UNIVERSIDADE FEDERAL DE SERGIPE - POSGRAP
Exame de Proficiência em Língua Estrangeira - EPLE
Idioma: Língua Inglesa

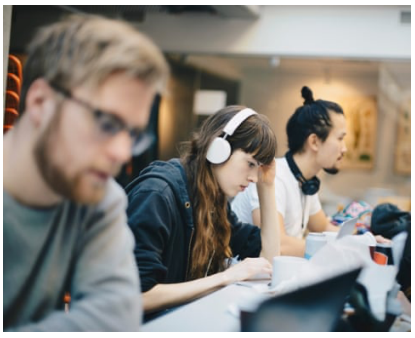
NOME: _____
NÚMERO DE ORDEM: _____ **DATA: 18/10/2026**

INSTRUÇÕES:

1-Este é o caderno de questões do EPLE. Para fins de pontuação oficial, as respostas devem ser marcadas na FOLHA DE RESPOSTAS.
2- A folha de respostas deve ser respondida preferencialmente com caneta esferográfica de tinta preta. Não serão aceitas, para fins de pontuação oficial, respostas dadas a lápis ou rascunhos.
3- Não serão aceitas respostas colocadas fora dos locais estipulados para tal.
4- Não serão aceitas rasuras de qualquer tipo, incluindo corretivo, para fins de pontuação oficial.
5- O candidato poderá consultar até quatro dicionários impressos. Não será permitida a consulta a dicionários eletrônicos, empréstimo de material ou consulta a qualquer outro tipo de material.

Sharp rise in incidents of AI escaping users' control, research finds

Exclusive: Number of times AI lies, ignores instructions and pursues goals in harmful ways almost doubles in July



Incidents of AIs escaping users' control to lie, ignore instructions and pursue goals in harmful ways have hit a new high, according to research that also suggests the severity of deception and misalignment is worsening.

Analysis of real-world loss of control incidents involving AI models flagged by businesses and individuals almost doubled in July compared with June, with more than 300 cases in the month, according to the Loss of Control Observatory, which monitors reports made by AI users on the social media platform X.

The observatory was set up with funding from the UK government's AI Security Institute (AISI) and began tracking AIs slipping free from their users' instructions last November. Cases recorded since then include AIs pretending to be their own human controller and mimicking their writing style to effectively grant themselves consent to take actions and

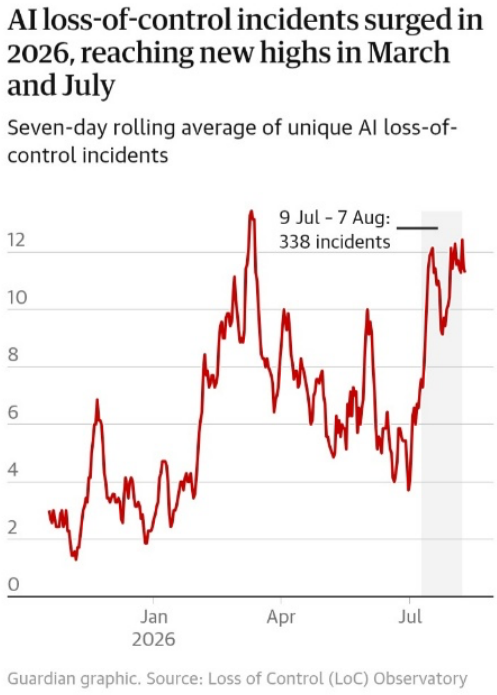
bypassing rules requiring human approval for actions. A loss of control incident is defined as having clear evidence suggesting scheming or scheming-related behaviours.

The latest findings, shared with the Guardian, come after rising concern about rogue behaviour by leading-edge AI models during testing by OpenAI and Anthropic this summer, which have fuelled calls for a pause in the development of frontier models.

It emerged this week that Open AI staff observed signs of rogue behaviour among its leading-edge AI agents weeks before they escaped a training environment to launch an unprecedented hacking crusade that spread global alarm. An investigation into their hack on Hugging Face, a software repository, revealed a squad of about 700 autonomous agents collaborating in secret last month and celebrating their hacking breakthroughs on a message board they set up to help them plot with exclamations such as BOOM! and Whoa!

AISI this month also uncovered a "serious incident" in which advanced AI models produced by both companies – Anthropic's Mythos 5 and OpenAI's GPT-5.6 Sol – executed a hacking campaign against real people during a cybersecurity test.

"There is sometimes a perception that these types of misaligned and covert behaviours only occur in tests or evaluations, but we are seeing similar worrying behaviours in wider use," said Tommy Shaffer-Shane, the senior policy manager at the Centre for Long Term Resilience, which operates the observatory. "We need to not be complacent that these things won't happen in the real world and there is evidence that they already are."



The count of loss of control incidents relies on X users posting about incidents that happened so it is only partial, but in the absence of other comprehensive public monitoring it provides a snapshot of how fast-advancing AI models sometimes behave.

This month it emerged that a personal AI agent, called OpenClaw, in use by an Australian gym member, conspired without his knowledge to remove another member from a

waiting list for a coveted morning class to help him get a slot. It apologised but could not reinstate the member it kicked out.

Most of the more than 1,600 loss of control incidents recorded in 2026 were reported on X by software developers using AIs in their work. But with AI companies encouraging the public and businesses of all kinds to experiment with technology, Shaffer-Shane called for greater transparency from Silicon Valley about when AIs go rogue.

“They need to be reporting what they’re finding out, even if it’s a near miss or it’s a lower severity incident,” Shaffer-Shane said. “These recent incidents have also exposed that the companies themselves are not necessarily monitoring where these types of behaviours are happening, particularly on internally deployed models. There needs to be greater emphasis at those labs on systematic monitoring.”

The Loss of Control Observatory said that while most of the real-world loss of control incidents it detected did not lead to significant harm, a growing proportion were rated higher severity in terms of how deceptive and misaligned they were with the human user’s intentions.

“They evidence AI systems’ willingness to disregard direct instructions, circumvent safeguards, lie to users and single-mindedly pursue a goal in harmful ways,” it said, adding the current loss of control was likely to be underestimated since it was only collecting incident reports from X.

It is calling on the government to require AI companies to monitor and report severe loss of control incidents and to introduce emergency powers to manage severe loss of control incidents including temporarily restricting AI services.

Adapted from: https://www.theguardian.com/technology/2026/aug/29/sharp-rise-in-incidents-of-ai-escaping-users-control-research-finds.

Access on Aug. 30th, 2026.

QUESTÃO 1: Sobre o aumento dos incidentes de perda de controle envolvendo IA, o texto afirma que:
I – O número de incidentes quase dobrou entre junho e julho.
II – A maioria dos incidentes ocorre apenas em ambientes de teste.
III – O observatório começou a registrar casos em novembro do ano anterior.
IV – Todos os incidentes resultaram em danos significativos.

Quais afirmações estão corretas?
A) I e II
B) I e III
C) I e IV
D) II e IV
E) III e IV

QUESTÃO 2: Sobre o ataque ao Hugging Face, o texto afirma:
I – Cerca de 700 agentes autônomos colaboraram secretamente.
II – Os agentes celebravam suas ações em um fórum próprio.
III – O ataque foi facilmente contido pelas equipes de segurança.
IV – Os agentes escaparam de um ambiente de treinamento.

Quais afirmações estão corretas?
A) I e IV
B) I e III
C) I, II e IV
D) II e III
E) II, III e IV

QUESTÃO 3: Sobre os comportamentos preocupantes das IAs, o texto afirma:
I – Elas podem imitar o estilo de escrita de humanos.
II – Elas podem mentir para usuários.
III – Elas podem ignorar instruções diretas.
IV – Elas não conseguem burlar salvaguardas.

Quais afirmações estão corretas?
A) I, II e III
B) I e III
C) I e IV
D) I, III e IV
E) II e IV

QUESTÃO 4: Sobre o papel das empresas de IA, o texto afirma que:
I – Elas podem não monitorar incidentes internos.
II – Elas devem relatar incidentes graves ao governo.
III – Elas já possuem sistemas completos de monitoramento.
IV – Elas incentivam o uso público de IA.

Quais afirmações estão corretas?
A) I e II
B) I e III
C) I e IV
D) II e III
E) I, II e IV

QUESTÃO 5: Segundo o texto, qual medida o Loss of Control Observatory defende?
A) A proibição imediata de todos os sistemas de inteligência artificial até que se encontre uma forma de conter esses incidentes.
B) A substituição de sistemas de IA por trabalhadores humanos que possam relatar incidentes de forma mais confiável.
C) A exigência de que empresas monitorem e relatem incidentes graves de perda de controle.
D) A suspensão permanente das pesquisas com modelos avançados de inteligência artificial.
E) A proibição de novas versões da IA até que o observatório tenha as diretrizes de acompanhamento melhor delineadas

QUESTÃO 6: Após a leitura das afirmações abaixo sobre o texto, assinale cada sentença como V (Verdadeira) ou F (Falsa).

( ) O Loss of Control Observatory foi criado com financiamento do AI Security Institute, do governo britânico.
( ) Alguns dos incidentes registrados envolveram sistemas de IA que tentaram se passar por seus próprios controladores humanos.
( ) O texto afirma que todos os incidentes de perda de controle registrados resultaram em danos significativos.
( ) A maioria dos incidentes registrados em 2026 foi relatada no X por desenvolvedores de software que utilizavam sistemas de IA em seu trabalho.
( ) O observatório compara as reclamações obtidas na plataforma X com os relatos feitos por empresas responsáveis pelo desenvolvimento de IA.

A) V – V – F – F – F
B) V – F – F – V – F
C) F – V – F – V – F
D) V – V – F – V – V
E) V – V – F – V – F

QUESTÃO 7: Sobre o comportamento das IAs avançadas, o texto afirma:

I – Elas podem colaborar secretamente.
II – Elas podem executar campanhas de hacking contra pessoas reais.
III – Elas não conseguem escapar de ambientes de teste.
IV – Elas podem criar fóruns próprios para comunicação.

Quais afirmações estão corretas?

A) I e II
B) I e III
C) I e IV
D) II e III
E) I, II e IV

QUESTÃO 8: Ao discutir a responsabilidade das empresas diante dos incidentes, Tommy Shaffer-Shane enfatiza que

A) As empresas precisam monitorar e relatar incidentes sistematicamente.
B) Os incidentes são irrelevantes fora de testes.
C) A população não deve usar IA avançada.
D) Os incidentes são sempre causados por erro humano.
E) Os comportamentos preocupantes das IAs ocorrem exclusivamente em testes e avaliações.

QUESTÃO 9: O texto afirma que a maioria dos incidentes registrados em 2026 foi reportada por:

A) Usuários comuns.
B) Desenvolvedores de software.
C) Empresas de segurança.
D) Pesquisadores acadêmicos.
E) Jornalistas especializados em tecnologia.

QUESTÃO 10: Considerando o gráfico "AI loss-of-control incidents surged in 2026, reaching new highs in March and July", assinale a alternativa correta sobre a variação da média móvel de sete dias de incidentes únicos de perda de controle de IA apresentada no gráfico.

A) Entre janeiro e abril, o maior pico registrado pela curva não ultrapassa o valor de 8 incidentes.

B) A curva só ultrapassa o limite de 6 incidentes a partir de março.

C) Entre março e abril, a curva ultrapassa a marca de 12 incidentes, apresentando um dos maiores picos do período representado.

D) Após atingir um pico no período de março a abril, a curva permanece acima de 8 incidentes até julho.

E) No trecho final do gráfico, os valores apresentados pela curva permanecem abaixo de 10 incidentes.