

Comment devenir un centaure

— version 2026

Ou pourquoi la plus belle promesse de 2018 s'est transformée, huit ans plus tard, en une question nettement moins romantique.

Prologue — Le centaure est né d'une humiliation

En mai 1997, Garry Kasparov a perdu aux échecs contre Deep Blue. Ce n'est pas le plus intéressant. Le plus intéressant, c'est ce qu'il a fait l'année suivante : au lieu de demander une revanche d'honneur, il a inventé les *Advanced Chess* — un format où chaque joueur humain dispose, en direct, d'un moteur de calcul. Humain plus machine. La légende du **centaure** était née.

En 2005, un tournoi online baptisé *Freestyle* posa la question qui intéressait vraiment Kasparov : un centaure est-il plus fort qu'un supercalculateur solo ? La réponse est devenue mythique dans le petit monde des essais d'Harvard Business Review et des keynotes TED : **oui**. Mieux : deux joueurs amateurs américains équipés de trois PC portables moyens et d'une discipline méthodologique rigoureuse ont battu des grands maîtres ET des supercalculateurs. Le message tenait en une ligne : l'intelligence n'est pas un axe unique, donc l'humain et la machine, spécialisés sur des dimensions différentes, gagnent ensemble.

En janvier 2018, Nicky Case a écrit l'essai définitif sur cette idée : *How To Become A Centaur*, publié par le MIT Journal of Design and Science, prix du concours *Resisting Reduction* organisé par Joi Ito. Case y ressuscitait l'**IA oubliée** — non pas Artificial Intelligence, mais **Intelligence Augmentation**, la lignée Vannevar Bush → Licklider → Engelbart → Alan Kay. Le message de Case tenait en un caractère typographique : dans « Humain + IA », le vrai sujet d'étude, c'est le **+**.

Huit ans plus tard, relisons. Avec tendresse. Mais sans complaisance.

Acte 1 — Le centaure est mort (aux échecs, en tout cas)

Commençons par le détail embarrassant que l'essai de Case n'aborde pas, parce qu'en 2018 on pouvait encore l'ignorer : **le centaure échiquéen n'existe plus**.

Tyler Cowen l'avait pressenti dès 2013, à partir de discussions sur les forums Rybka : à cadence de tournoi classique (90 min/partie), l'apport humain est devenu du bruit plus que du signal. Au-delà d'un certain seuil de puissance des moteurs, les « surpasse » de l'humain sur la machine se transforment statistiquement en erreurs. En 2017, Garry Kasparov lui-même affirmait encore qu'un bon opérateur humain pouvait battre le meilleur supercalculateur. En 2024, Marginal Revolution actait officiellement le décès : *Centaur chess is now run by computers*. Autrement dit, la nouvelle génération de « centaures » est composée d'un moteur principal (Stockfish, Leela, AlphaZero) **orchestré par un autre moteur** qui choisit entre plusieurs systèmes divergents.

L'humain a été prié de quitter la table.

Ce détail est fondamental, parce que Case a bâti tout son argumentaire sur le théorème du *No Free Lunch* — aucun algorithme ne peut être bon partout, donc la spécialisation est obligatoire, donc humains et machines gagnent à s'associer. Le théorème reste vrai. Mais il s'applique maintenant aux relations **entre moteurs**, pas entre humain et moteur. La diversité cognitive qui était l'alibi du centaure n'a pas disparu — elle a simplement migré entre espèces de silicium.

On me dira : les échecs sont un système fermé, fini, à information parfaite ; le travail d'un directeur des opérations, d'un officier NATO ou d'un négociateur commercial ne l'est pas. C'est exact, et cette asymétrie doit être reconnue d'emblée. Mais elle ne sauve pas la métaphore : elle la déplace. Dans un domaine ouvert, le verrou n'est plus la performance brute — c'est la **vitesse de capture décisionnelle**. Et cette vitesse, contrairement à la performance, est indifférente à la finitude du problème. Autrement dit : plus le domaine est vaste et incertain, plus l'incitation à déléguer la décision à la machine qui répond en 1,8 seconde est forte. Le centaure ouvre les 64 cases à l'infini, et se fait rattraper par la même logique d'absorption.

Moralité : dans le domaine borné des échecs, la fenêtre pendant laquelle humain et machine étaient complémentaires a duré environ **vingt ans** (1997-2017). Ensuite, le + humain est devenu un -.

C'est une nouvelle qui devrait nous intéresser, parce que les hyperscalers nous vendent depuis trois ans l'idée que **tout domaine professionnel** peut être modélisé comme un échiquier un peu plus grand.

Acte 2 — Il est ressuscité partout ailleurs (en moins bien)

Ce que Case n'avait pas prévu, et qu'aucun essai raisonnable de 2018 ne pouvait prévoir, c'est qu'en novembre 2022 le grand public rencontrerait ChatGPT. Et qu'en trois ans, l'humanité ferait son premier vrai **essai centaure à l'échelle planétaire** : plusieurs centaines de millions d'utilisateurs hebdomadaires de ChatGPT, auxquels s'ajoutent Claude, Gemini, Mistral, Copilot, Cursor, Perplexity, et la galaxie des assistants métier verticaux.

Comparé à tout ce que Case listait dans son essai (la mode algorithmique de Sung-Bae Cho en 2002, le quadrirotor évolutionnaire de Maurice Conti en 2016), c'est un changement d'échelle comparable à celui séparant le pigeon voyageur de TCP/IP. L'IA générative a réalisé, en deux ans, le programme de l'Intelligence Augmentation mieux que soixante ans d'Engelbart, Kay et Negroponte réunis.

Sauf que le centaure de 2026 ne ressemble pas du tout à celui que Case dessinait. Dans l'essai de 2018, le centaure est **symétrique** : l'humain apporte le sens, la stratégie, le jugement esthétique ; la machine apporte le calcul brut, la force de projection, la mémoire parfaite. Le + est une **interface de négociation** entre deux intelligences qui se respectent parce qu'elles se complètent.

Le centaure de 2026 est **asymétrique**. L'humain n'apporte plus un surplus cognitif — il apporte un *prompt*, un clic de validation et une carte bancaire. La machine apporte tout le reste. C'est moins un centaure qu'un **ridier Uber**. Et encore, Uber a l'honnêteté d'afficher le tarif à l'avance.

Acte 3 — Le MIT contre le MIT

L'ironie la plus féroce, c'est que c'est **la même institution qui a publié Case en 2018** — le MIT Media Lab — qui a publié en juin 2025 l'étude qui démonte la moitié de sa thèse.

Your Brain on ChatGPT: Accumulation of Cognitive Debt, signé Nataliya Kos'myna, Eugene Hauptmann et al., a suivi 54 participants pendant quatre mois, EEG à l'appui, répartis en trois groupes : LLM, moteur de recherche classique, et **brain-only** (rien, que les neurones). Les résultats, en résumé sans concession :

- **Connectivité neurale réduite jusqu'à 55 %** chez les utilisateurs LLM par rapport au groupe brain-only (différence de pic entre groupes, pas une moyenne universelle).
- **83 % des utilisateurs LLM incapables de citer le texte qu'ils venaient d'écrire** cinq minutes plus tôt.
- **Sentiment de propriété** (« c'est mon essai ») le plus bas dans le groupe LLM, le plus haut dans le groupe brain-only.
- Effet de rémanence : ceux qui ont commencé avec le LLM et ont ensuite dû s'en passer sont restés sous-engagés cognitivement. **La dette cognitive se capitalise.**

Je précise immédiatement les limites que les auteurs eux-mêmes revendiquent : 54 participants, tous dans la zone de Boston, 18-39 ans, ChatGPT seul testé, étude non encore peer-reviewée au moment où j'écris ces lignes. Les auteurs ont explicitement demandé à la presse de ne pas utiliser les mots *brain rot*, *dumb*, *stupid*. Ce n'est pas une preuve, c'est un signal. Mais ce signal converge avec plusieurs travaux indépendants sur la charge cognitive (Stadler, Bannert, Sailer, 2024), et il est suffisant pour enclencher un principe de précaution individuel et organisationnel. Ce n'est pas la peine d'attendre le nième réplicat pour commencer à s'interroger sur ses propres pratiques.

Ce que Case célébrait en 2018 — l'outil qui rend pensables des modes de pensée auparavant impossibles — se heurte en 2025 à l'observation contraire : **le LLM rend tellement facile de ne pas penser qu'il crée un nouveau mode de non-pensée jusqu'alors impossible**. Le centaure de 2018 se musclait le mollet humain en galopant. Le centaure de 2026 a délégué ses quatre pattes et commence à s'interroger, vaguement, sur l'atrophie de ses cuisses.

Acte 4 — Ce que Case n'a pas vu venir : la question de la propriété du cheval

Voici où l'essai de Case, malgré toutes ses qualités, trahit son époque : en 2018, il était encore possible de parler du couple humain+machine comme s'il s'agissait d'une relation **entre deux entités**. Un humain, une machine, et un trait d'union à dessiner proprement.

En 2026, ce cadrage est naïf. Parce que la moitié équine du centaure n'appartient pas au centaure. Elle appartient à une entreprise. Laquelle appartient à des actionnaires. Lesquels répondent à un cadre juridique et politique — le plus souvent américain ou chinois. Si vous utilisez un service hébergé par les quatre ou cinq hyperscalers dominants, votre cheval est **loué** — et son propriétaire peut à tout moment changer la selle, la race ou la direction de la course.

La métaphore symbiotique de Case (la mitochondrie et sa cellule hôte, datant d'il y a 1,5 milliard d'années) est mignonne mais omet un détail : les mitochondries ne sont pas cotées en bourse. Un centaure moderne qui travaille en environnement sensible — recherche médicale, dossier juridique, négociation commerciale, décision stratégique, analyse de défense — est statistiquement en train d'envoyer ses données à des serveurs soumis à des lois extraterritoriales : **CLOUD Act**, **FISA 702**, leurs équivalents chinois via la *National Intelligence Law* de 2017. C'est ce que je nomme, dans mes travaux, la **capture silencieuse** : pas une conquête, pas une invasion, pas un Waterloo. Juste un long glissement, un par un, de chaque décision vers un lieu que vous n'avez pas choisi. Et son corollaire plus récent, la **souveraineté représentationnelle** : la question de savoir quel corpus d'entraînement façonne, en amont, la manière dont une nation entière se raconte à elle-même.

Le problème n'est donc plus « comment bien travailler avec une IA ». Le problème est : **à qui appartient l'IA qui travaille avec vous ? Et, corollairement, que devient le jugement d'un pays dont toute la classe dirigeante, des étudiants aux ministres, réfléchit sur la même infrastructure cognitive propriétaire ?**

Case parlait de symbiose. Il faut aujourd'hui parler de **location-gérance cognitive**. Ce n'est pas moins poétique, c'est juste plus exact.

Acte 5 — Le nouveau « + »

Dans l'essai de 2018, le « + » était une question d'**ergonomie** : comment concevoir l'interface ? Quand céder la main à la machine, quand la reprendre ? Quel protocole de collaboration ?

En 2026, le « + » s'est dédoublé. Il est devenu simultanément **épistémique** et **géopolitique**.

Épistémique, parce que la vraie question n'est plus « comment bien utiliser le LLM » mais « quelle part de mon jugement suis-je en train de déléguer, et avec quelle conscience ? » L'étude MIT Media Lab de 2025 montre que les utilisateurs les plus compétents utilisent le LLM comme une caisse de résonance — ils pensent d'abord, consultent ensuite, corrigent après. Les utilisateurs moins compétents l'utilisent comme un substitut — ils consultent d'abord, et ce qui était censé être leur pensée devient, à leur insu, une variation stylistique de la pensée moyenne du corpus d'entraînement. **Le centaure de 2026 a un choix à faire entre être musclé par sa monture ou digéré par elle.**

Géopolitique, parce que l'infrastructure compte. Un avocat parisien qui utilise un service dont l'éditeur dépend du CLOUD Act n'est pas dans la même situation qu'un avocat parisien qui utilise une solution souveraine qualifiée. Un officier qui alimente une plateforme algorithmique sous droit tiers avec ses données de renseignement n'est pas dans la même situation qu'un officier qui alimente un système opéré sous droit européen. Le « + » recouvre désormais une question de **doctrine** : qui voit ce que je fais, qui s'en nourrit, qui peut l'interrompre ?

Le symbole typographique n'a pas changé. Sa facture, oui. En 2018 il coûtait une formation. En 2026, il coûte une **souveraineté**.

Six points pour (re)devenir un centaure en 2026

Case finissait son essai par quelques pistes de design. Je lui emboîte le pas, en musclant le propos — et en assumant le biais du doer : chaque point est un verbe, chaque point porte son indicateur. Trois axes, deux actions par axe.

Axe 1 — Discipline cognitive individuelle

1. Penser avant de produire. La séquence optimale est *brain* → *LLM* → *brain*, jamais *LLM* → *brain*. Les participants de l'étude MIT qui ont produit le meilleur travail sont ceux qui ont structuré leur pensée sans outil, puis consulté le LLM en enrichissement, puis repris la main pour l'arbitrage final. Ceux qui ont commencé par le LLM ne s'en sont jamais remis, même après retrait de l'outil. *Indicateur* : minutes de réflexion *brain-only* consignées avant le premier prompt, sur toute décision au-dessus d'un seuil de criticité défini à l'avance. *Objectif pratique* : 15 minutes minimum.

2. Mesurer le désaccord avec la machine. Un centaure qui approuve systématiquement son cheval n'est plus un centaure, c'est une selle. Si vous ne vous retrouvez pas en désaccord frontal avec votre LLM plusieurs fois par semaine, c'est qu'il pense à votre place, ou que vous avez arrêté de penser. Les deux hypothèses sont alarmantes. *Indicateur* : taux de rejet ou reformulation substantielle des sorties LLM. Sous 20 %, vous ne collaborez pas, vous validez.

Axe 2 — Doctrine organisationnelle

3. Établir une politique d'usage auditée. L'hygiène individuelle ne suffit pas à l'échelle d'une organisation. Il faut un cadre de gouvernance formel : charte d'usage différenciée par niveau de sensibilité des données, cartographie des flux, matrice RACI humain/agent par type de décision, analyse d'impact (DPIA au sens de l'article 35 RGPD). *Ancrages* : ISO/IEC 42001:2023 pour un système de management de l'IA, exigences EU AI Act applicables selon la catégorie (annexe III pour les systèmes à haut risque). *Indicateur* : pourcentage des processus critiques couverts par une politique formelle, documentée et testée.

4. Tracer la supply chain cognitive. Test opérationnel : pouvez-vous répondre en cinq minutes à la question « Quel modèle a traité ce document, depuis quel datacenter, sous quelle juridiction, avec quel sort réservé aux traces ? » Si non, ce n'est pas votre cheval, c'est votre aveuglement. *Ancrages exigibles selon contexte* : qualification SecNumCloud, label Cloud de confiance, EUCC, clauses contractuelles souveraines, audits tiers, garanties contractuelles d'absence de réutilisation pour entraînement. *Indicateur* : taux de couverture des usages par un fournisseur traçable au niveau requis.

Axe 3 — Architecture souveraine et résiliente

5. Modéliser la menace sur le cheval. Le LLM n'est pas seulement un outil — c'est une surface d'attaque. *Prompt injection*, empoisonnement de corpus, extraction de contexte, compromission de supply chain logicielle, exfiltration par sortie ambiguë. *Cadres* : OWASP Top 10 for LLM Applications, MITRE ATLAS, exigences EU AI Act pour les systèmes à haut risque. *Indicateur* : threat model documenté, red team annuel, plan de réponse à incident spécifique IA.

6. Préserver les sorties de secours. Un centaure dont la monture ne peut être remplacée est un otage poli. Diversité des fournisseurs (au moins deux modèles et deux juridictions pour les usages stratégiques). Plan de continuité opérationnelle dégradé mais fonctionnel sans LLM pendant 48 heures minimum. Et — discipline la plus sous-estimée — **entraînement brain-only**

hebdomadaire, à tous les niveaux hiérarchiques : un livrable par semaine, sans assistance, archivé. Le muscle cognitif fonctionne comme tout muscle : il s'atrophie quand on ne l'utilise pas. *Indicateur : exercice de continuité sans LLM réalisé au moins une fois par trimestre, avec mesure d'écart de productivité et de qualité.*

Épilogue — De *checkmate* à *teammate*... à *petmate* ?

Nicky Case concluait son essai par une formule imparable : ensemble, humains et IA peuvent passer de *checkmate* à *teammate*.

En 2018, c'était un programme optimiste, presque tendre. En 2026, il reste vrai — mais il lui manque une ligne. Parce qu'entre *teammate* et son petit cousin régressif *petmate*, la distance est de six mois de mauvais réflexes. Il y a huit ans, Case craignait que l'IA ne transforme les humains en animaux de compagnie. C'était une crainte métaphysique lointaine. Aujourd'hui, c'est un risque de produit.

Le centaure de 2026 n'est pas une catastrophe. C'est une discipline.

C'est refuser, à chaque usage, la facilité qui vous amputerait.

C'est savoir qui possède votre cheval, et sur quel terrain vous êtes autorisé à galoper.

C'est accepter que le « + » ne se donne pas : il se gagne, quotidiennement, contre la gravité de l'outil.

Case parlait de symbiose — deux espèces qui réussissent ensemble non **malgré** mais **grâce à** leurs différences. La formule tient toujours. Il suffit d'y ajouter, en 2026, une clause souveraine, une clause neurologique, et un plan B opérationnel.

Le centaure n'est pas mort. Il est simplement devenu plus exigeant à entretenir.

Comme nous, donc.

Cet essai est une relecture en dialogue avec l'article fondateur de Nicky Case, « How To Become A Centaur », Journal of Design and Science (MIT Media Lab / MIT Press), janvier 2018, disponible sous licence Creative Commons à l'adresse <https://jods.mitpress.mit.edu/pub/issue3-case>.

Références mobilisées

- Case, N. *How To Become A Centaur*, JoDS Issue 3, janvier 2018.
- Kos'myna N., Hauptmann E., Yuan Y.T., Situ J., Liao X.-H., Beresnitzky A.V., Braunstein I., Maes P. *Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task*, arXiv:2506.08872, juin 2025 (v2 décembre 2025), MIT Media Lab.
- Cowen, T. *Average is Over*, Dutton, 2013.
- Cowen, T. « What are humans still good for? The turning point in Freestyle chess may be approaching », *Marginal Revolution*, novembre 2013.
- Cowen, T. « Centaur chess is now run by computers », *Marginal Revolution*, février 2024.
- ISO/IEC 42001:2023 — *Artificial intelligence management system*.
- Règlement (UE) 2024/1689 établissant des règles harmonisées concernant l'intelligence artificielle (*EU AI Act*).

- OWASP Foundation, *Top 10 for Large Language Model Applications*.
- MITRE ATLAS, *Adversarial Threat Landscape for AI Systems*.
- Licklider, J.C.R. *Man-Computer Symbiosis*, 1960 ; Engelbart, D. *Mother of All Demos*, 1968 (mobilisés par Case, rappelés ici pour mémoire).