

#Blog 3: Hackers Are Using AI Right Now - Real Examples From the Summit

From the CSA Agentic AI Summit Series

By TekSol Global | Cybersecurity & AI Practice

You might think AI-powered cyberattacks are something from a sci-fi movie, a future threat, not a present one.

You'd be wrong.

At the CSA Agentic AI Summit, speakers shared real, documented examples of how AI is being used by criminals and nation-state hackers today. Some of it is surprisingly simple. Some of it is alarming. All of it is happening right now.

#Example 1: AI-Powered Phishing Emails That Actually Work

Most of us have seen a phishing email - those fake messages pretending to be from your bank or your boss, trying to trick you into clicking a link or handing over your password.

For years, these were easy to spot. Bad grammar. Weird spelling. Generic greetings like "Dear Customer." Anyone with a bit of common sense could sniff them out.

AI changed that.

Speakers from ThreatLocker demonstrated this live at the summit. They used an AI tool and asked it to create realistic phishing emails for security training purposes, and the results were striking.

The AI didn't just write a convincing email. It went out and **gathered publicly available information about the company**, the names of executives, office locations, company events, and used all of that to create a phishing email that looked completely legitimate.

In one example, it crafted an email impersonating a real company executive, referencing their actual Dublin office and a real data center expansion project. An employee unfamiliar with phishing tactics would have no reason to doubt it.

In another example, it generated a fake certification email tied to a real conference the company had hosted, tricking employees into downloading malware disguised as their exam results.

The key point: AI doesn't just write better fake emails. It researches its targets and personalizes the attack. That makes it dramatically more dangerous.

#Example 2: Nation-States Using AI to Probe Systems

At the summit, a speaker from Google's threat intelligence team explained something remarkable: the world's most sophisticated government hacking groups, known as **APT groups** (Advanced Persistent Threats), are actively experimenting with and using AI tools.

These are groups backed by countries like Russia, China, North Korea, and Iran. They have resources, patience, and a mission to steal secrets, disrupt infrastructure, or gain strategic advantages.

What are they using AI for?

- **Writing better malware**, code that's harder to detect
- **Translating and processing large amounts of stolen data** quickly
- **Scanning for vulnerabilities** in target systems
- **Creating convincing fake personas** online to conduct social engineering attacks
- **Generating spearphishing content** in multiple languages and cultural styles

The speaker was clear that AI isn't replacing these skilled human hackers, yet. But it's making them significantly more efficient. One skilled operator can now do the work that previously required a team.

#Example 3: AI Used in a Real Espionage Operation

One of the most discussed moments at the summit referenced a report that Anthropic (the company that makes Claude AI) published about their own tool being used as part of a real cyber espionage operation.

This wasn't theoretical. Someone had incorporated an AI assistant into an actual attack campaign, using it to help with research, code writing, and planning steps of a complex operation.

This is called **AI-assisted cyber espionage**, and it's a category that didn't really exist a few years ago.

What makes this particularly unsettling is that it shows AI tools, originally designed to be helpful, can be misused by bad actors without the tool's creators intending or approving it.

#Example 4: Agentic Attacks - AI That Hunts Autonomously

Here's where things get more advanced, and more concerning.

Speakers at the summit described emerging **agentic attack systems**: AI tools that don't just help a human hacker but can conduct parts of an attack on their own.

These systems can:

1. Scan a network to find all connected devices
2. Identify which devices have known software vulnerabilities
3. Test which vulnerabilities are actually exploitable
4. Launch the attack and attempt to gain access
5. If that fails, try a different approach automatically

This is sometimes called "attack orchestration." No human has to sit and manage every step. The AI agent just keeps going until it either succeeds or runs out of options.

One alarming data point shared at the summit: the average time between when an attacker first gets into a network and when they start causing serious damage ("**breakout time**") is now just **29 minutes**. The fastest observed was **27 seconds**.

Human security teams cannot respond that fast. That's why AI-powered defence has become so urgent.

#The "Phishing on Steroids" Problem

Let's tie this together with a simple picture of what a modern AI-powered attack can look like:

Step 1: The AI scans social media, company websites, and public records to build detailed profiles of employees at a target company.

Step 2: It generates personalized phishing emails for 1,000 employees simultaneously, each one tailored to that specific person's role, their manager's name, and recent company events.

Step 3: When an employee clicks and enters their password, the AI agent immediately uses those credentials to log into company systems.

Step 4: Once inside, another AI agent quietly maps the internal network, identifies where sensitive data is stored, and begins extracting it, all while appearing to be a normal user.

Step 5: By the time anyone notices, the data is gone.

This entire sequence, which once took a skilled team of hackers days or weeks, can now unfold in hours, with minimal human involvement.

#What Are Defenders Doing About This?

This question was central to the entire summit. The short answer is: they're fighting fire with fire.

Security teams are now using AI to:

- Detect unusual patterns in network behaviour that humans would miss
- Automatically block suspicious activity before damage is done
- Generate threat intelligence faster than any human analyst
- Simulate attacks on their own systems to find weaknesses first

The race between AI attackers and AI defenders is the defining security challenge of our time. And right now, it's very close.

Next up → Blog 4: AI Agents Inside Your Company - What Could Possibly Go Wrong?