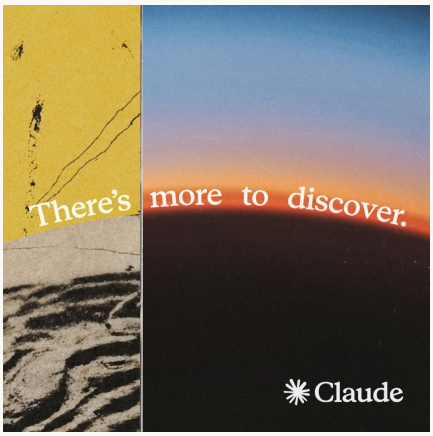
Claude BY ANTHROP\C

Opus 5.5 : un cran plus haut pour l'entreprise



Claude Opus 5.5 est le modèle Opus le plus performant d'Anthropic : il égale ou dépasse Fable 5.1 en code pour moins de la moitié du prix. Meilleur en code, en travail intellectuel et sur les tâches longues, Opus 5.5 marque un net progrès sur Opus 5. Il communique mieux : il rend compte de ce qu'il a fait, de ce qu'il a trouvé et de ce dont il a besoin.

Performance de pointe, coût réduit

À 4 $/20 $ par million de tokens (contre 5 $/25 $ pour Opus 5), Opus 5.5 fait plus avec moins de tokens, et ce nouveau tarif reverse ces gains aux clients. Des prix par token plus bas et des lectures de cache bien moins chères (0,20 $ par million de tokens, contre 0,50 $ pour Opus 5) s'ajoutent à ces gains d'efficacité. Facturé au token, il **coûtera 40 % de moins qu'Opus 5,** et comme Opus 5.5 consomme moins de tokens pour un meilleur travail, il est plus de 30 % plus rapide.

Agents

Orchestre avec fiabilité des tâches multi-outils complexes, planifie avec soin et fait avancer les tâches longues sous supervision minimale.

Code

Gère les tâches longues dans de grandes bases de code. Il trouve la cause racine avant de modifier quoi que ce soit, vérifie son travail en continu et explique ses changements simplement.

Flux de travail d'entreprise

Conçu comme le pilier de l'entreprise : il anime des agents qui mènent des projets de bout en bout, suit les consignes à la lettre et produit tableurs, présentations et documents de niveau professionnel.

Utilisation de l'ordinateur

Apporte vision haute résolution et raisonnement poussé à l'utilisation de l'ordinateur : tâches multi-étapes et multi-applications exigeant planification et jugement.

Analyse financière

Raisonne avec justesse sur rapports, modèles et graphiques, garde le contexte sur toute une opération ou un cycle de reporting, et résume clairement ses constats et sa démarche.

Vision

Lit avec une grande fidélité documents denses, graphiques, captures d'écran et schémas : fiable pour l'extraction de documents, l'analyse visuelle et l'interprétation d'images réelles complexes.

Quel modèle choisir

Modèle	Usages
Claude Fable 5.1	Problèmes de pointe : agents sur tâches longues, nouveaux usages, code et raisonnement ardus, autonomie totale
Claude Opus 5.5	Votre modèle quotidien du travail complexe. Entreprise et raisonnement poussé (juridique, analyse financière, travail intellectuel complexe), code complexe et en production
Claude Sonnet 5	Charges généralistes à grande échelle. Une intelligence capable et efficace pour le travail courant
Claude Haiku 4.5	Tâches fréquentes, sensibles à la latence ; sous-agents en orchestration

Selon les premiers utilisateurs, Claude Opus 5.5 égale ou dépasse Opus 5 avec une fraction des tokens, rédige plus clairement et plus concisément, avec peu de retouches, et mène seul de longues tâches sans surveillance, en vérifiant son travail en cours de route.

Nouveautés lancées avec Opus 5.5

Avec Opus 5.5, nous lançons de nouvelles fonctions et améliorations qui rendent Claude plus efficace sur les tâches longues :

Nouvelles options de compaction (bêta)

L'API peut désormais résumer les anciens échanges en gardant les récents intacts, en arrière-plan, sans mettre l'agent en pause.

Communication améliorée

Opus 5.5 communique nettement mieux ses avancées, mises à jour et résultats aux utilisateurs que les modèles Opus récents.

Claude Opus 5.5

	Opus 5.5	Fable 5.1	Opus 5	GPT-6 Astra	GPT-5.6 Sol
Agentic coding Terminal-Bench 4.0[1]	**66.4%**	55.8%	52.3%	57.9%	37.3%
Agentic coding FrontierCode v1.1 (Main)	**54.4%**	50.3%	48.0%	53.3%	47.5%
Agentic coding CursorBench 4.0	**57.8%**	51.8%	46.6%	—	41.7%
Knowledge work GDPval-AA v2.1	**1846**	1735	1708	1542	1588
Business workflows AutomationBench[2]	40.0%	31.4%	26.9%	**41.4%**	28.8%
Multidisciplinary reasoning Humanity's Last Exam	**67.7%** with tools	65.6% with tools	63.6% with tools	57.2% with tools	—
Agentic scientific research Terminal-Bench-Science 0.1[3]	58.7%	52.6%	29.0%	**64.6%**	22.4%
Computer use OSWorld 2.0	**81.8%** partial	80.7% partial	74.0% partial	—	—
Visual chart recognition Chartography	**89.0%** with tools	88.4% with tools	83.4% with tools	—	—

Unless otherwise noted, all Claude Opus 5.5 results use adaptive thinking at max effort. Terminal-Bench 4.0: Claude Opus 5.5 at xhigh effort and GPT-6 Astra as reported by OpenAI at high effort. Claude Opus 5.5 was evaluated with its production safeguards enabled. When they intervened, cybersecurity tasks were completed by Claude Opus 4.8, and biology and frontier LLM development tasks by Claude Opus 5. This likely reduces Claude Opus 5.5's performance on these benchmarks.

[1] **Terminal-Bench 4.0:** The standard error is ±2.6 pts for Claude Opus 5.5 and ±1.6–2 pts for the other Claude models. The public leaderboard (5 trials/task, Claude Code harness) reports Claude Opus 5 at 51.8%; our setup reproduces it at 52.3%, within noise. GPT-6 Astra and GPT-5.6 Sol figures are as reported by OpenAI.

[2] **AutomationBench:** AutomationBench results were run and reported by Zapier. Claude Opus 5.5 results come from Zapier's own evaluation during early access. Results for Opus 5, GPT-5.6 Sol, and GPT-6 Astra come from Zapier's public leaderboard.

[3] **Terminal-Bench-Science 0.1:** The standard error is ±3.5–5 pts per model. The public leaderboard (3 trials/task, Claude Code harness) reports Claude Opus 5 at 30.0%; our setup reproduces it at 29.0%, within noise. The GPT-6 Astra figure is as reported by OpenAI.