

Read This Website While You Still Can

Matteo Wong

The internet could be on the verge of a total meltdown. Yesterday, someone posted to a public forum that two major security flaws had been found in a programming language that undergirds the web. The language, called PHP, is used by as much as 70 percent of websites.

In order to protect themselves from hackers, a number of government, hospital, and school websites across Europe went dark this morning. The organization that maintains PHP quickly [confirmed](#) it was preparing to announce two major vulnerabilities alongside patches for them. Both weaknesses, officially disclosed today, appear to have existed in PHP for more than a decade but were only just [discovered](#) with the use of AI tools. Hackers have already been detected trying to exploit them, likely with the help of AI, and implementing the fixes to address the vulnerabilities could take companies days or weeks. In the meantime, tens of millions of servers, databases, and devices are exposed. “PHP is one of the fundamental pieces of technology that supports the internet,” Sam Rubin, a senior vice president at the cybersecurity firm Palo Alto Networks, told me. As of now, the websites of the Finnish health agency, the Port of Rotterdam, the Paris police force, 4chan, and an unknown number of others are all down while they address the PHP vulnerabilities.

[Read: How long until AI hacks everything?](#)

News of these weaknesses in PHP is the capstone in a weeks-long slew of devastating, AI-powered cybersecurity events. In recent weeks, criminals and state-backed attackers have used AI to [hack](#) into more than 100 companies, including John Deere and SharkNinja; to automate [hacks](#) on some 400 financial institutions in Latin America; and to [steal](#) hundreds of thousands of credit-card records from two companies. Since this summer, attackers have hit governments, data centers, and hospitals in Ukraine, Russia, Argentina, Bangladesh, and Indonesia. Also today, the developers of Node—a platform used by nearly half of software developers—disclosed multiple high-severity vulnerabilities. And that’s the abridged list.

For the better part of a year, cybersecurity experts have been [warning, with mounting alarm](#), about an [impending cyber-apocalypse](#). It may have begun. For decades, the internet has been riddled with holes and bugs. Code is finicky, and the standard IT approach has been to fix problems when they arise. For decades, the balance held: Skilled hackers are few, and finding and exploiting vulnerabilities takes time.

Recent advances in AI have totally upended that equilibrium. Top models have become very good at finding holes and writing malware, and can work much faster than even a team of humans. This is true of both open-source AI systems from China, which can be downloaded and used without restrictions, and the even more capable models from OpenAI and Anthropic, which are continually being jailbroken. The rate of software vulnerabilities identified per month has doubled over the past year. Meanwhile, the time that hackers need to begin exploiting those vulnerabilities—figuring out how to break in and out without being caught—has fallen, by some estimates, from weeks to hours. An actual intrusion, if automated with AI, can take as little as [39 seconds](#). Historical approaches to detecting and fixing vulnerabilities, or monitoring for suspicious internal activity, are insufficient.

AI has created new vulnerabilities, but it also promises to fix them: The vulnerabilities in PHP were discovered independently by a Google-owned cybersecurity company and another effort led by OpenAI. Google, OpenAI, and Anthropic, along with other tech firms, are all working with government agencies around the world to shore up their defenses. But, as numerous cybersecurity experts have told me, there is an asymmetry between attackers and defenders that AI is going to exacerbate: Hackers will use new technology with abandon, while companies are methodical and risk-averse—a hospital or electrical grid

can't risk that an AI security tool unintentionally takes a patient off life support or causes a brownout. And attackers need to find only one hole to succeed, whereas defenders need only one oversight to fail.

For a year, criminals and state-backed groups have been experimenting with how to use models to accelerate and automate their attacks. It was only a year ago that Anthropic [announced](#), with great consternation, what it declared to be the first instance of AI-orchestrated cyber-espionage. That still involved a lot of human oversight and required jailbroken Claude agents to work for many hours. Now that the agents are faster and more capable, hackers have gotten more adept at deploying them, and the automation is more extensive. Recent, widespread attacks on South Korean and Japanese websites were [reportedly](#) executed by a single Chinese cybercriminal who used AI to target more than 1,000 websites. "This certainly will not be the last large-scale vulnerability identified, because of what's possible with AI," Rubin said.

The past 24 hours represent a tipping point. Bulking up IT departments, implementing passkeys, and using two-factor authentication remain important. But the most urgent and practical advice being given now is to retreat: Prepare to simply go offline. The NSA [advised](#) companies that cannot immediately fix the PHP vulnerability to take down affected digital services. Hospitals, water utilities, and other crucial services are all preparing for hacks by relearning to operate analog. Store your data in fewer places. We're running out of time to prepare for the AI-hacking deluge. The dam may have already broken.