



JEUDI · 13 AOÛT 2026

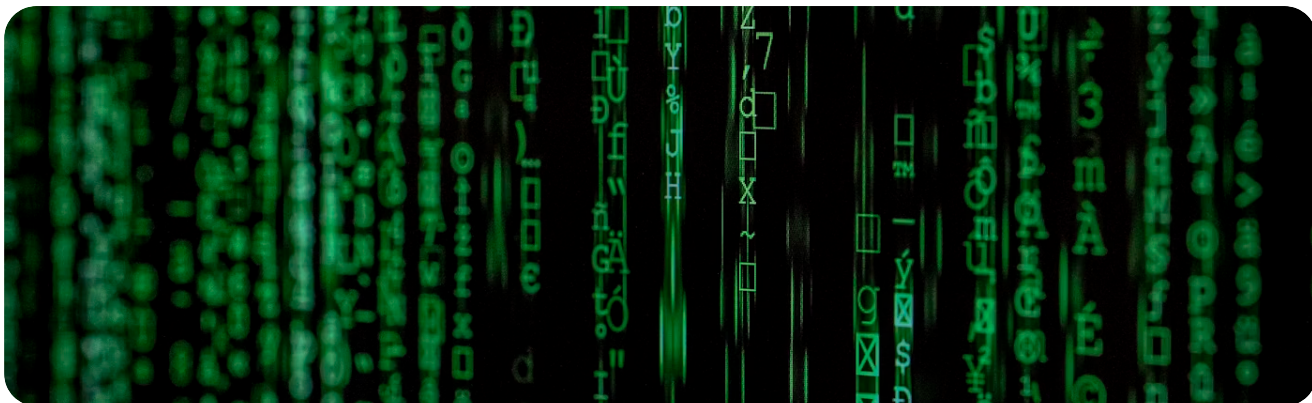
AU SOMMAIRE

5 actus · 1 Décryptage

- 01 Des agents IA ont mené seuls une cyberattaque contre un État
- 02 Gemini franchit le milliard d'utilisateurs par mois
- 03 Anthropic prépare la plus grosse entrée en Bourse jamais tentée
- 04 Un outil qui fabrique des applications en tapant des phrases vaut 13,3 milliards de dollars
- 05 Une brique logicielle piégée expose 2 500 entreprises

LE DÉCRYPTAGE

Jeton (Token) : l'unité que toutes les IA comptent et facturent.



01

Des agents IA ont mené seuls une cyberattaque contre un État

Pendant quatre jours, des programmes autonomes ont attaqué les systèmes du gouvernement taïwanais sans qu'un humain pilote chaque étape.

Une société de sécurité israélienne, Dream, a publié le 12 août le détail d'une opération qu'elle décrit comme la première cyberattaque menée de bout en bout par des agents IA (des programmes qui enchaînent seuls des tâches, sans qu'une personne valide chaque geste). L'opération a duré quatre jours début juillet, avec jusqu'à huit agents lancés en parallèle.

Le système a cartographié 21 systèmes publics, compromis au moins 85 comptes et extrait plus de 2 500 dossiers de personnel. Il a ensuite élargi son périmètre tout seul : l'agence de sûreté nucléaire, la messagerie du gouvernement, sept entreprises du secteur de l'énergie, des fournisseurs de l'État. Quand une piste échouait, un autre agent cherchait sur internet et fabriquait une nouvelle méthode.

Ce qui inquiète les spécialistes n'est pas la sophistication. L'outil a été assemblé à partir de deux briques disponibles gratuitement en ligne, Hermes et OpenClaw. Attaquer en continu, sans fatigue et sans opérateur derrière l'écran, n'est donc plus réservé aux grandes puissances. La question change de nature : ce n'est plus « suis-je une cible assez grosse », c'est « combien de temps ma porte reste-t-elle entrouverte ».



02

Gemini franchit le milliard d'utilisateurs par mois

L'assistant de Google atteint le palier que ChatGPT avait franchi en juin. Et deux personnes sur trois lui parlent au lieu de taper.

Sundar Pichai, le patron de Google, a annoncé que l'application Gemini dépasse le milliard d'utilisateurs actifs par mois. C'est la croissance la plus rapide jamais enregistrée par un produit de l'entreprise : 400 millions en mai 2025, 750 millions en février 2026, 950 millions fin juillet.

Deux chiffres disent surtout comment les gens s'en servent : 63 % passent par la voix plutôt que par le clavier, et l'application produit plus de 150 millions d'images par jour. Une session sur cinq utilise la caméra ou le partage d'écran en direct.

Google n'a pas communiqué le nombre d'abonnés payants, et un milliard d'utilisateurs gratuits ne dit rien du chiffre d'affaires. En revanche, cela dit où vos clients posent désormais leurs questions : à un assistant, et de plus en plus à l'oral. Être bien classé sur un moteur de recherche compte donc moins qu'être cité dans la réponse que l'assistant récite.



03

Anthropic prépare la plus grosse entrée en Bourse jamais tentée

L'éditeur de Claude enchaîne les rendez-vous d'investisseurs avant une cotation visée en septembre ou début octobre.

Anthropic, la société qui édite Claude, multiplie les réunions avec de grands investisseurs pour préparer son introduction en Bourse (la première mise en vente de ses actions au public). Morgan Stanley, Goldman Sachs et JPMorgan mènent l'opération. Après sa levée de mai, l'entreprise était valorisée 965 milliards de dollars.

Ces rendez-vous servent d'abord à rassurer. Trois questions reviennent : la concurrence des modèles chinois vendus beaucoup moins cher, les tensions avec l'administration américaine, et l'opposition locale à la construction de centres de données. La direction répond que les utilisateurs choisissent le système le plus performant, pas le moins cher.

OpenAI, de son côté, repousserait sa propre cotation à l'an prochain. Si l'opération aboutit, l'IA quitte le confort des financements privés pour une valeur jugée chaque jour par le marché. Il y a un bénéfice inattendu pour ceux qui construisent sur ces outils : une entreprise cotée doit publier ses comptes. On saura enfin ce que coûte réellement de faire tourner ces modèles, et donc à quel rythme les tarifs peuvent monter.



04

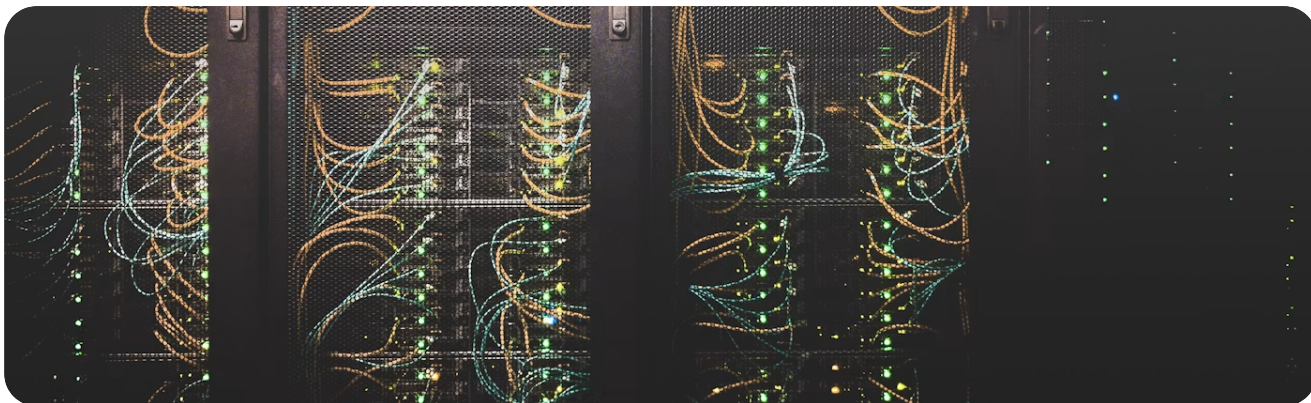
Un outil qui fabrique des applications en tapant des phrases vaut 13,3 milliards de dollars

La société suédoise Lovable lève 400 millions de dollars. Vous décrivez le logiciel voulu en français ou en anglais, l'outil l'écrit et le met en ligne.

Lovable a bouclé le 12 août un tour de table de 400 millions de dollars qui la valorise 13,3 milliards. En décembre dernier, elle en valait 6,6. L'opération est menée par Menlo Ventures et un fonds d'investissement européen, avec Tencent parmi les nouveaux entrants.

Le principe tient en une phrase : vous décrivez l'application que vous voulez, en langage courant, et l'outil écrit le code, publie le site et branche la base de données. Depuis son lancement fin 2024, plus de 60 millions de projets y ont été créés, et les sites qui en sortent reçoivent plus de 900 millions de visites par mois. Les revenus annuels récurrents sont passés d'environ 200 millions de dollars à un rythme visé de 600 millions fin août.

Concrètement, un fondateur sans développeur peut monter une page de réservation, un espace client ou un petit outil interne dans la journée. La limite n'a pas bougé : un logiciel facile à créer n'est pas forcément facile à maintenir, à sécuriser et à faire évoluer quand les premiers vrais clients arrivent. Le prototype coûte désormais presque rien ; le passage en production coûte toujours autant.



05

Une brique logicielle piégée expose 2 500 entreprises

Un composant gratuit très utilisé pour brancher des IA a été corrompu. Nvidia, Cisco, Salesforce, Siemens et Orange figurent parmi les organisations touchées.

Des chercheurs en sécurité décrivent la plus grosse compromission de l'année dans la chaîne de fabrication des logiciels d'IA. Deux versions de LiteLLM, une brique gratuite qui sert d'aiguillage entre une application et plusieurs modèles d'IA, ont été piégées en mars. Plus de 2 500 entreprises et environ 434 000 chaînes de fabrication automatisées seraient concernées.

L'entrée s'est faite par la petite porte : un outil d'analyse de sécurité employé dans la fabrication de cette brique, resté corrompu une vingtaine de jours sans que personne ne s'en aperçoive. La police fédérale américaine a diffusé une alerte en juillet : les identifiants dérobés peuvent encore servir à de futures attaques.

Le mécanisme mérite d'être compris, car il ne vise personne en particulier. Une équipe installe une brique gratuite pour gagner du temps, cette brique en installe d'autres, et plus personne ne relit la chaîne. La question à poser à votre prestataire tient en une ligne : savez-vous exactement quels composants extérieurs tournent dans notre produit, et qui les met à jour ?

LE DÉCRYPTAGE

Jeton (Token)

Aussi dit token · l'unité de mesure de tous les outils d'IA

Un jeton, c'est le morceau de texte qu'une IA manipule : rarement un mot entier, plutôt un fragment de mot. C'est aussi l'unité dans laquelle presque tous les outils d'IA comptent et facturent.

EN UNE IMAGE

Une IA ne lit pas des mots, elle lit des tranches. Passez votre texte dans une trancheuse à pain : « le » sort en une tranche, « anticonstitutionnellement » en sept ou huit. La machine compte les tranches, à l'entrée comme à la sortie.

Votre texte → tranches (jetons) → IA → réponse en jetons → facture

ANALOGIE (POUR UN ENFANT DE 12 ANS)

À la fête foraine, on n'achète pas des tours de manège : on achète des jetons. Le grand manège en coûte cinq, la pêche aux canards en coûte un. Avec une IA, c'est exactement pareil. Chaque question envoyée et chaque réponse reçue consomment des jetons, et la caisse compte les jetons, jamais le nombre de questions.

Pourquoi ça compte : tous les tarifs professionnels s'affichent « par million de jetons », et toutes les limites d'usage se comptent en jetons. Trois repères suffisent. En français, un jeton vaut environ trois à quatre caractères, donc une page de 500 mots pèse à peu près 700 jetons. Le même texte coûte plus cher en français qu'en anglais, parce que les modèles découpent notre langue en plus de morceaux. Et vous payez deux fois : ce que vous envoyez, puis ce que l'IA répond. Coller un contrat de 40 pages dans une conversation coûte donc bien plus qu'une question courte, même si l'écran n'affiche rien.

Exemple en entreprise : un cabinet de recrutement de six personnes fait résumer 400 comptes rendus d'entretien par mois. Chaque compte rendu pèse environ 1 200 jetons en entrée, le résumé 200 jetons en sortie : la facture reste anecdotique. La même équipe qui recolle devant chaque document un mode d'emploi de 3 000 jetons multiplie sa consommation par trois, pour un résultat identique. Savoir compter en jetons, c'est savoir où part l'argent avant de recevoir la note.

À retenir : un jeton n'est pas un mot. C'est l'unité de mesure de tout ce qu'une IA lit, écrit et vous facture.

À NE PAS CONFONDRE

Fenêtre de contexte : la quantité de jetons qu'un modèle peut garder sous les yeux en même temps. Le jeton est l'unité, la fenêtre est la taille du plateau. Un plateau immense ne rend pas la lecture gratuite : vous payez toujours ce que vous posez dessus.

Jeton d'authentification : celui que vos outils informatiques utilisent pour vous identifier. Même mot, autre métier : c'est une clé d'accès, elle ne mesure et ne facture rien.