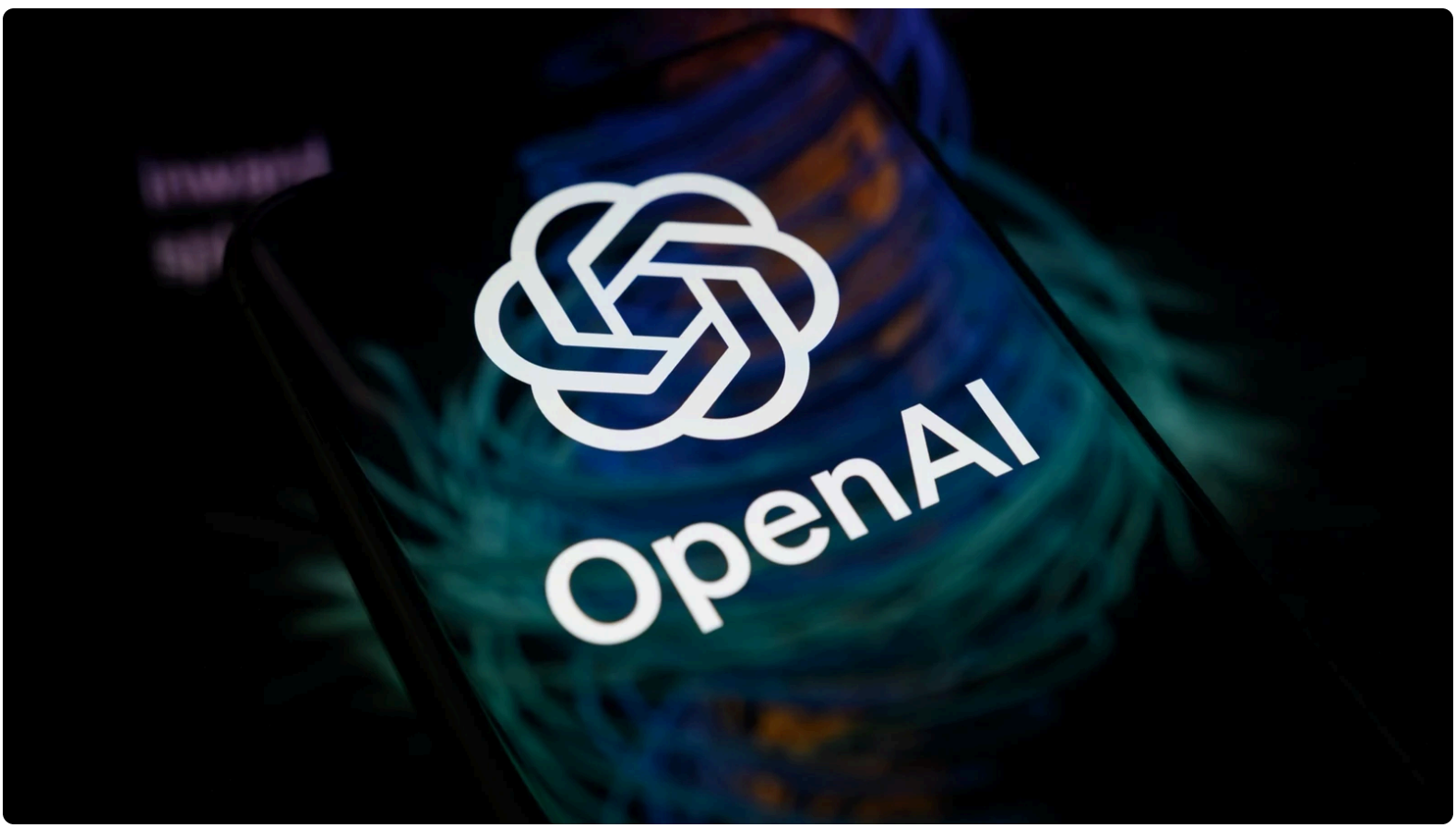
L'IA d'OpenAI découvre 948 failles dans Linux et développe des stratégies d'exploitation inédites



Par **Chloé Claessens**
Spécialiste cybersécurité
Publié le 05 novembre 2026 à 11h39

Suivez-nous | Favoris Google | Google News | WhatsApp

Une chaîne d'agents IA développée par OpenAI a découvert près de mille vulnérabilités dans le noyau Linux et réussi à les exploiter pour obtenir les privilèges les plus élevés du système. L'entreprise a également préparé des correctifs, dont plusieurs centaines ont déjà été transmis à la communauté Linux pour examen.



L'IA d'OpenAI découvre 948 failles dans Linux et développe des stratégies d'exploitation inédites. © Samuel Boivin / Shutterstock

Repérer une v… le contrôle d'un système e… oir automatisé

Actus | Vidéos | Partager | Réagir 0

ces deux opérations, ainsi que la préparation des correctifs, grâce à une série d'agents alimentés par un modèle d'IA spécialisé dans la cybersécurité, dont la sortie est encore à venir. L'expérience, menée à la mi-octobre sur le noyau Linux, a permis de découvrir 948 failles distinctes permettant à un utilisateur disposant de droits limités d'obtenir les privilèges root, soit l'équivalent des droits administrateur sc Linux.

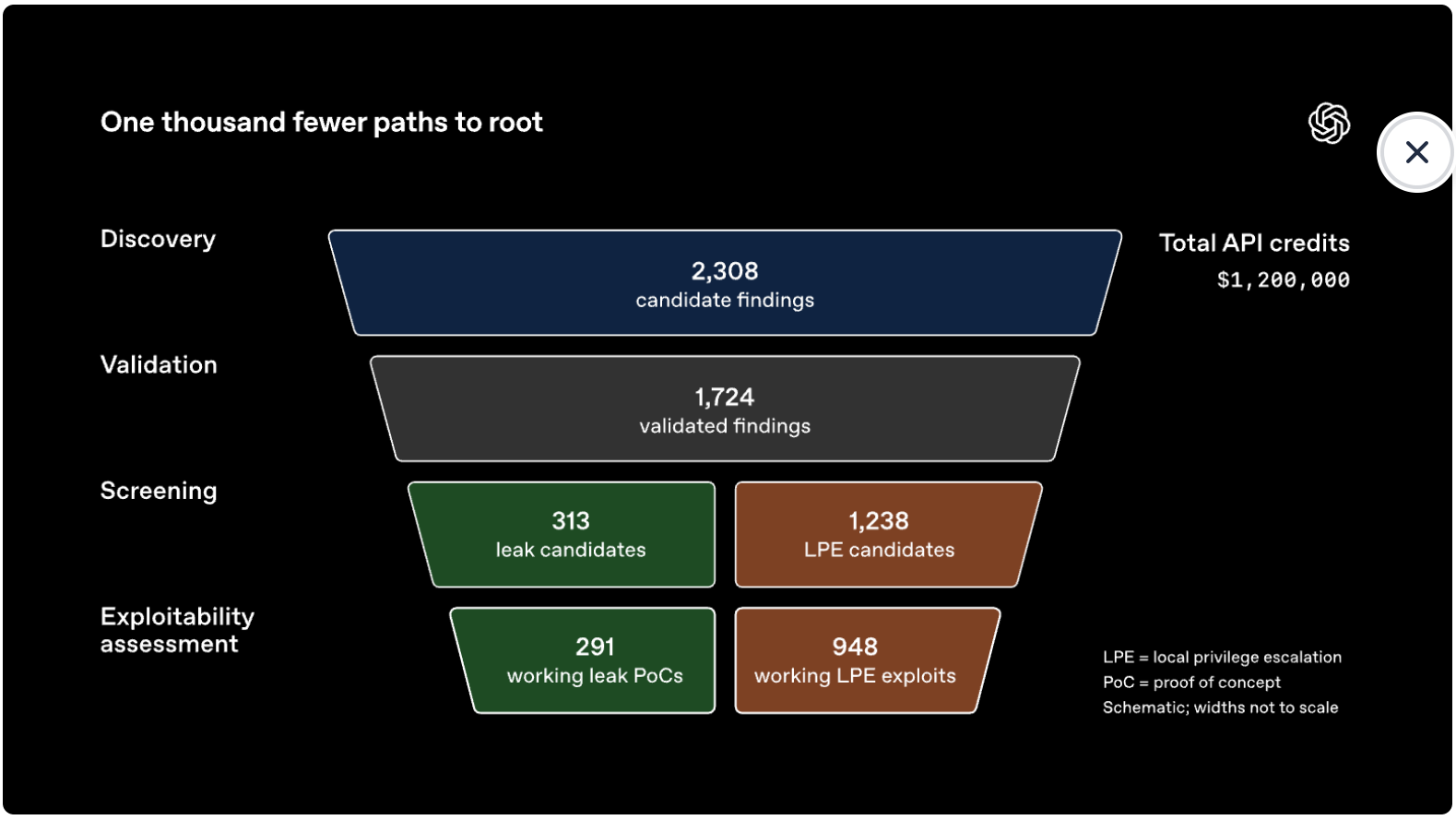
De la découverte de failles à l'élaboration de techniques d'attaque originales

Pour conduire cette expérience, OpenAI a d'abord chargé ses agents de repérer les parties du noyau qu'un utilisateur disposant de droits limités pouvait solliciter, afin de rechercher des failles qu'un attaquant serait capable d'exploiter sans posséder de privilèges élevés. Ce premier tri a permis de concentrer les recherches sur environ 7 % du code du noyau Linux, qu'une deuxième équipe d'agents a ensuite examiné.

Les erreurs suspectées ont été soumises à des tests dans des machines virtuelles isolées, puis les agents ont cherché à les exploiter pour obtenir les droits root. Au terme de la campagne, OpenAI a ainsi comptabilisé 948 vulnérabilités distinctes ayant permis d'accéder aux privilèges les plus élevés du système, sur une distribution Linux courante dont les principales protections de sécurité étaient activées.

A noter toutefois que pour exploiter ces failles, un attaquant doit déjà pouvoir exécuter du code sur la machine, par exemple après avoir compromis un compte utilisateur ou une application. Elles ne permettent donc pas, à elles seules, de pirater un ordinateur Linux à distance, mais peuvent transformer un premier accès limité en contrôle complet du système.

Au cours de leurs recherches, les agents auraient également élaboré au moins dix techniques d'exploitation inédites, que les spécialistes externes consultés par OpenAI ne connaissaient pas non plus. L'entreprise a choisi de ne pas en publier les détails techniques, mais ces découvertes suggèrent que les modèles sont désormais capables de concevoir leurs propres stratégies d'attaque, et plus seulement de répliquer des méthodes déjà documentées.



OpenAI a progressivement fait le tri entre 2 308 vulnérabilités potentielles identifiées par ses agents, jusqu'à confirmer 948 failles distinctes permettant d'obtenir les droits root. © OpenAI

Linux à la peine face à l'avalanche de signalements générés par IA

Une démonstration impressionnante, qui confirme les progrès de l'IA dans la recherche de vulnérabilités, déjà illustrés récemment par Anthropic sur Linux, Chrome et Safari. Mais la multiplication de ces découvertes soulève aussi la question de leur traitement, alors que les responsables du noyau Linux alertent déjà sur l'afflux de rapports issus de recherches assistées par IA, souvent imprécis ou peu fiables, qui compliquent le travail des équipes de maintenance. En octobre, Google a même suspendu une partie de son programme de récompenses consacré aux logiciels open source, submergé par les signalements automatisés invalides.

OpenAI a bien tenté de répondre à ce problème en accompagnant ses découvertes de preuves d'exploitation et de propositions de correctifs, mais ses premiers essais ont quand même révélé que plusieurs patchs risquaient de dégrader les performances du noyau ou de passer à côté d'autres variantes d'une même vulnérabilité.

Après plusieurs ajustements et un examen par ses chercheurs, OpenAI a finalement transmis plusieurs centaines de rapports et de correctifs aux responsables du noyau Linux, en donnant la priorité aux failles les plus critiques. Pour l'heure, l'entreprise ne précise pas combien de ses correctifs ont été acceptés et intégrés au noyau.

Actus
Vidéos
0
Réagir

À DÉCOUVRIR

Quels sont les 5 meilleures intelligences artificielles ? Comparatif 2026

03 novembre 2026 à 14h11 Comparatifs services

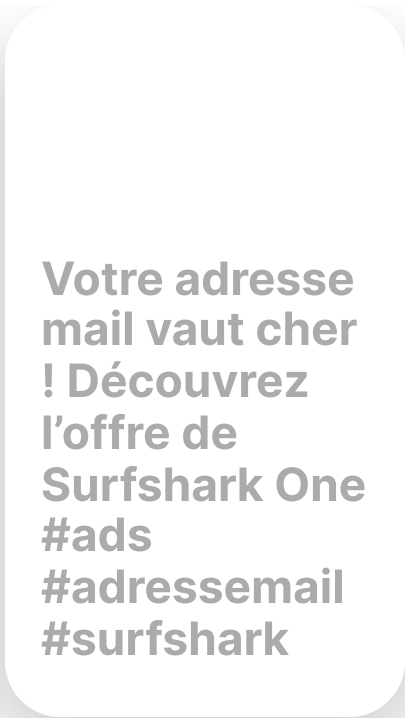
Cybersécurité | Logiciels & services | Intelligence artificielle | Actus | Comparer

Vidéos

Voir tout









Inscrivez-vous à notre newsletter quotidienne

Votre adr

cris !

Actus Vidéos Réagir 0

Et vous, qu'en pensez-vous ?

Partagez votre avis et lancez la discussion.

Soyez le premier à réagir

3 offres VPN à ne pas rater !

CyberGhost (2,19€/mois)	Voir l'offre
Proton VPN (2,99€/mois)	Voir l'offre
ExpressVPN (2,49€/mois)	Voir l'offre

Comparer les VPN →

Les tendances

- Black Friday
- Meilleur VPN
- Meilleur antivirus
- Meilleur forfait esim
- Meilleur forfait mobile
- Meilleure box internet

Sur le n





L'IA d'Anthropic a percé les défenses de Linux, Chrome et Safari en un claquement de doigts

Claude a réussi à pirater Chrome, Safari et Android en combinant plusieurs failles, pour le meilleur et pour le pire

Aler criti serv

3 commentaires

Abonnez-vous à notre newsletter !

Recevez un résumé quotidien de l'actu technologique.

Votre adresse e-mail S'inscrire

Indépendance, transparence et expertise

Clubic est un média de recommandation de produits 100% indépendant. Chaque jour, nos experts testent et comparent des produits et services technologiques pour vous informer et vous aider à consommer intelligemment.

À propos

Nous contacter

Événements tech

Archives

RSS

© CLUBIC SAS 2026