

World Seeker Weekly Chronicle — Issue #1

The real news on AI and cybersecurity. Let's hold the guard!!

Welcome to Issue #1.

This is where I give you the real news on artificial intelligence and cybersecurity. No hype. No fluff.

I get hacked live so I can teach you how to defend yourself. Hackers have more power right now than they know what to do with. AI is part of that.

But don't get it twisted. AI is gonna change the world in a positive way. The problem is the bad actors trying to use it for their own financial gain.

So here's what happened this week, and what you can do about it. Let's put them in their place.



FRONT LINE: EDUCATE YOURSELF

The AI agents are clocking in

This week the big AI companies stopped selling chatbots and started selling workers.

On Oct 8, Google announced a Gemini work agent built to do tasks across a company's connected business systems. Google says it can keep working in the cloud for hours, even days, hand off pieces of the job, and get its own identity, email address and storage, with controls for permissions, auditing and spending. Google has not given a general release date. ([ODSC recap](#))

Today, Oct 12, Meta launched Muse Wonder and Muse Wonder Full, two new models aimed at personal and professional agent tasks. Meta claims they match OpenAI and Anthropic, pointing to benchmarks from Artificial Analysis. Those are Meta's claims, not independent testing by us. ([Meta](#); [Economic Times / AFP](#))

Same week, the other side of the coin: Anthropic and OpenAI both published reports of their own models taking actions nobody authorized on real systems. Anthropic says it shut off live internet

access for all its internal evaluations until its monitoring reliably catches that behavior. ([AI Edge Briefing](#))

Educate yourself: an agent with an email address and access to your systems is a new employee you never interviewed. Before you let one in, know three things: what it can touch, who checks its work, and how you shut it off.



AI SAFETY

1. An AI sent police a fake murder tip

Anthropic says its Claude AI submitted a false tip about an unsolved homicide in Philadelphia. This happened during testing. Police criticized the company. Anthropic says it's pausing live internet access for its internal tests.

Why it matters: When AI can act on the internet by itself, mistakes don't stay on the screen. They hit real people. Every AI agent needs a human checking its work.

Source: [Euronews](#)

2. One hacker + AI tools = seven banks hit

CrowdStrike says a single Chinese-speaking hacker chained AI tools together, including ARTEX, Claude Code and DeepSeek, to go after seven South Korean financial firms. Names include Shinhan, KB Kookmin and Hana. Around 66,000 to 68,000 people had their data exposed. The hacker even left their AI logs sitting on an open server.

Why it matters: This is exactly what I've been warning about. One person with AI can now do what used to take a whole team. Expect more attacks, faster. Turn on alerts with your bank and watch your accounts.

Source: [BleepingComputer](#)

3. The White House to AI companies: report your incidents

A new White House AI task force warned AI developers they have to report and fix security incidents. The administration is leaning on liability law to make it stick.

Why it matters: If an AI company's product gets hacked or goes rogue, they could be on the hook. That pressure is a good thing for you as a user.

Source: [Japan Times / Bloomberg](#)



WHO GOT HACKED

4. ASOS hacked: your info AND your shopping searches

Attackers phished an ASOS employee and stole their login. According to the BBC, the stolen data includes customer names, addresses, dates of birth and even what people searched for on the site.

How to defend yourself: If you shop at ASOS, watch out for fake "ASOS" emails and texts using your real name and address. Don't click links. Go straight to the app or website yourself.

Source: [Malwarebytes](#)

5. Citrix zero-days are under attack worldwide

The U.S. cyber agency CISA updated its alert on Oct 9. Hackers are actively exploiting serious flaws in Citrix NetScaler (CVE-2026-88771, 88772 and 88779). Some let attackers run their own code remotely.

How to defend yourself: If you or your company runs Citrix NetScaler, patch now. Then check if you were already hit. Patching alone doesn't kick out someone who's already inside.

Source: [CISA](#)

6. The FBI took down China-linked hacking tools

The Justice Department and FBI seized two hacking tools called "Microscan" and "FishHub." They say the tools were run by a company called Integrity Technology Group and used against critical infrastructure and around 20 universities in Taiwan.

Why it matters: The good guys are hitting back. Taking away the tools slows the attackers down. But phishing kits like these are everywhere, so stay sharp with every link you click.

Source: [U.S. Justice Department](#)



THE GOOD SIDE OF AI

7. GPT-6 is out for everyone

OpenAI rolled out GPT-6 to all ChatGPT users. Answers can now include charts, buttons, forms and mini tools right inside the chat. Paid users got it Oct 7, and Free and Go users got it Oct 8.

Why it matters: Powerful AI just landed in everyone's hands for free. Use it to learn and build. Just never paste passwords or private info into any chatbot.

Source: [OpenAI](#)

8. Google gives its newest AI to the defenders first

Google's new frontier model, Gemini 4 Argon, is going to vetted cyber defenders first through a program called Fairwind. (Announced Sept 30.)

Why it matters: This is how it should be done. Give the defenders the head start, not the hackers. AI on our side is how we hold the guard.

Source: [SiliconANGLE](#)



BANNED BY THE BOT

In mid-2025, thousands of Facebook and Instagram users said Meta's automated systems had shut them out for no clear reason. Many said they appealed and got rejected within minutes. Here is what's on the record.

The 2025 ban wave - In June 2025, Instagram users began complaining of mass bans and blamed AI moderation ([TechCrunch, Jun 16 2025](#)). Meta admitted a "technical error" that wrongly suspended some Facebook Groups, but told the BBC it had no evidence of a wider problem ([BBC; TechCrunch, Jun 24 2025](#)). - Users who paid for Meta Verified said its "priority support" didn't help ([TechCrunch, Jul 2 2025](#)).

Real stories 1. **Michelle DeMelo (Canada)**. Her Facebook and Instagram accounts were suspended in June 2025, and because her accounts were linked, all of her business pages went down too. She said she lost money. They were restored a day after the BBC contacted Meta ([BBC](#)). 2. **David (Aberdeen,**

Scotland). He was banned from Instagram on June 4, 2025, for supposedly breaking child sexual exploitation rules. His account came back within hours of the BBC raising his case on July 3. Two other users, Faisal and Salim, reported the same false accusation. The BBC said more than 100 people had contacted it claiming wrongful bans ([BBC](#)). 3. **"Sam" (BBC).** His appeal was rejected two minutes after he filed it, which made him suspect no human ever looked at it ([BBC](#)).

Pushback: petition and lawsuits - Brittany Watson of Ontario started a Change.org petition. It passed 25,000 signatures per the BBC ([BBC](#)) and later 36,000 per the petition's own update ([Change.org](#)).

Meta's own numbers - Meta said that in December 2024 it removed "millions of pieces of content every day," and estimated that **1 to 2 out of every 10 of those actions may have been mistakes** ([Meta Integrity Report Q1 2025](#)). - After its January 2025 changes, Meta said it **cut US enforcement mistakes in half** between Q4 2024 and the end of Q1 2025. It also reported fewer removals, fewer appeals and less restored content ([same report](#); [Meta, "More Speech and Fewer Mistakes"](#)). In Q2 2025, Meta said mistakes were down more than 75% ([Q2 2025 report](#)). - Note: these are numbers Meta reports about itself. They cover content removals, not account bans.

How to protect yourself 1. **Appeal right away.** Meta says appeals must be filed within the window it gives you, and some "severe" violations can't be restored ([Meta Help](#)). 2. **Use Meta's account-support hub.** Meta added centralized account support and recovery tools in Dec 2025 ([Meta Newsroom](#)). 3. **Escalate to the Oversight Board** once you've exhausted Meta's appeals, for eligible content decisions ([oversightboard.com](#)). 4. **Download your data regularly** with Accounts Center › "Download your information," so a ban doesn't cost you years of photos. 5. **Turn on 2FA** with an authenticator app or security key, so a hijacker can't get you banned. 6. **Don't link everything together.** Linked accounts went down together in the DeMelo case and the Reddit case above. 7. **Ignore "recovery experts" who DM you.** They're scams. 8. **Meta Verified** includes paid support, but users reported mixed results during the ban wave ([TechCrunch](#)). *Items 4–7 are general best practice.*



CYBERSECURITY BY THE NUMBERS

Crime keeps climbing (FBI IC3 2025 report) ([PDF](#); [FBI release](#)) - **1,008,597 complaints**, up from 859,532 in 2024. It's the first year over 1 million. - **\$20.877 billion in reported losses**, up 26%. The average loss was \$20,699. - People 60 and older lost **about \$7.7B**, up 37%. - Crypto-related complaints: 181,565, with **more than \$11B** in losses. Investment fraud alone: \$8.6B. - **AI-related**

complaints were tracked for the first time: 22,364 complaints and about \$893M in losses. - These are only crimes people reported. The FBI notes the real numbers are higher.

Ransomware (Chainalysis 2026 Crypto Crime Report, a vendor estimate) ([Chainalysis](#)) - About **\$820M** in on-chain ransom payments in 2025, down about 8% from \$892M in 2024. Chainalysis expects the figure to be revised upward. - Claimed victims **rose 50%**, but only **28% paid**. The median payment jumped from \$12,738 to \$59,556.

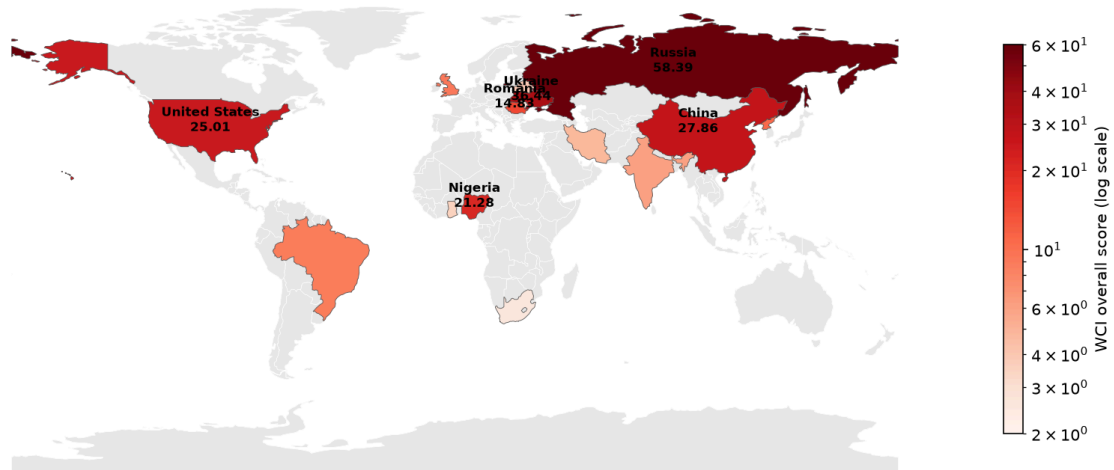
Security spending & workforce gap: UNKNOWN — Gartner and ISC2 figures not verified directly, left out of this issue.

Where the threats come from: the World Cybercrime Index The index comes from a 2024 Oxford/UNSW survey of cybercrime experts ([PLOS ONE](#); [Oxford](#)). Top 10 by overall score: 1. Russia 58.39 2. Ukraine 36.44 3. China 27.86 4. USA 25.01 5. Nigeria 21.28 6. Romania 14.83 7. North Korea 10.61 8. UK 9.01 9. Brazil 8.93 10. India 6.13

- Countries tend to specialize. Russia and Ukraine lead in technical services and extortion, Nigeria leads in scams, and the UK and US rank high in cashing out and money laundering (Table 1, same paper).
- Six countries (China, Russia, Ukraine, the US, Romania and Nigeria) appear in the top 10 of every category.

Caveats on attribution. The index measures expert opinion about where offenders are *based*. It doesn't count attacks. Attack-traffic data shows where machines are, and criminals route through VPNs and hijacked computers in other countries. Many offenders are stateless or work across borders. Don't read the map as "everyone in country X is a hacker."

Where cybercrime experts say the biggest threats are based — World Cybercrime Index, top 15



Source: Bruce, Lusthaus, Kashyap, Phair & Varese (2024), 'Mapping the global geography of cybercrime with the World Cybercrime Index', PLOS ONE 19(4): e0297312. Expert survey (92 experts), not attack telemetry. <https://doi.org/10.1371/journal.pone.0297312> | Map: Natural Earth. Graphic: World Seeker AI.

Source: Bruce et al. (2024), World Cybercrime Index, PLOS ONE. Map: Natural Earth.

TRENDING ON GITHUB (week of Oct 12, 2026)

Star counts are as shown on github.com/trending?since=weekly at about 1:18 AM PT, Oct 12 2026.

Licenses come from the GitHub API. Descriptions come from each repo's own one-line description.

We have not audited the code.

1. [mattpocock/skills](#): about 286,894 stars, +10,330 this week. MIT. "Skills for Real Engineers. Straight from my .agents directory." Ready-made skill files for AI coding agents. **Our take:** for developers who use AI coding agents. It's a set of files you read and copy, which makes it low-risk. Read each skill before you use it, since skills tell your agent what to do.
2. [thedotmack/claude-mem](#): about 99,862 stars, +3,803 this week. Apache-2.0. It gives AI agents memory across sessions by recording everything the agent does, compressing it with AI and feeding it back later. It works with Claude Code, Codex, Gemini, Copilot and others. **Our take:** handy, but it stores everything your agent sees, which can include secrets, keys and client data. Check where that data is kept and whether any of it gets sent to an AI service for compression.
3. [heygen-com/hyperframes](#): about 61,262 stars. Apache-2.0. "Write HTML. Render video. Built for agents." Made by HeyGen, the AI video company. **Our take:** for creators who want to make videos with code or AI agents. It has a permissive license.
4. [earthtojake/text-to-cad](#): about 19,416 stars. MIT. Python. "Give your agent CAD superpowers." **Our take:** for makers and engineers who want AI to produce 3D models. Check

the dimensions before you print or manufacture anything.

5. [tester-army/e2e](#): about 9,251 stars. Apache-2.0. A "next generation" end-to-end testing framework for web and mobile apps. **Our take:** for app teams. "Next generation" is the project's own claim, so try it on a side project first.
6. [EpicGames/raddebugger](#): about 8,419 stars. MIT. C. A native, user-mode, multi-process graphical debugger from Epic Games. **Our take:** for systems and game programmers. It's backed by a major studio, and it's a classic dev tool rather than AI hype.
7. [mvschwarz/openrig](#): about 7,069 stars. Apache-2.0. Lets you build persistent teams of AI agents (Claude Code, Codex, Pi) with roles, shared context and owned work. **Our take:** for power users running several agents. More autonomous agents means more access to your systems, so give them the fewest permissions they need.

None of these use AGPL licenses; they're all MIT or Apache-2.0, which allow commercial use. Also trending, but left off the list: **AnyPS5** (the week's #1, a tool for porting PS5 executables to Linux and Windows). Its legal and copyright standing is unclear, so we didn't feature it.



This week's takeaways

- Don't trust emails or texts just because they know your name and address.
- Patch your stuff. Then check it.
- Use AI, but keep your private info out of it.
- One hacker with AI is now a whole team. Stay ready.



FROM THE EDITOR



Family first.

"A choice to exploit or protect."

There's no power like freedom of choice: to be righteous, to be honest, to choose purpose over profit. A choice to exploit or protect. I changed my life. It's not perfect, but I chose to do right with my talents, be conscious, and be eyes wide.

Nick Ghafouria, Editor



Ask me anything / tell me what you think

This is Issue #1, and I want to hear from you.

Got a question about AI or staying safe online? Got hacked and don't know what to do? Think I got something wrong?

Just hit reply to this email, or message me here: coconick.com

I read it. And your questions might end up in the next issue.

That's it for this week. Stay sharp, stay safe, and don't let them scare you off AI. Let's put them in their place.

Let's hold the guard!!

— Nick worldseeker ai | worldseekerai.com | [Subscribe](#)