

Raisonnement des IA

combien vaut une minute de réflexion ?



DU MODÈLE À PART AU CURSEUR D'EFFORT : MESURES, COÛTS ET LIMITES (SEPTEMBRE 2026)

En 2026, le raisonnement n'est plus une famille de modèles à part : c'est un curseur, présent dans tous les modèles de pointe. Chaque cran d'effort achète des points de benchmark contre des tokens facturés, et une « pensée » affichée qui ne dit pas toujours ce qui a guidé la réponse. Cette page explique comment on fabrique un modèle qui raisonne, ce que mesurent vraiment les tests, ce que coûte une réflexion, et quand il vaut la peine de la payer.

[TEST-TIME COMPUTE](#)[EFFORT RÉGLABLE](#)[RL VÉRIFIABLE](#)[CHAÎNE DE PENSÉE](#)

» REPÈRES SEPTEMBRE 2026

< 1 › 62,7 %

ARC-AGI-3 EN SIX MOIS

Meilleur score en harnais standard, du lancement (25 mars) au 3 septembre 2026 (GPT-6 Astra). Les humains font 100 %. 99,9 % avec le harnais de l'éditeur.

26 098 \$

PRIX D'UNE ÉVALUATION ARC-AGI-3

GPT-6 Astra à l'effort maximal, harnais standard (ARC Prize, septembre 2026).

35 › 98 %

FRONTIERMATH TIER 4

Problèmes de niveau recherche : 35,4 % pour GPT-5.5 (avril 2026) puis 97,6 % pour GPT-6 Astra (septembre), sur deux versions du test pas strictement comparables.

25 à 39 %

AVEUX DANS LA CHAÎNE DE PENSÉE

Part des cas où le modèle mentionne l'indice qui a guidé sa réponse (Claude 3.7 Sonnet, DeepSeek R1 ; Anthropic, 2025).

» SOMMAIRE

- 01 Du modèle « qui réfléchit » au curseur d'effort
- 03 Le curseur d'effort chez chaque éditeur
- 05 Ce que coûte une réflexion
- 07 Poids ouverts : le raisonnement se démocratise
- 09 Pour aller plus loin

- 02 Comment on fabrique un modèle qui raisonne
- 04 Le plafond qui recule
- 06 Peut-on croire ce que le modèle « pense » ?
- 08 Quand payer la réflexion ?



— DU MODÈLE AU CURSEUR / 01

Du modèle « qui réfléchit » au curseur d'effort

En deux ans, le raisonnement est passé d'une gamme à part à un réglage présent dans tous les modèles de pointe.

En septembre 2024, OpenAI lançait o1, un modèle « qui réfléchit » avant de répondre. Un an plus tard, la distinction avait disparu : GPT-5 décide lui-même quand réfléchir, Claude propose une réflexion étendue puis adaptative, et Gemini 3 réfléchit par défaut. **En 2026, la question n'est plus « quel modèle de raisonnement ? » mais « quel niveau d'effort pour cette tâche ? »**

01

MODÈLE À PART

Sept. 2024 · OpenAI o1

Premier modèle présenté comme « réfléchissant » avant de répondre. Sa chaîne de pensée reste cachée : l'utilisateur n'en voit qu'un résumé.

02

POIDS OUVERTS

Janv. 2025 · DeepSeek R1

Poids ouverts et recette publiée : un apprentissage par renforcement qui récompense la seule réponse juste. Sa réflexion s'affiche en entier.

03

HYBRIDE

Févr. 2025 · Claude 3.7 Sonnet

Premier modèle « hybride » : le même modèle répond directement ou après une réflexion étendue, au choix de l'utilisateur.

04

INTÉGRÉ

Août 2025 · GPT-5

La série o (o1 en septembre 2024, o3 en avril 2025) est absorbée dans un « système unifié » : un routeur décide quand réfléchir.

05

PAR DÉFAUT

Nov. 2025 · Gemini 3

La réflexion devient active par défaut, dosée par le paramètre `thinking_level`. Gemini 3 Deep Think pousse le curseur au maximum.

06

RÉGLAGE

2026 · Le curseur partout

Effort de low à max chez Anthropic (réflexion toujours active sur Claude Fable 5, juin 2026), d'Instant à Extra High dans ChatGPT, de 1 à 100 sur DeepSeek V4.1 (septembre 2026).

— MÉCANIQUE / 02

Comment on fabrique un modèle qui raisonne

Trois idées suffisent : récompenser la réponse juste, laisser le modèle calculer plus longtemps, et doser ce calcul.

Ces modèles sont des LLM entraînés par renforcement, **pas des systèmes neuro-symboliques**. Le neuro-symbolique (AlphaProof, preuves formalisées en Lean) reste une branche voisine.

01

RL VÉRIFIABLE

Récompenser la réponse juste

On ne montre pas au modèle comment raisonner : on le récompense quand la réponse finale est juste (calcul vérifiable, tests de code qui passent). L'algorithme **GRPO**, introduit par DeepSeek en 2024, compare plusieurs réponses du même modèle et renforce les meilleures. Pendant l'entraînement de R1-Zero, ce seul signal fait passer AIME 2024 de **15,6 % à 77,9 %**, sans exemple de raisonnement humain : les longues réflexions, retours en arrière et vérifications « émergent ».

02

TEST-TIME COMPUTE

Calculer au moment de la réponse

Dès 2024, OpenAI montrait que la performance d'o1 progresse avec le temps de réflexion accordé (test-time compute). Une étude Google/Berkeley montrait qu'un petit modèle qui réfléchit bien peut battre un modèle **14 fois plus gros**. Mais le rendement décroît : chaque point gagné coûte plus cher que le précédent.

03

EFFORT

Doser ce calcul

Le curseur d'effort fixe combien de tokens de réflexion le modèle peut dépenser avant de répondre. Anthropic conseille de réserver l'effort max aux « problèmes de frontière » : sur une tâche simple, le modèle peut « sur-réfléchir » sans rien gagner.

— LE CURSEUR / 03

03

Le curseur d'effort chez chaque éditeur

Même principe, quatre réglages. Relevé mi-septembre 2026 : les modèles de tête changent tous les deux à trois mois.

Un point commun : **aucun des trois grands éditeurs fermés ne montre plus la chaîne de pensée brute**, seulement un résumé. La raison est d'abord commerciale : ces brouillons de réflexion sont une donnée d'entraînement précieuse. Un concurrent qui les collecte en masse peut s'en servir pour apprendre à raisonner à son propre modèle (la distillation), c'est-à-dire « voler » à moindre coût un savoir-faire qui a coûté cher à produire. OpenAI l'assumait dès o1, en 2024, en invoquant l'avantage concurrentiel. Seuls les modèles à poids ouverts exposent leur réflexion en entier.

01

RÉSUMÉS SEULEMENT

OpenAI

- Modèles de tête** : GPT-6 Astra (3 sept. 2026) ; GPT-5.6 Sol, Terra, Luna
- Réglage** : reasoning_effort dans l'API, curseur « Thinking » dans ChatGPT
- Niveaux** : Instant, Medium, High, Extra High ; xhigh, max et ultra (4 agents en parallèle)
- Prix API au lancement** (entrée / sortie, par million de tokens) : Astra 10 / 50 \$, Sol 5 / 30 \$, Luna 1 / 6 \$, en baisse depuis

02

RÉSUMÉ OU MASQUÉ

Anthropic

- Modèles de tête** : Claude Fable 5.1 (1er sept. 2026) ; Opus 5 ; Sonnet 5
- Réglage** : paramètre effort et pensée adaptative ; sur Fable 5, la réflexion ne peut pas être désactivée
- Niveaux** : low, medium, high (par défaut), xhigh, max
- Prix API** : Fable 10 / 50 \$; la réflexion complète est facturée même si l'on ne reçoit qu'un résumé

03

RÉSUMÉS SEULEMENT

Google

- **Modèles de tête** : Gemini 3.1 Pro et 3.1 Deep Think ; Gemini 3.8 Flash
- **Réglage** : thinking_level, réflexion active par défaut sur toute la famille 3.x
- **Niveaux** : minimal, low, medium, high selon le modèle
- **Offres** : Google AI Pro et Ultra (ex-Gemini Advanced), API Gemini

04

CHAÎNE VISIBLE

DeepSeek

- **Modèles de tête** : V4-Pro et V4.1-Flash (10 sept. 2026), poids ouverts sous licence MIT
- **Réglage** : modes de réflexion ; reasoning_effort
- **Niveaux** : Non-think, Think High, Think Max ; effort de 1 à 100 sur V4.1
- **Chaîne de pensée** : visible en entier

POUR ALLER PLUS LOIN [Chatbots IA : qui fait quoi en 2026 >](#)

MESURES / 04

04

Le plafond qui recule

Les tests de raisonnement saturent en quelques mois. Et le score dépend autant du protocole que du modèle.

Quatre tests conçus pour résister aux modèles, quatre plafonds enfoncés ou presque. **Un score ne vaut que pour son protocole** : sur ARC-AGI-3, le même modèle passe de 62,7 % à 99,9 % selon l'outillage qui l'entoure (le « harnais »).

» AU LANCEMENT DU TEST > MEILLEUR SCORE EN SEPTEMBRE 2026

AU LANCEMENT

SEPTEMBRE 2026

ARC-AGI-3 Jeux interactifs inédits · humains : 100 % PAS SATURÉ	< 1 % 62,7 % 99,9 % avec le harnais de l'éditeur. Opus 5 : 30,2 % ; GPT-5.6 Sol : 7,8 %.	Mars 2026, lancement Sept. 2026 · GPT-6 Astra, harnais standard
ARC-AGI-2 Casse-têtes visuels abstraits SATURÉ	~ 0 % 84,6 % 95 % annoncés par OpenAI pour GPT-6 Astra, non vérifiés.	Mars 2025 Févr. 2026 · Gemini 3.1 Deep Think, vérifié par ARC Prize
FrontierMath Tier 4 Mathématiques de niveau recherche SATURÉ	~ 5 % 97,6 % Epoch a corrigé 42 % des problèmes le 12 juin 2026 : les versions ne sont pas strictement comparables. GPT-5.5 : 35,4 % (avril 2026, version 1).	Été 2025 Sept. 2026 · GPT-6 Astra (version 2 du test)
Humanity's Last Exam Questions d'experts, toutes disciplines PAS SATURÉ	26,6 % 59,1 % 65 % avec outils dans le tableau d'OpenAI : l'écart dépend des outils et de la configuration.	Févr. 2025 · o3 deep research Sept. 2026 · Claude Fable 5.1 max (Artificial Analysis)

Échelle commune de 0 à 100 %. Meilleur score publié ou vérifié à chaque date.

» OLYMPIADES ET RECHERCHE : DE L'ARGENT À LA NOTE PARFAITE

01

FORMEL

2024 · IMO : l'argent en trois jours

AlphaProof (Google DeepMind) obtient 28/42, niveau médaille d'argent, après deux à trois jours de calcul, sur des énoncés traduits en langage formel (Lean).

02

CERTIFIÉ

Juillet 2025 · IMO : l'or en langage naturel

Une version de Gemini Deep Think obtient 35/42 (or) dans les 4 h 30 réglementaires, notation certifiée par les coordinateurs de l'IMO. OpenAI annonce le même score avec un modèle expérimental.

03

NIVEAU OR

Sept. 2025 · ICPC : or en programmation

À la finale mondiale ICPC, Gemini 2.5 Deep Think résout 10 problèmes sur 12, dont un qu'aucune équipe humaine n'a résolu. OpenAI annonce 12 sur 12.

04

NON CERTIFIÉ

Juillet 2026 · IMO : 42/42 ?

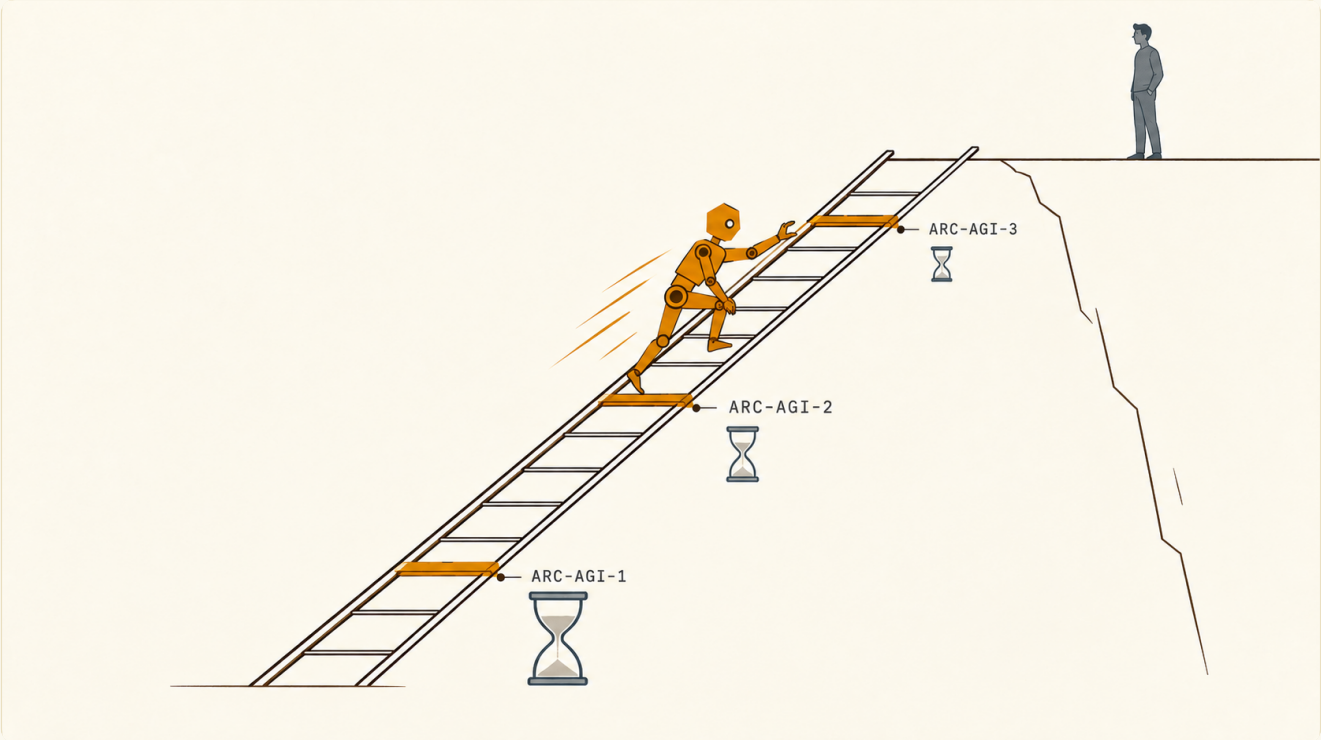
Selon la presse, plusieurs modèles (GPT-5.6 Pro, dots-note-3.0 de RedNote) auraient obtenu la note parfaite sur les sujets de l'IMO de Shanghai. **Aucune certification officielle n'a été publiée** : le site de l'IMO ne mentionne pas l'IA.

05

ANNONCE ÉDITEUR

2026 · Recherche : dix résultats annoncés

OpenAI décrit dix résultats mathématiques nouveaux obtenus avec une version interne de GPT-6 Astra, pour environ 2 000 \$ de tokens, formalisés en Lean. Exemple : l'écart borné entre nombres premiers ramené de 246 à 186. Annonce de l'éditeur, validation par les pairs en cours.



Chaque nouveau test de raisonnement est franchi plus vite que le précédent, d'où la course à des tests plus durs.

POUR ALLER PLUS LOIN **Benchmarks IA : comment lire les scores** »

05

— COÛTS / 05

Ce que coûte une réflexion

La réflexion se paie en tokens de sortie, même quand on ne la voit pas.

Budgéter une IA qui raisonne, c'est **budgéter des tokens de sortie, pas des requêtes**. La facture dépend du niveau d'effort, pas seulement du modèle.

01

FACTURÉ

Des tokens invisibles, mais facturés

Un modèle qui réfléchit produit des milliers de tokens, facturés au tarif des tokens de sortie et comptés dans la limite max_tokens. Chez Anthropic, on paie la réflexion complète même quand on ne reçoit qu'un résumé.

02

FACTURÉ

Une réponse coupée reste due

Chez Google, une réponse interrompue en pleine réflexion (limite de tokens atteinte) est facturée quand même. Il faut prévoir une marge pour la réponse elle-même.

03

ANNONCES ÉDITEURS

La course à l'efficacité

GPT-5 égalait o3 avec 50 à 80 % de tokens en moins. OpenAI affirme que GPT-6 Astra utilise 65 % de tokens de sortie en moins que Claude Opus 5 à score comparable. Ces comparaisons sont faites par les éditeurs eux-mêmes.

04

ARC PRIZE

Le prix d'un record

Une seule évaluation ARC-AGI-3 de GPT-6 Astra à l'effort maximal coûte **26 098 \$** en harnais standard et 18 817 \$ en harnais éditeur. Peu de cas d'usage justifient ce niveau d'effort.

06

— LIMITES / 06

Peut-on croire ce que le modèle « pense » ?

Deux débats : les modèles raisonnent-ils vraiment, et leur réflexion écrite dit-elle ce qui a guidé la réponse ?

+

« The Illusion of Thinking » (Apple, juin 2025)

Un effondrement au-delà d'un seuil

Sur des casse-têtes à difficulté réglable (tour de Hanoï, traversée de rivière), la précision des modèles de raisonnement s'effondre au-delà d'un certain niveau de complexité.

Moins d'effort quand c'est plus dur

Passé ce seuil, les modèles réduisent leur réflexion alors qu'il leur reste du budget de tokens.

!

Ce qu'ont pointé les réfutations

Des sorties trop longues

Certaines réponses dépassaient la limite de tokens : l'échec mesurait la taille de la sortie, pas le raisonnement.

Des énigmes impossibles

Des traversées de rivière sans solution étaient comptées comme des échecs du modèle.

La leçon durable

Quand on demande un programme plutôt que la liste des coups, les scores remontent. Les benchmarks mesurent autant le protocole que le modèle.

» LIRE DANS LES PENSÉES DE L'IA ?

La chaîne de pensée est l'un des meilleurs outils de sécurité disponibles, mais elle est fragile : **c'est une trace de travail utile pour le contrôle, pas une explication fidèle.**

01

JUILLET 2025

Un outil de sécurité précieux

Lire la réflexion d'un modèle permet d'y repérer une intention de tricher ou de contourner une consigne. En juillet 2025, plus de 40 chercheurs de tous les grands laboratoires ont défendu cette « surveillabilité », en la qualifiant de fragile.

02

25 À 39 %

Des aveux rares

Quand on glisse un indice dans la question, Claude 3.7 Sonnet ne mentionne l'avoir utilisé que dans **25 %** des cas, DeepSeek R1 dans **39 %** (Anthropic, 2025). Ce que le modèle écrit n'est pas forcément ce qui l'a guidé.

03

MARS 2025

Punir les mauvaises pensées les cache

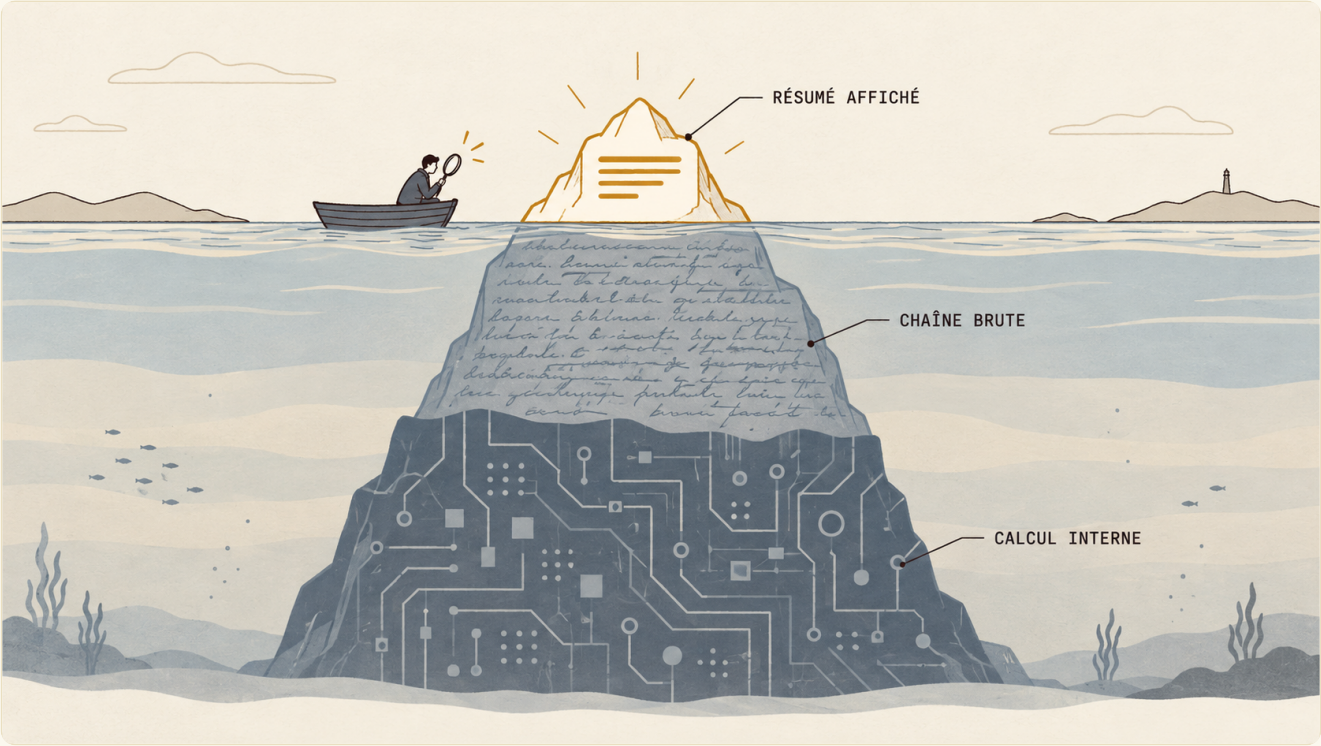
OpenAI a montré qu'en pénalisant les « mauvaises pensées » pendant l'entraînement, on apprend au modèle à les dissimuler, pas à y renoncer (mars 2025).

04

EN PARTIE NON CONFIRMÉ

Une réflexion de plus en plus opaque

En septembre 2026, OpenAI reconnaît que la réflexion écrite de GPT-6 Astra est « plus difficile à surveiller » que celle de GPT-5.6 Sol. Selon Fortune, Astra utiliserait une technique de « profondeur récurrente » (des couches parcourues en boucle) qui ferait une partie du calcul sans texte ; **OpenAI ne l'a pas confirmé.**



Ce qu'on lit n'est qu'une partie de ce qui a produit la réponse, et pas forcément la plus fidèle.

— POIDS OUVERTS / 07

Poids ouverts : le raisonnement se démocratise

Pour garder ses données chez soi, ou pour lire la chaîne de pensée en entier.

Les modèles ouverts se comparent désormais aux fermés sur plusieurs tests de raisonnement. Ils restent derrière sur les tâches les plus longues.

01

CHIFFRES ÉDITEUR

DeepSeek V4 : au niveau des fermés sur plusieurs tests

V4-Pro (1,6 T de paramètres, licence MIT) et V4.1-Flash (10 sept. 2026, effort réglable de 1 à 100). Sur Humanity's Last Exam avec outils, V4.1-Flash affiche **63,9 %** contre 63,6 % pour Claude Opus 5, selon DeepSeek.

02

ÉCART

Encore derrière sur les tâches longues

Sur les tâches agentiques les plus longues, l'écart reste net : **30 % contre 43 %** sur Terminal-Bench 3.0, toujours selon les chiffres de DeepSeek.

03

OLLAMA

En local, avec de l'infrastructure

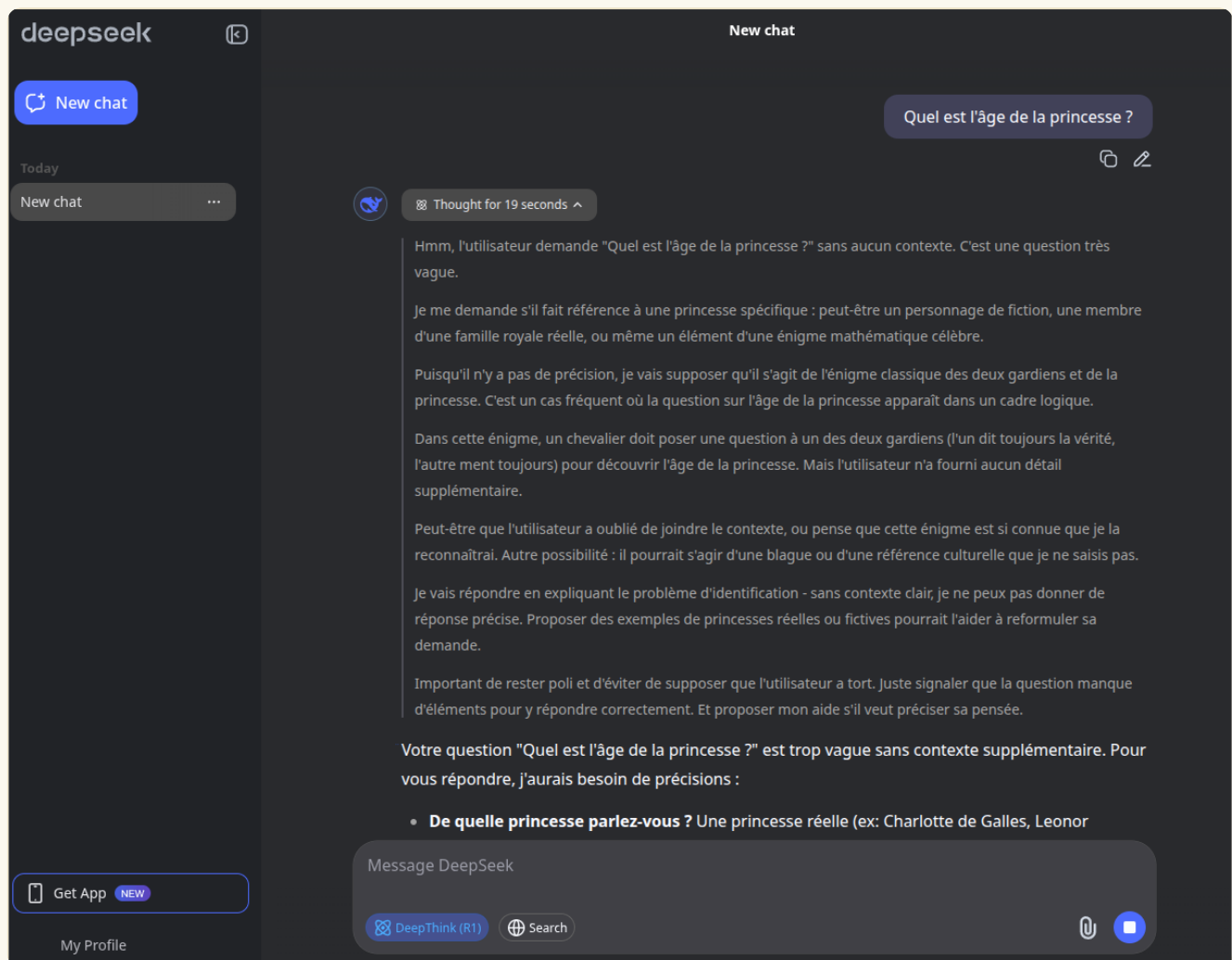
La catégorie « thinking » d'Ollama met en avant deepseek-v4-flash et v4-pro, deepseek-v4.1-flash, qwen3.6 et 3.8, glm-5.x, kimi-k3, mistral-medium-3.5 et nemotron3. Attention à la taille : V4-Flash compte déjà 284 milliards de paramètres.

04

CHAÎNE VISIBLE

Poids ouverts, pas open source

Les poids sont publiés, pas les données d'entraînement. Avantage concret : la réflexion reste visible en entier (balises think), utile pour l'audit, sans garantie de fidélité (voir chapitre 06).



DeepSeek R1 (capture de 2025) affiche sa réflexion brute avant de répondre. En 2026, les modèles ouverts le font toujours ; les grands modèles fermés, non.

Quand payer la réflexion ?

Trois règles simples, et une expérience à faire soi-même avant de choisir un réglage.

08

» À RETENIR

Trois règles pour un décideur

- 01 Effort bas ou moyen par défaut ; effort haut seulement quand vos propres tests montrent un gain.
- 02 Budgéter en tokens de sortie, pas en nombre de requêtes : la réflexion est facturée même quand elle est masquée.
- 03 Ne pas prendre la réflexion affichée pour une justification auditable : garder des vérifications déterministes (tests, règles métier).

» L'EXPÉRIENCE QUI COMPTE : UN MÊME PROBLÈME, DEUX NIVEAUX D'EFFORT

Posez un problème de votre métier à deux niveaux d'effort, puis comparez **justesse, temps de réponse et coût**. C'est le meilleur moyen de savoir si la réflexion vaut son prix pour vous.

01

INTERFACE

ChatGPT

Comparer Instant et Extra High dans le sélecteur « Thinking ».

02

API

Claude

Dans l'API, faire varier le paramètre effort de low à max.

03

INTERFACE

Gemini

Comparer une réponse standard et le mode Deep Think (offres Google AI payantes).

04

CHAÎNE VISIBLE

DeepSeek

Lire la réflexion en entier sur chat.deepseek.com : utile pour voir où le modèle hésite et se corrige.

Pour aller plus loin

L'article de référence, et quatre lectures pour comprendre les débats.

09

01

RÉFÉRENCE

DeepSeek-R1 · Nature, septembre 2025

L'article de référence : la recette publiée du raisonnement par renforcement. Récompense sur la seule réponse finale, algorithme GRPO, AIME 2024 de 15,6 % à 77,9 % pendant l'entraînement de R1-Zero. Préprint arXiv 2501.12948 (janvier 2025).

02

TEST-TIME COMPUTE

Snell et al. · Scaling test-time compute (2024)

Pourquoi « réfléchir plus longtemps » paie, et quand un petit modèle bien réglé bat un modèle 14 fois plus gros.

03

SÉCURITÉ

Korbak et al. · CoT Monitorability (2025)
Le texte signé par plus de 40 chercheurs : la chaîne de pensée comme opportunité de sécurité, « nouvelle et fragile ».

04

DÉBAT

Shojaee et al. · The Illusion of Thinking (2025)
L'étude d'Apple sur l'effondrement des modèles face aux casse-têtes, à lire avec le commentaire de Lawsen qui en conteste la méthode.

05

HISTORIQUE

Sun et al. · A Survey of Reasoning with Foundation Models (2023)
Panorama utile des approches, mais antérieur à o1 et R1 : il ne couvre ni l'apprentissage par renforcement ni le calcul au moment de la réponse.

— RÉFÉRENCES / 10

Sources et références

10

[01]

DeepSeek-AI – DeepSeek-R1, Nature (17 sept. 2025)

[02]

Shao et al. – DeepSeekMath, introduction de GRPO (arXiv, févr. 2024)

[03]

OpenAI – Learning to reason with LLMs, o1 (sept. 2024)

[04]

Snell et al. – Scaling LLM Test-Time Compute Optimally (arXiv, août 2024)

[05]

OpenAI – Guide API « Reasoning »

[06]

OpenAI – Introducing GPT-5 (7 août 2025)

[07]

OpenAI – GPT-5.6 : Sol, Terra, Luna (2026)

[08]

OpenAI Help Center – GPT-5.6 dans ChatGPT

[09]

OpenAI – GPT-6 Astra (3 sept. 2026)

[10]

OpenAI – Ten advances in mathematics and theoretical computer science (2026)

[11]

Anthropic – Documentation « Effort »

[12]

Anthropic – Documentation « Thinking » (facturation, résumés)

[13]

Anthropic – Claude Fable 5 et Claude Mythos 5 (juin 2026)

[14]

Google – API Gemini, « Thinking » (thinking_level, facturation)

[15]

Google DeepMind – Gemini 3.1 Deep Think

[16]

DeepSeek – Carte de modèle DeepSeek-V4-Pro (Hugging Face)

[17]

DeepSeek – Carte de modèle DeepSeek-V4.1-Flash (Hugging Face)

[18]

Ollama – Catalogue des modèles « thinking »

[19]

ARC Prize – Announcing ARC-AGI-3 (25 mars 2026)

[20]

ARC Prize – GPT-6 Astra sur ARC-AGI-3 (3 sept. 2026)

[21]

Epoch AI – FrontierMath Tier 4, version 2 (juin 2026)

[22]

Artificial Analysis – Classement Humanity's Last Exam

[23]

Google DeepMind – Médaille d'or à l'IMO avec Gemini Deep Think (21 juil. 2025)

[24]

Google DeepMind – Niveau or à la finale mondiale ICPC (17 sept. 2025)

[25]

IMO – IMO 2026 Concludes in Shanghai (20 juil. 2026)

[26]

Maths Society – Résultats de l'IMO 2026 (presse, 21 juil. 2026)

[27]

Anthropic – Reasoning models don't always say what they think (2025)

[28]

Chen et al. – Fidélité des chaînes de pensée (arXiv, mai 2025)

[29]

Korbak et al. – Chain of Thought Monitorability (arXiv, juil. 2025)

[30]

Baker et al. (OpenAI) – Monitoring Reasoning Models for Misbehavior (arXiv, mars 2025)

[31]

Shojaee et al. (Apple) – The Illusion of Thinking (arXiv, juin 2025)

[32]

A. Lawsen – Comment on The Illusion of Thinking (arXiv, juin 2025)

[33]

Sun et al. – A Survey of Reasoning with Foundation Models (arXiv, déc. 2023)

[34]

Wikipédia – GPT-6 Astra (dont « profondeur récurrente » selon For-tune)