xaviers.ai
L'information IA pour votre Comex

HEBDOMADAIRE
Édition n° 1
mardi 3 novembre 2026
Faits du 24 octobre au 3 novembre 2026
Prochaine édition : dimanche 8 novembre, 19 h
Xavier Schällebaum

Le marché rationne l'IA : correction boursière, dette chez Oracle et Microsoft, 895 milliards d'engagements chez Anthropic, et une première ligne européenne sur le droit d'auteur

Cette semaine, l'argent de l'intelligence artificielle se fait plus cher et le droit se précise. Le Nasdaq 100 est entré en correction, Intel a perdu 29 % en une séance, Oracle emprunte à 8,4 % et Microsoft place 25 milliards d'obligations ; Anthropic affiche une croissance hors norme mais 895 milliards de dollars d'engagements de calcul ; Apple paie la pénurie de mémoire que créent les centres de données. À Luxembourg, l'avocat général de la Cour de justice distingue l'entraînement, licite par défaut, de la restitution d'articles, à risque. Et les filigranes invisibles deviennent la norme chez xAI comme chez Anthropic. Voici ce qu'un comité de direction doit en retenir, et ce qu'il peut décider.

L'ESSENTIEL EN TROIS POINTS

1. **Le marché ne finance plus l'IA sans preuve de retour.** Le Nasdaq 100 a reculé de 10,4 % depuis son record du 6 octobre, Intel a chuté de 29 % le 30 octobre, Oracle emprunte 20 milliards de dollars dont des obligations à environ 8,4 %, et OpenAI aurait réduit sa levée de 30 à 20 milliards (rapporté). Le gouverneur de la Banque de France, Emmanuel Moulin, évoque « un effet de contagion en Europe » (Maddyness, 2 novembre, rapporté).
2. **Le calcul coûte, et il se voit dans les comptes.** Anthropic affiche près de 17 milliards de dollars de revenus au troisième trimestre (16,96 milliards, +47 % en un trimestre), mais 9,9 milliards de coûts de calcul et 895,4 milliards d'engagements d'achat (prospectus amendé du 2 novembre) ; John Ternus, le nouveau directeur général d'Apple, juge les prix de la mémoire « unprecedented » (2 novembre).
3. **Le droit et la traçabilité se précisent.** L'avocat général de la Cour de justice de l'Union européenne valide l'entraînement au titre de la fouille de textes mais juge que les réponses de Gemini peuvent violer le droit voisin des éditeurs (3 novembre) ; l'API Grok filigrane ses textes par défaut dès le 7 novembre et Anthropic étend ses filigranes à quatre modèles anciens ; OpenAI, Google, Microsoft et Anthropic adoptent un même signal robots.txt de refus d'entraînement.

En chiffres

895,4 Md$
engagements d'achat de calcul et d'infrastructure d'Anthropic au 30 septembre 2026, dont environ 447 Md$ exigibles avant fin 2030 (prospectus amendé du 2 novembre)

−10,4 %
recul du Nasdaq 100 au 2 novembre par rapport à son record du 6 octobre ; plus de 2 000 Md$ de capitalisation technologique effacés en octobre

16,96 Md$
revenus d'Anthropic au troisième trimestre 2026, +47 % sur le trimestre précédent, pour une perte d'exploitation de 567 M$ (prospectus amendé du 2 novembre)

+259 %
hausse sur un an des exportations sud-coréennes de semi-conducteurs en octobre (56,4 Md$), portées par la mémoire pour l'IA (Reuters, 1er novembre)

12 %
taux de réussite du meilleur modèle de pointe sur des tâches expertes de plusieurs heures, selon le nouveau banc d'essai GDPx de SemiAnalysis (2 novembre)

I **L'essentiel** People · Planet · Profit · Corporate · Dates clés
II **Publications signalées** Les études, les textes et les points de vue qui méritent le temps d'un dirigeant
III **Nouvelles IA, nouveaux usages** Ce qui change dans les assistants et les outils
IV **Orientations stratégiques** Anticiper · Surveiller · Protéger · Développer
Annexe **Les sociétés à suivre** Valorisation, cours, chiffre d'affaires, résultat, signal de la semaine

Tous droits réservés · XAVIER S. AI

Lire cette édition en ligne et la transférer :
xaviers.ai/editions/2026-11-03

L'essentiel

Les faits de la période, classés en quatre étiquettes. People : dirigeants, emploi, déclarations. Planet : régulation, géopolitique, énergie, souveraineté. Profit : résultats, financements, prix. Corporate : produits, incidents, contentieux. Tout est sourcé ; « rapporté » signale une information de presse sans communiqué.

PEOPLE Dirigeants, emploi, déclarations

Goldman Sachs retire les agents de code aux banquiers et aux traders. Selon Bloomberg, des consignes internes mises à jour en octobre réservent Claude Code et GitHub Copilot aux quelque 12 000 ingénieurs de la banque, après un examen de productivité et de risques ; banquiers et traders les utilisaient depuis plusieurs mois (rapporté). Premier grand groupe financier à restreindre les agents aux profils techniques : un précédent pour la gouvernance des outils agentiques dans les fonctions métier (aiweekly.co).

Les licenciements « IA » laissent une cicatrice durable, selon Goldman Sachs. Sur plus de 15 000 personnes licenciées entre 2016 et 2025, celles qui occupaient des postes exposés à l'IA mettent environ un mois de plus à retrouver un emploi et sont, un an après, dix points plus souvent encore au chômage ; les 35 ans et plus mettent cinq semaines de plus. Challenger, Gray & Christmas attribue à l'IA 120 136 suppressions de postes aux États-Unis jusqu'en septembre, première cause invoquée. Un argument pour la requalification interne plutôt que la séparation (fortune.com).

Ce que disent les patrons : Ternus, Altman, Musk. John Ternus, directeur général d'Apple depuis le 1er septembre : « We don't want to ship AI for AI's sake », l'utilisateur ne doit pas avoir à savoir que l'IA est à l'œuvre (Wired, 31 octobre). Sam Altman juge « quite terrible », du point de vue de l'allocation mondiale, que chacun ait de la mémoire vive dans son ordinateur, utilisée 5 % du temps : il la veut dans le nuage (30 octobre). Elon Musk confirme la réduction des quotas de Grok : « The cost of serving Grok Bot was unsustainably high relative to revenue » (2 novembre) (techcrunch.com, pcgamer.com, finance.yahoo.com).

OpenAI recrute un ancien secrétaire à l'Armée ; Cognizant affecte 25 000 personnes à ses outils. Dan Driscoll, secrétaire à l'Armée de terre jusqu'en septembre, devient conseiller principal d'OpenAI pour la sécurité nationale, chargé d'aider les alliés des États-Unis à adopter l'IA (3 novembre). Cognizant crée un groupe de plus de 25 000 collaborateurs pour déployer Codex et ChatGPT chez ses clients, onze jours après en avoir affecté 18 000 à Gemini (30 octobre). Rappel : Microsoft propose un départ volontaire à 60 000 salariés, environ 27 % de ses effectifs (thenextweb.com, thenextweb.com).

PLANET Régulation, géopolitique, énergie, souveraineté

Élections de mi-mandat : l'IA et les centres de données entrent dans la campagne. D'après les données de la commission électorale fédérale compilées le 3 novembre, les comités d'action politique liés à l'IA ont dépensé 79 millions de dollars dans 68 circonscriptions, pour des candidats des deux partis. Près de 200 milliards de dollars de projets de centres de données ont été retardés, suspendus ou bloqués au premier semestre par des oppositions locales (Data Center Watch), et 41 mesures locales les concernant figuraient sur les bulletins de dix États ; 77 % des Américains craignent une hausse de leur facture d'électricité (Reuters/Ipsos). Selon Axios, OpenAI prépare pour la session parlementaire du 9 novembre une campagne en faveur de « limites de vitesse » obligatoires pour les systèmes les plus puissants (rapporté). Aucun résultat n'était connu à l'heure de la rédaction (citationneeded.news, techcrunch.com, investorplace.com).

Nvidia imposerait à ses clients cloud un plafond de 49 % d'activité avec des clients chinois. Selon le Financial Times, Nvidia a demandé verbalement à ses partenaires de centres de données, pour les nouveaux contrats, de réaliser au moins 51 % de leurs services auprès d'entreprises non chinoises ; Pékin dit ne pas être au courant (rapporté). En toile de fond, le New York Times affirme, d'après le renseignement américain, qu'une dirigeante d'un client singapourien de Nvidia aurait renseigné Pékin ; Nvidia et la Chine démentent, Donald Trump dit examiner le dossier (rapporté). Les fournisseurs de calcul en Asie, en Europe et au Moyen-Orient devront documenter la nationalité de leurs clients (thenextweb.com, euronews.com).

La mémoire pour l'IA devient un fait macroéconomique en Corée. Exportations sud-coréennes d'octobre : 115,3 milliards de dollars, +93,7 % sur un an, plus forte hausse depuis 1974 ; semi-conducteurs +259,4 % ; vers les États-Unis

Tous droits réservés · XAVIER S. AI

+185,8 % (Reuters). Le président Lee Jae Myung consacre 21 300 milliards de wons du budget 2027 aux puces, aux centres de données et à l'IA physique, plus 4 600 milliards à un modèle souverain. Sam Altman a rencontré une cinquantaine de dirigeants de Samsung le 31 octobre. Pour tout acheteur d'équipements, c'est l'explication de la hausse des prix de la mémoire (finance.yahoo.com, koreaherald.com).

Énergie : les centres de données apprennent à s'effacer, et l'État russe se sert. OpenAI s'engage à réduire la consommation de ses campus quand le réseau est tendu : 1 000 MW délestables sur son site de 3 200 MW en Géorgie, soit 31 % de la pointe (programme Grid Protectors, 2 novembre). Au Royaume-Uni, Drax, Brookfield et Radiant annoncent 20 milliards de livres pour jusqu'à 1 GW sur le site d'une ancienne centrale, première phase de 100 MW au premier trimestre 2028 (3 novembre). Vladimir Poutine oblige les centres de données privés à réserver 10 % de leur capacité à l'État, mobilisable sous 24 heures (décret du 2 novembre). Près de Memphis, une juge fédérale laisse la NAACP poursuivre xAI pour 35 turbines exploitées sans permis (31 octobre) (thenextweb.com, constructionbriefing.com, themoscowtimes.com, techcrunch.com).

Abou Dhabi confie à OpenAI l'instruction réglementaire, hébergée aux Émirats. Quatre accords signés le 2 novembre avec Sam Altman : traduction des réglementations en format lisible par machine pour l'évaluation de conformité, premier pilote en production avec l'autorité agroalimentaire, modèles exécutés sur une infrastructure située aux Émirats, décisions finales humaines ; déploiement de la plateforme Frontier dans l'administration, les entreprises et l'éducation. Un modèle d'administration « native de l'IA » que d'autres États observeront (thenationalnews.com).

Rappels de la semaine précédente : la Suède et Bruxelles. Les ministères suédois bloquent les versions navigateur de ChatGPT et de Claude, Copilot restant recommandé sous accord de sécurité (premier précédent européen de liste blanche) ; la Commission européenne a prolongé au 10 novembre sa consultation sur le droit d'auteur et l'IA (euronews.com, ec.europa.eu).

PROFIT Résultats, financements, prix

Le marché corrige les valeurs de l'IA et commence à trier les bilans. Le 30 octobre, Intel a perdu environ 29 % après une prévision de chiffre d'affaires de 15 à 16 milliards de dollars pour le quatrième trimestre et l'annonce d'un recul des livraisons de PC d'environ 12 % en 2026, faute de mémoire, absorbée par les centres de données. Le 2 novembre, le Nasdaq 100 est entré en correction, à 10,4 % sous son record du 6 octobre ; plus de 2 000 milliards de dollars de capitalisation technologique ont été effacés en octobre, et les sept géants, qui pèsent ensemble 22 830 milliards, ont perdu de l'ordre de 1 100 milliards en cinq séances (calcul de la rédaction d'après les clôtures). Oracle a lancé le même jour 10 milliards d'obligations, dont des titres 2066 à environ 8,4 %, et enregistré 10 milliards de ventes d'actions ; ses contrats de couverture de défaut s'échangent autour de 300 points de base, contre moins de 60 pour Microsoft, qui a placé 25 milliards d'obligations avec plus de 100 milliards d'ordres. Le gouverneur de la Banque de France, Emmanuel Moulin, juge possible qu'une correction américaine se propage en Europe, les investissements étant « de plus en plus financés par de la dette » (Maddyness, rapporté) (itv.com, finance.yahoo.com, thestar.com.my, maddyness.com).

Anthropic amende son prospectus : près de 17 Md$ de revenus au trimestre, 895 Md$ d'engagements, une valorisation revue à la baisse. Revenus du troisième trimestre de 16,96 milliards de dollars (+47 % sur le trimestre précédent, 33,2 milliards sur neuf mois), perte d'exploitation de 567 millions attribuée à 5 milliards de calcul acheté à court terme, perte nette de 23,3 milliards essentiellement comptable (104 milliards sur neuf mois, dont 94,3 de réévaluation non monétaire d'instruments convertibles) ; coûts de calcul et d'infrastructure de 9,9 milliards, environ 58 % des revenus ; engagements d'achat de calcul d'environ 895,4 milliards, dont 447 avant fin 2030 et 69,8 à douze mois. Le prospectus ajoute que les estimations internes pourraient « imminemment » atteindre la définition d'un progrès « dramatiquement accéléré » attribuable à l'IA, sans qu'aucun modèle actuel ne soit désigné comme tel. Selon le Wall Street Journal, les investisseurs institutionnels visent 1 500 à 1 600 milliards de valorisation contre 2 000 évoqués, avec une tournée possible dès la semaine du 9 novembre (rapporté). OpenAI, de son côté, aurait ramené sa levée de 30 à 20 milliards, à 1 250 milliards de valorisation (Financial Times, rapporté) (stocktitan.net, uk.finance.yahoo.com, techcrunch.com, techcrunch.com).

Apple : un trimestre record, mais la mémoire et Siri AI pèseront sur la marge. Chiffre d'affaires du trimestre clos le 26 septembre de 114,5 milliards de dollars (+12 %), bénéfice net de 30,7 milliards ; iPhone +17 %, Mac +24 %, services 31,6 milliards (record) ; exercice 2026 à 478,8 milliards. Pour le trimestre de décembre, Apple attend une marge brute de 45 à 46 % contre 47,9 %, pénalisée par le coût de la mémoire, jugé « unprecedented » par John Ternus, et par le lancement de Siri AI, dont les fonctions exécutées sur les serveurs d'Apple seront soumises à des limites d'usage, avec un accès supplémentaire payant. Même Apple paie la pénurie et commence à facturer l'IA à l'usage (macrumors.com, apple.com).

Tous droits réservés · XAVIER S. AI

Deux signaux de demande : Palantir et Marvell. Palantir : chiffre d'affaires du troisième trimestre de 2,31 milliards de dollars (+96 %), le commercial américain (984 millions, +148 %) dépassant pour la première fois le gouvernement ; bénéfice net de 1,13 milliard ; prévision 2026 relevée à 8,53 milliards. Marvell rachète Lumentum pour environ 115 milliards de dollars en actions, prime de 23,9 %, pari sur la connectivité optique des centres de données ; le marché a sanctionné l'acheteur (−12,7 % le 2 novembre). La demande des entreprises reste forte, mais les paris jugés trop chers sont punis (fortune.com, finance.yahoo.com).

Les abonnements se resserrent : quotas unifiés chez OpenAI, Grok rationné, Google double ses tarifs le 1er décembre. À partir de décembre, les offres Plus, Pro et Business de ChatGPT partageront une même allocation de jetons entre ChatGPT et Codex, dont OpenAI ne publie pas la taille ; l'usage des agents de code pèsera donc sur le quota des collaborateurs. xAI réduit les quotas de Grok Bot pour des raisons de coût. Chez Anthropic, les offres Max (100 $) et Ultra (200 $, sans limite sur cinq heures pour les comptes vérifiés) incluent des crédits d'API. Le tarif de lancement de Gemini 4 Argon double le 1er décembre : à intégrer dans les budgets 2027 (cellcog.ai, support.claude.com).

CORPORATE Produits, incidents, contentieux

Droit d'auteur : l'avocat général de la Cour de justice distingue l'entraînement, licite, de la restitution, à risque. Dans l'affaire Like Company c. Google (C-250/25), première affaire européenne sur l'IA générative et le droit d'auteur, l'avocat général Maciej Szpunar estime le 3 novembre que l'entraînement sur des articles relève en principe de l'exception de fouille de textes et de données, sous condition d'accès licite et de respect d'une réservation de droits lisible par machine ; mais que la récupération d'un article pour composer une réponse n'est pas couverte, et que la reproduction d'une expression reconnaissable peut constituer une contrefaçon, le droit voisin des éditeurs pouvant être violé « dans certaines circonstances ». Les conclusions ne lient pas la Cour ; arrêt attendu dans les prochains mois. Décryptage en rubrique II (thenextweb.com, dataconomy.com).

Les filigranes invisibles deviennent la norme, et un signal commun de refus d'entraînement apparaît. À partir du 7 novembre, l'API de Grok appliquera par défaut un filigrane statistique invisible à ses textes, au nom de l'article 50 de l'AI Act, désactivable par les administrateurs sous leur responsabilité. Selon un e-mail adressé à ses clients et relayé par la presse spécialisée, Anthropic étend depuis le 3 novembre ses filigranes textuels à Claude Mythos 5, Opus 4.7, Opus 4.6 et Sonnet 4.6, l'outil de détection restant réservé à des tiers habilités. Le 30 octobre, OpenAI, Google, Microsoft et Anthropic se sont engagés à respecter un même jeton robots.txt de refus d'entraînement, issu d'un projet de norme IETF rédigé chez Anthropic : Anthropic l'applique déjà, OpenAI « dans les prochains mois », Microsoft au 1er février 2027, Google en 2027. Il ne concerne ni l'exploration ni les réponses générées (docs.x.ai, gigazine.net, blogs.bing.com, datatracker.ietf.org).

Cybersécurité : les attaquants utilisent les mêmes agents de code que les entreprises. Sophos attribue au groupe du rançongiciel Aur0ra un usage « extensif » d'agents de code (Cursor, modèles d'Anthropic, puis xAI et Moonshot) pour la reconnaissance, le vol de mots de passe et l'élévation jusqu'à administrateur de domaine, les garde-fous étant contournés en présentant l'intrusion comme un test autorisé ; Check Point suit une campagne de faux installateurs « GPT-6 Astra » et « Claude Opus » qui chargent une bibliothèque malveillante à côté d'un programme signé. OpenAI a corrigé une faille permettant à un lien partagé de détourner une session ChatGPT (prime de 5 000 $). HTTP2Shell et la faille du noyau Linux CVE-2026-98432 sont exploitées activement ; la semaine précédente avait déjà cumulé WordPress 7.1.4, PHP2Shell et une fausse extension Claude Code chez jusqu'à 5 000 développeurs (research.checkpoint.com, sophos.com, simonwillison.net).

Contentieux : le New York Times attaque Brave, OpenAI rassemble vingt soutiens au fair use. Le 2 novembre, le New York Times puis trois titres du groupe Tribune ont assigné Brave devant la cour fédérale de Manhattan pour un index de plus de 30 milliards de pages, constitué par un robot non identifiable et revendu à des entreprises d'IA nommées, dont Anthropic, Perplexity, Mistral et xAI. Dans le procès du Times contre OpenAI, vingt mémoires favorables au fair use ont été déposés à la date limite du 30 octobre, dont celui du ministère de la Justice ; les répliques sont attendues le 20 novembre. Le même jour, OpenAI a restreint pour les mineurs les informations pratiques sur les armes et annoncé un rapport de sécurité semestriel (Reuters, rapporté). Au Royaume-Uni, Google conteste des écrans de choix étendus aux assistants ; aux États-Unis, le ministère de la Justice s'informe sur le rachat de Hugging Face par Nvidia (rapporté) (techcrunch.com, ipwatchdog.com).

Mesurer ce que les modèles savent vraiment faire : 12 % sur des tâches expertes. SemiAnalysis lance GDPx, un banc d'essai de tâches de plusieurs heures dans onze domaines à forte valeur (conception de puces, cybersécurité, biologie, finance, droit) : Claude Opus 5.5 est premier avec 12 % de réussite, Gemini 4 Carbon seul dans le même ordre de grandeur ; finance et droit donnent près de zéro à tous les modèles. Google, de son côté, a fait tourner Gem-

Tous droits réservés · XAVIER S. AI

ma 4 sur une puce TPU en orbite, un jalon symbolique plus qu'une capacité (newsletter.semianalysis.com, blog.google).

DATES **Dates clés à venir**

3-4 nov.	Résultats d'AMD (3, après clôture), de Super Micro (3) et d'Arm (4) ; poids complets de Kimi K3.1 Pro promis par Moonshot AI (4)	Demande de calcul ; premier grand modèle chinois publié sous le nouveau régime d'enregistrement.
7 nov.	Filigrane invisible activé par défaut sur l'API Grok	Tout texte produit par l'API devient détectable, sauf désactivation sous responsabilité de l'administrateur.
9 nov.	AI Act : fin des commentaires sur l'acte d'exécution des articles 75c et 75d ; reprise de la session du Congrès américain ; tournée d'introduction en bourse d'Anthropic possible (rapporté)	Procédure d'amendes visant les fournisseurs de modèles ; régulation fédérale ; valorisation d'Anthropic.
10 nov.	Clôture de la consultation européenne sur le droit d'auteur et l'IA ; résultats de CoreWeave ; le 12, nouvelle politique d'usage d'Anthropic	Rémunération des créateurs ; santé financière des fournisseurs de calcul.
15 nov.	Chine : fin de la consultation sur l'export des technologies d'IA ; Bartz c. Anthropic : date limite des paiements	Disponibilité des modèles chinois ; indemnisation des auteurs.
17-20 nov.	Résultats de Nvidia (17) ; Apply AI Summit de la Commission (17) ; répliques dans le procès du New York Times contre OpenAI (20)	Investissement mondial ; adoption en Europe ; première décision sur le fair use en vue.
30 nov.-2 déc.	Fin du mode rapide de Claude Opus 4.8 (30) ; doublement du tarif de Gemini 4 Argon (1er) ; AI Act : fin de la transition pour le marquage des contenus générés (2)	Migrations techniques et budgets 2027 ; marquage lisible par machine obligatoire.
15 déc. - 1er fév.	Bing : préférences d'entraînement et d'usage dans les réponses (15 déc.), entrée en vigueur le 1er février 2027	Décider ce que l'entreprise autorise sur ses contenus publics.

II Publications signalées

Huit textes de la période, chacun avec son décryptage pour un comité de direction, et les points de vue qui les éclairent.

Conclusions de l'avocat général dans l'affaire Like Company c. Google (C-250/25)

Cour de justice de l'Union européenne, avocat général Maciej Szpunar · 3 novembre 2026

Saisie par un tribunal de Budapest à la demande d'un éditeur hongrois, la Cour examine pour la première fois l'IA générative au regard du droit d'auteur, à propos des résumés d'articles produits par Gemini. L'avocat général propose une grille en deux temps : l'entraînement des grands modèles sur des articles protégés relève en principe de l'exception de fouille de textes et de données (article 4 de la directive de 2019), à condition d'un accès licite et du respect d'une réservation de droits lisible par machine ; en revanche, la récupération d'un article pour construire une réponse n'est pas couverte par cette exception, et la reproduction d'une expression reconnaissable dans la réponse peut constituer une contrefaçon, le critère étant de savoir si l'utilisateur peut se dispenser de consulter la publication d'origine. Les faits ne sont pas protégés, seule leur expression l'est.

> *Le même jour, Nvidia, OpenAI, Google et Microsoft venaient d'adopter un jeton commun de réservation de droits dans robots.txt : la « réservation lisible par machine » que l'avocat général pose en condition devient opérationnelle.*

Décryptage. Si la Cour suit son avocat général, entraîner sera licite par défaut, restituer sera à risque. Pour une entreprise, cela touche d'abord les projets qui font lire des contenus de presse ou d'éditeurs à un assistant pour répondre : il faudra des licences, ou des réponses qui ne dispensent pas de consulter la source. Pour les éditeurs, c'est l'occasion de poser la réservation de droits lisible par machine sur leurs sites. Les conclusions ne lient pas la Cour.

Texte intégral : curia.europa.eu · **Analyses :** thenextweb.com · p4sc4l.substack.com

Tous droits réservés · XAVIER S. AI

Progress on provable inference for large language models (FlashProve)

Anthropic, Frontier Red Team, avec l'EPFL et le MIT · 2 novembre 2026

Un prototype permet à un client de vérifier cryptographiquement qu'une réponse provient bien du modèle annoncé, sans accès aux poids. Démonstration sur un modèle ouvert de 400 milliards de paramètres : le prouveur génère environ 110 jetons par seconde sur un seul processeur graphique, la preuve pèse 18,8 Mo pour 1 024 jetons et se vérifie en 17 secondes ; le surcoût de calcul reste de 70 à 95 fois celui d'un service classique. Les propriétés sont prouvées formellement ; le code est publié sous licence Apache. Limites signalées : progrès intermédiaire, pas un produit ; contexte limité à 1 024 jetons ; ne prouve pas l'absence d'un autre traitement en parallèle.

Décryptage. C'est la première brique crédible pour exiger d'un fournisseur ou d'un intermédiaire la preuve du modèle réellement servi : audit indépendant, détection d'une substitution de modèle par un routeur, vérification de règles de déploiement. À inscrire dès maintenant dans les clauses d'auditabilité des contrats, même si la technique n'est pas mûre.

Texte intégral : anthropic.com · rapport technique (PDF) · code

Speeding up planning decisions with AI (PlanAssist)

Google DeepMind, avec le gouvernement britannique, Google Cloud et Faculty · 31 octobre 2026

Un outil fondé sur Gemini prend en charge les vérifications routinières des permis de construire des particuliers : validation des documents, contrôle des contraintes patrimoniales et des politiques locales, rédaction de rapports sourcés modifiables. Pilote depuis juin dans trois autorités locales : réduction jusqu'à 58 % du temps passé par instructeur sur ces dossiers, plus de 70 % des rapports générés approuvés sans modification ; extension à quinze autorités pour un essai de douze mois, puis à tous les conseils anglais en 2027 si concluant. Aucun protocole statistique publié ; l'instructeur conserve la décision.

Décryptage. Un cas d'usage public à grande échelle, chiffré, d'un agent qui instruit des dossiers réglementés sous contrôle humain : transposable à tout back-office de conformité, de sinistres ou d'achats. Les 30 % de rapports retouchés rappellent que le gain vient de la relecture, pas de la délégation.

Texte intégral : deepmind.google

As AI Reshapes SaaS, Companies Need Pivot-Ready Pricing

BCG (J. Pineda), avec AWS et Metronome · 30 octobre 2026

L'IA agentique déstabilise le modèle économique du logiciel par abonnement : environ 2 000 milliards de dollars de capitalisation perdus par les éditeurs américains début 2026 ; 57 % des entreprises en hypercroissance ont déjà un modèle hybride abonnement et usage (enquête Stripe, plus de 2 000 entreprises) ; 67 % d'entre elles ont changé de prix au moins trois fois en deux ans ; 66 % des acheteurs prévoient de dépenser plus en IA en réduisant ailleurs. Les auteurs recommandent une « taxonomie de la valeur », de ne pas indexer le revenu sur le coût des jetons et de préparer l'arrivée d'acheteurs qui seront eux-mêmes des agents. Limites : coécrit par trois acteurs qui vendent des services de tarification ; échantillons non précisés.

Décryptage. Lu du côté de l'acheteur : les contrats logiciels vont se renégocier plus souvent, vers des modèles hybrides où l'usage des agents fera varier la facture. Exiger des plafonds, des indicateurs d'usage lisibles et des clauses de sortie, et considérer l'accès par API et MCP comme un poste de coût à part.

Texte intégral : bcg.com

How AI agents shape experts' judgment calls

Anthropic (A. Herdağdelen, H. Joren, A. Tamkin) · 27 octobre 2026

Cinquante analystes expérimentés devaient conseiller une coopérative agricole fictive, avec Claude Code dont la position sur chaque choix avait été tirée au sort et inscrite dans une consigne cachée. Ils ont suivi l'agent dans 86 % des décisions, contre 50 % attendus sans effet ; quand l'agent contredit un choix déjà exprimé, celui-ci change dans 60 % des cas ; 77 % reprennent ses arguments, 15 % seulement le citent. Ni l'expérience ni la confiance dans l'IA ne prédisent cette déférence. En contrepoint, six des onze participants qui ont refusé l'agent ont commis des erreurs de méthode invalidantes, contre trois sur cinquante avec lui.

> *Marc Andreessen (a16z, 26 octobre) : comme le tableur et l'e-mail, les agents entrent dans l'entreprise par les salariés, dans le désordre ; il ne faut ni les mettre en cage ni les laisser sans limite, mais leur construire un « habitat » fait de droits d'accès, de conventions et d'audits (a16z.com).*

Tous droits réservés · XAVIER S. AI

Décryptage. Un agent d'analyse n'est pas neutre : il a suffi qu'il défende une position pour que des analystes expérimentés la suivent dans 86 % des décisions, souvent sans le dire. Sur les sujets à enjeu, exiger qu'il expose les options concurrentes et garder la trace de ses apports. Cette semaine, Goldman Sachs a tranché autrement, en retirant les agents aux fonctions métier.

Texte intégral : anthropic.com · article (PDF)

GDPx : measuring frontier model capabilities on expert tasks

SemiAnalysis · 2 novembre 2026

Un banc d'essai de tâches allant jusqu'à huit heures, dans onze domaines à forte valeur et des environnements fournis par quatorze partenaires. Claude Opus 5.5 est premier avec 12 % de réussite ; Gemini 4 Carbon est le seul modèle dans le même ordre de grandeur ; GPT-6 Astra domine l'auto-amélioration et les matériaux ; la finance et le droit donnent près de zéro à tous les modèles en notation binaire ; les classificateurs de sécurité bloquent certaines tâches de cybersécurité et de biologie.

> *OpenAI publie le même jour LifeVal, son propre banc d'essai de 169 tâches de la vie courante, où ses modèles progressent de 49 % à 81 % en dix mois, sans publier ni les tâches ni les grilles (news.aibase.com).*

Décryptage. L'écart entre les démonstrations et les tâches expertes réelles reste considérable : 12 % au mieux. Pour calibrer un projet d'agents, viser les tâches courtes, vérifiables et encadrées ; et se méfier des bancs d'essai que chaque laboratoire construit pour lui-même.

Texte intégral : newsletter.semianalysis.com

The End of the Internet, et la réponse de Nathan Lambert

Ben Thompson, Stratechery · 26 octobre 2026 ; Nathan Lambert, Interconnects · 30 octobre 2026

Ben Thompson part d'une semaine de failles (PHP, Node.js, GitLab, Java, Cisco, LiteLLM, Railway, autorités de certification) pour conclure que l'IA donne aux attaquants un temps illimité et remplace l'ingéniosité par l'itération : l'Internet ouvert, où chacun expose son serveur au monde, touche à sa fin.

> *Nathan Lambert (Interconnects, 30 octobre), « Was I wrong about open models & cyber risks? » : après un « octobre noir », il maintient que les modèles ouverts rééquilibrent attaque et défense, mais reconnaît avoir mal formulé sa position d'avril : une période de turbulence coûteuse est à prévoir, et les modèles fermés ont joué un rôle plus important que les ouverts dans les incidents récents (interconnects.ai).*

Décryptage. Que l'on suive Thompson ou Lambert, la conclusion opérationnelle est la même : tout service exposé sera sondé en continu par des outils automatisés. Inventaire de ce qui est exposé, réduction, authentification forte, hébergement géré ; et, cette semaine l'a montré avec Aur0ra, supervision des agents de code que les attaquants utilisent aussi.

Texte intégral : stratechery.com · interconnects.ai

ChatGPT's expanding world of work

OpenAI, équipe de recherche économique · 26 octobre 2026

Sept trimestres, près de 100 pays, 165 activités de travail : aux États-Unis, le groupe « planification et décisions » passe de 35,8 % à 39,6 % des requêtes professionnelles entre début 2025 et le troisième trimestre 2026, porté par la pensée créative (de 11,8 % à 18,5 %) ; les tâches informatiques reculent de 11,7 % à 6,8 %. Limites : comptes individuels seulement, parts et non volumes.

> *Torsten Slok (Apollo, 24 octobre) : hors des sept géants de la technologie, le chiffre d'affaires par salarié et les marges des grandes entreprises américaines stagnent ; l'avenir de l'IA se jouera dans ces entreprises ordinaires (apollo.com).*

Décryptage. L'usage professionnel glisse de l'exécution vers la conception et le conseil : l'indicateur utile devient la qualité des idées et des décisions produites avec l'outil, plus que le taux d'adoption.

Texte intégral : openai.com · **Données :** openai.com/signals

Introducing the AI Chip Imports explorer

Epoch AI (L. Melito, I. Juniewicz) · 27 octobre 2026

Tous droits réservés · XAVIER S. AI

Sur 27,4 millions d'équivalents H100 arrivés dans le monde entre 2022 et juin 2026, les États-Unis en ont reçu 72 %, la Chine 10 %, l'Union européenne 5 % ; la part américaine atteint 78 % au premier semestre 2026. Limites : codes douaniers mêlant puces d'IA et autres ; contrebande hors mesure.

Décryptage. Les entreprises européennes resteront largement dépendantes de capacités situées hors d'Europe : sécuriser l'accès au calcul, l'intégrer au coût des projets et savoir où les données sont traitées. La règle des 51 % que Nvidia imposerait à ses partenaires montre que cet accès peut aussi se politiser.

Texte intégral : epoch.ai

L'hypothèse du temps exponentiel réfutée : deux laboratoires, douze heures d'intervalle

OpenAI (26 octobre 2026) et Anthropic (article signé « Claude », 27 octobre 2026)

OpenAI annonce une réfutation, trouvée par ses modèles et vérifiée par machine, d'une hypothèse de 1999 sur le problème logique 3-SAT ; Anthropic publie douze heures plus tard un algorithme voisin trouvé par un modèle interne. Aucune validation par la communauté à ce stade ; portée purement asymptotique ; aucun effet annoncé sur la cryptographie.

> *Jack Clark (Import AI, 26 octobre) : un chercheur seul, aidé par l'IA, a entraîné en trois mois et demi, pour moins de 10 millions de dollars de calcul, un modèle qui demandait une équipe de dix personnes un an plus tôt (importai.substack.com).*

Décryptage. Pour les directions de la recherche, la vérification devient la compétence clé ; deux annonces à douze heures d'intervalle imposent d'attendre l'avis des pairs.

Texte intégral : openai.com · article d'Anthropic

Également parues. OpenAI, programme Grid Protectors (2/11) et LifeVal (3/11) ; Epoch AI, les grands modèles de langage mentionnés dans 17 % des prépublications de biologie, presque autant que les outils spécialisés (30/10) ; BCG, le war gaming stratégique par agents multiples (3/11) ; Richard Gray dans ChinaTalk, le krach coréen et la chaîne de la mémoire (1/11) ; Exponential View, en octobre, 57 % des publicités politiques américaines mentionnant les centres de données s'y opposaient, 3 % les soutenaient (2/11) ; Simon Willison, la chaîne d'attaque qui prenait le contrôle de comptes ChatGPT, corrigée (31/10) ; Goldman Sachs, note sur les licenciés de postes exposés à l'IA (30/10) ; The Batch, les peurs de l'IA pour Halloween (30/10) ; MIT, prix Pritzker pour Regina Barzilay et Jim Collins, antibiotique découvert par IA (2/11) ; UK AI Security Institute, filtrage des données et reconstruction d'outils biologiques (26/10) ; OpenAI, Model Spec (29/10) ; GitHub, Octoverse 2026 (28/10) ; Zilan Qian (GovAI, ChinaTalk), la part réelle des mini-séries dans les jetons consommés en Chine, environ 7,4 % (30/10).

III Nouvelles IA, nouveaux usages

Ce qui a changé dans les assistants les plus utilisés, avec pour chacun une opportunité à saisir et un point de vigilance, puis les applications professionnelles qui méritent un essai. Un « agent » est un assistant capable d'agir seul pour le compte de l'utilisateur ; le « MCP » est le protocole standard qui relie un assistant aux outils et aux données d'une entreprise.

CLAUDE · ANTHROPIC

Des conversations plus rapides pour les agents, une offre Ultra, et une date à retenir : le 30 novembre

Depuis le 2 novembre, l'API accepte des **fils de messages** : le client n'envoie plus toute la conversation à chaque tour, ce qui réduit d'environ 35 % le délai de première réponse sur les très longs contextes, sans changer le coût ; le fil est conservé chiffré au plus une heure. L'offre Max 20x devient **Ultra** (200 $ par mois, sans limite sur cinq heures pour les comptes vérifiés, crédits d'API inclus). Le **mode rapide de Claude Opus 4.8 sera supprimé le 30 novembre** : les intégrations qui l'utilisent doivent migrer. Claude Code peut désormais **interdire toute écriture hors des répertoires de travail**, quel que soit le mode. Les clients Enterprise disposent depuis le 28 octobre de rapports d'usage par équipe ; en France, analyser les conversations de salariés suppose de les informer et de consulter le CSE.

Opportunité. Un éditeur qui fait tourner un agent de traitement de documents longs gagne en réactivité sans surcoût grâce aux fils de messages ; un directeur des opérations repère, grâce aux rapports d'usage, les tâches à transformer en compétences partagées.

Vigilance. Les fils de messages gardent un état côté serveur, à vérifier au regard des exigences de résidence et de journalisation ; se servir des rapports d'usage pour noter des salariés exposerait l'entreprise à un contentieux social et à une sanction au titre du RGPD.

platform.claude.com · support.claude.com · github.com/anthropics

CHATGPT · OPENAI

Tous droits réservés · XAVIER S. AI

Un quota unique pour ChatGPT et Codex en décembre, un modèle de biologie à accès contrôlé

À partir de décembre, les offres Plus, Pro et Business partageront **une même allocation de jetons entre ChatGPT et Codex** ; les modèles avancés consomment davantage, le chat rapide est exclu, et OpenAI ne publie pas la taille des allocations. Sur mobile, deux comptes peuvent rester connectés simultanément, utile pour séparer compte personnel et compte professionnel (30 octobre). **GPT-Rosalind-6**, mêmes poids que GPT-6 Astra mais moins de refus sur le double usage biologique, est réservé à des organisations approuvées après vérification, hors du ChatGPT grand public. Le modèle **GPT-6.1 Terra**, qui pilote un navigateur et un bureau, reste proposé à la moitié du prix du modèle supérieur.

Opportunité. Une direction de la recherche en santé peut demander l'accès à GPT-Rosalind-6 pour la conception de molécules, avec des garanties de revue humaine ; un commercial confie à GPT-6.1 Terra la saisie de ses comptes rendus dans un logiciel accessible seulement par navigateur.

Vigilance. Revoir les budgets Business avant décembre, car l'usage des agents de code pèsera sur le quota des collaborateurs ; les données de santé et génétiques relèvent de l'article 9 du RGPD, et OpenAI peut restreindre ou révoquer l'accès à Rosalind, un risque de continuité. Un lien partagé a permis de détourner des sessions ChatGPT jusqu'à une correction récente.

cellcog.ai · deploymentsafety.openai.com · i-scoop.eu

COPILOT · MICROSOFT

Brancher Copilot sur ses propres données

Depuis le 27 octobre, les administrateurs de Microsoft 365 Copilot créent et déploient leurs propres **connecteurs MCP** vers les données de l'entreprise ; Researcher puise dans Canva et PowerPoint reprend le style d'une présentation existante. Chez GitHub, **Copilot Fusion** choisit et combine plusieurs modèles selon l'effort demandé, avec une remise temporaire de 30 % ; la ligne de commande Copilot gère l'authentification Kerberos derrière un proxy d'entreprise (2 novembre). Aucune nouveauté majeure de Microsoft 365 Copilot dans la période.

Opportunité. Une direction juridique relie Copilot à sa base de contrats par un connecteur MCP maison : chaque juriste l'interroge en langage courant sur les échéances ou les clauses de résiliation, sans export ni copie de fichiers.

Vigilance. Un connecteur mal paramétré ouvre à tous les utilisateurs des documents réservés : reproduire et tester les droits d'accès avant tout déploiement. Les mesures d'usage de Viva Insights, si elles comparent des salariés, relèvent du contrôle de l'activité.

learn.microsoft.com · github.blog

GEMINI · GOOGLE

Créer et modifier des documents dans la conversation ; un instructeur de permis de construire

Depuis le 29 octobre, **Gemini crée, affiche et modifie des fichiers** dans un éditeur plein écran, PDF, documents, tableurs et présentations, exportables en PowerPoint ou vers Drive, pour tous les comptes, en France aussi. Dans Sheets, « Fill with Gemini » cite la source web de chaque cellule remplie. **Gemini 4 Carbon** reste réservé à des spécialistes de la cyberdéfense. Au Royaume-Uni, PlanAssist instruit les permis de construire des particuliers avec jusqu'à 58 % de temps gagné (rubrique II). Aucune autre nouveauté de l'application dans la période.

Opportunité. Un contrôleur de gestion transforme un tableau de résultats en présentation pour le comité de direction, la retouche dans la conversation, puis l'exporte en PowerPoint.

Vigilance. Une cellule sourcée n'est pas une cellule exacte : les chiffres remplis automatiquement restent à vérifier avant toute décision ; la création de fichiers étant ouverte aux comptes personnels, une règle interne doit encadrer les données de l'entreprise.

blog.google · deepmind.google

MISTRAL AI ET XAI

Un grand modèle ouvert européen, et un filigrane par défaut chez Grok dès le 7 novembre

Mistral Large 4, environ 1 000 milliards de paramètres, est disponible en poids ouverts depuis le 29 octobre : on peut le faire tourner sur ses propres serveurs ; usage commercial autorisé, accord séparé pour revendre le modèle en service ; API à 0,70 $ par million de jetons en entrée et 2,10 $ en sortie ; aucune nouveauté de l'assistant Vibe dans la période. Chez xAI, **l'API Grok appliquera par défaut, à partir du 7 novembre, un filigrane statistique invisible** à tous ses textes, y compris les sorties structurées, au nom de l'article 50 de l'AI Act ; un administrateur peut le désactiver modèle par modèle en attestant d'une exception et en assumant la responsabilité, avec journal d'audit. Un outil de détection est fourni sur demande.

Tous droits réservés · XAVIER S. AI

Opportunité. Une banque, un hôpital ou un industriel tenu au secret peut faire tourner Mistral Large 4 chez lui ou chez un hébergeur européen, et l'adapter à ses documents internes sans qu'aucune donnée ne quitte son périmètre.

Vigilance. Si vos produits réécrivent ou republient du texte produit par Grok, décider avant le 7 novembre si vous conservez le filigrane ou si une exception s'applique ; désactiver, c'est assumer. Revendre l'accès à Mistral Large 4 en service sans l'accord séparé serait une violation de licence.

mistral.ai · docs.x.ai

ASSISTANTS CHINOIS

Un agent de code autonome chez DeepSeek, un modèle de 1 600 milliards de paramètres chez Meituan

DeepSeek publie en source ouverte **dsh-superagent**, prototype d'agent qui surveille un dépôt de code, transforme les demandes en tickets exécutés par des agents isolés et rend des propositions de modification, avec une interface pour les points qui exigent un humain ; les auteurs préviennent qu'il a « un accès complet à l'utilisateur du système » et que son code a été largement écrit par des agents (2 novembre). Meituan lance **LongCat-2.5**, modèle multimodal de 1 600 milliards de paramètres entraîné sur environ 50 000 accélérateurs chinois, accessible par API, poids ouverts annoncés (2 novembre). Les poids complets de Kimi K3.1 Pro sont promis pour le 4 novembre ; DeepSeek entraînerait son prochain modèle sur 160 000 puces Huawei (non confirmé). Doubao Work (ByteDance) réunit Kimi, GLM et DeepSeek sur les mêmes crédits.

Opportunité. Pour une entreprise implantée en Chine, Doubao Work réunit plusieurs grands modèles chinois dans un même outil ; dsh-superagent montre vers quoi vont les agents de développement : des files de tickets traitées en continu.

Vigilance. Un agent qui a tous les droits de l'utilisateur du système, dans un code écrit par des agents, s'isole strictement avant tout essai ; les services hébergés en Chine sont soumis à la loi sur le renseignement de 2017 : aucun document client sans validation de la direction informatique.

github.com/deepseek-ai · technode.com

Applications à essayer

1. **Zoom : les résumés de réunion envoyés automatiquement, disponible dans l'Union européenne.** Depuis le 30 octobre, un nœud « Envoyer les données de réunion » des Workflows de ZoomMate envoie des résumés personnalisés vers plusieurs applications en une étape, avec des modèles de résumé déployés par les administrateurs. Vérifier les destinations autorisées : un résumé de réunion peut contenir des décisions confidentielles. support.zoom.com
2. **Slack : un assistant qui agit sur événement.** Depuis le 28 octobre, Slackbot lance une tâche quand survient un message, une réaction ou un canal créé, après accord de l'utilisateur ou seul ; les administrateurs valident un catalogue de compétences partagées. Réserver l'action automatique aux tâches réversibles. slack.com
3. **Oracle et OpenAI : piloter son progiciel de gestion depuis ChatGPT.** Des modules permettent d'interroger en langage courant les progiciels Fusion et NetSuite et de lancer des actions approuvées ; ni date ni prix publics. Avant tout essai, demander par écrit où sont traitées les données financières. oracle.com
4. **Factorial : un agent qui prépare le comité de direction.** One construit une présentation pour le comité de direction à partir des données RH (effectifs, absences, heures supplémentaires), avec une recommandation d'effectifs et l'origine de chaque chiffre ; il rédige aussi des avis de retard, qui peuvent constituer une sanction disciplinaire en droit français : vérifier chaque cas avant l'envoi. factorial.fr
5. **Lattice : les réunions d'équipe préparées et résumées par l'IA.** Lattice Meetings réunit ordre du jour, notes et actions de toutes les réunions, entretiens individuels compris ; l'assistant peut rejoindre l'appel pour en capter le contenu. De 4 à 10 $ par personne et par mois ; prévenir les salariés et le CSE avant de l'activer. lattice.com
6. **Otter : un agent qui prépare et prolonge les réunions.** Le mode « Work », pour tous les comptes payants, enchaîne notes de préparation, e-mails de suivi et mises à jour dans Salesforce ou Jira, chaque action étant validée. En France, enregistrer une réunion suppose le consentement des participants. computerworld.com
7. **Apple Creator Studio : l'IA dans Keynote, Pages et Final Cut Pro.** Annoncé pour novembre, un assistant fondé sur des modèles d'Apple et d'OpenAI dans les logiciels de création, 12,99 $ par mois aux États-Unis ; l'assistant ne fonctionne qu'en anglais américain. apple.com

Tous droits réservés · XAVIER S. AI

IV Orientations stratégiques

Trois mouvements se croisent cette semaine. Le capital se rationne : le marché sanctionne les investissements sans preuve de retour, les fournisseurs se financent par la dette et la dilution, et le superviseur français s'inquiète d'une contagion. Le coût réel du calcul remonte jusqu'aux produits : Apple rogne sa marge et facturera l'IA à l'usage, OpenAI unifie ses quotas, xAI rationne Grok, Google double ses tarifs en décembre. Et le cadre se précise : une ligne européenne sur le droit d'auteur, des filigranes par défaut, un signal commun de réservation de droits, des agents de code aux mains des attaquants comme des défenseurs.

Pour un dirigeant, l'impact est triple. Les budgets d'IA 2027 doivent intégrer des prix en hausse, des quotas et des fournisseurs dont la solidité financière diverge. Les projets qui font lire des contenus de tiers à un assistant, et ceux qui publient des textes générés, relèvent désormais d'une décision juridique autant que technique. Et la gouvernance des agents, qu'ils analysent ou qu'ils codent, devient un sujet de comité d'audit : Goldman Sachs les a retirés à ses banquiers, l'étude d'Anthropic montre qu'ils orientent les décisions, Sophos montre qu'ils servent aux rançongiciels.

XAVIER S. AI
AI
L'INFORMATION IA POUR VOTRE COMEX
TOUS DROITS RÉSERVÉS · XAVIER SCHÄLLEBAUM

Tous droits réservés · XAVIER S. AI

ANTICIPER

Des budgets d'IA 2027 sous contrainte de prix, de quotas et de solidité des fournisseurs. Quota unifié ChatGPT-Codex en décembre, tarif de Gemini doublé le 1er décembre, Grok rationné, Siri AI à l'usage, Oracle à 300 points de base de prime de risque. Prévoir des clauses de prix plafonnés, des indicateurs d'usage lisibles, et une revue de la dépendance à chaque fournisseur de calcul, y compris de sa capacité à se financer.

La ligne européenne sur le droit d'auteur : entraîner, oui ; restituer, à risque. Si la Cour suit son avocat général, les projets qui font lire des contenus de presse ou d'éditeurs à un assistant pour répondre devront s'appuyer sur des licences ou sur des réponses qui ne dispensent pas de consulter la source. Recenser dès maintenant les usages de type RAG sur des contenus de tiers, et poser la réservation de droits lisible par machine sur ses propres contenus.

Le filigrane devient la norme : décider ce que l'entreprise publie. Grok filigrane par défaut dès le 7 novembre, Anthropic étend ses filigranes, l'AI Act impose le marquage des contenus générés le 2 décembre. Tout texte produit ou retouché par ces modèles devient traçable par des tiers habilités : fixer une politique de communication, de relecture et de mention de l'IA avant que des tiers ne la détectent.

SURVEILLER

Les clauses de nationalité sur le calcul. La règle des 51 % que Nvidia imposerait à ses partenaires cloud, le dossier du renseignement américain, la rencontre prévue entre Donald Trump et Xi Jinping : les entreprises exposées à la Chine, ou clientes de fournisseurs de calcul asiatiques et moyen-orientaux, doivent s'attendre à des clauses nouvelles sur l'identité des clients et des usages.

L'acceptabilité des centres de données, et leur délai. Près de 200 milliards de dollars de projets freinés par des oppositions locales aux États-Unis, 41 mesures sur les bulletins de dix États, un campus britannique de 1 GW dont la première phase n'arrive qu'en 2028, une réquisition de 10 % des capacités en Russie. La capacité de calcul en Europe restera rare et politique : contractualiser tôt, et prévoir l'effacement comme condition d'implantation.

L'emploi exposé à l'IA : des choix qui laissent des traces. Goldman Sachs mesure une cicatrice durable chez les licenciés de postes exposés, surtout après 35 ans ; Challenger attribue 120 136 suppressions à l'IA depuis janvier aux États-Unis ; Microsoft propose 60 000 départs. Privilégier la requalification interne, documentée, et préparer la communication sociale avant toute réorganisation « IA ».

PROTÉGER

Les agents de code sont aussi des outils d'attaque. Sophos attribue à un groupe de rançongiciel l'usage d'agents de code pour toute la chaîne d'intrusion, garde-fous contournés par un simple prétexte de test autorisé. Inventaire des agents, des extensions et des clés ; droits limités ; sortie réseau bloquée par défaut ; journalisation ; et provenance vérifiée de tout installateur d'outil d'IA.

Un cycle de correctifs quotidien, pas hebdomadaire. HTTP2Shell et la faille du noyau Linux sont exploitées, après WordPress, PHP2Shell et la fausse extension de la semaine précédente ; un modèle ouvert sait transformer un correctif publié en attaque. Corriger dans la journée, analyser son propre code à chaque modification, et considérer tout correctif publié comme une piste d'attaque.

Les données sensibles hors des outils sans contrat. GPT-Rosalind-6 attire les équipes de recherche vers des données de santé relevant de l'article 9 du RGPD ; les assistants chinois restent soumis à la loi sur le renseignement ; la Suède a tranché par une liste blanche. Une liste d'outils couverts par un contrat, des règles d'usage connues de tous, et un modèle téléchargé traité comme un logiciel non signé.

Tous droits réservés · XAVIER S. AI

DÉVELOPPER

Organiser la contradiction face aux agents d'analyse. La déférence est massive (86 %), et ni l'expérience ni la méfiance envers l'IA ne protègent ; Goldman Sachs a choisi la restriction, d'autres choisiront l'encadrement. Sur toute décision à enjeu préparée avec un agent, exiger deux options argumentées, faire défendre l'option contraire et noter ce qui vient de la machine. C'est une règle de gouvernance, pas une précaution technique.

Choisir ses cas d'usage sur preuve chiffrée. PlanAssist fait gagner jusqu'à 58 % de temps d'instruction avec 70 % de rapports acceptés tels quels ; Palantir voit son commercial américain croître de 148 % ; mais le meilleur modèle ne réussit que 12 % des tâches expertes longues. Viser les tâches courtes, répétitives, vérifiables et encadrées, mesurer le temps gagné et le taux de reprise, et renoncer aux promesses d'autonomie sur les tâches expertes.

Exiger la preuve du modèle servi et l'auditabilité. FlashProve montre qu'un client pourra vérifier qu'une réponse vient bien du modèle annoncé ; le jeton robots.txt commun donne à chaque éditeur un signal unique à activer. Inscrire l'auditabilité du modèle dans les contrats, et décider, avant le 1er février 2027, ce que l'entreprise autorise sur ses contenus publics : entraînement, réponses générées, ou ni l'un ni l'autre.

ANNEXE Les sociétés à suivre

Capitalisation boursière ou dernière valorisation connue, cours à la dernière clôture disponible (2 novembre pour les marchés américains, 3 novembre pour l'Asie) avec la variation du jour puis sur cinq séances, chiffre d'affaires et résultat net du dernier trimestre publié, et le fait marquant de la semaine. Les sept géants américains pèsent ensemble 22 830 Md$. Les sociétés non cotées sont signalées en italique.

SOCIÉTÉ	VALORISATION	COURS	CHIFFRE D'AFFAIRES	RÉSULTAT	SIGNAL DE LA SEMAINE
Apple États-Unis	4 940 Md$	338,32 $ (2/11, −1,8 %) ; −1,6 % sur 5 séances	114,5 Md$ (T4 ex. 2026, +12 %)	30,7 Md$	Trimestre record, mais marge brute attendue à 45-46 % en décembre à cause du coût de la mémoire, jugé « unprecedented » ; Siri AI facturée à l'usage au-delà de limites ; Apple repasse devant Nvidia en capitalisation. macrumors.com · stockanalysis.com
Nvidia États-Unis	4 880 Md$	202,06 $ (2/11, −1,3 %) ; −6,5 %	96,2 Md$ (T2 ex. 2027, +106 %)	59,7 Md$	Plafond de 49 % de clients chinois demandé à ses partenaires cloud (Financial Times, rapporté) ; garanties de dette envisagées pour le centre de Nscale (rapporté) ; résultats le 17 novembre. thenextweb.com · stockanalysis.com
Microsoft États-Unis	3 770 Md$	507,12 $ (2/11, −2,8 %) ; −4,8 %	92,3 Md$ (T1 ex. 2027, +19 %)	36,6 Md$	25 Md$ d'obligations placées avec plus de 100 Md$ d'ordres, première émission américaine depuis 2017 ; départ volontaire proposé à 60 000 salariés ; 10 Md$ de revenus liés à OpenAI comptabilisés sur le trimestre. thestar.com.my · stockanalysis.com
Alphabet États-Unis	3 430 Md$	278,99 $ (2/11, −0,3 %) ; −11,9 %	127,6 Md$ (T3 2026, +25 %)	26,5 Md$ (−25 %)	Pire séance depuis son introduction en bourse le 29 octobre (−12,6 %) : bénéfice sous les attentes, 51,4 Md$ d'investissements, trésorerie disponible négative ; émission obligataire de 6 Md€ le 2 novembre ; Google Cloud +103 %. malaymail.com · stockanalysis.com
Amazon États-Unis	2 660 Md$	245,63 $ (2/11, +2,2 %) ; −1,4 %	201,9 Md$ (T3 2026, +12 %)	20,7 Md$	AWS à 46 Md$ (+39 %) ; trésorerie disponible de −42,4 Md$ sur douze mois, environ 220 Md$ d'investissements prévus en 2026 ; rebond le 2 novembre après des résultats sous les attentes. constellationr.com · stockanalysis.com
Meta États-Unis	1 550 Md$	606,73 $ (2/11, −1,3 %) ; −3,8 %	64,6 Md$ (T3 2026, +26 %)	9,9 Md$ (après une charge de 9,9 Md$)	S&P Global passe la perspective à négative (30 octobre) ; émission obligataire d'au moins 20 Md$ après l'augmentation de capital d'environ 37,5 Md$ (rapporté) ; dépenses 2026 relevées à 175-179 Md$; enquête de la CMA britannique sur l'usage des contenus de presse. stockanalysis.com
SpaceX-xAI États-Unis	1 850 Md$	136,37 $ (2/11, +1,2 %) ; −7,2 % ; sous le	14,6 Md$ (T3 2026, +161 %)	4,4 Md$	xAI, absorbée en février, et Cursor, rachetée en août, sont dans le périmètre ; quotas de Grok réduits, leur coût étant « unsustainably high » selon Elon Musk ; rumeur de fusion avec Tesla (Reuters,

Tous droits réservés · XAVIER S. AI

SOCIÉTÉ	VALORISATION	COURS	CHIFFRE D'AFFAIRES	RÉSULTAT	SIGNAL DE LA SEMAINE
		prix d'introduction de juin (135 $)			rapporté) ; plainte de la NAACP sur les turbines de Memphis maintenue. finance.yahoo.com · stockanalysis.com
Oracle États-Unis	361 Md$	119,22 $ (2/11, −8,8 %) ; −5,4 %	19,3 Md$ (T1 ex. 2027, +30 %)	4,7 Md$	20 Md$ levés le 2 novembre, dont des obligations à environ 8,4 % ; prime de risque de crédit autour de 300 points de base ; titre à plus de 60 % sous son pic de septembre 2025 ; trésorerie disponible positive visée au second semestre de l'exercice 2028. finance.yahoo.com · stockanalysis.com
Palantir États-Unis	429 Md$	178,34 $ (2/11, −3,1 %) ; −10,3 % ; −6 % en séance le 3/11	2,31 Md$ (T3 2026, +96 %)	1,13 Md$	Le commercial américain (+148 %) dépasse pour la première fois le gouvernement ; prévision 2026 relevée à 8,53 Md$; mais le titre baisse : croissance séquentielle attendue de 15 % et multiple de bénéfices d'environ 250. fortune.com · 247wallst.com
Anthropic États-Unis	1 500 à 1 600 Md$ visés (Wall Street Journal, rapporté)	*non cotée ; symbole ANT demandé*	16,96 Md$ (T3 2026, +47 %)	−23,3 Md$ (surtout comptable) ; exploitation −567 M$	Prospectus amendé le 2 novembre : 895,4 Md$ d'engagements de calcul, tournée possible dès le 9 novembre, cotation visée avant Thanksgiving (rapporté) ; note « BB » anticipée après la cotation (Financial Times, rapporté). stocktitan.net · techcrunch.com
OpenAI États-Unis	852 Md$ (dernier tour clos) ; 1 250 Md$ visés (rapporté)	*non cotée*	n. d.	n. d.	Levée ramenée de 30 à 20 Md$ (Financial Times, rapporté) ; quota unifié ChatGPT-Codex en décembre ; recrutement de Dan Driscoll ; accords avec Abou Dhabi ; programme Grid Protectors ; cotation évoquée pour 2027 (rapporté). techcrunch.com
Mistral AI France	21 Md€ (série D, septembre 2026)	*non cotée*	n. d.	n. d.	Mistral Large 4 en poids ouverts depuis le 29 octobre, distribué par Microsoft Foundry ; citée comme client de l'index de Brave dans la plainte du New York Times. orrick.com · mistral.ai
DeepSeek Chine	≈ 74 Md$ avant levée (rapporté)	*non cotée*	n. d.	n. d.	Tour d'environ 16 Md$ avec le fonds d'État, Tencent, CATL et Meituan (The Information, rapporté) ; agent de code dsh-superagent publié ; prochain modèle entraîné sur des puces Huawei (non confirmé). aiweekly.co · github.com/deepseek-ai
Samsung Electronics Corée du Sud	≈ 1 064 Md$ (1 429 000 Md₩)	219 500 ₩ (3/11, −4,2 %) ; −1,1 %	195 300 Md₩ (T3 2026, ≈ 145 Md$)	Rés. d'exploitation 107 400 Md₩ (≈ 80 Md$, ×8,8)	Première entreprise technologique à dépasser 100 000 Md₩ de résultat d'exploitation trimestriel ; −4 % le 3 novembre après l'avertissement d'Apple ; usine de Taylor réservée à des composants HBM pour AMD et Nvidia (rapporté). news.samsung.com · finance.yahoo.com
SK hynix Corée du Sud	≈ 734 Md$ (985 800 Md₩)	1 398 000 ₩ (3/11, −5,0 %) ; +4,3 %	104 600 Md₩ (T3 2026, +328 %)	Rés. d'exploitation 83 330 Md₩ (marge ≈ 80 %)	100 000 Md₩ de retour aux actionnaires ; mémoire à haut débit vendue jusqu'en 2029, plus de la moitié de la capacité 2027-2028 sous contrats longs. koreatimes.co.kr · finance.yahoo.com
Broadcom États-Unis	1 540 Md$	322,23 $ (2/11, −1,2 %) ; −2,1 %	29,6 Md$ (T3 ex. 2026, +86 %)	13,1 Md$	63 Md$ de puces pour Anthropic financés par une société ad hoc menée par Blackstone ; −5,5 % le 30 octobre dans le sillage d'Intel ; résultats vers le 10 décembre. stockanalysis.com
AMD États-Unis	762 Md$	466,84 $ (2/11, −7,5 %) ; −12,1 %	11,5 Md$ (T2 2026, +50 %)	2,3 Md$	Résultats du troisième trimestre le 3 novembre après la clôture ; deux séances de forte baisse après les prévisions d'Intel ; Morgan Stanley relève son objectif à 510 $. stockanalysis.com
TSMC Taïwan	1 850 Md$ (ADR)	398,63 $ (ADR, 2/11, −1,9 %) ; −6,1 %	46,7 Md$ (T3 2026, +51 %)	≈ 23,9 Md$ (763 Md TWD)	Marge brute de 66,4 % ; prévision de 52 à 54 Md$ pour le quatrième trimestre ; repli sectoriel sans nouvelle propre. investor.tsmc.com · stockanalysis.com
CoreWeave États-Unis	27,7 Md$	50,15 $ (2/11, −8,6 %) ; −9,1 %	2,58 Md$ (T2 2026, +112 %)	−626 M$	Deux séances à −9 % sur fond de craintes sur l'endettement et le financement des « néoclouds » ; partenariat de sécurité avec CrowdStrike ; résultats le 10 novembre. stockanalysis.com
	252 Md$ et 456 Md$			10,5 Md¥ et 56,0 Md¥	Apple Intelligence attendu en Chine en décembre avec Alibaba (SCMP, rapporté) ; Joe Tsai défend

Tous droits réservés · XAVIER S. AI

SOCIÉTÉ	VALORISATION	COURS	CHIFFRE D'AFFAIRES	RÉSULTAT	SIGNAL DE LA SEMAINE
Alibaba et Tencent Chine		99,55 HKD (+0,1 %) et 397,20 HKD (+1,3 %) le 3/11	269 Md¥ (T1 ex. 2027) et 204,8 Md¥ (T2 2026)		l'open source à Wuzhen ; Tencent parmi les souscripteurs du tour de DeepSeek et partenariat Mastercard sur le commerce par agents. stockanalysis.com · stockanalysis.com
Unitree Chine	≈ 37,9 Md$ (252,6 Md¥)	624,58 ¥ (3/11, −5,7 %) ; ×4,1 depuis l'introduction d'août	665 M¥ (T3 2026, +70 %)	113 M¥ (−17 %)	Croissance sous l'attente et bénéfice en recul, recherche et développement multipliée par 3,5 ; multiple de bénéfices supérieur à 400. finance.yahoo.com

Données de sources publiques citées, à la date indiquée ; elles peuvent avoir évolué depuis. Cette annexe ne constitue pas un conseil en investissement.

Faites connaître xaviers.ai

Cette lettre est gratuite et vit de ses lecteurs. Si elle vous a été utile, transférez-la au membre de votre comité de direction qui pilote le sujet, partagez-la sur LinkedIn, ou inscrivez un collègue. Vous pouvez aussi télécharger l'édition en PDF, avec son sceau, pour la lire ou la faire circuler en réunion. Et répondez à ce message : quel sujet voulez-vous voir décrypté dimanche prochain ?

- **S'inscrire** : sur xaviers.ai, ou en envoyant un e-mail à contact@xaviers.ai avec l'objet « Inscription à xaviers.ai » ; vous recevrez un e-mail de confirmation à valider.
- **Transférer** ce PDF à un membre de votre Comex, ou **partager** l'édition en ligne : xaviers.ai/editions/2026-11-03.
- **Répondre** : vos remarques et vos sujets à contact@xaviers.ai.

Un collègue qui reçoit cette lettre de votre part n'est pas inscrit pour autant : il s'inscrit lui-même sur xaviers.ai et confirme son adresse. Aucune liste n'est partagée ni vendue.

Flashez pour vous inscrire

xaviers.ai · L'information IA pour votre Comex · Édition n° 1 · 3 novembre 2026 · Édité par BYZANCE S.A.S., 5 rue du Général Delanne, 92200 Neuilly-sur-Seine · RCS Nanterre 519 870 711 · Directeur de la publication : Xavier Schällebaum · contact@xaviers.ai · Ce document d'information et d'analyse ne constitue ni un conseil en investissement, ni un conseil juridique ou fiscal personnalisé ; les données financières sont reprises de sources publiques citées, à la date indiquée.

L'INFORMATION IA POUR VOTRE COMEX
TOUS DROITS RÉSERVÉS · XAVIER SCHÄLLEBAUM

Tous droits réservés · XAVIER S. AI