March 2026

F-Alert

The latest U.S. cyber security threat updates from F-Secure threat intelligence experts

F-Secure

Europe Cracks Down on Social Media—Will the U.S. Follow?

WHERE: All States

WHAT: The Netherlands is the latest country to propose banning social media for children. This raises a key question: could similar age restrictions take hold in the United States? For now, that appears doubtful. First Amendment protections make it difficult to implement a broad social media ban.

KEY FACTS:

- France has already approved a social media ban for children under 15, set to take effect in September, while the UK government is considering a ban for children under 16. However, the United States faces significant legal barriers to adopting comparable restrictions.
- According to a Harvard Law Review article, age-verification requirements for social media may conflict with First Amendment protections covering both users' access rights and platforms' editorial control. Regulatory authority is largely limited to platform design features rather than blanket limits tied solely to age or content.
- While the intention may be well meaning—and the harms social media can pose to children are well documented—such bans come with tradeoffs. ID checks carry privacy risks, and moderating content in a consistently fair way is nearly impossible.

EXPERT INSIGHT:

"While a nationwide ban is unlikely in the U.S. any time soon, governments proposing social media restrictions are multiplying across Europe. It almost goes without saying that children will seek—and likely find—ways to circumvent such bans. If policymakers want to make a meaningful impact, they may be better off focusing on potentially harmful platform features, such as infinite scrolling, rather than broad access restrictions."

Joel Latto
Threat Advisor
Helsinki, Finland

AI Platform OpenClaw Goes Viral, Ignoring Basic Security

WHERE: All States

WHAT: OpenClaw (formerly Clawdbot, then Moltbot) is an open-source AI agent platform that went viral after launching Moltbook, a social network for agents. Both lack meaningful security controls—ignoring decades of cyber security lessons and enabling bots to rapidly scale early internet-era mistakes.

KEY FACTS:

- OpenClaw ships with sandboxing disabled, no authentication required, and full shell access enabled, leading to predictable security breaches such as credential exposure and account takeover.
- Many of the agent "skills" that make OpenClaw highly extensible also contain serious vulnerabilities—or are outright malicious—enabling silent data exfiltration and backdoor installation.
- Moltbook, the "vibe coded" social network for OpenClaw and other AI agents, likewise prioritized virality over security, exposing 1.5 million authentication tokens within hours of launch.

EXPERT INSIGHT:

"We've seen how "ship first, secure later" plays out. It took decades to harden early internet protocols. But OpenClaw is replaying that same old movie at 100x speed with autonomous agents that never sleep, can infect one another, and process irreversible financial transactions."

Dr Megan Squire
Threat Intelligence Researcher
North Carolina, USA

Trending Scam

Tax Season Is Here—and Scammers Are Targeting Gen Z

WHERE: All States

WHAT'S HAPPENING:

- Many assume tax scams mainly target older adults, but according to the F-Secure U.S. Scam Intelligence & Impacts Report, Gen Z is the most exposed: 18–25-year-olds report the highest scam victimization rate.
- As tax season ramps up, scammers exploit the stress, confusion, and financial pressure younger adults often face—in many cases for the first time.
- Tax scams follow a predictable pattern: trigger emotion (panic, fear, excitement about a refund), apply pressure ("act now"), disrupt attention (links, calls, login prompts), capture value (credentials, money, data), then escalate (refund theft, account takeover). Once emotion takes hold, rational checks often disappear.

WHAT TO DO:

- The simplest defense: pause. If a tax-related message triggers fear, urgency, or even relief, stop. Never click links in tax-related texts or emails, call numbers included in those messages, or share sensitive information with online "helpers."
- Always verify independently via official websites accessed directly—not through email links. Log in to known tax software accounts separately and take your time. Urgency is the scammer's advantage.

Breach That Matters

Popular 'Chat & Ask AI' Mobile App Exposes 300M Messages

WHERE: All States

WHAT'S HAPPENING:

- 300 million messages from 25 million users of 'Chat & Ask AI'—a mobile app that connects users to AI models like ChatGPT and Gemini—were exposed in a major data breach. The incident stemmed from a basic Google Firebase misconfiguration, with access controls left open rather than a hack.
- An independent security researcher discovered he could access millions of chat histories, timestamps, user settings, and more. No usernames or passwords were leaked, but many messages were deeply personal, with users confiding in the AI as they would a trusted friend.
- The breach highlights the risks of "wrapper apps"—mobile apps that repackage existing AI services without matching their security standards.

WHAT TO DO:

- This type of exposure is not uncommon. Wrapper apps often rush to market with limited security oversight. Be cautious with third-party AI apps and consider using official platforms like ChatGPT or Gemini instead.
- Review privacy policies and user reviews before downloading third-party AI apps and avoid sharing sensitive information.

AI Agents Are Now Hiring Humans to Do What They Can't

WHERE: All States

WHAT: RentAHuman is a new platform that lets AI agents hire real people for tasks they can't complete—such as picking up packages, taking photos, or attending meetings. It promises flexible income in an automated future, but accepting tasks from AI agents means entering a system with little oversight and no identity verification.

KEY FACTS:

- A request to collect a package in a remote location could be a setup. Delivering a message could make you part of harassment or stalking. In each case, your identity and physical safety are tied to a system with no clear accountability.
- Agents could also split criminal operations into small, innocent-looking tasks outsourced to different workers. One registers a country-code domain. Others send promotional emails. Once traffic builds, the site is swapped for a phishing page. The workers' real identities are tied to the scam while the agent remains untraceable—and may disappear without paying.
- The risk runs both ways: a hired human providing technical services could steal credentials, install malware, or inject malicious code.

EXPERT INSIGHT:

"As AI agents increasingly operate online, trust is the missing piece. There's no reliable way to verify whether a request is legitimate or who is responsible if something goes wrong. If a task turns out to be criminal, the worker may be the only traceable link. My advice is this: if you can't see the full context of what you're contributing to or any worker protections, don't accept the task."

Khalid Alnajjar
Senior AI & Data Scientist
Helsinki, Finland

Google's AI Is Becoming More Personal—and More Invasive

WHERE: All States

WHAT: Google has announced the beta launch of its new 'Personal Intelligence' feature for Gemini, saying it will make the AI model "more personal, proactive, and powerful." The feature connects apps like Gmail, Google Photos, and Google Search to build a detailed user profile and personalize the experience based on a person's digital footprint.

KEY FACTS:

- Google's AI is becoming increasingly personalized, now able to infer appointments from calendar entries and identify license plates from photos. The upside? The better assistants know users, the more helpful they can be. The downside: deeper personalization requires broader data access—and increases privacy risk.
- Google has previously faced scrutiny over voice recordings captured by its AI-powered Google Assistant, activated by text and voice commands ("Hey Google") to manage tasks such as making phone calls and setting reminders.
- The digitalization of biometrics—such as fingerprints, facial features, and voice—has improved authentication. But when combined with AI, these immutable identifiers can be exploited for identity theft, financial fraud, and other forms of abuse.

EXPERT INSIGHT:

"The problem is that the whole is more dangerous than any single piece. A set of photos seems harmless. A shared detail seems harmless. Recorded audio seems harmless. But once combined, they create a detailed profile that can be used for identity fraud, deepfakes, or targeted manipulation in ways no single element could. If a criminal gains access to a deeply personalized Gemini account, they unlock a rich pool of sensitive data."

Timo Salmi
Senior Product Marketing Manager
Oulu, Finland

illuminate

> “Illuminate, F-Secure’s research function, brings together experts to explore the human, social, and technical aspects of security. We identify emerging threats, prototype new protection systems, and anticipate future risks to keep consumers safe. By staying ahead of the curve, we navigate a constantly evolving digital world and ensure F-Secure delivers trusted, reliable, and innovative cyber security solutions.”

Laura James
Vice President, Research
F-Secure

About F-Secure

F-Secure makes every digital moment more secure, for everyone. We deliver brilliantly simple, frictionless security experiences that make life easier for the tens of millions of people we protect and our 200+ partners.

For more than 35 years, we've led the cyber security industry, inspired by a pioneering spirit born out of a shared commitment to do better by working together.

For the latest news and updates visit f-secure.com or follow us on our social channels.

F-Secure | illuminate