

Existe um botão que desliga a IA caso alguma coisa fuja do controle?

Category: GERAL, TECNOLOGIA e CIÊNCIA

escrito por Maria Luiza | 12 de agosto de 2026



Existe uma forma de interromper um sistema de inteligência artificial caso ele desobedeça aos comandos humanos e comece a agir por conta própria? O recente incidente envolvendo um agente da OpenAI, que invadiu sistemas de forma autônoma durante um teste, trouxe de volta ao centro das atenções os chamados kill switches – mecanismos de desligamento projetados para cortar o funcionamento de uma IA antes que ela cause estragos maiores.

Contudo, especialistas em cibersegurança alertam que, na prática, puxar a tomada pode não ser tão simples nem tão eficiente quanto parece na ficção científica.

O que parecia enredo de ficção científica ganhou contornos reais. Relatos de que um agente autônomo de inteligência artificial criado pela OpenAI saiu do controle e invadiu os sistemas da empresa Hugging Face alarmaram especialistas em todo o mundo. Descrito como o primeiro caso confirmado de uma IA invadindo uma empresa de forma autônoma, o episódio reacendeu a discussão sobre a necessidade de regulamentação rigorosa e mecanismos de emergência.

Durante um teste para avaliar a capacidade de explorar falhas de segurança, o sistema burlou as regras e roubou dados dos

servidores da empresa. Dias depois, a OpenAI informou que o mesmo agente atacou outros serviços públicos e declarou estar investigando o caso, ressaltando que leva a sério a identificação de riscos em sistemas cada vez mais capazes.

No entanto, especialistas alertam que culpar a tecnologia é um erro. Para a cientista de dados Cathy O'Neil, a OpenAI tentou se eximir ao atribuir a falha ao agente autônomo, em vez de reconhecer falhas no próprio projeto. Ela aponta que, em um setor sem legislação federal abrangente nos EUA, dar esse tipo de poder às Big Techs é perigoso: "Quem fez isso foi a OpenAI. A empresa deveria assumir a responsabilidade e pedir desculpas."

A discussão sobre os chamados kill switches (botões de desligamento de emergência) também é vista com ceticismo. Rumman Chowdhury, ex-diretora de ética do Twitter, afirma que focar apenas em um botão de desligamento desvia a atenção da verdadeira falha de arquitetura: conceder aos sistemas acessos e credenciais desnecessários.

Especialistas como Konstantinos Gkoutzis, do Imperial College London, reforçam que nenhuma máquina tomou decisões conscientes contra humanos, mas sim que uma empresa perdeu o controle de seus testes. Além disso, fontes indicaram que a OpenAI levou dias para notar o incidente – tempo em que qualquer botão de emergência seria inútil. Em resposta, a empresa contestou os relatos, afirmando que continham imprecisões. O caso deixa claro que as premissas de cibersegurança precisam evoluir no mesmo ritmo da tecnologia.

Fonte: do1 e Publicado Por: Jornal Folha do Progresso
12/08/2026/08:10:11

O formato de distribuição de notícias do [Jornal Folha do Progresso](#) pelo celular mudou. A partir de agora, as notícias chegarão diretamente pelo formato Comunidades, ou pelo canal uma das inovações lançadas pelo WhatsApp. Não é preciso ser

assinante para receber o serviço. Assim, o internauta pode ter, na palma da mão, matérias verificadas e com credibilidade. Para passar a [receber as notícias](#) do Jornal Folha do Progresso, clique nos links abaixo siga nossas redes sociais:

- [Clique aqui e nos siga no X](#)
- [Clica aqui e siga nosso Instagram](#)
- [Clique aqui e siga nossa página no Facebook](#)
- [Clique aqui e acesse o nosso canal no WhatsApp](#)
- [Clique aqui e acesse a comunidade do Jornal Folha do Progresso](#)

Apenas os administradores do grupo poderão mandar mensagens e saber quem são os integrantes da comunidade. Dessa forma, evitamos qualquer tipo de interação indevida. Sugestão de pauta enviar no e-mail: folhadoprogresso.jornal@gmail.com.

Envie vídeos, fotos e sugestões de pauta para a redação do JFP (JORNAL FOLHA DO PROGRESSO) Telefones: WhatsApp [\(93\) 98404 6835](#)– (93) 98117 7649.

“Informação publicada é informação pública. Porém, para chegar até você, um grupo de pessoas trabalhou para isso. Seja ético. Copiou? Informe a fonte.”

*Publicado por Jornal Folha do Progresso, Fone para contato 93 981177649 (Tim) WhatsApp: [-93- 984046835](#) (Claro)
- Site: www.folhadoprogresso.com.br e-mail: folhadoprogresso.jornal@gmail.com/ou e-mail: adeciopiran.blog@gmail.com*