November 2026

F-Alert

The latest US cyber security threat updates from F-Secure threat intelligence experts



F-Secure



EXPERT INSIGHT:

“The technology behind a voice cloning scam may be sophisticated, but the way criminals collect money can be surprisingly simple; collecting cash in person makes stolen money harder to trace and recover. This shows why scam protection can’t focus solely on suspicious links or online payments and is a reminder why we must help consumers recognize the warning signs of manipulation, even when the voice they hear sounds familiar.”

Timo Salmi
Senior Product Manager
Oulu, Finland

AI Scam Shows Why Payment Protection Isn’t Enough

WHERE: Orlando, Florida

WHAT: An Orlando woman lost thousands of dollars after scammers impersonated a friend needing urgent help getting out of jail, using what she suspected was an AI-cloned voice. Rather than requesting a digital payment, the scammers sent couriers to her home to collect cash in person, bypassing the fraud detection measures typically used by payment providers.

KEY FACTS:

- The caller claimed to be a friend who had been arrested following a car accident. A second caller, posing as a public defender, asked for money for his release.
- The victim was instructed to place cash in an envelope alongside a case number and paperwork, then couriers arrived at her home to collect the money.
- This is an established scam tactic; the FBI has previously warned about criminals arranging in-person cash collections as part of impersonation and emergency scams. AI voice cloning can make these scams even more convincing.

Meta's Muse Agent Grows Fast, Raises Privacy Concerns

WHERE: All States

WHAT: Meta's Muse personal AI agent launched on September 8 and became one of the fastest-downloaded apps on record. But its rapid rise was accompanied by serious privacy incidents: within its first month, the agent reportedly read a journalist's private Mac messages and shared a Facebook Marketplace seller's home address with a stranger.

KEY FACTS:

- Some estimates put Muse at 5 million downloads in just 22 days. Others calculate 1.8 million iOS downloads for Muse in the first 12 days, compared with 1.3 million for ChatGPT at its app launch.
- Journalist Jason Aten reported that Muse accessed private Mac messages without his permission, while Meta maintains that Muse can only access messages if the user enables specific permissions. Apple has since announced plans for tighter permission controls, warning that risks "will grow substantially" as agents become more autonomous.
- YouTuber Matt Robb reported that during a routine Facebook Marketplace transaction, Muse accepted a lowball offer and shared his home address without permission. Meta says the agent had requested permission, which Robb disputes.



EXPERT INSIGHT:

"Muse shows what happens when an AI agent reaches millions of consumers before its permission model is ready. Consumers should give agents the least amount of access possible, avoid 'always allow' settings, and regularly check which connectors are enabled. Platforms also have a responsibility, and Apple's response to make high-risk permissions harder to grant is a step in the right direction."

Dr Megan Squire
Principal Threat Intelligence Researcher
North Carolina, United States

Trending Scam

Fake ChatGPT Ads on Google Lead to Malware Traps

WHERE: All States

WHAT'S HAPPENING:

- Cyber criminals are targeting people searching for ChatGPT on Google, using sponsored ads that appear in search results. Clicking these ads takes users to a malicious custom GPT hosted on the legitimate ChatGPT domain.
- Victims then encounter a fake security verification instructing them to update their subscription or use a backup domain. Clicking the link prompts users to run a command that installs malware on their device.
- This is a new twist on a longstanding tactic known as malvertising, where cyber criminals use online ads to spread malware. As advertising expands into AI chatbot experiences, similar tactics could create new opportunities for deception.

WHAT TO DO:

- Access ChatGPT directly through chatgpt.com rather than sponsored search results and be cautious of unexpected requests to visit alternative websites.
- Question unusual requests. Never run commands on your device unless you're certain what they'll do, especially when prompted by a link or pop-up.

Breach That Matters

Hackers Steal Sensitive Records of Millions of Students

WHERE: California (statewide)

WHAT'S HAPPENING:

- Hackers stole the personal information of millions of current and former California public school students in a data breach involving Educational Testing Service (ETS).
- Stolen data includes names, addresses, dates of birth, school information, and sensitive disability and federal assistance records. Social Security numbers, financial details, and passwords were not exposed.
- Attackers used a compromised login credential to access and copy student records between May and September 2026. Some affected records belong to former students who are now adults.

WHAT TO DO:

- Stolen personal information could be used for identity theft, fraud, and targeted phishing scams. Be cautious of unexpected calls, emails, or messages, even if the sender knows your name, address, or other personal details.
- Parents should consider placing a free credit freeze for children under 16, while affected adults can check their eligibility for free identity protection services.



EXPERT INSIGHT:

“While Tavus is taking a responsible approach to safety before release, these AI personas will only become more convincing. We should expect bad actors to exploit similar technology for cyber crime, especially impersonation scams, as it becomes more widely available. The challenge will be helping consumers recognize deception when even a live video conversation is no longer reliable proof that they're speaking to a real person.”

Joel Latto
Threat Advisor
Helsinki, Finland

AI Personas Can Now Gesture and Interrupt in Real Time

WHERE: All StatesI

WHAT: AI company Tavus has unveiled Griffin-Lite, a preview of its new Griffin model, which generates realistic AI personas that respond to people in real time during video conversations. These personas could transform customer service and other digital interactions, but they could also make impersonation scams harder to recognize.

KEY FACTS:

- Humanlike AI personas are nothing new, but the ability to respond naturally in real time takes this technology to another level. Griffin can react without waiting for someone to finish speaking, using facial expressions, gestures, and vocal interruptions to create more life-like interactions.
- In Tavus’s own study, 48% of 54 participants believed they were speaking to a real person during a one-minute video call. However, the findings have not been independently replicated.
- Griffin-Lite is currently available only to selected testers. Tavus says further safety measures are needed before wider release, while work on a more advanced model is already underway.

F-Secure Launches Human-Centered Research Hub

WHERE: All States

WHAT: F-Secure has launched its human-centered Illuminate Research Hub, showcasing original studies on people and technology as they actually interact. The hub brings together expert-led research and thought leadership exploring emerging consumer risks and the realities of trust, deception, and online resilience in the AI era.

KEY FACTS:

- The Illuminate Hub brings together research across five core areas: consumer AI harms; scams, fraud, and deception; trust and the information environment; consumer digital lives; and the agentic AI ecosystem.
- Among the hub's new flagship studies are the AI Harms Taxonomy and the Scam Encyclopedia. The taxonomy maps 26 ways AI can harm consumers, connecting each harm to contributing layers and potential defenses, while the encyclopedia documents more than 50 scams and 50 warning signs from a victim-centered perspective.
- In addition to the hub launch, F-Secure's Fall 2026 Digital Trust Report draws on Illuminate's research to explore how consumers use AI chatbots, what their behaviors reveal about digital trust, and what it means for service providers.



EXPERT INSIGHT:

"Helping people stay resilient and thrive online means understanding the messy reality of their digital lives, not just the threats they face. Illuminate combines research on human behavior, threats, harms, and consumer technology, using modern research methods and creative approaches. This lets us see risks that others miss and build secure products that people actually want. I'm proud of the care and rigor the team brings to researching these complex, fast-moving areas."

Dr Laura James
VP, Research
Liverpool, United Kingdom



“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 is a human-first, AI-powered consumer cyber security experience company with 38 years of expertise in tackling digital threats. We help Digital Service Providers turn trust into a high-value growth engine — protecting their customers while enabling them to live their best digital lives in a world of relentless, AI-driven scams.

With billions of digital interactions secured each year, tens of millions of consumers protected globally, and over $10bn in partner value created, we deliver proven impact at scale.

To find out more visit f-secure.com/partners or follow us on our social channels.

F-Secure | illuminate