# Cybercriminal group UNC6032 exploits AI hype with malware-laden fake ads on Facebook and LinkedIn



A group of cybercriminals, identified as UNC6032, has been leveraging the rising interest in artificial intelligence by promoting malicious ads on social media platforms, thereby endangering users' sensitive information such as credentials and credit card details. According to a report by Mandiant, a subsidiary of Google focused on cybersecurity, thousands of fraudulent advertisements were detected on Facebook, with a smaller number appearing on LinkedIn since November 2024. These ads lead to over 30 deceptive websites that disguise themselves as legitimate AI video generation tools, such as Luma AI and Canva Dream Lab, which falsely advertise capabilities for converting text and images into videos.

Once users unwittingly click on a "Start Free Now" button on these spoofed websites, they encounter a faux video-generation interface that appears legitimate. After engaging with the interface and triggering a phony loading sequence, users are prompted to download a ZIP file containing malware. This malware, once executed, compromises the victim's device by establishing a backdoor, logging keystrokes, and scanning for password managers and digital wallets. The scale of UNC6032's operations is staggering, with Mandiant noting that their malicious ads have reached more than two million users on Facebook and LinkedIn, although the firm clarified that this reach doesn’t equate to the number of successful infections.

Meta, Facebook's parent company, has responded to this evolving threat landscape by removing the malicious ads, blocking associated URLs, and shutting down the rogue accounts responsible for disseminating these frauds. A Meta spokesperson acknowledged the ongoing challenges posed by cybercriminals, asserting that "Cyber criminals constantly evolve their tactics to evade detection.” Despite their efforts to mitigate the risks posed by these campaigns, the company admits it is difficult to determine the exact number of victims.

In conjunction with Mandiant's findings, reports from other cybersecurity analysts reveal similar patterns of deception used by different actors, including the promotion of fake AI services like OpenAI's ChatGPT and DALL-E through suspect Facebook pages and advertisements. Such engagements lead unsuspecting individuals into compromising positions where they download malware designed to exfiltrate sensitive data, often sold on dark web marketplaces or used to conduct further online fraud.

These malicious campaigns are not merely isolated incidents; they form a broader, increasingly sophisticated ecosystem of cybercrime. Reports indicate that as of 2023, Meta had disrupted nearly ten new malware strains and blocked over 1,000 dangerous URLs linked to fraudulent activities. This proactive stance underscores the ongoing battle between social media platforms and cybercriminals, which exploits both popular trends and established trust in AI technologies.

Further scrutiny reveals that UNC6032 actively rotates the domains featured in its ads to circumvent detection, creating new advertisements daily to maintain pressure on online security measures. This relentless cycle of adaptation is augmented by the use of advanced malware families capable of performing sophisticated tasks, such as keylogging and extracting sensitive data, which heightens the stakes for potential victims.

Mandiant has also highlighted the ease with which such deceptive AI websites can draw in users, pointing out their infrastructure of fake promotions and hijacked accounts that continuously lure in those seeking legitimate tools. This highlights a critical vulnerability in online safety and the urgent need for heightened vigilance among users, particularly as the allure of AI technologies continues to grow.

In summary, the threat posed by UNC6032 and similar factions illustrates the darker side of increased reliance on digital tools, especially within the AI landscape. As organisations like Meta strive to shore up defences, individuals must remain acutely aware of the potential risks associated with seemingly innocuous online interactions to protect their personal and financial information.

## Reference Map:

* Paragraph 1 – [[1]](https://www.theregister.com/2025/05/27/fake_social_media_ads_ai_tool/), [[2]](https://www.bleepingcomputer.com/news/security/fake-facebook-midjourney-ai-page-promoted-malware-to-12-million-people/), [[6]](https://candid.technology/meta-facebook-ai-ads-malware/)
* Paragraph 2 – [[1]](https://www.theregister.com/2025/05/27/fake_social_media_ads_ai_tool/), [[3]](https://about.fb.com/news/2023/05/how-meta-protects-businesses-from-malware), [[7]](https://thehackernews.com/2023/05/meta-takes-down-malware-campaign-that.html)
* Paragraph 3 – [[1]](https://www.theregister.com/2025/05/27/fake_social_media_ads_ai_tool/), [[4]](https://www.bleepingcomputer.com/news/security/fake-ai-editor-ads-on-facebook-push-password-stealing-malware/), [[5]](https://cloud.google.com/blog/topics/threat-intelligence/threat-actors-generative-ai-limited)
* Paragraph 4 – [[1]](https://www.theregister.com/2025/05/27/fake_social_media_ads_ai_tool/), [[2]](https://www.bleepingcomputer.com/news/security/fake-facebook-midjourney-ai-page-promoted-malware-to-12-million-people/), [[7]](https://thehackernews.com/2023/05/meta-takes-down-malware-campaign-that.html)
* Paragraph 5 – [[1]](https://www.theregister.com/2025/05/27/fake_social_media_ads_ai_tool/), [[2]](https://www.bleepingcomputer.com/news/security/fake-facebook-midjourney-ai-page-promoted-malware-to-12-million-people/), [[6]](https://candid.technology/meta-facebook-ai-ads-malware/)
* Paragraph 6 – [[1]](https://www.theregister.com/2025/05/27/fake_social_media_ads_ai_tool/), [[2]](https://www.bleepingcomputer.com/news/security/fake-facebook-midjourney-ai-page-promoted-malware-to-12-million-people/), [[4]](https://www.bleepingcomputer.com/news/security/fake-ai-editor-ads-on-facebook-push-password-stealing-malware/)
* Paragraph 7 – [[1]](https://www.theregister.com/2025/05/27/fake_social_media_ads_ai_tool/), [[5]](https://cloud.google.com/blog/topics/threat-intelligence/threat-actors-generative-ai-limited), [[6]](https://candid.technology/meta-facebook-ai-ads-malware/)

Source: [Noah Wire Services](https://www.noahwire.com)

## Bibliography

1. <https://www.theregister.com/2025/05/27/fake_social_media_ads_ai_tool/> - Please view link - unable to able to access data
2. <https://www.bleepingcomputer.com/news/security/fake-facebook-midjourney-ai-page-promoted-malware-to-12-million-people/> - Cybercriminals have been using Facebook advertisements and hijacked pages to promote fake AI services, such as MidJourney, OpenAI's SORA, ChatGPT-5, and DALL-E, to infect users with password-stealing malware. These malicious campaigns involve creating fraudulent Facebook communities that impersonate popular AI tools, tricking users into downloading malware-laden software. The malware targets Windows computers, stealing sensitive information like passwords, credit card details, and cryptocurrency wallet data. The stolen data is then sold on dark web markets or used by attackers to breach online accounts and commit fraud. ([bleepingcomputer.com](https://www.bleepingcomputer.com/news/security/fake-facebook-midjourney-ai-page-promoted-malware-to-12-million-people/?utm_source=openai))
3. <https://about.fb.com/news/2023/05/how-meta-protects-businesses-from-malware> - Meta's security researchers actively track and take action against hundreds of threat actors worldwide. In 2023, they detected and disrupted nearly ten new malware strains, including those posing as ChatGPT browser extensions and productivity tools. These malware families target individuals through email phishing, malicious browser extensions, ads, mobile apps, and various social media platforms, aiming to run unauthorized ads from compromised business accounts across the internet. Meta has identified these malware operations at different stages of their lifecycle and has seen rapid adversarial adaptation in response to their detection. Since March 2023, Meta has blocked and shared with industry peers more than 1,000 malicious links from being shared across its technologies and reported a number of browser extensions and mobile apps to peer companies. They have also issued cease and desist letters to individuals behind certain malware operations and referred them to law enforcement. ([about.fb.com](https://about.fb.com/news/2023/05/how-meta-protects-businesses-from-malware?utm_source=openai))
4. <https://www.bleepingcomputer.com/news/security/fake-ai-editor-ads-on-facebook-push-password-stealing-malware/> - A Facebook malvertising campaign targets users searching for AI image editing tools and steals their credentials by tricking them into installing fake apps that mimic legitimate software. The attackers exploit the popularity of AI-driven image-generation tools by creating malicious websites that closely resemble legitimate services and trick potential victims into infecting themselves with information-stealer malware. The attacks start with phishing messages sent to Facebook page owners or administrators, which will send them to fake account protection pages designed to trick them into providing their login information. After stealing their credentials, the threat actors hijack their accounts, take control of their pages, publish malicious social media posts, and promote them via paid advertising. Users who click the URL promoted in the malicious ad are sent to a fake web page impersonating legitimate AI photo editing and generating software, where they are prompted to download and install a software package. However, instead of AI image editing software, the victims install the legitimate ITarian remote desktop tool configured to launch a downloader that automatically deploys the Lumma Stealer malware. The malware then quietly infiltrates their system, allowing the attackers to collect and exfiltrate sensitive information like credentials, cryptocurrency wallet files, browser data, and password manager databases. This data is later sold to other cybercriminals or used by the attackers to compromise the victims' online accounts, steal their money, and promote further scams. ([bleepingcomputer.com](https://www.bleepingcomputer.com/news/security/fake-ai-editor-ads-on-facebook-push-password-stealing-malware/?utm_source=openai))
5. <https://cloud.google.com/blog/topics/threat-intelligence/threat-actors-generative-ai-limited> - Mandiant has observed information operations leveraging AI-generated and AI-manipulated video technologies to promote desired narratives since 2021. As other forms of AI-generated video technology improve, such as those involving the impersonation of individuals, they expect to see them increasingly leveraged. In May 2023, Mandiant reported to customers on DRAGONBRIDGE’s use of an AI-generated 'news presenter' in a short news segment-style video. They assessed that this video was likely created using a self-service platform offered by the AI company, D-ID. ([cloud.google.com](https://cloud.google.com/blog/topics/threat-intelligence/threat-actors-generative-ai-limited?utm_source=openai))
6. <https://candid.technology/meta-facebook-ai-ads-malware/> - Cybersecurity researchers from Trend Micro uncovered a scheme where threat actors manipulated paid Facebook promotions spotlighting large language models (LLMs) to disseminate malicious code. Their ultimate goal was to implant a treacherous browser add-on and steal users’ credentials. The hackers employed seemingly open tools like URL shorteners for URL redirection, Google Sites for hosting web content, and cloud storage such as Google Drives and Dropbox to house their malicious files. When the researchers shared their findings with Meta, they tracked the rogue actor and their tactics and deleted the deceitful pages and ads. Meta also pledged to bolster its detection mechanisms to root out analogous fraudulent content, incorporating insights gleaned from internal and external threat intelligence. The modus operandi of the threat actor entails leveraging paid Facebook promotions to allure unsuspecting victims. The ads exhibit faux profiles of marketing companies or departments, readily distinguishable by telltale signs like artificially inflated follower counts, counterfeit reviews, and a scant online history. These ads, adorned with the allure of productivity enhancement, amplified outreach, revenue augmentation, or even AI-driven learning, are designed to entice users into their trap. Some ads dangle access to the elusive Google Bard or Meta AI before users’ eyes, promising a world of AI-aided benefits. Once a user succumbs to temptation and clicks the ad link, they are steered to a simplistic webpage extolling the merits of LLMs. Ostensibly innocuous, this page harbours a link tempting users to download the coveted ‘AI package’. The devious malefactors employ encrypted archives with rudimentary passwords, like ‘999’ or ‘888’, to evade antivirus detection. ([candid.technology](https://candid.technology/meta-facebook-ai-ads-malware/?utm_source=openai))
7. <https://thehackernews.com/2023/05/meta-takes-down-malware-campaign-that.html> - Meta took steps to take down more than 1,000 malicious URLs from being shared across its services that were found to leverage OpenAI's ChatGPT as a lure to propagate about 10 malware families since March 2023. The development comes against the backdrop of fake ChatGPT web browser extensions being increasingly used to steal users' Facebook account credentials with an aim to run unauthorized ads from hijacked business accounts. Meta has identified these malware operations at different stages of their lifecycle and has seen rapid adversarial adaptation in response to their detection. Since March 2023, Meta has blocked and shared with industry peers more than 1,000 malicious links from being shared across its technologies and reported a number of browser extensions and mobile apps to peer companies. They have also issued a cease and desist letter to individuals behind certain malware operations and referred them to law enforcement. ([thehackernews.com](https://thehackernews.com/2023/05/meta-takes-down-malware-campaign-that.html?utm_source=openai))