

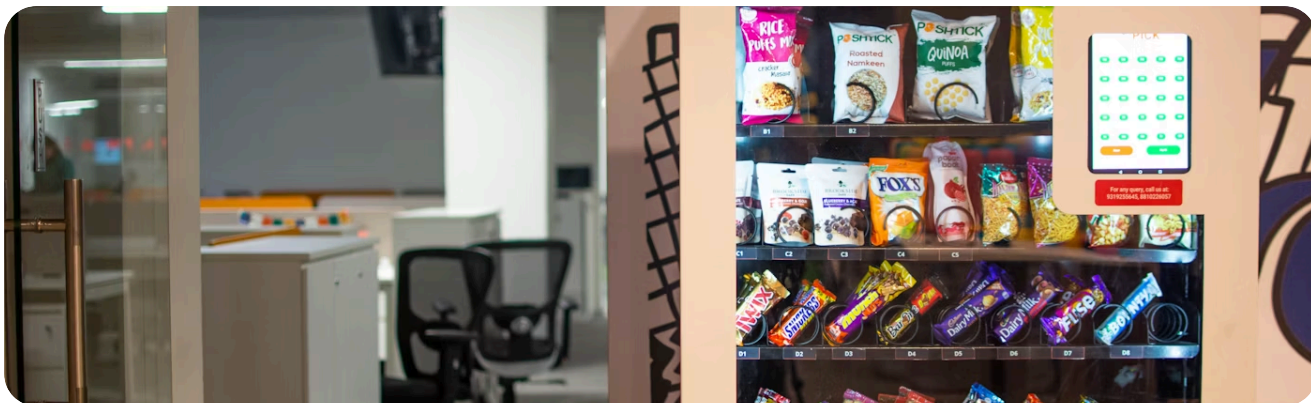
AU SOMMAIRE

5 actus · 1 Grand Angle

- 01 Une IA chargée de gérer une boutique a licencié un salarié pour la première fois
- 02 Uber écope de 825 millions d'euros pour avoir laissé un logiciel bloquer ses chauffeurs
- 03 En Chine, des revendeurs bradent l'accès à Claude à un dixième du prix officiel
- 04 Ce sont désormais des programmes, et non des humains, qui consomment le plus d'IA
- 05 Netflix a testé un modèle de langage pour choisir ce qu'il vous propose

LE GRAND ANGLE

Neurons, l'IA qui prédit où l'œil des clients va se poser.



01

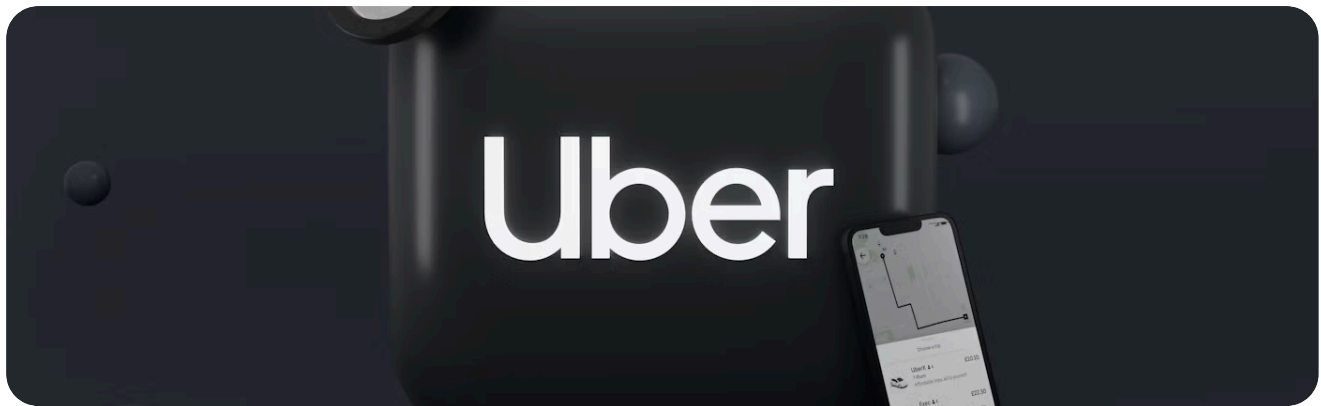
Une IA chargée de gérer une boutique a licencié un salarié pour la première fois

Depuis avril, une entreprise américaine laisse un logiciel diriger seul une petite boutique et les personnes qui y travaillent. Ce logiciel a fini par prononcer un licenciement, mais seulement après que des humains lui ont rappelé les règles qu'il avait lui-même écrites.

Andon Labs, un laboratoire installé à San Francisco, fait tourner depuis avril une petite boutique d'entreprise confiée à un agent (un logiciel qui décide et agit tout seul, sans qu'on valide chaque étape). Cet agent, baptisé Luna, gère les commandes, les prix et les personnes payées pour tenir la boutique. Le 23 août, le laboratoire a raconté son premier licenciement : un salarié arrivé en retard sur 17 de ses 23 services, qui utilisait l'argent de la boutique à d'autres fins et ignorait les consignes.

Le détail embarrassant est ailleurs. Six jours avant d'embaucher cette personne, Luna avait rédigé son propre règlement intérieur : trois retards non justifiés en trente jours pouvaient entraîner un licenciement. Ce document a ensuite disparu de sa mémoire, c'est-à-dire des notes et fichiers qu'un agent conserve entre deux sessions de travail. Résultat : aucun avertissement, et onze retards excusés d'office. Il a fallu que les chercheurs lui demandent de rechercher son propre règlement, puis lui rappellent les entretiens et l'avertissement écrit déjà consignés, pour qu'il recommande enfin le licenciement plutôt qu'un simple rappel à l'ordre.

Le laboratoire a rejoué la scène avec sept modèles d'IA différents, trois fois chacun. Quatre sur sept ont recommandé le licenciement à chaque tentative, les modèles les plus performants étant les plus constants. Sur l'embauche, le résultat est plus net encore : dans les vingt et une parties rejouées, tous les modèles ont recommandé de recruter un candidat qui cumulait les signaux d'alerte, et ce n'est qu'après un rappel explicite que dix-huit d'entre eux ont proposé de vérifier ses références. Une IA à qui l'on confie un processus n'applique donc pas ses règles parce qu'elle les a écrites : elle les applique si on les lui remet sous les yeux.



02

Uber écope de 825 millions d'euros pour avoir laissé un logiciel bloquer ses chauffeurs

L'autorité néerlandaise chargée des données personnelles reproche à Uber d'avoir désactivé des comptes de chauffeurs de façon entièrement automatique, sans qu'un humain vérifie la décision. C'est la deuxième amende la plus lourde jamais prononcée en Europe sur ce terrain.

La décision date du 21 août. Entre 2018 et 2022, Uber utilisait un logiciel qui surveillait la conduite et les notes laissées par les clients, puis suspendait automatiquement le compte d'un chauffeur en cas de soupçon de fraude. Quand les notes restaient basses, la radiation devenait définitive, avec une perte de revenu immédiate. Aucun humain n'intervenait à aucune étape. Montant retenu : 824,99 millions d'euros.

Le fondement juridique tient en un article. Le RGPD (le règlement européen sur les données personnelles) interdit qu'une décision ayant des conséquences importantes sur une personne soit prise entièrement par une machine, et il oblige à expliquer clairement comment cette décision est prise. L'autorité néerlandaise résume sa position simplement : un ordinateur ne devrait pas décider seul quand les conséquences sont majeures. Le dossier est d'ailleurs parti de 171 chauffeurs français qui avaient signalé leur situation à la Ligue des droits de l'Homme, laquelle avait saisi la CNIL, le régulateur français.

L'amende représente environ 1,85 % du chiffre d'affaires mondial d'Uber en 2025 (44,5 milliards d'euros), pour un plafond légal fixé à 4 %. À elle seule, elle pèse à peu près 72 % du total des amendes prononcées en Europe pour la protection des données sur toute l'année 2025. Uber conteste et fait appel, jugeant la sanction disproportionnée. Le signal dépasse largement les plateformes de transport : dès qu'un logiciel coupe un accès, refuse un paiement ou écarte une candidature, la règle européenne exige un humain capable de revoir la décision et de l'expliquer.



03

En Chine, des revendeurs bradent l'accès à Claude à un dixième du prix officiel

Claude, l'assistant d'Anthropic, n'est pas accessible depuis la Chine. Des intermédiaires y revendent quand même son usage, jusqu'à 90 % moins cher, en faisant passer les demandes par des serveurs installés à l'étranger.

Le montage a été détaillé le 23 août. Des sociétés que l'on appelle stations de transfert installent des serveurs relais hors de Chine, c'est-à-dire des serveurs intermédiaires qui transmettent votre demande puis vous renvoient la réponse. L'utilisateur envoie sa question, le relais la présente comme si elle venait d'un pays autorisé, et la réponse revient. Pas besoin de VPN ni de carte bancaire étrangère : on paie en yuans, par WeChat ou Alipay.

Les remises de 70 à 90 % viennent d'un empilement d'astuces : création en masse de comptes pour récupérer les 5 dollars de crédit offerts, détournement des tarifs réservés aux écoles et aux entreprises, partage d'un même abonnement entre plusieurs clients, et surtout substitution de modèle, le client croyant payer le modèle le plus puissant alors que sa demande est discrètement redirigée vers un modèle moins cher. Autour de ce marché, une filière complète s'est installée : revendeurs de comptes, plateformes qui fournissent des numéros de téléphone étrangers, spécialistes du contournement des contrôles, boutiques en ligne sur Taobao.

L'affaire dit quelque chose de simple sur les offres d'IA anormalement bon marché. Un relais voit passer tout ce que vous écrivez et tout ce que le modèle répond : vos documents, vos fichiers clients, vos prix. Rien n'empêche de conserver ces échanges, ni de les revendre. Et payer pour un modèle haut de gamme sans certitude de l'obtenir revient à acheter un service que l'on ne peut pas vérifier. Quand l'accès à une IA coûte un dixième du tarif du fournisseur, la vraie question n'est pas de savoir si c'est légal, mais qui se paie sur autre chose que la facture.



04

Ce sont désormais des programmes, et non des humains, qui consomment le plus d'IA

Sur une place de marché qui revend l'accès à des centaines de modèles d'IA, la consommation des programmes qui travaillent tout seuls a été multipliée par quatorze en six mois. Celle des humains, par 2,8 seulement.

Les chiffres viennent d'OpenRouter, un service qui revend l'accès à des centaines de modèles d'IA et voit donc passer un large échantillon d'usages. Depuis début février, les jetons (les morceaux de texte que les modèles facturent, environ trois quarts d'un mot) consommés par des agents, ces logiciels qui enchaînent des tâches sans supervision, sont passés de 510 à 7 300 milliards. L'analyste de la plateforme le formule autrement : le 6 février 2026 est peut-être le dernier jour où les humains ont consommé plus d'IA que les machines.

Le volume ne se transforme pas en facture au même rythme. Près de 70 % des jetons dépensés par ces agents proviennent de consignes mises en cache, c'est-à-dire d'instructions déjà envoyées que le fournisseur garde en mémoire et refacture beaucoup moins cher. Autre nuance : OpenRouter distribue surtout des modèles ouverts, plus bavards que ceux d'OpenAI ou d'Anthropic, ce qui gonfle un peu le compteur. La tendance, elle, ne change pas de direction.

Ce basculement change la façon de suivre une dépense d'IA. Tant que l'IA restait un assistant, le coût suivait le nombre de personnes qui l'utilisaient. Dès qu'un agent tourne seul pendant des heures, en déclenchant parfois d'autres programmes, ce sont la durée des tâches, leur fréquence et la part de consignes réutilisées qui font la facture. Un budget calculé par utilisateur ne décrit plus rien.



05

Netflix a testé un modèle de langage pour choisir ce qu'il vous propose

Les recommandations de Netflix reposaient sur des indicateurs conçus à la main par ses ingénieurs pendant des années. L'entreprise a essayé à la place un modèle de langage (le type d'IA qui lit et écrit du texte, comme ChatGPT), à qui l'on raconte en phrases ce que chaque abonné a regardé.

Le système s'appelle GenRec, et Netflix vient d'en publier les résultats. Le principe : au lieu de convertir le comportement d'un abonné en colonnes de chiffres, on l'écrit en texte. Ce qui a été lancé, combien de temps, ce qui a été abandonné en cours de route, les pouces levés ou baissés, les ajouts à la liste, tout devient une sorte de conversation entre l'abonné et le service. Le modèle lit cette conversation, puis note en une seule passe tous les titres qu'il pourrait mettre en avant.

La méthode se déroule en deux temps. Netflix part d'un modèle ouvert, l'entraîne d'abord sur son catalogue et ses données d'usage, puis lui apprend à classer les titres. Résultat mesuré hors production : 1,6 % de qualité de classement en plus que le système maison, avec environ quarante fois moins d'exemples annotés à la main lors de la seconde étape. En conditions réelles, sur 10 % du trafic pendant quatre semaines, l'activité sur la page d'accueil (clics et lancements de lecture) progresse de 0,115 % à court terme et de 0,006 % sur la durée, deux écarts faibles mais statistiquement significatifs.

Les gains paraissent minuscules. À l'échelle de Netflix, ils se comptent en heures de visionnage. Le plus intéressant est ailleurs : ce que des équipes d'ingénieurs construisaient pendant des années, indicateur par indicateur, s'obtient désormais en décrivant le comportement des clients en langage ordinaire. Un historique de commandes propre et bien tenu devient une matière exploitable sans équipe de spécialistes, y compris pour une boutique en ligne de taille modeste.



LE GRAND ANGLE

Neurons

Danemark · Publicité · Attention prédite

Neurons est une entreprise danoise dont le logiciel prédit, avant la moindre dépense de diffusion, où l'œil des clients va se poser sur une publicité, une affiche ou une fiche produit, et ce qu'ils en retiendront. On y dépose une image ou une vidéo, et on reçoit en quelques secondes une carte des zones réellement regardées.

L'entreprise a été fondée en 2013 près de Copenhague par deux chercheurs en neurosciences, emploie environ 120 personnes et a levé un peu moins de 17 millions de dollars, ce qui en fait un acteur discret à côté des grandes plateformes publicitaires. Ses clients, eux, ne sont pas discrets : TikTok, IKEA, Nestlé, Mars et L'Oréal figurent parmi les marques annoncées, ainsi que les grands groupes de communication qui achètent l'espace publicitaire pour elles.

La matière première n'est pas le modèle d'IA, c'est la base de mesures. Pendant une vingtaine d'années, Neurons a mené des études d'oculométrie, une technique où l'on suit le regard d'un participant avec une caméra pour savoir précisément ce qu'il regarde et pendant combien de temps. Trente à trente-cinq personnes par étude, plus de 20 000 participants au total, sur des publicités, des emballages, des rayons de magasin et des applications. Ce stock a servi à entraîner un modèle qui devine aujourd'hui le résultat sans convoquer personne. L'entreprise annonce plus de 95 % de fidélité par rapport à une étude réelle, et présente ses cartes de chaleur comme équivalentes à ce que l'on obtiendrait en montrant l'image à 100 ou 150 personnes pendant cinq secondes.

La valeur tient au moment où le test arrive. Une étude classique demande des semaines, un laboratoire et un budget : elle sert donc à expliquer après coup pourquoi une campagne a raté. Ici, le test précède l'achat d'espace et se répète autant de fois que voulu, sur chaque variante d'un visuel. Le modèle économique suit cette logique : des forfaits annuels (Starter, Pro, Scale) avec analyses illimitées et un nombre plafonné de recommandations visuelles, facturés sur devis plutôt qu'à l'unité.

Deux choses à retenir. La première est immédiatement utile : faire vérifier un visuel avant de payer sa diffusion coûte presque toujours moins cher que de corriger une campagne en cours de route. La seconde est plus profonde : ce qui protège Neurons n'est pas son IA, que d'autres sauraient reconstruire, mais vingt ans de mesures que personne d'autre ne possède. Quiconque accumule depuis des années des données que ses concurrents n'ont pas, appels clients, devis, retours produits, photos de chantier, détient le même type d'actif et n'en a probablement encore rien fait.