

KLEAA

Lire en ligne

Une voix de patron clonée par IA a coûté 129 millions de dollars

Bon lundi à tous.

Une question traverse cette édition : l'IA vaut-elle vraiment ce qu'on la paie ? Les investisseurs revoient à la baisse le prix d'OpenAI et d'Anthropic, les entreprises consomment plus d'IA sans que leur facture suive, et un nouveau test montre que les meilleures IA qui écrivent du code ne finissent seules qu'1 tâche sur 4. Côté risques, une bande a obtenu 129 millions de dollars de virements en clonant la voix de patrons. Le Grand Angle montre comment une start-up new-yorkaise entraîne les salariés à déjouer ces faux appels.

Quelqu'un devrait lire ça ?

AU SOMMAIRE

01 La justice américaine poursuit 3 hommes qui ont cloné la voix de patrons pour obtenir des virements

02 OpenAI et Anthropic acceptent d'être valorisés moins cher que prévu pour lever de l'argent

03 Les entreprises utilisent de plus en plus l'IA, mais leur facture n'augmente presque plus

04 Les IA qui écrivent du code ne finissent correctement qu'1 tâche sur 4 dans un test proche du réel

05 Plus de 800 outils de piratage pilotés par IA traînent sur internet sans mot de passe



01

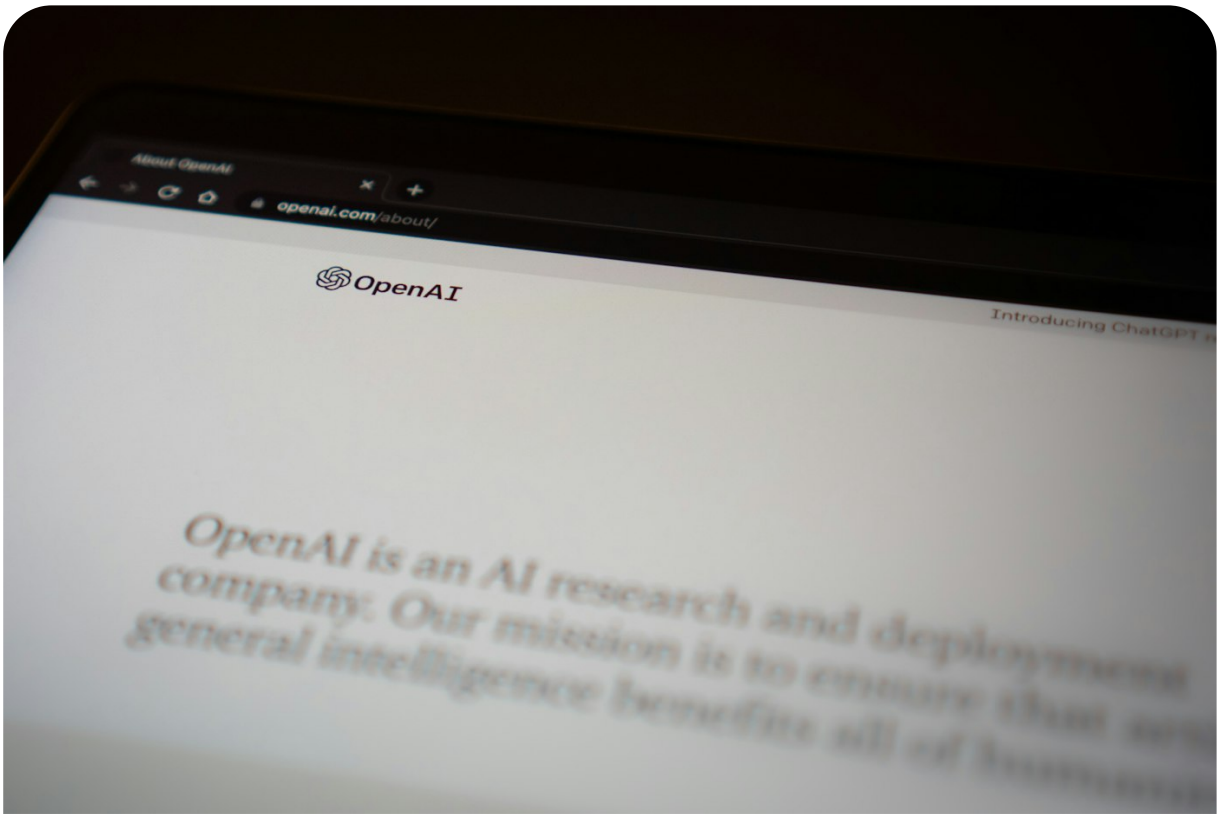
La justice américaine poursuit 3 hommes qui ont cloné la voix de patrons pour obtenir des virements

Selon les procureurs de New York, la bande appelait des comptables en se faisant passer pour leur dirigeant, avec une voix copiée par IA. Hôpitaux, banques et entreprises ont viré au total environ 129 millions de dollars.

Les charges ont été rendues publiques jeudi 5 novembre à New York. Le scénario se répète d'une victime à l'autre. Un message, en apparence envoyé par le PDG, demande à un salarié habilité à faire de gros virements de l'aider sur une opération confidentielle. Viennent ensuite les appels : la voix du patron, et parfois son visage en visio, sont fabriqués par IA. Un faux avocat d'un grand cabinet international prend le relais et donne les ordres de paiement.

La liste des victimes montre qu'aucune taille n'est à l'abri : une fintech (une start-up de services financiers) britannique a perdu 11,6 millions de dollars, un fabricant de nourriture pour animaux de Floride 5,2 millions, 2 réseaux d'hôpitaux du Missouri et de Californie près de 8 millions à eux deux, une banque du Rhode Island 2,7 millions. L'argent transitait par des comptes dans 9 pays, puis était converti en cryptomonnaie pour brouiller les pistes. Les 3 suspects, arrêtés en Espagne et au Royaume-Uni, restent présumés innocents.

Une société d'investissement new-yorkaise a bloqué un virement de 2,85 millions de dollars au dernier moment. C'est la vraie leçon de l'affaire : quelques secondes d'enregistrement suffisent désormais pour copier une voix, donc reconnaître celle de son patron au téléphone ne prouve plus rien. La parade la plus fiable n'est pas technique, c'est une règle interne : tout virement inhabituel se confirme par un second canal, en rappelant soi-même un numéro déjà connu.



02

OpenAI et Anthropic acceptent d’être valorisés moins cher que prévu pour lever de l’argent

Les créateurs de ChatGPT et de Claude cherchent des dizaines de milliards de dollars. Face à des investisseurs devenus prudents, ils revoient leur prix à la baisse.

OpenAI a reçu vendredi 6 novembre environ 11 milliards de dollars, selon Bloomberg. C’est la première partie d’une levée de fonds qui pourrait atteindre 20 milliards, contre 30 milliards visés au départ. L’entreprise est valorisée 1 250 milliards de dollars, au lieu des 1 400 milliards espérés. Sa valorisation (le prix total que les investisseurs attribuent à une entreprise) recule donc de plus de 10 %.

Anthropic, le créateur de Claude, maintient son entrée en Bourse avant Thanksgiving (le 26 novembre aux États-Unis), selon le Financial Times. Elle veut lever entre 80 et 100 milliards de dollars, sur une valeur de 1 250 à 1 750 milliards, alors que les plus optimistes espéraient 2 000 milliards. Ses revenus annuels atteignent pourtant environ 81 milliards de dollars, contre 65 milliards fin juillet.

Le contexte explique ce coup de frein. Les taux d’emprunt montent, les dépenses prévues dans l’IA paraissent démesurées, et la victoire des démocrates aux élections de mi-mandat, qui promettent de freiner les centres de données, a fait chuter de 6,4 % l’indice Nasdaq 100 (les 100 plus grosses valeurs tech américaines) jeudi. Ces 2 entreprises fixent le prix des outils que des millions de personnes utilisent chaque jour : moins d’argent facile pour elles, c’est plus de pression pour rentabiliser chaque abonnement.



03

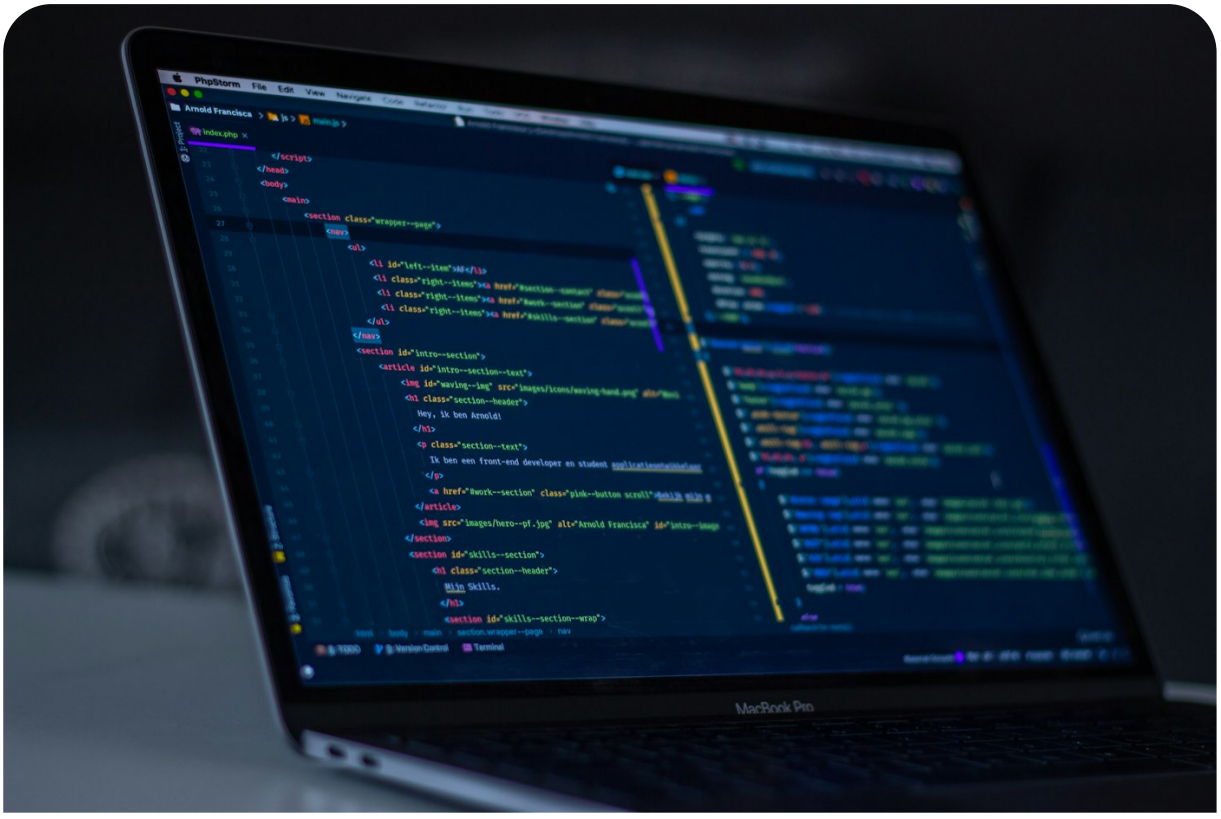
Les entreprises utilisent de plus en plus l’IA, mais leur facture n’augmente presque plus

Ramp, une société américaine qui fournit des cartes de paiement à des milliers d’entreprises, constate que leurs dépenses d’IA ont stagné cet automne. La quantité d’IA consommée, elle, continue de grimper.

Ramp voit passer les factures de ses clients, ce qui lui donne une vue directe sur ce que les entreprises paient vraiment pour l’IA. En juillet, ces dépenses avaient bondi de 42 % en un mois. Sur août et septembre réunis, elles n’ont progressé que de 1,5 %, puis de 8,9 % le mois suivant. Pendant ce temps, le volume de jetons (les morceaux de texte que l’IA lit et écrit, et qui servent à la facturer) n’a jamais cessé d’augmenter.

3 raisons expliquent l’écart. Les prix baissent : OpenAI a divisé par 2 le tarif de certains modèles en septembre, et Anthropic a aussi réduit les siens, ce qui lui a coûté environ 600 millions de dollars de croissance selon ses propres documents. Les modèles moins chers ou gratuits se multiplient. Et les entreprises serrent la vis : budgets par salarié, blocage des modèles les plus chers pour les tâches longues, routeurs (des outils qui envoient chaque demande vers l’IA la moins chère capable de la traiter).

Les écarts restent énormes. L’entreprise médiane dépense 12 à 15 dollars par salarié et par mois, quand les 1 % de plus gros clients approchent 10 000 dollars. Selon Ara Kharazian, économiste chez Ramp, presque toute la dépense vient de jeunes entreprises tech financées par des fonds. Sa formule résume la tendance : « On peut utiliser plus d’IA, mais la payer moins. »



04

Les IA qui écrivent du code ne finissent correctement qu'1 tâche sur 4 dans un test proche du réel

Des chercheurs d'Alibaba ont confié à 10 IA de vraies demandes de modification d'applications web. La meilleure ne livre un travail entièrement correct que dans 26 % des cas.

Le test s'appelle Co-Design Bench. Au lieu d'exercices d'école, il reprend de vraies demandes d'utilisateurs qui voulaient modifier une application web existante : changer la mise en page, ajouter une fonction, lire le montant d'un ticket de caisse et mettre à jour les totaux. Une première IA teste ensuite l'application dans un navigateur, une seconde vérifie qu'aucun résultat n'a été truqué. Seules les fonctions qui marchent vraiment rapportent des points.

Résultat : GPT-6 Astra, le meilleur des 10 modèles testés, remplit en moyenne 70 % de ce qui est demandé, mais ne livre un travail complet que dans 26,4 % des cas. Claude Opus 5 et Qwen (le modèle d'Alibaba) suivent. Tous décrochent quand une demande mélange 3 ou 4 types de changements à la fois (design, contenu, interactions). Sur les anciens tests, ces mêmes modèles frôlent pourtant la note parfaite.

Le même week-end, Andrej Karpathy, cofondateur d'OpenAI aujourd'hui chercheur chez Anthropic, a montré l'autre face. Avec Claude Code, il a monté en un week-end un site qui sélectionne 6 454 vidéos à partir de 43 millions de commentaires publiés en 20 ans sur un forum d'informaticiens, puis un second site en 3 minutes. Les 2 réalités coexistent : l'IA va très vite sur un projet neuf et bien cadré, mais retoucher un outil existant demande encore, 3 fois sur 4, un humain pour finir le travail.

Le regard KLEAA

Ce test mesure une situation précise : une seule consigne, puis l'IA laissée seule jusqu'au bout. C'est justement là qu'elle échoue. Karpathy, lui, sait exactement ce qu'il veut, découpe le travail et juge le résultat à chaque étape. Le gain de temps est bien réel, mais il revient d'abord à ceux qui savent vérifier. La méthode qui marche le mieux aujourd'hui reste simple : de petites demandes, testées une par une, plutôt qu'une longue consigne qu'on espère parfaite.

Si l’IA fait surtout gagner du temps à ceux qui savent déjà juger son travail, qui formera ceux qui ne savent pas encore ?



05

Plus de 800 outils de piratage pilotés par IA traînent sur internet sans mot de passe

Artex, un logiciel gratuit qui mène des cyberattaques presque sans humain, a été retrouvé ouvert à tous sur des centaines de serveurs. Certains contenaient encore des données volées à des entreprises.

La découverte vient de Censys, une société qui scanne internet à la recherche de systèmes mal protégés, et a été rapportée par TechCrunch le 6 novembre. Artex est un outil chinois de test d'intrusion (un logiciel qui cherche les failles d'un réseau comme le ferait un pirate), publié gratuitement en juillet. Microsoft le décrit comme un ensemble d'agents IA (des programmes qui enchaînent seuls des actions) capables de mener une attaque du repérage jusqu'au vol de données, avec très peu d'intervention humaine.

Censys a compté plus de 800 installations d'Artex accessibles depuis internet sans aucun mot de passe. Toutes ne sont pas malveillantes : des experts en sécurité s'en servent aussi pour tester leurs clients. Mais l'une d'elles, liée aux attaques contre des banques sud-coréennes en septembre, contenait des données de sociétés de paiement et de jeux vidéo, et restait ouverte depuis des mois. Artex a aussi servi en octobre contre des milliers d'organisations dont les fichiers sont hébergés sur le cloud de Microsoft.

Le point commun de ces attaques n'a rien de sophistiqué : des mots de passe faciles à deviner et l'absence de double authentification (un code envoyé sur le téléphone en plus du mot de passe). L'IA permet aux pirates d'aller plus vite et de viser plus large, mais elle entre surtout par les portes laissées ouvertes. Artex va devenir payant, mais des dizaines de copies de son code circulent déjà.



LE GRAND ANGLE

Adaptive Security entraîne les salariés à déjouer les faux appels de leur patron

États-Unis (New York) · Cybersécurité · Formation des équipes

Adaptive Security est une start-up new-yorkaise qui piège les salariés de ses clients, pour de faux : elle leur envoie des e-mails, des SMS et même des appels avec la voix clonée de leur dirigeant, puis forme aussitôt ceux qui se font avoir.

L'affaire des 129 millions de dollars le montre : l'arnaque ne passe plus par un e-mail bourré de fautes, mais par un appel qui a la voix du patron. Adapti Security, fondée en 2023 à New York, est née de ce constat. Son fondateur, Brian Long, a dirigé pendant 7 ans et demi Attentive, une grande plateform de marketing par SMS. La société a levé 136 millions de dollars, dont une partie auprès du fonds d'OpenAI, qui signait là son premier investissement dans la cybersécurité.

Le logiciel commence par faire ce que ferait un escroc : il rassemble ce qui est public sur chaque salarié (profil LinkedIn, communiqués, vidéos), soit plu de 1 000 informations par personne. Il fabrique ensuite de fausses attaques sur mesure : un SMS urgent du directeur financier, l'e-mail d'un faux fournisseur, ou un appel où la voix du dirigeant, copiée à partir de quelques secondes d'enregistrement, réclame un virement. Selon Brian Long, 2 minutes de vidéo suffisent pour créer un avatar (une copie animée d'une personne) capable de tenir une visio.

Ceux qui tombent dans le piège reçoivent tout de suite une courte formation adaptée à leur poste et à l'erreur commise, avec vidéo et quiz générés par IA. La direction voit quels salariés sont les plus visés et quelles failles reviennent. La société revendique plus de 2 000 clients, dont PayPal, Figma, Ram et Bose. Face aux acteurs historiques de la formation à la sécurité, elle mise sur une idée simple : les vieux exercices d'e-mails piégés ne préparent pas à un appel cloné.

Le principe se copie sans acheter de logiciel. Une entreprise peut tester elle-même ses équipes avec un faux message urgent du patron, ou un appel q demande de changer un RIB. Surtout, elle peut fixer une règle que l'IA ne sait pas imiter : aucun virement inhabituel sans rappel sur un numéro déjà connu, ou sans mot de code convenu à l'avance.