[redacted]

INTEL471

[redacted]

ANNOUNCEMENT

Executive Intelligence Update – Feb. 4, 2025

Published: Today 21:52:05 UTC

Intel 471's Executive Intelligence Update gives CISOs and security teams insight into trends, threats, risks and regulations. In this issue:

- Does DeepSeek pose a cybersecurity risk?
- Cracked, Nulled forums shut down by authorities
- Google: How Iranian and Chinese APTs use Gemini AI
- Funksec ransomware group noisy, erratic

1. Does DeepSeek pose a cybersecurity risk?

What's happening: The release of DeepSeek's R1 large language model (LLM) upended the world of artificial intelligence (AI) due to its high performance vis-à-vis other LLMs. But security companies warn it is vulnerable to jailbreaks and that safety guardrails are already being removed from versions of it.

Producing nefarious content: DeepSeek comes from High Flyer, which is a Chinese quantitative hedge fund that was interested in using AI for investment decisions. Although DeepSeek has been cast in the West as an unknown, it has very much been on the radar of China watchers as one of many companies and projects in the country with a horse in the AI race. The company claims it developed a high-performing LLM on par with OpenAI and others at a fraction of the cost and possibly without the benefit of the best Nvidia high-performing chips that are under export restrictions. To prove it, DeepSeek showed its homework, both in a research paper and by making the R1 model open weight, which means its training parameters can be accessed (this doesn't include training data, however). The term "open weight" is conflated with open source (although DeepSeek also calls R1 open source on its website). R1 is accessible through a web-based portal and a mobile app, DeepSeek AI assistant. Users can also download the model locally and run it on modest laptop software. Security companies have probed it and found R1 is apparently vulnerable to various types of jailbreaks that bypass safety guardrails. Palo Alto Networks found DeepSeek is vulnerable to two types of jailbreaks it developed, Deceptive Delight and Bad Likert Judge, and a third one, Crescendo, which was developed by Microsoft. Palo Alto says DeepSeek "enabled explicit guidance on malicious activities, including keylogger creation, data exfiltration and even instructions for incendiary devices, demonstrating the tangible security risks." Threat intelligence firm Kela also had a go at DeepSeek, finding that it was vulnerable to "a wide range of scenarios, enabling it to generate malicious outputs, such as ransomware development, fabrication of sensitive content, and detailed instructions for creating toxins and explosive devices."

Intel 471's view: There are a variety of ways to process what the advent of DeepSeek means. The bird's eye view of this development is that DeepSeek's work shows China is not very far behind in the AI race, which close AI watchers already knew. It demonstrates how AI remains unpredictable in the sense that new breakthroughs are always on the horizon, which means a shifting security landscape.

The jailbreaks Kela and Palo Alto Networks uncovered will invariably get remedied in future official versions. But it's the unofficial versions of DeepSeek's LLMs that appear in places such as Hugging Face that may be a worry. Observers have already noticed uncensored versions of DeepSeek appear where the guardrails have been removed by a process called abliteration. Cybercriminals are using generative AI now for productivity gains, so the proliferation of uncensored models means new opportunities for developers to offer illicit services that leverage unrestricted generative AI.

Aside from using DeepSeek for malicious ends, there are risks in using the model itself. DeepSeek's home base in China prompts the usual adversarial concerns. Queries inputted into a model are logged, including all IP addresses, device telemetry and the rest of the usual surveillance economy data that fuels the online advertising industry. In theory, downloading the model locally and running it without internet access would be a safer way to test it. However, this should only be done in a strict test environment that does not have any access to an enterprise network or resources.

2. Cracked, Nulled forums shut down by authorities

What's happening: Cracked and Nulled, two long-running cybercrime forums that attracted millions of users over the years to buy hacking tools, credentials and malware hosting, have been taken out by law enforcement as part of Operation Talent.

Administrator charged: Operation Talent is an international effort coordinated by the European Union Agency for Law Enforcement Cooperation (Europol) and led by German authorities in collaboration with agencies from Australia, France,

[redacted]

Greece, Italy, Romania and the U.S. Authorities said Cracked had more than 4 million users and generated more than US $4 million in revenue through the sale of stolen login credentials and other hacking tools. More than 17 million people in the U.S. were affected, including a woman whose credentials were bought on the site and who was later cyberstalked and threatened, according to the U.S. Department of Justice. The FBI also allegedly seized the domains used by the MySellix e-commerce platform, which was the payment processor for Cracked, and the StarkRDP remote desktop protocol (RDP) virtual hosting provider. Nulled had more than 5 million users and generated more than US $1 million in annual revenue through sales of stolen login credentials, stolen identity documents and hacking and fraud tools. The U.S. also unsealed charges against Lucas Sohn, 29, an Argentinian man living in Spain, for reportedly administering Nulled. Sohn reportedly operated an escrow function for transactions involving stolen credentials. Escrow services hold funds paid by one party to another for goods and services, and the funds are released when the purchaser is happy with what they've bought.

Intel 471's view: Cracked and Nulled were some of the longest running, easily accessible clear web cybercrime forums. Cracked had been running since 2018 and Nulled ran since 2016. The forums were notable as a place where people with an interest in cybercrime and taking it further could get their feet wet. Both contributed to a cybercriminal ecosystem that caused great harm. As with law enforcement's consistent efforts to take down other forums, such as Raid Forums and BreachForums, we'd expect someone might try to revive one or both of these forums as they are money spinners, even though the era of running a big clearnet cybercrime forum seems excessively risky.

3. Google: How Iranian and Chinese APTs use Gemini AI

What's happening: A new report from Google sheds light on how adversaries are using and abusing its Gemini generative AI tool but notes that none of the uses marked a sea change.

Helpful in defense, too: Google's report very much aligns with ones released last year by OpenAI and its partner Microsoft into how North Korea, China, Iran and Russia were using generative AI tools. Generative AI is a useful tool for attackers, but "it is not a game-changer it is sometimes portrayed to be" and there is no evidence that gen AI has been used to develop "novel capabilities," Google writes. Google analyzed prompts entered into Gemini by advanced persistent threat (APT) groups and information actors (IO) that are tracked by Google Threat Intelligence Group. These actors used Gemini for a variety of tasks: infrastructure research, vulnerability research, payload development, scripting and evasion techniques, content generation, developing personas and messaging, and translation. IO actors from Iran were the heaviest users of Gemini, Google says, while Russian and Chinese IO actors used it for general research and content creation. Russian threat actors were noted to use it for technical tasks, such as porting malware to other coding languages and adding encryption features. North Korean actors employed it for finding free hosting providers, target reconnaissance and creating cover letters and researching jobs. The country has been outed as trying to place workers at Western companies, both to earn income and for potential attacks, such as within the cryptocurrency industry. Rather than marking a big sea change in attacker tactics, Google writes that generative AI allows them to move faster, likening it to the use of Metasploit and Cobalt Strike for exploitation activity. But it warned the "AI landscape is in constant flux, with new AI models and agentic systems emerging daily." It also benefits defenders, Google's vice president of threat intelligence told the Wall Street Journal: "AI is not yet a panacea for threat actors and may actually be a far more important tool for defenders. The real impact here is they are gaining some efficiency. They can operate faster and scale up."

Intel 471's view: Google's observations coincide with other assessments that threat actors are very much using generative AI to bolster their productivity. This is not a heart-stopping threat as pointed out earlier, but it does allow for increased scale and more polished execution of spear-phishing attacks, for example.

4. Funksec ransomware group noisy, erratic

What's happening: The **Funksec** ransomware-as-a-service (RaaS) group arrived with a significant amount of fanfare and claimed breaches. However, its true impact is still being evaluated.

Racist messages and threats: The **Funksec** group first was seen in the underground in October 2024 after the actor **Desertstorm** attributed leaked datasets and access offers to the group on the BreachForums cybercrime forum. The group officially emerged in December 2024 following advertisement campaigns the actors **Scorpionlord** aka **scorpion** and **el_farado** promoted on underground forums. The group's activity primarily focuses on compromising government websites and web defacement attacks, database breaches and exfiltration, offers of unauthorized access and victim extortion via threatening and racist messages on compromised assets or via **Funksec's** data leak site (DLS) dubbed FunkForum, which also acts as a cybercrime forum. The group's operators developed a range of tools including the FunkLocker ransomware, the FDDOS script designed to perform distributed denial-of-service (DDoS) attacks, the funkgenerate tool and the JQRAY remote access tool. The group told one interviewer it uses AI for some of its development work, a claim that Check Point Research concluded was likely accurate. The **Funksec** threat actors appear to have financial- and political-based motivations, targeting U.S. resources that support Israel. As of Jan. 30, 2025, the group had added 91 organizations to its DLS. However, Check Point Research found that some datasets the group leaked already were leaked in other hacktivist campaigns.

Intel 471's view: The **Funksec** group's frequent strategic shifts and scattered focus suggest that while it is experimenting, it also may be uncertain about its direction. The group's overall reliability is questionable due to significant doubts about its authenticity and competency. Analysis of leaked datasets revealed most of **Funksec's** victims had their public-facing websites compromised or databases dumped rather than being victims of a typical network intrusion followed by ransomware deployment. No evidence has surfaced of the group's ransomware being used in real attacks at

the time of this report. The group's apparent lack of technical expertise, careless operational security (OPSEC) and exaggerated claims appear to be offset by its significant effort and determination, with new features for different aspects of its operations announced steadily. Its use of AI is notable, reinforcing the technology's growing role in both legitimate and malicious activities. For more information and this complete report, please contact Intel 471.