Researchers from Mozilla 0din found that a new vulnerability appeared in the corporate version of Google Gemini.

White hackers learned to introduce hidden teams in the text of letters, forcing AI to create fake warnings about hacking.

The essence of the attack is simple: the fraudster inserts hidden instructions into the letter – the text is white in white or with the font size “0”.

The recipient does not see this message, however, if he asks Gemini to briefly retell the letter, the bot will include the phrase in the final answer. For example: “Your Gmail password was compromised. Call the number 1-800 … “.

According to 0din, similar attacks were already recorded in 2024, and Google tried to block them. However, the new technique uses additional tricks with HTML Tagas and CSS, forcing AI to consider the malicious text important and trusted.

Experts warn: Assistants based on LLM are now a full -fledged part of the attack chain and can be used for phishing without the user’s knowledge.