

entreprise

GPT-5.4 contre Claude Opus 4.6 et Sonnet 4.6 : ce que chaque famille fait vraiment mieux, et le critère qui pèse plus lourd que les benchmarks pour équiper ton équipe.

BLYX-AGENCY.COM · 3 AVRIL 2026 · PAR MELVIN BRETON

ChatGPT garde l'avantage sur la polyvalence et la prise en main : présentations, tableurs, réponses rapides, c'est le produit le plus installé du marché. Claude prend l'avantage dès que le travail devient long et exigeant : dossiers volumineux, rédaction structurée, code, tâches menées de bout en bout.

Le critère qui décide vraiment n'est pourtant ni le classement ni les benchmarks : c'est l'usage précis que tu équipes, et ce que le contrat dit de tes données. Comparatif vérifié le 30/08/2026.

1. Les modèles comparés et leur date de sortie

Chaque éditeur met en avant un modèle vitrine et une gamme derrière. Le duel se joue sur trois modèles, listés ici avec leur positionnement affiché (vérifié le 30/08/2026).



GPT-5.4	OpenAI	5 mars 2026	Le plus capable et efficace pour le travail professionnel
Claude Opus 4.6	Anthropic	5 février 2026	Le haut de gamme : code, raisonnement long, gros dossiers
Claude Sonnet 4.6	Anthropic	17 février 2026	La performance de pointe à un coût d'usage plus léger

Un point de méthode avant d'aller plus loin : un comparatif de modèles se périme en quelques semaines, parce que les deux éditeurs publient de nouvelles versions à un rythme soutenu. C'est pourquoi chaque affirmation liée à une version porte ici une date de vérification. Les numéros changent vite, la grille de lecture reste : ce que chaque famille fait mieux, et le critère qui départage vraiment.

2. Ce que ChatGPT fait mieux

ChatGPT reste le produit le plus mûr et le plus répandu. C'est l'outil que tes collaborateurs connaissent déjà, souvent parce qu'ils l'utilisent chez eux, et cette familiarité pèse beaucoup dans une adoption d'équipe.

Sur le fond, OpenAI positionne GPT-5.4 comme son modèle le plus performant pour le travail professionnel (vérifié le 30/08/2026). Ses points forts revendiqués :

- ✓ les livrables de bureau : présentations, tableurs, documents ;
- ✓ la polyvalence : bon presque partout, du courrier client à l'analyse rapide ;



La force de ChatGPT tient en une phrase : il est bon presque partout et utilisable immédiatement par tout le monde. Pour une équipe qui découvre l'IA, cet avantage est réel et il ne faut pas le sous-estimer.

3. Ce que Claude fait mieux

Claude, édité par Anthropic, prend l'avantage dès que la tâche devient longue, structurée ou technique.

Opus 4.6, le modèle haut de gamme, est mis en avant sur quatre terrains (vérifié le 30/08/2026) :

- ✓ le code : plus fiable, y compris dans de gros projets existants, en relecture comme en correction ;
- ✓ la planification : il tient un plan sur des tâches en plusieurs étapes sans se perdre en route ;
- ✓ les dossiers volumineux : il raisonne sur des contextes très longs, jusqu'à un million de tokens en version bêta, soit un très gros volume de documents traité d'un coup ;
- ✓ l'endurance : il mène des tâches déléguées de bout en bout, sur la durée.

Sonnet 4.6 est peut-être le modèle le plus stratégique du lot : il rapproche la performance du haut de gamme d'un coût d'usage qui permet de le mettre dans toutes les mains. Ce sont souvent ces modèles intermédiaires qui font basculer un marché.



Pour une PME, la traduction est simple. Rédaction exigeante, analyse

 **Blyx** de cent pages, développement d'un outil interne, tâche  confiée à l'IA du début à la fin : c'est le terrain de Claude.

4. Le tableau de décision par usage

Ton usage dominant	L'outil qui prend l'avantage
Présentations, tableurs, documents de bureau	ChatGPT
Questions rapides du quotidien, polyvalence	ChatGPT
Prise en main immédiate par une équipe peu formée	ChatGPT
Rédaction longue et structurée	Claude
Analyse de dossiers volumineux	Claude
Code et développement d'outils internes	Claude
Tâches déléguées de bout en bout	Claude

Ce tableau tranche le duel de marque, pas la question voisine : quel outil pour quel métier de ton entreprise. Cette question a son propre comparatif, qui inclut Copilot : [Claude, ChatGPT ou Copilot selon le métier](#).

5. Les benchmarks et leurs angles morts



OpenAI appuie le positionnement de GPT-5.4 sur plusieurs scores

Blyx voici tels que communiqués par l'éditeur (vérifiés le 30/08/2026).



Benchmark	Score GPT-5.4	Ce que ça mesure
GDPval	83,0 %	Tâches de travail intellectuel sur 44 métiers
OSWorld-Verified	75,0 %	Pilotage d'un ordinateur, navigation dans un système
BrowseComp	82,7 %	Navigation web complexe
Humanity's Last Exam (avec outils)	52,1 %	Raisonnement de très haut niveau
SWE-Bench Pro (Public)	57,7 %	Résolution de bugs réels dans du code

Quatre angles morts avant d'en tirer une décision :

- ✓ ce sont des chiffres d'éditeur : OpenAI mesure et publie ses propres scores, sans contre-expertise indépendante dans cette communication ;
- ✓ Anthropic, sur cette génération, communique surtout en qualitatif (fiabilité du code, endurance, planification), sans batterie de scores équivalente ;
- ✓ un benchmark mesure des tâches types, pas ton métier : un score de 83,0 % sur 44 métiers ne dit rien de ta manière de faire des devis ;





Conclusion pratique : les benchmarks situent les familles, ils ne départagent presque jamais une décision d'équipement. Deux semaines d'essai sur tes propres cas t'apprendront plus que tous les classements.

6. Les deux écosystèmes autour des modèles

Le modèle n'est qu'une partie de la décision. Autour, chaque éditeur construit un environnement de travail complet, et c'est souvent lui qui fait la différence à l'usage.

Côté OpenAI, l'avantage est la maturité : ChatGPT est un produit poli, diffusé massivement, avec des outils intégrés et une distribution grand public et professionnelle sans équivalent. GPT-5.4 est disponible dans ChatGPT, par API et dans Codex, l'outil de développement maison (vérifié le 30/08/2026).

Côté Anthropic, l'avantage est l'accélération. Les briques principales de l'écosystème Claude, vérifiées le 30/08/2026 :

Brique	Ce que c'est
Claude Code	L'outil de développement en agent : terminal, éditeur, Slack, web
Cowork	Claude qui travaille à côté de toi sur des tâches en plusieurs étapes



MCP	Un standard ouvert pour brancher Claude sur tes outils et tes données
Skills	Des capacités sur mesure ajoutées à Claude, propres à ton entreprise
Intégrations	Slack, Excel, PowerPoint, Chrome, application de bureau

S'ajoute un avantage de perception : la position d'Anthropic sur la sûreté et l'éthique lui donne une image de sérieux qui compte dans les secteurs sensibles.

La lecture qui domine encore le marché, un écosystème mûr chez OpenAI face à un bon modèle isolé chez Anthropic, est périmée. Le sujet n'est plus de savoir si Claude a un écosystème, mais à quelle vitesse cet écosystème monte, et cette vitesse est élevée.

7. Le critère oublié : tes données

Un comparatif de performance répond à la mauvaise question si tu comptes mettre l'outil entre les mains de toute une équipe. Avant le classement, trois questions contractuelles :

- ✓ tes échanges servent-ils à entraîner les futurs modèles de l'éditeur, et peux-tu le désactiver ;
- ✓ où tes données sont-elles hébergées, et sous quel droit ;





Ces trois questions pèsent plus lourd que dix points de benchmark, parce qu'un devis, un contrat ou un fichier client collé dans une conversation engage ton entreprise. Le volet formules est détaillé dans [ChatGPT gratuit ou payant en entreprise](#), et si ton activité exige un hébergement maîtrisé, le sujet complet est traité dans [le guide de l'IA souveraine](#).

Si cette question des données est décisive chez toi, les [trois scénarios pour sortir du CLOUD Act](#) la traitent en profondeur.

8. Combiner les deux sans se disperser

Beaucoup d'entreprises finissent avec les deux outils, et ce n'est pas un problème en soi. Ça le devient quand personne n'a décidé qui utilise quoi, et que l'entreprise paie deux abonnements pour le même usage.

Trois règles suffisent :

- ✓ un outil principal par personne, choisi selon son travail dominant, pas selon la préférence du moment ;
- ✓ un essai en conditions réelles avant tout engagement : deux semaines sur tes propres documents, tes propres courriers, ton propre code ;
- ✓ une réévaluation à date fixe, une ou deux fois par an, parce que le duel bouge vite.



Et avant de conclure qu'il te faut la formule supérieure, vérifie d'abord la façon dont l'outil est utilisé. Un abonnement Claude bien piloté couvre  **Blyx** beaucoup plus qu'on ne le croit : les réflexes sont dans [faire durer ton quota Claude](#).

La répartition complète, rôle par rôle et au-delà de ces deux familles, est le sujet de [comment choisir ton IA](#).

9. Ce que ça change pour ta PME

Le duel ChatGPT contre Claude est réellement relancé : OpenAI garde la maturité et la diffusion, Anthropic monte vite sur la qualité des modèles et l'environnement de travail. C'est une bonne nouvelle pour toi, la concurrence tire les deux familles vers le haut.

Mais la décision d'équipement ne se prend pas dans ce sens. On ne choisit pas un modèle dans l'absolu pour ensuite lui chercher des usages : on part des tâches réelles de l'entreprise, on qualifie celles qui valent un équipement, puis on choisit l'outil qui les sert le mieux, avec l'équipe formée pour s'en servir. Ce travail de qualification, c'est exactement ce que produit un [audit IA](#) : l'inventaire de tes usages, leur priorisation, et l'outil adapté à chacun.

Retiens la hiérarchie : l'usage d'abord, le contrat sur les données ensuite, le classement des modèles en dernier. C'est la seule grille qui restera vraie à la prochaine version.

Questions fréquentes

[ChatGPT ou Claude pour une PME qui démarre avec l'IA](#)



ChatGPT est le point d'entrée le plus simple : prise en main immédiate,

 **Blyx** et une familiarité que ton équipe a souvent déjà. Claude 

devient le bon premier choix si tes usages dominants sont longs et structurés, comme l'analyse de dossiers, la rédaction exigeante ou le code. Dans les deux cas, la vraie première étape est l'inventaire de tes usages, pas le choix de l'outil.

Claude est-il vraiment meilleur que ChatGPT pour le code

Sur la génération comparée ici, vérifiée le 30/08/2026, Claude fait référence sur le code, en particulier dans les projets existants de grande taille, et son écosystème avec Claude Code est construit pour ça. GPT-5.4 reste sérieux sur ce terrain, avec 57,7 % affichés sur SWE-Bench Pro selon OpenAI. Le plus fiable reste de tester les deux sur ton propre projet.

Les benchmarks suffisent-ils pour choisir entre ChatGPT et Claude

Non. Les scores disponibles sont publiés par les éditeurs eux-mêmes, mesurent des tâches types et se périment à chaque nouvelle version. Ils situent les deux familles l'une par rapport à l'autre, mais ils ne remplacent ni un essai sur tes cas réels ni la lecture du contrat sur tes données.

Faut-il payer les deux abonnements en même temps

Rarement au départ. La règle la plus saine est un outil principal par personne, choisi selon son travail dominant. Les deux abonnements se justifient quand deux usages distincts le demandent réellement, p



Combien de temps ce comparatif reste-t-il valable

Pour les chiffres, quelques semaines : les deux éditeurs publient de nouvelles versions à un rythme soutenu, c'est pourquoi chaque affirmation datée porte ici la mention vérifié le 30/08/2026. La grille de décision, elle, tient dans la durée : ChatGPT pour la polyvalence installée, Claude pour le travail long et le code, et le contrat sur tes données comme critère final.

