

Forum:[Hack Forums](#)

Thread: [Artificial Intelligence PromptLock: First AI-Powered Ransomware Emerges](#)

MadScientist - September 2, 2025, 9:18 am

According to ESET, which discovered the threat, the AI-powered ransomware is only proof-of-concept (PoC) or work-in-progress for now, but appears to be designed with all the functionality of traditional ransomware.

Dubbed PromptLock, the malware is written in GoLang and relies on OpenAI's GPT-OSS:20b, an open-weight model that can be used without proprietary restrictions.

The threat, ESET explains in a series of posts on social media, relies on hard-coded prompts to generate Lua scripts on the fly, and uses these scripts to perform operations such as filesystem enumeration, file inspection, data exfiltration, and encryption.

Both Windows and Linux variants of Promptlock have been observed, and the generated Lua scripts are cross-platform compatible, ESET warns. The ransomware uses the SPECK 128-bit algorithm for file encryption

"Based on the detected user files, the malware may exfiltrate data, encrypt it, or potentially destroy it. Although the destruction functionality appears to be not yet implemented," the cybersecurity firm notes.

While the idea of AI-powered ransomware roaming around sounds frightening, Promptlock attacks would require several pre-conditions that are unlikely to be met in typical networks.

First, Promptlock uses the GPT-OSS:20b model locally via the Ollama API, which means that Ollama needs to be running on the victim's system. This would require resources that typical computer systems do not have.

ESET observed Promptlock sending requests on the local network, hypothesizing it is reaching either to a locally running Ollama server, or to an internal proxy redirecting those requests to an external Ollama server.

As security researchers have pointed out, the success of a Promptlock attack also depends on the victim having poor network segmentation and failing to implement prompt guardrails, or allowing outgoing traffic aimed at LLM ports and protocols.

However, ESET has pointed out that the malware appears to be only a concept and not fully operational, and that it has not been deployed in the wild yet.

Source - <https://www.securityweek.com/promptlock-first-ai-powered-ransomware-emerges/>