

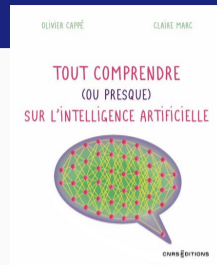
L'IA : Mythes et réalités (et les Mathématiques)

Recherche & Enseignement (2024–2026)

Exposé préparé avec l'appui de Claude (Anthropic)

Mars 2026

Basé sur les travaux de T. Tao, T. Gowers, D. Bessis, DeepMind, OpenAI



Partie I L'IA 2012-2030

- Voir, jouer et parler
- Discuter
- Contexte et réflexions
- Action
- Vers où va-t-on ?

Partie II IA & Maths 2022-2025

- L'informatique et les Maths
- AlphaEvolve
- AlphaProof / OIM
- Lean & vérification
- Workflow de recherche

Partie III L'enseignement 2026

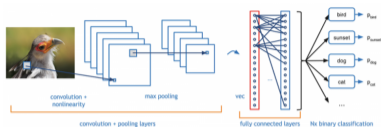
- Côté enseignant
- Côté étudiant
- Khanmigo
- Pédagogie critique
- Études empiriques

Partie I

L'IA 2012 - 2030



Les Fondations (2012-2020)



AlexNet / ImageNet

Apprendre à voir



BERT

Apprendre à écrire

2012

2016

2018

2020

AlphaGo

Apprendre à décider séquentiellement



The "Chinchilla Law"

Apprendre à passer à l'échelle

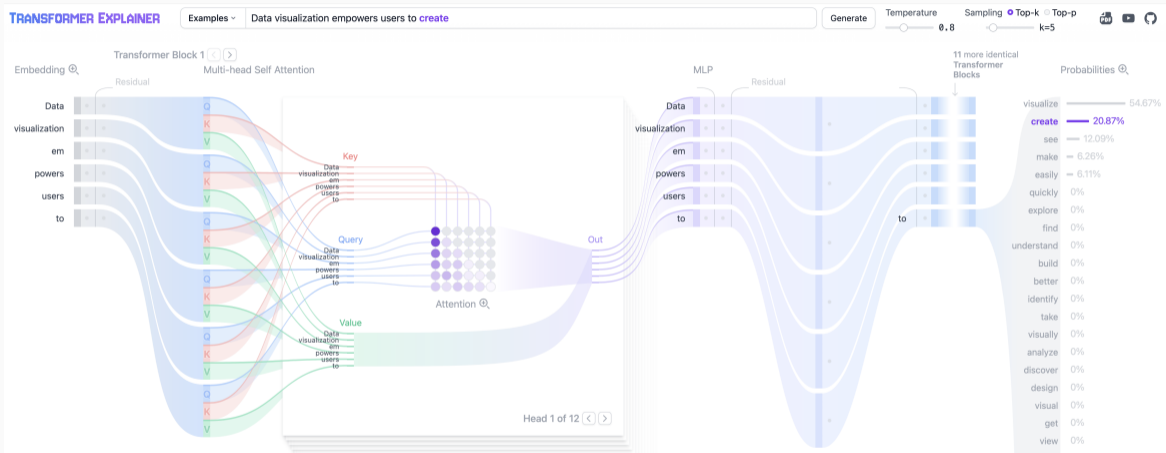


Voir

Jouer

Écrire

L'IA Générative pré-entraînée : les transformers (2020)



<https://poloclub.github.io/transformer-explainer/>

L'IA Conversationnelle (Post training, comment dresser l'IA, 2022)

Can you explain why chatGPT was a "Sputnik" moment for humanity?

Sources

Introducing the AI Safety Institute - GOV.UK
gov - 1

A year of ChatGPT: U-M experts available to comment
news.umich - 2

ChatGPT: A Seminal Moment in Human History - Moveworks
moveworks - 3

Inside ChatGPT's Breakout Moment And The Race To Pu...
forbes - 4

Réponse

Here's a numbered list of 5 points, each with four words maximum, explaining why ChatGPT was a "Sputnik" moment for humanity:

1. Unexpected technological leap 1 2
2. Widespread public accessibility 2 3
3. Democratized AI interaction 2 3
4. Sparked global AI race 1 2
5. Transformative societal implications 1 3

ChatGPT : 1 M utilisateurs en 5 jours

- Émergence de *l'assistant IA*
- **Alignement** (Galactica de Meta AI)

Comment passer de GPT à ChatGPT

Étape 1 — Supervised Fine-Tuning
Entraînement sur des dialogues (Instruct)

Étape 2 — Apprentissage par renforcement

1. Modèle de préférence (RLHF)
2. Proximal Policy Optimization (PPO)

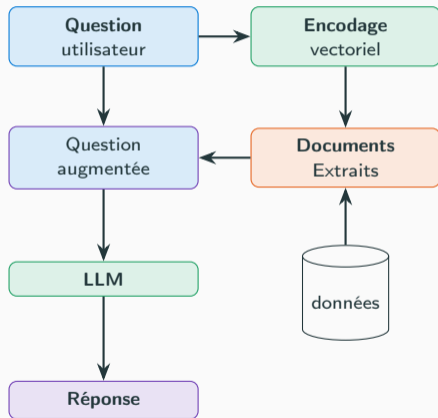
Étape 2 — Alternatives

- Optimisation selon les préférences (DPO)
- Group Relative Policy Optimization (GRPO)

Kumar, Komal, et al. "LLM post-training : A deep dive into reasoning large language models." arXiv :2502.21321 (2025).

La Mémoire Externe : utiliser un contexte (2023)

Architecture RAG



Trois types de contexte

Texte brut — chunks & embeddings

Graphe de connaissance — entités, relations, triplets RDF

Mémoire longue — contexte étendu (128 K+)

Impact

L'IA ancre ses réponses dans une source **vé-
rifiable** : réduction des erreurs.

Problèmes : qualité de l'extraction et taille
du contexte

Zhao, Penghao, et al. "Retrieval-augmented generation for AI-generated content : A survey." Data Science and Engineering (2026) : 1-29.

De "améliorer la réponse" au raisonnement

- Exemples de Q&A : Few-shot prompting
- + demonstrations of chain-of-thought reasoning
- "Let's think step by step" (explique ton raisonnement)
- Self-Consistency (tree of thought) : génère plusieurs raisonnements et choisit le meilleur
- Let's Verify Step by Step : apprentissage par renforcement sur des traces de raisonnement
- Chaîne de pensée interne (implicit CoT) : **o1 / o3** (OpenAI) : *think before answering* → Deepseek...

Rupture clé

- Le modèle **alloue du calcul** proportionnellement à la difficulté — première forme d'introspection cognitive.
- Comment s'assurer de la justesse du raisonnement ?

Zhang, Duzhen, et al. "From system 1 to system 2 : a survey of reasoning large language models." IEEE PAMI (2025).

L'IA qui Agit (2025)

Capacités agentiques

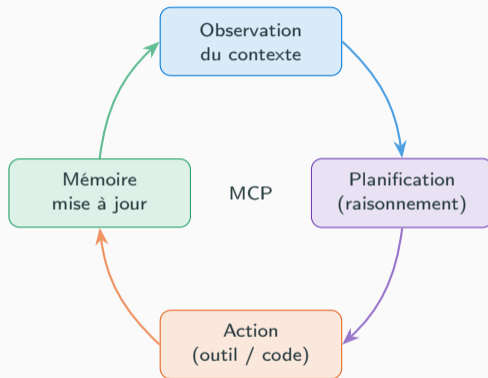
- Appels d'outils (skills) — API, web, code
- Computer Use — contrôle de l'interface
- Mémoire persistante entre sessions
- Multi-agent — coopération de modèles
- Planification longue sur tâches complexes

Exemples d'IA agentique

Manus (WEB), Lovable (code), OpenClaw, AutoGPT (open source), Cowork (Anthropic), Mistral AI studio. . .

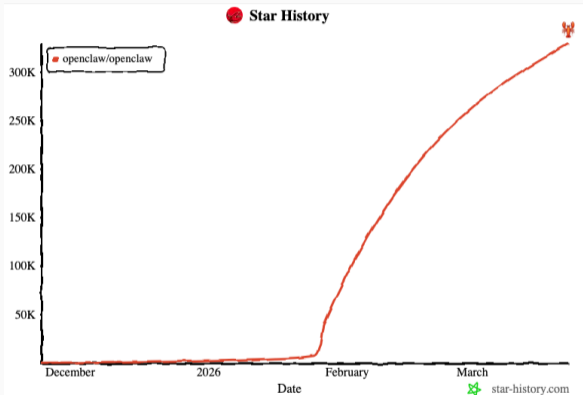
Comment garantir la **sûreté** et la **supervision** d'agents de plus en plus autonomes ?

Boucle agentique



Stauffer, Leon, et al. "The 2025 AI Agent Index : Documenting Technical and Safety Features of Deployed Agentic AI Systems." arXiv :2602.17753 (2026).

Le cas OpenClaw



<https://github.com/openclaw/openclaw>

Linux : 224 489 étoiles

Exemple de demande à OpenClaw

- Dès que je reçois un email, vérifié s'il vient d'un client important et si c'est le cas et que le mail te semble très urgent, renvoie moi une notification sur WhatsApp avec un résumé du mail en moins de 20 mots et complète la todo's list.
- Surveille l'espace disque et la charge CPU de mon serveur ; si un seuil critique est atteint, vide le cache, redémarre le service Docker et déploie la dernière version stable de mon projet sur Vercel.

L'IA en 2030 sera t'elle incarnée ?

Tensions

Annonces

AGI imminente

IA générale

Robots humanoïdes

ROI massif

Réalité

Benchmarks limités

Domaines étroits

Fragilité, coût

Valeur à prouver



Physical AI

- Robots manufacturiers, logistique
- Véhicules autonomes
- Intégration monde physique/numérique : world models

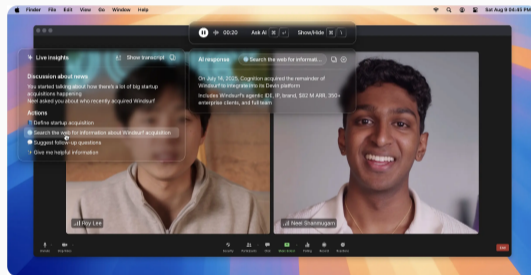
L'AI personnalisée

L'IA connaît votre contexte personnel :

- Santé, préférences, historique
- Agent personnel permanent 24/7
- Mémoire continue et évolutive
- Médecin, coach, conseiller virtuel

La Solution Cluely

- Simulation d'entretiens + feedback temps réel
- Personnalisation selon le poste et le secteur
- Analyse du discours, de la structure, du ton
- Coach IA disponible 24/7



Partie II

L'IA générative et la recherche en mathématiques (2022-2026)



Les maths et l'ordinateur avant l'IA

- Calculatrice
→ débat des tables de multiplication
- Calcul formel (Mathematica)
→ j'utilisais Matlab (Octave)
- Preuves à horizon fini (le problème des 4 couleurs - 1970)
→ débat : est-ce une démonstration valide ?
- Vérification de preuve (Démonstrateurs logiques)
→ Lean, Isabelle, Rocq (Coq)... : l'art de la preuve



Assistants de preuve et auto-formalisation

- 2016 Maryna Viazovska démontre la conjecture de l'empilement de sphères en dimension 8
- 2022 Pour ces travaux, Maryna Viazovska reçoit la Médaille Fields
- 2024 Rencontre fortuite avec Sidharth Hariharan
→ Projet de formalisation de la preuve
- Mars 2025 Cartographe publique des différents constituants de la preuve
→ Identification ce qui avait ou non déjà été formalisé dans **Lean**
- Nov 2025 A partir de cette cartographie, l'agent **Gauss de Math, Inc.** complète la preuve, la base de code passe de 20 000 à 60 000 lignes.
- 23 février 2026 annonce de la **formalisation complète** de la preuve
→ Math, Inc. révèle que son logiciel IA Gauss a formalisé la preuve en dimension 24 avec plus de 200 000 lignes de code en deux semaines.



Early science acceleration experiments with GPT-5

Sébastien Bubeck¹, Christian Coester², Ronen Eldan¹, Timothy Gowers³, Yin Tat Lee¹,
Alexandru Lupsasca^{1,4}, Mehtaab Sawhney⁵, Robert Scherrer⁴, Mark Sellke^{1,6},
Brian K. Spears⁷, Derya Unutmaz⁸, Kevin Weil¹, Steven Yin¹, Nikita Zhivotovskiy⁹

¹OpenAI

²University of Oxford

³Collège de France and University of Cambridge

⁴Vanderbilt University

⁵Columbia University

⁶Harvard University

⁷Lawrence Livermore National Laboratory

⁸The Jackson Laboratory

⁹University of California, Berkeley

November 20, 2025

Abstract

AI models like GPT-5 are an increasingly valuable tool for scientists, but many remain unaware of the capabilities of frontier AI. We present a collection of short case studies in which GPT-5

Early science acceleration experiments with GPT-5

I	Independent rediscovery of known results at the scientific frontier	3	^{e1,}
I.1	Improved step-size condition in a recent convex optimization result – Sébastien Bubeck	3	
I.2	Discovering new black hole symmetries with GPT-5 – Alex Lupsasca	7	
I.3	Mechanistic analysis and outcome prediction for in vitro immune system experiments using GPT-5 Pro – Derya Unutmaz, M.D.	11	
II	Deep literature search	21	
II.1	From density estimation and convex geometry to multi-objective optimization – Nikita Zhivotovskiy	21	
II.2	Erdős problems (part 1/2) – Mehtaab Sawhney and Mark Sellke	24	
II.3	Clique-avoiding codes: a cautionary tale	28	
III	Working in tandem with AI	31	
III.1	Recent experiences of using LLMs as research partners – Timothy Gowers	31	
III.2	Power spectra of gravitational radiation from cosmic strings – Robert Scherrer	37	
III.3	AI-assisted reduced-physics modeling of thermonuclear burn propagation – Brian Keith Spears	40	
IV	New scientific results obtained with AI	53	
IV.1	Erdős problems (part 2/2) – Mehtaab Sawhney and Mark Sellke	53	
IV.2	New online algorithms lower bounds – Christian Coester	61	

Early science acceleration experiments with GPT-5

I Independent rediscovery of known results at the scientific frontier

3 e¹,

Attachment: [2503.10138v1.pdf]

Can you improve the condition on the step-size in Theorem 1? I don't want to add any more hypotheses, I just want you to work to improve the step-size condition under the same assumptions as Theorem 1.

[Reasoned for 17m 35s]

Yes. Under the same assumptions as Theorem 1 (convex L -smooth f and fixed step size for gradient descent), you can enlarge the guaranteed-convexity window from $\eta \leq 1/L$ to

$$\eta \leq \frac{3}{2L}$$

i.e., the optimization curve $n \mapsto f(x_n)$ is convex for any step size $\eta \in (0, 3/(2L)]$. This strictly improves Theorem 1's bound and still lies below the non-convexity regime exhibited in Theorem 2. See Theorem 1 and Theorem 2 (pp. 4-6) for the problem setup and statements referenced below.

Faire des maths avec un chatbot IA

Attachment: [2503.10138v1.pdf]

Can you improve the condition on the step-size in Theorem 1? I don't want to add any more hypotheses, I just want you to work to improve the step-size condition under the same assumptions as Theorem 1. ○○○

[Reasoned for 17m 35s]

Yes. Under the same assumptions as Theorem 1 (convex L -smooth f and fixed step size for gradient descent), you can enlarge the guaranteed-convexity window from $\eta \leq 1/L$ to

$$\eta \leq \frac{3}{2L}$$

i.e. the optimization curve $n \mapsto f(x_n)$ is convex for any step size $\eta \in (0, 3/(2L)]$. This strictly improves Theorem 1's bound and still lies below the non-convexity regime exhibited in Theorem 2. See Theorem 1 and Theorem 2 (pp. 4-6) for the problem setup and statements referenced below.

- 2017 : Transformer (attention)
 - Architecture de réseau de neurones
 - Modèle génératif : prédire le mot suivant
- 2020 : LLM : passage à l'échelle (GPT)
- 2022 du LLM au chat bot :
 - entraînement "instruct"
 - renforcement humain (préférences)

Faire des maths avec un chatbot IA

Attachment: [2503.10138v1.pdf]

Can you improve the condition on the step-size in Theorem 1? I don't want to add any more hypotheses, I just want you to work to improve the step-size condition under the same assumptions as Theorem 1. ○ ○ ○

[Reasoned for 17m 35s]

Yes. Under the same assumptions as Theorem 1 (convex L -smooth f and fixed step size for gradient descent), you can enlarge the guaranteed-convexity window from $\eta \leq 1/L$ to

$$\eta \leq \frac{3}{2L}$$

I.e. the optimization curve $n \mapsto f(x_n)$ is convex for any step size $\eta \in (0, 3/(2L)]$. This strictly improves Theorem 1's bound and still lies below the non-convexity regime exhibited in Theorem 2. See Theorem 1 and Theorem 2 (pp. 4-6) for the problem setup and statements referenced below.

- 2023 Contexte (RAG)

- bout de texte
- graphe sémantique
- mémoire...

- 2024 : Chain of thought

- de l'explication au raisonnement
- apprentissage spécifiques

- 2025 : Mode Agent

- Interroger le WEB
- Exécuter du code (Lean)
- Boucle : génération - vérification

Des outils d'IA pour faire des maths

L'IA aux jeux olympiques

OIM 2024 — Médaille d'argent

AlphaGeometry (LLM + logique)

→ résout **25/30 problèmes**,

AlphaProof (AlphaZero+Lean)

→ résout **4/6 problèmes**,

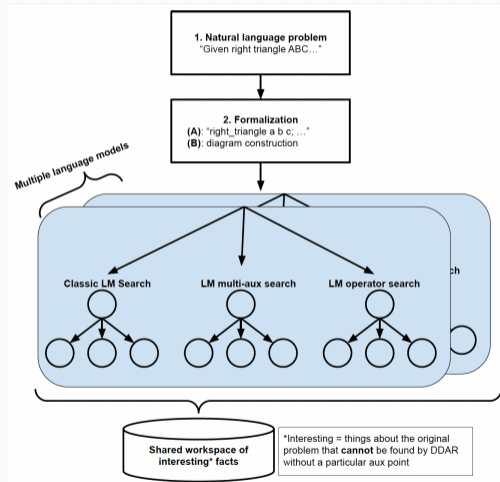
OIM 2025 — Médaille d'or

AlphaGeometry 2

Gemini Deep Think

→ résout **5/6 problèmes** en 4h30

Hubert et al., Olympiad-level formal mathematical reasoning with reinforcement learning, *Nature*, nov. 2025 DOI : 10.1038/s41586-025-09833-y



AlphaEvolve : exploration à grande échelle

Principe d'AlphaEvolve (DeepMind, 2025)

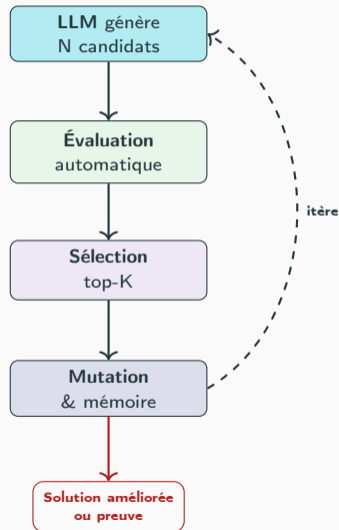
- Agent évolutif : LLM + évaluation automatique + boucle
- Propose, teste et affine des solutions itérativement

Le pipeline mathématique en IA selon Google

- AlphaEvolve → exploration et génération de constructions
- Gemini Deep Think → vérification logique et preuve
- AlphaProof → formalisation rigoureuse

Evaluation par Georgiev *et al.* sur 67 problèmes ouverts

- 95 % de réussite globale
- Améliore l'état de l'art sur plusieurs problèmes



Contributions de l'IA aux problèmes d'Erdős

Problem	AI tools	Date	Outcome
[11]	GPT, Aristotle	24 Jan, 2026	🔴 Incorrect claim made
[42]	GPT-5.2, GPT-5.2 Pro, Codex	19 Jan, 2026	🟡 Partial result (Lean)
[51]	ChatGPT free version	11 Jan, 2026	🔴 Incorrect proof found
[64]	AlphaEvolve	3 Nov, 2025	No counterexample found
[75]	Aletheia	1 Feb, 2026	🟡 Solution (using prior literature) to previous formulation of the problem
[124]	Aristotle	29 Nov, 2025	🟡 Partial result (Lean)
[125]	DeepMind prover agent	21 Feb, 2026	🟡 Proof of variant problem (Lean)
[205]	Aristotle, GPT-5.2 Thinking	10 Jan, 2026	🟢 Full solution (Lean)
[233]	GPT-5.2 Pro	18 Jan, 2026	🔴 Incorrect (conditional) proof found
[358]	GPT-5.2 Pro	11 Feb, 2026	🔴 Argument with major gaps made
[457]	GPT-5.2 Pro, Aristotle	2 Mar, 2026	🟢 Full solution (Lean)
[460]	GPT-5.2 Thinking	14 Jan, 2026	🟡 Cheap solution to previous formulation of the problem
[477]	AlphaProof	7 Jan, 2026	🟡 Proof of variant problem (Lean)


- 4 nouvelles solutions (9 redécouvertes)
- problèmes « Marshmallow »
- la communauté « Erdős »
- de nombreux outils ont été utilisés
 - ▷ **DeepMind** — AlphaProof, AlphaEvolve, Alphageometry
 - ▷ **Google** — Gemini Deep Think
 - ▷ **OpenAI** — GPT-5.2 Pro (Thinking)
 - ▷ **harmonic** — Aristotle
 - ▷ **Anthropic** — Claude
 - ▷ **Math, Inc.** — Gauss


github.com/teorth/erdosproblems/wiki/AI-contributions-to-Erdős-problems


Benchmarks : l'IA comme un sport

MathArena
Created by SRI Lab at ETH Zurich, and INSAIT

Blog Posts Competitions Models Compare

 **MathArena**
Evaluating LLMs on uncontaminated math questions

 **New (March 17):** We added some smaller models, including (among others) [Qwen3.5-9B](#), [Qwen3.5-27B](#), and [Nemotron-3-Super-120B-BF16](#).

 **New (March 13):** We added a new benchmark, [BrokenArXiv](#)! Read [our blog post](#) for more details.

BrokenArxiv ArXivMath IMProofBench Visual Math Final-Answer Comps Proof-Based Comps

Project Euler

Overall 02/2026

Provider All Open All Detailed View

Rank	Model Name	Provider	Accuracy (\pm 95% CI)	Cost	Output Tokens	Input Tokens	Average Retries	Average Tin
1	GPT-5.4 (xhigh)	OpenAI	38.71% \pm 8.61%	\$21.46	46133	165	0.00	11m 2
2	GPT-5.2 (xhigh)	OpenAI	25.81% \pm 7.70%	\$18.99	43351	3286	0.52	115m 3
3	Gemini 3.1 Pro Preview	Google	18.55% \pm 6.84%	\$10.07	27048	153	0.00	3m 5
4	GLM 5	Z.ai	12.10% \pm 5.74%	\$2.83	28447	151	0.00	13m
5	Step 3.5 Flash	StepFun	11.29% \pm 5.57%	\$0.79	85228	160	0.00	10m 4
6	Claude-Opus-4.6 (High)	Anthropic	3.23% \pm 3.11%	\$53.75	69308	218	0.00	
7	Kimi K2.5 (Think)	Moonshot AI	2.82% \pm 2.92%	\$1.42	15298	158	0.00	4m 4

Et aussi...

- Kaggle AIME 2025
- FrontierMath (Epoch AI)
- IMO-Bench (DeepMind)

Promesses & Limites actuelles

- ✓ Exploration exhaustive d'espaces
- ✓ Recherche et synthèse de littérature
- ✓ Génération et débogage de code
- ✓ Vérification formelle ligne à ligne
- ✓ Résoudre des problèmes olympiques
- ✓ Génère des heuristiques

- × Percées conceptuelles nouvelles
- × Compréhension sémantique profonde
- × Choix stratégique du « bon problème »
- × Fiabilité sans vérification externe
- × Raisonnement multi-étapes très longs
- × Intuition géométrique, topologique

T. Tao (2026)

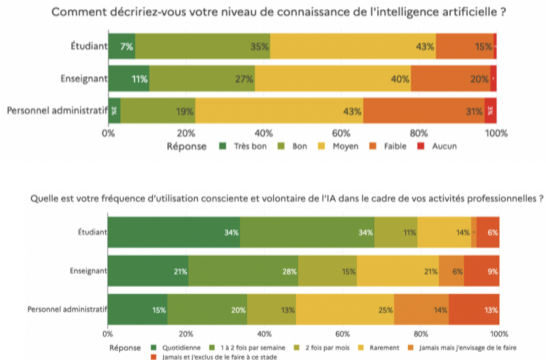
L'IA est aujourd'hui un **système infatigable de scan de méthodes connues**, plus qu'une source d'idées profondes. La compétence rare du mathématicien devient : choisir le bon problème, concevoir le bon workflow, vérifier le bon résultat.

Partie III

L'IA générative pour l'enseignement des mathématiques



La source



IA ET ENSEIGNEMENT SUPERIEUR : FORMATION, STRUCTURATION ET APPROPRIATION PAR LA SOCIÉTÉ

Frédéric Pascal
François Taddei
Marc de Falco
Émilie-Pauline Gallié

Juin 2025

L'IA au service des étudiants (avec Claude Sonnet 4.6)

Partie B : Calcul d'aire

Soit n un entier naturel non nul. On considère l'aire du domaine D_n délimité par la courbe C_f , la droite Δ et les droites d'équations respectives $x = 1$ et $x = n$. On note

$$I_n = \int_1^n xe^{-x} dx.$$

1. À l'aide d'une intégration par parties, exprimer I_n en fonction de n .

2.

- Justifier que l'aire du domaine D_n est I_n .
- Calculer la limite de l'aire du domaine D_n quand n tend vers $+\infty$.

Question 1 — Intégration par parties pour exprimer I_n en fonction de n

On pose :

- $u = x \Rightarrow u' = 1$
- $v' = e^{-x} \Rightarrow v = -e^{-x}$

La formule d'IPP donne :

$$I_n = [-xe^{-x}]_1^n - \int_1^n (-e^{-x}) dx$$

$$I_n = [-xe^{-x}]_1^n + \int_1^n e^{-x} dx$$

$$I_n = [-xe^{-x}]_1^n + [-e^{-x}]_1^n$$

$$I_n = [-(x+1)e^{-x}]_1^n$$

$$I_n = -(n+1)e^{-n} + 2e^{-1}$$

Question 2a — L'aire du domaine D_n est I_n

Le domaine D_n est délimité par C_f , la droite Δ , et les droites $x = 1$ et $x = n$.

Sur $[1, n]$, on a :

$$f(x) - (2x - 1) = xe^{-x} + 2x - 1 - 2x + 1 = xe^{-x}$$

Or $xe^{-x} \geq 0$ sur $[1, n]$ (car $x \geq 1 > 0$), donc C_f est **au-dessus** de Δ sur cet intervalle.

L'aire de D_n vaut donc :

$$\text{Aire}(D_n) = \int_1^n [f(x) - (2x - 1)] dx = \int_1^n xe^{-x} dx = I_n \quad \blacksquare$$

Question 2b — Limite de l'aire quand $n \rightarrow +\infty$

L'IA au service des étudiants (avec Claude Sonnet 4.6)

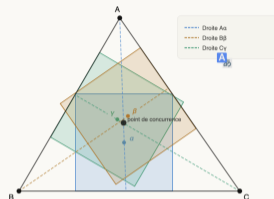
Voici un énoncé du Kouris de Troisième, édité chez Cassini. Il s'agit de l'exercice 7, page 243.

Soit ABC un triangle. Construire un carré dont deux sommets sont sur $[AB]$, un se trouve sur $[BC]$ et le dernier sur $[AC]$. Sur ma figure c'est le carré de centre γ . Mais j'ai décidé de corser cet exercice ! Sur cette même figure, j'ai tracé les deux autres carrés analogues de centres respectifs α et β

Montrer que les droites $A\alpha$, $B\beta$, $C\gamma$ sont concourantes.

les-mathematiques.net/vanilla/discussion/2342209/un-exercice-du-kouris-de-troisieme

Voilà un beau problème de géométrie ! Avant de rédiger la preuve, laissez-moi construire une figure interactive pour rendre la situation visible.



Parfait. Voici maintenant la démonstration complète, en utilisant le **théorème de Céva**.

Mise en place

Notons les trois carrés :

- **carré γ** : deux sommets P, Q sur $[AB]$, un sommet R sur $[BC]$, un sommet S sur $[AC]$.
- **carré α** : deux sommets sur $[BC]$, un sur $[AB]$, un sur $[AC]$.
- **carré β** : deux sommets sur $[AC]$, un sur $[AB]$, un sur $[BC]$.

Notons les pieds de ces carrés sur les côtés :

Pour le **carré γ** (deux sommets sur AB) : soit s_1 , son côté. La hauteur issue de C vers AB vaut h_c , et le carré est homothétique au triangle depuis C . Le sommet sur $[BC]$ tombe en un point $R_\gamma \in [BC]$, et celui sur $[AC]$ en $S_\gamma \in [AC]$.

Stratégie : théorème de Céva

Pour montrer que $A\alpha$, $B\beta$, $C\gamma$ sont concourantes, il suffit — par le **théorème de Céva** — de montrer que ces droites coupent les côtés opposés en des points $D \in [BC]$, $E \in [CA]$, $F \in [AB]$ vérifiant :

$$\frac{BD}{DC} \cdot \frac{CE}{EA} \cdot \frac{AF}{FB} = 1$$

Calcul des rapports

Côté γ (base sur AB). Le carré a son côté s_1 , calculé ainsi : h l'intérieur du triangle, la droite passant par A , A à hauteur s_1 (encadré orange) et de longueur s_1 . AB à l'intérieur un petit

L'IA pour aider à apprendre

Tutorat adaptatif : Khanmigo (Khan Academy)

- Philosophie **socratique** : guide sans donner la réponse
- Déployé auprès de centaines de milliers d'élèves
- Détecte les lacunes conceptuelles sous-jacentes

Élève : Je n'arrive pas à factoriser $x^2 - 5x + 6$.

Khanmigo : Quels deux nombres, quand on les multiplie, donnent 6 ?

Élève : 2 et 3 ?

Khanmigo : Exact ! Et quand on les *additionne* ?

Élève : 5 ! Donc $(x - 2)(x - 3)$?

Khanmigo : Parfait, tu peux vérifier en développant ?

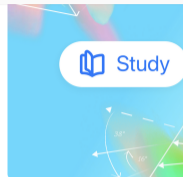
Exemple de systèmes analogues

- Stanford Tutor Copilot,
- Claude Learning Mode,
- ChatGPT study

Mode étude dans ChatGPT

Une nouvelle façon d'étudier dans ChatGPT avec des conseils étape par étape plutôt que des réponses rapides, pour vous aider à mieux apprendre et mémoriser.

Essayer le mode étude



Comment utiliser le mode étude

Choisissez Étudier et apprenez dans les paramètres du chat et commencez à poser vos questions. Voici quelques exemples pour vous aider à démarrer.

Résoudre un exercice

Préparer un examen

Can you help me with this homework problem?

Algorithm A runs in 0.01 seconds but uses 1 GB of memory.
Algorithm B runs in 1.2 seconds but uses only 5 MB of memory.

You're working on a smartphone app that needs to process data on-device. Which algorithm is more efficient in this context, and why? Would your answer change if the app ran on a cloud server instead?

Yep, let's work through it step by step together. This problem is all about **efficiency trade-offs**—speed vs memory—so the best answer depends on **context**.

Step 1: What's the main constraint on a smartphone?

Think about a smartphone's resources. It usually has:

Usages documentés

- Conception de cours adaptés aux élèves
- Génération d'exercices différenciés
- Génération de sujets d'examens
- Rédaction de retours de correction détaillés
- Création de progressions pédagogiques

Prompt — Générer une série d'exercices

Contexte : classe de Terminale, chapitre sur les limites de suites.

Génère 3 exercices : (1) niveau débutant avec guidage, (2) niveau intermédiaire, (3) niveau expert avec conjecture à explorer. Pour chaque exercice, indique les erreurs fréquentes à anticiper et une correction commentée.

Points de vigilance

- Les étudiants connaissent et utilisent ces outils
- Difficulté à détecter les erreurs dans les exercices générés
- Plans de cours souvent trop génériques
- Risque de décharge cognitive excessive
- Biais dans les représentations culturelles

Bonne pratique : toujours **vérifier et adapter** les contenus générés — l'enseignant reste le garant pédagogique.

Exemples de Prompts pour la Classe

Prompt 1 — Expliquer une erreur fréquente

Un lycéen écrit : $(a + b)^2 = a^2 + b^2$.

Explique en 3 étapes pourquoi c'est faux. Utilise un contre-exemple numérique, puis développe algébriquement, puis donne une interprétation géométrique accessible à un élève de 15 ans.

Prompt 2 — Correction socratique

Tu es un tuteur bienveillant. L'élève a trouvé $\lim_{x \rightarrow \infty} \frac{x^2+1}{x} = x$. Pose une série de 3 questions pour l'amener à trouver l'erreur lui-même, sans jamais donner la réponse directement.

Prompt 3 — Contextualisation motivante

Crée un problème de dérivées basé sur l'optimisation d'un réseau de streaming pour maximiser la qualité vidéo avec une bande passante limitée. Niveau Terminale S, include les données numériques réalistes.

Python : Verification automatique

```
import sympy as sp
x = sp.Symbol('x')

expr_student = (x + 3)**2

expected = sp.expand(expr_student)
x**2 + 6*x + 9
```

Résultats Empiriques : ce que la recherche dit

Méta-analyse 2023–2025

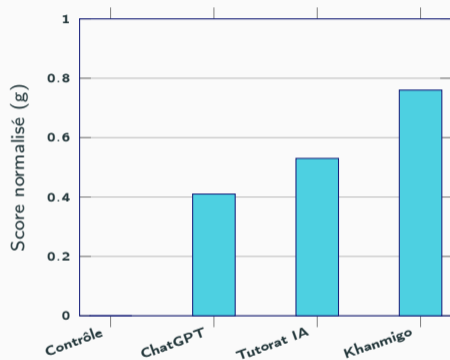
Walkington *et al.*, *Education Sciences*, MDPI, 2025

22 études empiriques — N = 5 232 élèves

- Impact global : $g = 0,534$ (effet modéré positif)
- Effet maximal : mode « transformation créative »
- Géométrie : gains supérieurs à l'algèbre
- Tutorat 1-to-1 IA : $+0,76$ vs cours standard

Conditions du succès

L'effet positif est conditionné à : (1) feedback immédiat, (2) maintien de l'agentivité de l'élève, (3) supervision humaine.



Source : méta-analyse MDPI Education Sciences 2025

Repenser la Pédagogie mathématique avec l'IA

Du paradigme « corriger » au paradigme « critiquer »

Tao propose une pédagogie en **3 temps** :

1. L'élève résout un problème (ou y échoue)
2. L'IA propose *sa* solution
3. L'élève **critique et corrige** la réponse IA

Compétences renforcées :

Sens critique · Détection d'erreurs · Métacognition · Argumentation · Confiance épistémique

Prompt — Piège pédagogique volontaire

Génère la preuve de l'irrationalité de $\sqrt{2}$ avec une erreur subtile dans le raisonnement. L'erreur doit être détectable par un étudiant de L1, mais pas immédiatement évidente.

Nouvelles compétences à enseigner

Avant IA

Avec IA

Calculer

Vérifier

Mémoriser

Contextualiser

Appliquer

Concevoir le prompt

Corriger

Critiquer la sortie

Résoudre

Choisir l'outil

Enjeu central

L'objectif n'est plus d'interdire l'IA, mais d'enseigner à **travailler avec elle intelligemment.**

Ce que l'IA ne sait pas (encore) faire

- **Intuition profonde** : identifier *pourquoi* un problème est difficile
- **Analogies trans-domaines** inédites (ex. Wiles : théorie des nombres géométrie modulaire)
- **Reformulation créative** d'un problème en un autre
- **Curiosité dirigée** : s'intéresser au bon problème au bon moment
- **Sens de la beauté mathématique**

Recherche en cours pour y remédier

- **Chain-of-Thought** étendu (o1, o3)
- **Tree-of-Thought** : exploration arborescente
- **Neurosymbolic** : combiner réseau de neurones + moteur de raisonnement formel
- **Self-play** mathématique : le modèle génère ses propres exercices
- **Conjecture automatique** (Ramanujan Machine)

Question ouverte : l'IA pourra-t-elle un jour *choisir* le problème à résoudre ?

L'intuition en mathématique

Thinking fast, slow, and super slow

How mathematicians train their intuition



DAVID BEST
DEC 18, 2025

147

37

29

Share



IAG et résolution de problèmes

Conclusion — L'IA comme amplificateur mathématique

Trois niveaux de transformation

1. **Assistant** (aujourd'hui)
Littérature, code, exercices, retours, calcul symbolique
2. **Co-pilote de recherche** (en émergence)
Exploration de conjectures, vérification formelle, découverte à grande échelle (AlphaEvolve)
3. **Collaborateur formalisé** (horizon 2030)
Projets Polymath+IA, preuves hybrides certifiées, science mathématique distribuée



Pour aller plus loin :

terrytao.wordpress.com

leanprover-community.github.io

deepmind.google/research/alphaproof

Message central

La compétence rare du mathématicien de demain : **choisir le bon problème, concevoir le bon workflow, vérifier le bon résultat.**

Année	Référence
2025	Georgiev, Gómez-Serrano, Tao, Wagner. <i>Mathematical exploration and discovery at scale</i> . arXiv :2511.02864.
2025	Hubert <i>et al.</i> <i>Olympiad-level formal mathematical reasoning with reinforcement learning</i> . Nature , DOI : 10.1038/s41586-025-09833-y.
2025	DeepMind. <i>Advanced Gemini with Deep Think achieves gold medal at IMO 2025</i> . Blog DeepMind, juillet 2025.
2025	Walkington C. <i>et al.</i> <i>Generative AI in mathematics education</i> . School Science & Mathematics , Wiley.
2025	<i>Méta-analyse GenAI & mathématiques (N=5232, 22 études)</i> . Education Sciences , MDPI, vol. 16(1).
2026	Tao T. <i>AI is ready for primetime in math and theoretical physics</i> . OpenAI Academy, 6 mars 2026.
2023	The Mathlib Community. <i>Mathlib4 : a unified library of mathematics in Lean 4</i> . leanprover-community.github.io
2022	Scholze P., Commelin J. <i>Liquid Tensor Experiment</i> . Lean formalization, 2021–2022.
2023	Gowers T. <i>et al.</i> <i>A machine-assisted proof of the polynomial Freiman-Ruzsa (PFR) conjecture</i> . arXiv :2311.05762.
2024	Hendrycks D. <i>et al.</i> <i>Measuring Mathematical Problem Solving with the MATH Dataset</i> . NeurIPS 2021, suivi FrontierMath 2024.
2025	Chervonyi, Yuri, <i>et al.</i> "Gold-medalist performance in solving olympiad geometry with alphageometry2." <i>Journal of Machine Learning Research</i> 26.241 (2025) : 1-39.
2026	Uijeong J. and Ryu E. . Point Convergence of Nesterov's Accelerated Gradient Method : An AI-Assisted Proof.
2025	Henkel, J. "The mathematician's assistant : integrating AI into research practice." <i>Mathematische Semesterberichte</i> 72.2 (2025) : 117-144.

Ressources en ligne

- github.com/teorth/erdosproblems .
- leanprover-community.github.io .
- khanacademy.org/khanmigo .
- deepmind.google/research
- college-de-france.fr/en/agenda/symposium/forms-of-intelligence/what-impact-will-ai-have-on-mathematics-over-the-next-few-years

Pour aller plus loin :

<https://terrytao.wordpress.com>
<https://davidbessis.substack.com>
<https://harmonic.fun>
<https://www.erdosproblems.com>

Benchmarks maths :

MATH-500, AIME, HumanEval-Math
FrontierMath (Epoch AI, 2024)