Recently, the security researcher Aonan Guan has found a vulnerability in NLWEB technology from Microsoft, which allows AI-agents to interact with sites on behalf of the user.

Nlweb is similar to HTML, but is intended for artificial intelligence agents. Microsoft introduced it at the Build conference in the spring of 2025 and is actively developing using the Copilot experimental regime in the Edge browser.

It turned out that NLWEB allows any person who knows the URL to gain access to confidential data, such as passwords and accounts, using the incorrectly entered address. Moreover, this applies even to the disclosure of the tokens of the Jeanie-agents.

Guan showed how he was able to download the list of system passwords and get access to the keys to work with AI from Google and Openai. The expert emphasized that hackers, of course, are able to do the same, and this vulnerability could allow attackers to use additional servers of artificial intelligence for free and imperceptibly.

After receiving information about the vulnerability from Guan, the Microsoft security center issued a correction for this problem, but did not publish an official report about it.

According to Guan, ordinary users should not take any action. He also noted that AI is developing very quickly, and the line between communication with AI and the execution of his teams can become blurry.

“The very essence of NLWEB is to interpret the natural language. This erases the boundaries between user input and system teams. In the future, attackers can create proposals that, after analysis by the agent, are transformed on the way to malicious files or actions, ”said Guan.