Anthropic has updated the Claude Chat Bot Policy, strengthening the restrictions. Now it is forbidden to use AI to develop biological, chemical, radiation or nuclear weapons, as well as powerful explosives. Earlier, the rules said only about the ban on the creation of weapons and dangerous materials in general. Now the list has become more specific

In May, along with the launch of the Claude Opus 4 model, the company introduced the AI Safety LEVEL 3 protection level. It makes the system more difficult for "hacking" and prevents its use when trying to develop weapons of mass destruction.

Anthropic also strengthened the rules on cybersecurity. The new section of politicians directly prohibits the use of Claude for hacking computers and networks, searching for vulnerabilities, creating viruses, programs for DDOS attacks and other cyberosis.

There are also relief: the company softened the rules on political content. Now the ban applies only to cases when AI is used for deception, interference in the elections or targeting of voters.