to make that disclosure. Did they only get the bare minimum that they needed which would have better protected the company providing that access?

This excess privileges is not a new problem. Remember that BPOs were once called MSPs. We talked about that years ago, Managed Service Providers. We covered that story of a dental services MSP which had been compromised by a ransomware group. This group struck gold because the way the MSP operated was to require full access to their clients' networks. The ransomware group took advantage of this unfettered network access to install ransomware and encrypt the PCs and other equipment of every one of the MSP's customers. It was a widespread disaster for the MSP and for every one of the dental offices it served.

There was no defensible reason for the MSP to have a fully privileged network connection to each of its clients' internal networks. They didn't need that, but that was the easy path that was taken. If the access had been strictly transactional against a service provided and running on the client's side, far less, if any, damage could have ever been done. So, philosophically, this is what must change. Any organization wishing to outsource services must consider the consequences of that service provider becoming a hostile entity. Maybe not by design, maybe by mistake, maybe by compromise. Maybe by an insider, you know, accepting a bribery. It doesn't matter how. The question is what happens if they become a hostile entity. So instead the way to solve that is to design and provide an API linkage that will protect their interests under any circumstances, no matter what their contractors might do.

A familiar example of this sort of function - because we know how to do this; right? A familiar example is an HSM, the hardware security module, whose internal write-only private key and machinery can be employed to sign a file, while at the same time nothing and no one can exfiltrate and steal its secrets. The analogy is not perfect, but the point I want to make is that designing with the concept of least privilege is what should always be done. Always. In the HSM example, there was no need to allow the device's internal private key to ever be exposed, no matter how much the user of that key might be implicitly trusted. Thus the key should never be exposable, not because it would be stolen, but because it could be.

I've talked a lot about not exposing any non-public service to the public Internet. This is another example where "least privilege" comes into play. When I've said that authentication doesn't work, I've meant that it must not be depended upon to work. I've asked why someone in North Korea, whom you almost certainly don't intend to have accessing your enterprise's network, should even be given the opportunity to challenge your network's authentication system. If you were monitoring every incoming connection, one by one, to the publicly exposed management interface of your enterprise's firewall, and a connection attempt was inbound from North Korea, would you not choose to drop its packets? Of course you would. If North Korea is being allowed to connect to your cloud services, that's not least privilege.

So my point is, even though the concept of least privilege could hardly be simpler and more easily explained - it is a trivial concept - it turns out it's not trivial to actually deploy it in every instance, so it's not something that is robustly deployed in the real world. But it needs to be. I believe it's the only way forward.

Through the years of this podcast I've broadly divided problems into two categories; right? We've got mistakes that are made, it's going to happen; and also, second category, policies that are deliberate. AI-driven code-checking reasonably promises, as we talked about last week, to finally enable us to deliver bug-free code. I would argue AI fixing human errors. We're in a whole different world if it's AI code from the start. I don't put that in the same class at all.

AI fixing human mistakes like we talked about last week, that seems like a near certainty to have happen. And while that's terrifically exciting, it won't cure all our ills because failures to implement least privilege, they're not mistakes. They're policies. They're the

result of decisions that were made. This means that to further improve our delivered security moving forward, we need to make the decision to far more robustly design for least privilege operations. That's how we get where we want to go from a security standpoint and, you know, stop having just breach du jour.

Leo: Is it related to zero trust? It's kind of like the idea of zero trust; right?

Steve: Yeah. Yeah.

Leo: It's basically, you know, give it - but it's just fundamental in security, give it as little as it needs and no more.

Steve: I know. Except that people don't.

Leo: Right. It's a lot easier.

Steve: Some company is in a hurry to get their help desks set up and say hey, here, you know, here's a credential that lets you log onto our database so that you're able to look up our customers. Except if that person, if that contractor goes bad, you've just lost your database.

Leo: They don't even have to go bad. They just - you're only as good as the weakest security practice of any contractor.

Steve: Right.

Leo: Right?

Steve: Exactly. And they have no loyalty to you.

Leo: Right.

Steve: They keep succumbing to bribery because it's like, hey, how much money? Okay.

Leo: But what - I dimly remember the story of a - it was I think an electric company that still had open remote access ports to a former contractor who had left - they never took away their privilege for remotely accessing the system. And, well, of course that's a recipe for disaster. That's just...

Steve: Yeah, yeah.

Leo: I mean, we could do a lot better.

Steve: It's funny, too, because in movies you see people having their credentials revoked the moment that, you know, it's like, you know, give us your passkey and your parking pass.

Leo: Security guard comes, here's your box, put your stuff in there. Yeah.

Steve: Yup. And you absolutely at that point, you know, you want their password to no longer work.

Leo: You know, when we've had to terminate employees in the past, we've done that. And it can be very painful. And the few times that we didn't, we deeply regretted it. And it wasn't out of maliciousness, I don't think. It was more out of just, you know, not paying attention or whatever. And stuff disappeared. And I don't know, it's just...

Steve: Yeah, least privilege is the easiest thing to say, but it's so easy not to do it.

Leo: Yeah, yeah, hardest thing to do.

Steve: But it's so easy not to do it, yes. Least...

Leo: Because we want to trust. We want to be - we want convenience, but we also want to be trusting. But when it comes to security, trust no one; right? Steve taught us that.

Steve: Exactly right.

Leo: Steve Gibson's at GRC.com. That's his website, proudly stuck in the 1990s. But it's fast, and there's no JavaScript. GRC.com. Actually, there's a little JavaScript here and there for a few things you have to do. But...

Steve: Yup.

Leo: Only when absolutely necessary. A few things you might want to check out there, of course SpinRite, the world's best mass storage maintenance, recovery, and performance-enhancing utility. Version 6.1 is out.

Steve: I will say there's no JavaScript library. I wrote it all by hand, of course.

Leo: Key, yes, that's the key.

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

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