

L'IA QUI NE PARLE PAS

Jev, de la jeune entreprise californienne TypeSafe AI, ne répond pas par des phrases mais par des probabilités. Trois semaines après sa sortie, Vercel et Netlify le proposent à leurs clients, OpenAI et Amazon lancent des produits semblables, et l'entreprise négocierait une valorisation de 10 milliards de dollars. Nous avons vérifié ses promesses.

COMMENT ÇA MARCHE

Des nombres au lieu des mots

AU CRIBLE

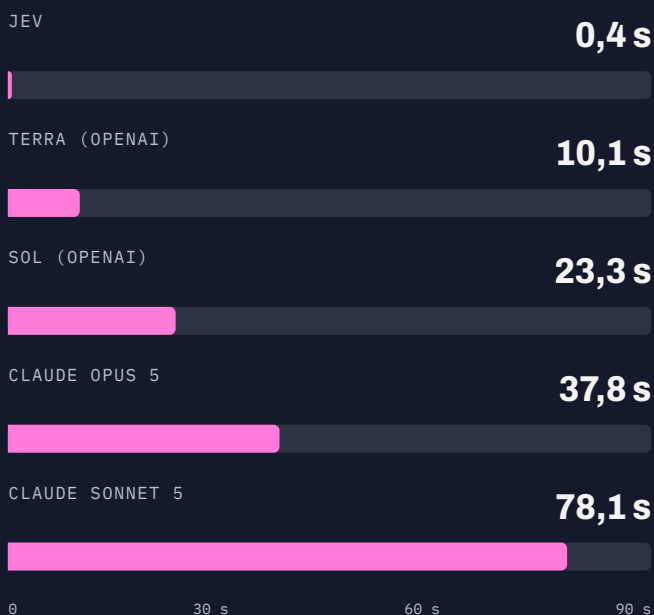
« 444 fois moins cher », vraiment ?

LA SCIENCE

Ce que veut dire une probabilité « calibrée »

LA COURSE

OpenAI, Amazon et un Européen dans la roue



Temps moyen pour traiter un dossier dans quatre tâches d'entreprise (alertes de sécurité, factures, service client, contrôle d'agents). Mesures publiées par TypeSafe, non reproduites par un tiers.

UNE IA QUI NE DIT PAS UN MOT

Une IA qui ne fait pas de phrases ? C'est le pari de Jev, présenté le 15 septembre par la jeune entreprise californienne TypeSafe AI¹. Pas de blabla : il répond avec des nombres, des probabilités, et c'est tout.

En trois semaines, ça s'est emballé : les hébergeurs Netlify et Vercel le proposent déjà^{18, 17}, Amazon a publié un modèle inspiré de Jev³⁶, et l'entreprise discuterait d'une valorisation de plus de 10 milliards de dollars³⁸. Ça fait beaucoup de bruit en très peu de temps.

Alors on a fait le tri entre les promesses et ce qui se vérifie. Et on t'explique comment ça marche, sans jargon.

La rédaction de **SOURCE**

SOURCE, C'EST QUOI ?

- 1 Un sujet, un numéro.**
On prend un sujet dont tout le monde parle, et on le creuse à fond.
- 2 Chaque fait a sa source.**
Les petits numéros renvoient à la liste en fin de numéro : tu peux tout vérifier toi-même.
- 3 Des verdicts clairs.**
Les tampons Vrai, À nuancer, Exagéré ou Faux disent ce qu'on a trouvé, quel que soit le camp.
- 4 On dit ce qu'on ignore.**
Ce qui reste flou est rangé dans « Angles morts ».

DANS CE NUMÉRO

01 Avant de lire

Une IA pour les logiciels, pas pour les humains

02 Comment ça marche

Un état, des questions, des probabilités

03 Chronologie

De 2017 à trois semaines de folie

04 En chiffres

Rapide, bon marché, mais pas le plus juste

05 Au crible

Six phrases vérifiées

06 La science

Calibration, Kahneman et Jevons

07 Face à face

Convaincus et sceptiques

08 La course

Les concurrents en trois semaines

09 Bon à savoir

Ce que Jev ne sait pas faire

10 Lexique

Huit mots pour suivre

11 Angles morts

Ce qu'on ne sait pas encore

12 Méthode et sources

Nos 39 références

UNE IA POUR LES LOGICIELS, PAS POUR LES HUMAINS

ChatGPT, Claude ou Gemini écrivent des phrases pour des humains. Jev reçoit un texte et des questions fermées, et renvoie des nombres : une probabilité, un choix, une note. Pas un mot de plus⁵.

TypeSafe AI l'a présenté le 15 septembre 2026 comme le premier « modèle Système 1 », fait pour décider vite à l'intérieur d'un programme¹. Exemple : trier le message d'un client vers le bon service, en une fraction de seconde et pour presque rien.

1 Rapide

70 à 500 millisecondes par réponse, selon TypeSafe¹.

2 Bon marché

0,042 dollar par million de jetons envoyés ; les réponses sont gratuites⁹.

3 Prévisible et mesuré

Il ne peut répondre qu'une des options prévues, avec des probabilités censées dire à quel point il est sûr de lui^{1, 8}.

QUI EST DERRIÈRE

LE FONDATEUR

Diogo Almeida

- Ex-chercheur d'OpenAI, coauteur principal d'InstructGPT (2022), l'ancêtre de ChatGPT²². Parti il y a deux ans fonder TypeSafe³⁴.

L'ENTREPRISE

TypeSafe AI, San Francisco

- Cofondée avec Erik Gafni et Sasha Sheng³. 40 millions de dollars levés auprès du fonds DCVC, selon la presse³⁸.

UN ÉTAT, DES QUESTIONS, DES PROBABILITÉS

Le programme envoie un « état », c'est-à-dire le texte ou les données à juger, et une liste de questions typées. Jev répond à toutes en même temps.

L'ÉTAT ENVOYÉ À JEV

Au secours ! Mes versements échouent depuis 3 jours.

Jev

des probabilités, pas de phrases

NOUL • OUI OU NON	oui		0,95
is_urgent	La probabilité, entre 0 et 1, que la réponse soit oui ⁶ .		
Ce message exprime-t-il une urgence ?			
CHOICE • UN CHOIX	facturation		0,88
department	technique		0,12
Quelle équipe doit s'en occuper ?	ventes		0,00
Une option parmi une liste fixée, jusqu'à 255. Choix retenu : facturation, avec une confiance de 0,81 ¹² .			
SCORE • UNE NOTE	calme		0,00
frustration	frustré		0,95
À quel point le client est-il frustré ?	très en colère		0,05
Une échelle de 2 à 10 niveaux décrits en mots. Ici 1,05 entre 0 (calme) et 2 (très en colère) : la note peut tomber entre deux niveaux. Confiance : 0,92 ¹² .			

Trois exemples de la documentation de l'API de TypeSafe, sur le même message, traduits de l'anglais (modèle jev-1.13.0). La documentation pose chaque question séparément ; on peut aussi les poser ensemble, en un seul appel^{12,6}.

POURQUOI C'EST SI RAPIDE

UN LLM

écrit mot après mot

- Chaque morceau de mot (un « jeton ») dépend du précédent. Il faut ensuite lire et vérifier le texte produit pour s'en servir dans un programme¹.
- Entraîné surtout pour plaire à des évaluateurs humains (RLHF) ou réussir des tâches vérifiables. Prix d'entrée : 0,20 à 10 dollars par million de jetons, et la sortie coûte environ cinq fois plus, selon le relevé de TypeSafe¹.

JEV

répond à tout d'un coup

- Toutes les probabilités sortent en un seul passage. Ajouter des questions change à peine le temps de réponse^{1,6}. Dans un exemple de TypeSafe, poser 13 questions sur l'article Wikipédia consacré au RGPD en un seul appel coûte 12,2 fois moins cher et va 10 fois plus vite qu'en 13 appels¹³.
- Entraîné pour donner des probabilités honnêtes, avec une méthode maison baptisée RLCD, non publiée en détail¹.

CINQ ANS DE RECHERCHE, TROIS SEMAINES DE FOLIE

Jev descend d'une technique qui a rendu les chatbots aimables. Sa sortie a déclenché une course en quelques jours.

LES RACINES

12 juin 2017

▲ ÉTUDE

Paul Christiano et cinq coauteurs montrent qu'on peut entraîner une IA à partir des préférences d'humains qui comparent deux comportements. C'est la base de ce qu'on appellera le RLHF²⁰.

4 mars 2022

▲ ÉTUDE

OpenAI publie InstructGPT : un modèle de 1,3 milliard de paramètres, affiné avec des retours humains, est préféré par les évaluateurs à GPT-3, cent fois plus gros. Diogo Almeida est l'un des six auteurs principaux²².

2024

— PRESSE

● DÉCLARATION

Almeida quitte OpenAI et fonde TypeSafe AI, déçu que des modèles si doués pour la conversation servent si peu à automatiser^{34, 1}.

LA SORTIE

15 septembre 2026

■ ENTREPRISE

● DONNÉE

TypeSafe présente Jev, en accès anticipé sur liste d'attente¹. Le même jour, l'annonce récolte 1 989 points et 520 commentaires sur Hacker News, le forum de référence des développeurs¹⁹.

17 septembre

■ ENTREPRISE

Netlify, grand hébergeur de sites web, ajoute Jev à sa passerelle d'IA. Son concurrent Vercel le propose aussi, et élargit l'accès le 21 septembre^{18, 17}. Selon Vercel, cité par Arcturus Labs, aucun modèle n'y a été adopté aussi vite³¹.

18 septembre

— PRESSE

TechCrunch rapporte que la demande a été si forte que TypeSafe a brièvement été incapable de servir ses clients³⁴.

22 septembre

▲ TEST

« Jev en 25 lignes de Python » : la société NobodyWho imite le principe avec un petit modèle gratuit. Le texte est très partagé³⁰.

24 septembre

— PRESSE

Selon le site The Information, TypeSafe discute d'une levée de plus d'un milliard de dollars, sur la base d'une valorisation de plus de 10 milliards³⁸.

LA RIPOSTE

29 septembre

■ ENTREPRISE

OpenAI annonce son API Decisions, bâtie sur son modèle Luna : des réponses choisies parmi des options définies à l'avance, à partir de texte ou d'images¹⁴.

1^{er} octobre

— PRESSE

Amazon publie en libre accès Strands Decider 2B, un modèle de décision inspiré de Jev, né d'un projet personnel de l'ingénieur Marc Brooker³⁶.

6 octobre

— PRESSE

La société Musubi publie PolicyLM-1.7B, un modèle de décision ouvert dédié à la modération de contenus³⁷.

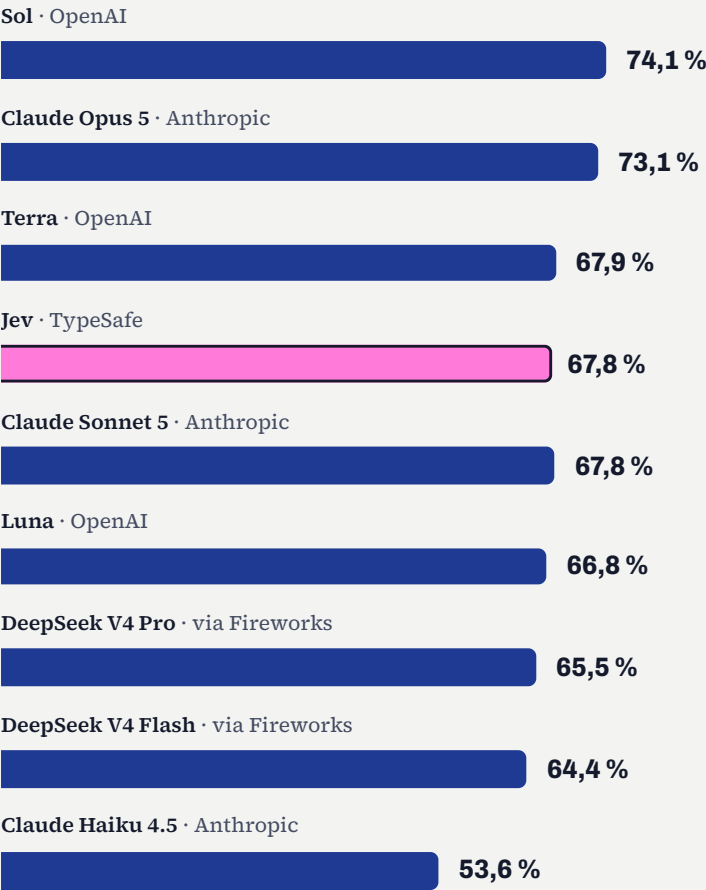
RAPIDE, BON MARCHÉ, MAIS PAS LE PLUS JUSTE

TypeSafe a publié ses propres tests : quatre tâches d'entreprise découpées en petites questions, soumises à Jev et à huit autres modèles. Voici ce qu'ils montrent, et ce qu'ils ne montrent pas.

ACCORD AVEC LA RÉFÉRENCE	TEMPS PAR DOSSIER	COÛT POUR 1 000 DOSSIERS
67,8 % Sol d'OpenAI fait 74,1 %	0,4s contre 10,1 s pour le plus rapide des autres	0,40\$ contre 3,30 \$ pour le moins cher des autres

Moyennes des quatre tâches, chaque modèle utilisé dans un circuit de questions découpées. Source : TypeSafe, page des évaluations consultée le 7 octobre 2026⁴.

Qui donne les mêmes réponses que les meilleurs ?



Jev n'est pas le plus juste. Deux modèles bien plus chers, Sol d'OpenAI et Claude Opus 5 d'Anthropic, s'accordent plus souvent avec la référence. Jev fait jeu égal avec Terra et Claude Sonnet 5⁴.

La référence n'est pas la vérité. TypeSafe ne compare pas les réponses à des étiquettes vérifiées par des humains, mais à la moyenne de deux modèles de pointe, GPT-6 Astra et Claude Fable 5.1. L'entreprise reconnaît que ce choix avantage les modèles d'OpenAI et d'Anthropic, et que les tâches ont été écrites par sa propre équipe^{4,1}.

Découper aide tout le monde. Selon les mêmes tests, chaque modèle est plus juste, moins cher et plus rapide quand la tâche est découpée en petites questions que quand on lui donne tout d'un bloc⁴.

Combien coûtent 1 000 dossiers ?

Jev · TypeSafe

0,40 \$

Luna · OpenAI

3,30 \$

DeepSeek V4 Flash · via Fireworks

5,90 \$

Claude Haiku 4.5 · Anthropic

19,50 \$

Terra · OpenAI

30,40 \$

DeepSeek V4 Pro · via Fireworks

41,30 \$

Sol · OpenAI

83,60 \$

Claude Sonnet 5 · Anthropic

117,40 \$

Claude Opus 5 · Anthropic

176,10 \$

Là où Jev écrase la concurrence : le prix. À précision égale avec Terra d'OpenAI, il coûte environ 76 fois moins cher et va environ 25 fois plus vite. Calcul de SOURCE sur les valeurs arrondies publiées par TypeSafe⁴.

Ces écarts expliquent l'enthousiasme des développeurs : une question qui coûtait des centimes coûte désormais une fraction de centime. Un programme peut alors en poser des milliers par seconde.

VÉRIFICATION

SIX PHRASES AU CRIBLE

NOS VERDICTS

VRAI

Les données le confirment.

À NUANCER

Vrai dans un périmètre précis, pas en général.

EXAGÉRÉ

Le sens est juste, la formule est gonflée.

INVÉRIFIABLE

Les informations nécessaires ne sont pas publiques.

NOS TYPES DE SOURCES

ENTREPRISE

Annonce, documentation ou prix publiés par l'entreprise.

DONNÉE PUBLIQUE

Chiffre consultable par tous.

ÉTUDE OU TEST

Recherche, ou test publié avec sa méthode.

PRESSE

Article ou enquête journalistique.

DÉCLARATION

Ce qu'une personne affirme, sans preuve publiée.

« 193,6 fois plus rapide, 444,6 fois moins cher. »

PAGE D'ACCUEIL DE TYPESAFE

Ces chiffres viennent des tests maison de TypeSafe, qui les juge elle-même plutôt dans le haut de ce qu'on peut espérer en pratique^{2,1}. Sa démonstration vidéo utilise une requête « très simplifiée » et un texte court, ce qui avantage son modèle, reconnaît-elle¹. Des tiers confirment de gros écarts, plus modestes : 5 à 18 fois plus rapide que Luna d'OpenAI chez Vercel³⁴.

À NUANCER

DÉMONSTRATION DE
TYPESAFE

0,114s

contre 8,566 s pour le
LLM comparé

COÛT DE LA MÊME
REQUÊTE

0,000081\$

contre 0,013880 \$

typesafe.ai/blog/introducing-system-one-models-an

- The numbers for LLMs are from OpenRouter i.e., there almost certainly is bias here: more complex queries might be routed to better models.
- Our number is not empirical. Schema matching is guaranteed, thus we can confidently add 0% into the plots.

Annnonce de TypeSafe, 15 septembre 2026, consultée le 8 octobre 2026. Surlignage de SOURCE¹.

« Zéro hallucination. »

PAGE D'ACCUEIL DE TYPESAFE

Vrai au sens strict : Jev ne renvoie qu'une des réponses prévues, il ne peut pas inventer d'option. Le « 0 % » de TypeSafe n'est pas une mesure mais une conséquence de la construction, comme le dit l'annonce elle-même : « Notre chiffre n'est pas empirique » (traduit de l'anglais)^{1,2}. Faux au sens courant : Jev peut choisir la mauvaise option. Sa documentation liste ses erreurs connues et rappelle qu'une bonne calibration ne garantit pas qu'une réponse isolée soit juste^{10,5}.

À NUANCER

« Un prix d'entrée 238 fois plus bas que Claude Fable 5.1. »

PAGE D'ACCUEIL DE TYPESAFE

Anthropic facture Claude Fable 5.1 10 dollars par million de jetons en entrée et 50 dollars en sortie¹⁶. Jev coûte 0,042 dollar en entrée, et rien en sortie⁹. Dix divisé par 0,042 donne bien 238.

Mais Fable écrit, raisonne et programme ; Jev ne fait que trancher entre des options¹¹.

VRAI

CLAUDE FABLE 5.1

10\$

par million de jetons en
entrée

JEV 1.13

0,042\$

par million de jetons en
entrée

CHRISTIANO ET AL.,
2017

6 auteurs

Almeida n'en fait pas
partie

INSTRUCTGPT, 2022

20 auteurs

Almeida est 4^e, parmi
les auteurs principaux

« Diogo Almeida a co-inventé le RLHF. »

SITE DE TYPESAFE, REPRIS PAR TECHCRUNCH

Le RLHF, l'apprentissage par renforcement à partir de retours humains, a été décrit en 2017 par Paul Christiano et ses collègues, puis appliqué aux résumés en 2020. Diogo Almeida ne signe aucun de ces articles^{20,21}. Il est en revanche l'un des six auteurs principaux d'InstructGPT (2022), qui a appliqué la méthode aux assistants et mené à ChatGPT²². TypeSafe dit qu'il a « co-inventé le RLHF et InstructGPT » : la seconde moitié tient mieux que la première³.

À NUANCER

« TypeSafe vaut 10 milliards de dollars. »

TITRES DE PRESSE, FIN SEPTEMBRE

The Information parle de discussions, pas d'une levée bouclée ; Forbes estimait la valorisation initiale à environ 200 millions³⁸. Aucun chiffre officiel n'a été publié.

INVÉRIFIABLE

« Ses prix sont subventionnés. »

SOUPÇON RÉPANDU CHEZ LES DÉVELOPPEURS

TypeSafe écrit qu'elle ne peut pas prouver le contraire et s'attend à ce que ses prix baissent¹. Ses comptes ne sont pas publics.

INVÉRIFIABLE

UNE PROBABILITÉ, ÇA SE VÉRIFIE

Tout l'intérêt de Jev tient dans un mot : calibré. Quand un modèle bien calibré annonce 80 %, il doit avoir raison environ huit fois sur dix, et 20 % doit se réaliser environ deux fois sur dix⁸. Cela se mesure sur un grand nombre de réponses, jamais sur une seule⁵.

Les réseaux modernes sont trop sûrs d'eux

Le constat de départ.

Université Cornell. Les réseaux de neurones récents, contrairement à ceux d'il y a dix ans, sont mal calibrés. Une correction très simple, le « temperature scaling », suffit souvent à les remettre d'aplomb sur les jeux de données testés, des tâches de classement d'images et de textes²³.

SOLIDITÉ DE LA MÉTHODE



Revue par les pairs, nombreux jeux de données

Guo et al., ICML 2017

Plaire aux humains dégrade la calibration

L'ironie de l'histoire.

OpenAI. Avant son affinage, GPT-4 était très bien calibré sur un extrait d'un test de connaissances à choix multiples (MMLU). Après l'entraînement qui le rend agréable à utiliser, dont le RLHF, sa calibration est « nettement » dégradée, écrit OpenAI. C'est l'argument central de TypeSafe : la méthode qu'Almeida a contribué à populariser rend les modèles moins fiables pour décider^{24,1}.

SOLIDITÉ DE LA MÉTHODE



Rapport d'entreprise, un seul jeu de test

OpenAI, rapport technique GPT-4, 2023

InstructGPT

D'où vient le RLHF des chatbots.

OpenAI, 20 auteurs. Des humains classent les réponses d'un modèle, puis on l'entraîne à produire celles qu'ils préfèrent. Résultat sur les demandes testées : un modèle de 1,3 milliard de paramètres est préféré à GPT-3, qui en compte 175 milliards²².

SOLIDITÉ DE LA MÉTHODE



Prépublication détaillée, évaluations humaines

Ouyang et al., arXiv, 2022

Système 1, Système 2

L'origine du nom « Système 1 ».

Daniel Kahneman, Prix Nobel d'économie mort en 2024, décrit deux façons de penser : un Système 1 rapide, intuitif et émotionnel, un Système 2 plus lent et plus logique²⁷. TypeSafe reprend l'image et reconnaît que le Système 1 a la réputation de se tromper souvent ; elle promet de démontrer le contraire pour ses modèles¹.

SOLIDITÉ DE LA MÉTHODE



Livre grand public, synthèse de travaux

Kahneman, Thinking, Fast and Slow, 2011

Le paradoxe de Jevons

L'origine du nom « Jev ».

En 1865, l'économiste anglais William Stanley Jevons observe que des machines à vapeur plus économes n'ont pas réduit la consommation de charbon : « C'est le contraire qui est vrai. »²⁵ TypeSafe parie qu'une IA cent fois moins chère sera utilisée bien plus de cent fois plus¹. Une revue de 2009 rappelle que ce paradoxe est très difficile à tester : les preuves sont « loin d'être concluantes », même si l'effet rebond semble plus fort qu'on ne le pense d'habitude²⁶.

SOLIDITÉ DE LA MÉTHODE



Revue critique, revue par les pairs

Sorrell, Energy Policy, 2009

CONVAINCUS ET SCEPTIQUES

Ceux qui l'ont essayé dans leurs programmes, ceux qui s'en méfient, et la réponse du fondateur.

LES CONVAINCUS

« C'est le seul qui renvoie une vraie probabilité, ce qui le rend idéal pour automatiser des workflows »

Nikhil Mudholkar, directeur technique de Bryo AI, cité par TechCrunch, 18 septembre 2026. Dans son test sur des courriels professionnels, Gemini était un peu plus juste mais 10 à 20 fois plus cher³⁴.

Vercel utilisait Luna d'OpenAI pour vérifier la sûreté de commandes. En le remplaçant par Jev, l'entreprise a obtenu des résultats 5 à 18 fois plus rapides et plus justes. *(propos résumés)*

Pranit Sharma, ingénieur chez Vercel, cité par TechCrunch, 18 septembre 2026³⁴

« Jev a jugé tout ce que j'ai écrit en 0,7 seconde »

Mike Taylor, responsable des évaluations chez Every, titre de son test, 15 septembre 2026²⁸

LES SCEPTIQUES

« Une régression encore plus poussée vers les systèmes d'apprentissage automatique en boîte noire »

Simon Willison, développeur et blogueur, 21 septembre 2026. Il espère que personne ne s'en servira pour classer des candidats à l'embauche, à cause des biais invisibles²⁹.

Tout le monde présente Jev comme la prochaine frontière de l'IA. « Nous ne le pensons pas vraiment. » Suit une imitation en 25 lignes. *(propos résumés, sauf la citation)*

NobodyWho, éditeur de logiciels d'IA locale, 22 septembre 2026³⁰

OpenAI utilise depuis des années ses modèles comme classificateurs implicites. Si elle reproduit l'entraînement, elle pourra copier Jev rapidement. *(propos résumés)*

Arcturus Labs, cabinet de conseil en IA, 21 septembre 2026³¹

LA RÉPONSE DU FONDATEUR

Diogo Almeida, à TechCrunch, fin septembre 2026

« Rapide et pas cher, c'est très facile. Si vous voulez vraiment rapide et pas cher, prenez des dés. L'intelligence, c'est la partie difficile. »³⁵ Pour lui, les copies viennent de spécialistes séduits par une architecture plus que d'équipes décidées à rendre l'intelligence utile. *(propos résumés)*³⁶

Les phrases entre guillemets sont traduites de l'anglais par nos soins, au plus près du texte lu dans la source citée. Les propos résumés sont signalés.

LES CONCURRENTS EN TROIS SEMAINES

Jev a créé une catégorie que TechCrunch appelle déjà les « modèles de décision ». Géants et petites équipes ont suivi en quelques jours³⁷.

1 OpenAI, API Decisions

Annoncée le 29 septembre, bâtie sur Luna. Elle accepte texte et images, ce que Jev ne fait pas, et répond environ 10 fois plus vite que l'API classique d'OpenAI, selon l'entreprise. En bêta publique au 7 octobre^{14, 15, 9}.

2 Amazon, Strands Decider 2B

Publié en libre accès le 1^{er} octobre, assez petit pour tourner sur un ordinateur. Il est bâti sur un modèle ouvert chinois, Qwen3.5-2B³⁶.

3 Kev, projet ouvert

Une famille de petits modèles sur la même base Qwen. Le plus gros serait à moins de trois points de Jev, ou devant, dans 9 catégories de tests sur 11, selon ses auteurs, qui précisent que la comparaison n'est pas contrôlée³³.

4 Musubi, PolicyLM-1.7B

Publié le 6 octobre, ouvert, spécialisé dans la modération de contenus³⁷.

LE TEST D'UNE ENTREPRISE ALLEMANDE

Privatemode, qui héberge des modèles d'IA en Europe, a transformé un modèle ouvert, GLM-5.3-Flash, en modèle de décision, puis l'a comparé à Jev sur 28 jeux de données publics. Attention : c'est un concurrent qui teste son propre produit³².

PRÉCISION

égalité

chacun meilleur sur 10 jeux ;
écart médian de 0,7 point
pour Jev, non significatif

COÛT D'UN MILLION DE DÉCISIONS

16 €

pour Jev, contre 62 € pour
GLM-5.3-Flash

LA DISTANCE COMPTE

GLM-5.3-Flash · depuis l'Allemagne

180 ms

Jev · depuis l'Allemagne

264 ms

Jev · depuis les États-Unis

164 ms

GLM-5.3-Flash · depuis les États-Unis

299 ms

« On peut refaire Jev en 25 lignes de code. »

EXAGÉRÉ

NOBODYWHO, 22 SEPTEMBRE 2026

Elles reproduisent la forme : un petit modèle ouvert dont on lit les probabilités pour chaque option³⁰. Pas la qualité, qui dépend de l'entraînement. Des copies plus sérieuses s'en approchent, selon leurs auteurs^{32, 33}.

CE QUE JEV NE SAIT PAS FAIRE

TypeSafe publie la liste des faiblesses connues de sa version 1.13. Elle est plus franche que la page d'accueil¹⁰.

1 Compter et calculer

Jev ne compte pas de façon fiable, et l'erreur grandit avec ce qu'il faut compter. Conseil de TypeSafe : laisser le calcul au code¹⁰.

2 Comparer des dates

Mieux vaut lui faire extraire le jour, le mois et l'année, puis comparer dans le programme¹⁰.

3 Lire entre les lignes

Il répond à la question écrite, pas à celle qu'on avait en tête. Les négations et les conditions implicites sont prises au pied de la lettre¹⁰.

4 Ignorer l'ordre des options

Changer l'ordre des choix peut changer la réponse. TypeSafe conseille de vérifier en les mélangeant¹⁰.

5 Voir, écouter, écrire

Texte uniquement : ni image, ni son, ni vidéo. Et aucune génération de texte : il ne remplace pas le modèle d'un assistant de programmation^{9, 11}.

LE SAVIEZ-VOUS ?

7 \$

Le coût d'une heure de jeu pour une démonstration où Jev joue à Doom, à raison de 10 questions par seconde, selon TypeSafe¹.

64 000

Le nombre maximal de jetons par requête, questions comprises. Au-delà, il faut découper⁹.

2012

L'année où une autre entreprise nommée Typesafe, fondée par les créateurs du langage Scala, levait 14 millions de dollars. Aucun lien avec TypeSafe AI³⁹.

ASTUCE • LIRE UNE RÉPONSE

LA CONFIANCE N'EST PAS LA PROBABILITÉ

Pour un choix ou une note, Jev donne aussi une « confiance » calculée à partir des probabilités : 1 quand tout se concentre sur une option, 0 quand tout est réparti également⁷. Une confiance basse signale une question ambiguë ou un texte qui ne contient pas la réponse. C'est le moment de passer la main à un humain.

ASTUCE • AVANT DE L'ADOPTER

TESTEZ SUR VOS PROPRES DONNÉES

TypeSafe le recommande elle-même : commencer avec des seuils prudents, mesurer sur ses propres cas, puis ajuster⁷. Un test de Simon Willison montre pourquoi : interrogé sur le fait de savoir si chaque ville de la baie de San Francisco était une « bonne ville », Jev a placé la riche Cupertino en tête et East Palo Alto en dernier²⁹.

HUIT MOTS POUR SUIVRE

« LLM »

Grand modèle de langage. Un programme entraîné sur d'immenses quantités de textes, qui produit du texte mot après mot : ChatGPT, Claude, Gemini.

« Calibration »

L'accord entre les probabilités annoncées et la réalité : ce qui est annoncé à 80 % doit arriver environ 80 fois sur 100⁸.

« RLHF »

Apprentissage par renforcement à partir de retours humains : on entraîne le modèle à produire les réponses que des humains préfèrent²².

« Modèle ouvert »

Un modèle dont les paramètres sont téléchargeables, comme Qwen. La plupart des copies de Jev en partent^{33, 36}.

« Jeton »

Le morceau de mot que lit un modèle. Les prix de l'IA se comptent en jetons, par million ou, pour Jev, par milliard⁹.

« Confiance »

Chez TypeSafe, un indice entre 0 et 1 qui résume à quel point les probabilités d'une réponse sont concentrées⁷.

« RLCD »

Le nom donné par TypeSafe à sa propre méthode d'entraînement, centrée sur des probabilités honnêtes. Elle n'est pas publiée en détail¹.

« Effet rebond »

Quand une ressource devient moins chère à utiliser, on en consomme davantage, ce qui grignote l'économie réalisée. Le paradoxe de Jevons en est la forme extrême²⁶.

ANGLES MORTS

CE QU'ON NE SAIT PAS ENCORE

? Comment Jev est construit

Un transformeur, comme les LLM, mais sans publication technique ; des observateurs soupçonnent une base de modèle ouvert. Il serait entraîné uniquement sur des données fabriquées par TypeSafe, selon Diogo Almeida³⁴.

? Ses biais

Une probabilité sans explication ne dit pas ce qui l'a fait pencher. Aucun audit de biais n'a été publié²⁹.

? Les données des utilisateurs européens

TypeSafe indique que son service est hébergé sur la côte ouest des États-Unis¹. Nous n'avons pas analysé ses conditions d'utilisation au regard du droit européen.

? Sa calibration, mesurée par d'autres

La promesse centrale. Nous n'avons trouvé aucune mesure indépendante et à grande échelle ; les tests de tiers portent sur peu de cas ou viennent de concurrents.

? Son vrai coût et l'argent levé

Comptes non publics ; les montants des levées ne viennent que de la presse³⁸.

COMMENT CE NUMÉRO A ÉTÉ FAIT

- 1

Nous avons lu l'annonce de TypeScript, sa documentation technique et ses grilles de prix dans leur version du 7 octobre 2026. Ces pages évoluent vite.
- 2

Les chiffres des graphiques « En chiffres » ont été relevés sur la page des évaluations de TypeScript. Ce sont des mesures de l'entreprise, non reproduites.
- 3

Pour le parcours de Diogo Almeida, nous sommes remontés aux articles scientifiques et à leur liste d'auteurs sur arXiv.
- 4

Les tests de tiers ont été lus dans leur version originale. Quand leur auteur vend un produit concurrent, nous le signalons.
- 5

Les citations sont traduites de l'anglais au plus près ; les autres propos sont résumés et signalés.

LIMITE DE CE NUMÉRO

Quatre articles de Forbes, celui du Wall Street Journal du 4 octobre et celui de The Information du 24 septembre n'ont pas pu être lus, faute d'accès. Les montants de la levée et la valorisation sont donc cités d'après Sovereign Magazine, qui résume ces deux derniers médias. Le test d'Every n'a été lu qu'en partie, la suite étant réservée aux abonnés. Un billet du blog de TypeScript ne s'affiche pas pour les lecteurs automatiques.

DOCUMENTS DES ENTREPRISES

- 1

Almeida D., « Introducing System One Models & Jev », blog de TypeScript AI, 15 septembre 2026. [typesafe.ai](#)
- 2

TypeScript AI, page d'accueil, consultée le 7 octobre 2026. [typesafe.ai](#)
- 3

TypeScript AI, page « Team », consultée le 7 octobre 2026. [typesafe.ai/team](#)
- 4

TypeScript AI, « Workflow evals », page des évaluations, consultée le 7 octobre 2026. [evals.typesafe.ai](#)
- 5

Documentation TypeScript, « System One ». [docs.typesafe.ai](#)
- 6

Documentation TypeScript, « Primitives (Questions) » et « Noul ». [docs.typesafe.ai](#)
- 7

Documentation TypeScript, « Confidence ». [docs.typesafe.ai](#)
- 8

Documentation TypeScript, « AI primer ». [docs.typesafe.ai](#)
- 9

Documentation TypeScript, « Models » (Jev 1.13 : prix, limites, contexte). [docs.typesafe.ai](#)
- 10

Documentation TypeScript, « Jev 1.13 jaggedness » (faiblesses connues). [docs.typesafe.ai](#)
- 11

Documentation TypeScript, « Jev with coding agents ». [docs.typesafe.ai](#)
- 12

Documentation TypeScript, « API reference ». [docs.typesafe.ai](#)
- 13

Documentation TypeScript, recette « Parallel questions » (13 questions sur l'article Wikipédia du RGPD). [docs.typesafe.ai](#)
- 14

OpenAI, « DevDay 2026 Recap », 29 septembre 2026, rubrique API Decisions. [openai.com](#)
- 15

OpenAI, documentation « Decisions », consultée le 7 octobre 2026. [developers.openai.com](#)
- 16

Anthropic, « Pricing », documentation de la plateforme Claude, consultée le 7 octobre 2026. [platform.claude.com](#)

- 17

Taneja R., Zheng J., « AI Gateway now supports TypeScript clients and an HTTP API for Jev », Vercel, 21 septembre 2026. [vercel.com](#)
- 18

Netlify, « TypeScript Jev now available in AI Gateway », 17 septembre 2026. [netlify.com](#)

DONNÉE PUBLIQUE

- 19

Hacker News, fil « Introducing System One Models and Jev », 15 septembre 2026 : points et commentaires relevés le 7 octobre 2026 via l'API publique. [news.ycombinator.com](#)

ÉTUDES ET OUVRAGES

- 20

Christiano P., Leike J., Brown T. B., Martic M., Legg S., Amodei D. Deep reinforcement learning from human preferences. arXiv, 12 juin 2017. [doi.org/10.48550/arXiv.1706.03741](#)
- 21

Stiennon N. et al. (9 auteurs). Learning to summarize from human feedback. arXiv, 2 septembre 2020. [arxiv.org/abs/2009.01325](#)
- 22

Ouyang L., Wu J., Jiang X., Almeida D. et al. (20 auteurs). Training language models to follow instructions with human feedback. arXiv, 4 mars 2022. [doi.org/10.48550/arXiv.2203.02155](#)
- 23

Guo C., Pleiss G., Sun Y., Weinberger K. Q. On Calibration of Modern Neural Networks. ICML 2017. [doi.org/10.48550/arXiv.1706.04599](#)
- 24

OpenAI. GPT-4 Technical Report, 2023, figure 8 (calibration avant et après l'affinage). [doi.org/10.48550/arXiv.2303.08774](#)
- 25

Jevons W. S. *The Coal Question*, Macmillan, 1865, chapitre 7 « Of the Economy of Fuel » (extraits publiés par Yale Energy History). [energyhistory.yale.edu \(PDF\)](#)
- 26

Sorrell S. Jevons' Paradox revisited: The evidence for backfire from improved energy efficiency. *Energy Policy*, 37 (4), 1456-1469, 2009 (résumé lu). [doi.org/10.1016/j.enpol.2008.12.003](#)

27 Kahneman D. *Thinking, Fast and Slow*, Farrar, Straus and Giroux, 2011 (présentation de l'éditeur Penguin Random House). penguinrandomhouse.com

▲ TESTS ET ANALYSES DE TIERS

-
- 28 Taylor M., « Mini-Vibe Check: TypeScript's Jev Judged Everything I've Written in 0.7 Seconds », Every, 15 septembre 2026 (lu en partie). every.to
-
- 29 Willison S., « Jev introduces a new shape of LLM – System One, aka Decision Models », 21 septembre 2026. simonwillison.net
-
- 30 O. Carmo D., « Jev in 25 lines of Python », NobodyWho, 22 septembre 2026. nobodywho.ai
-
- 31 Arcturus Labs, « Will OpenAI Eat Jev's Lunch? », 21 septembre 2026. arcturus-labs.com
-
- 32 Hötter J., « Turn GLM-5.3-Flash into a Jev-like System One model », Privatemode, 24 septembre 2026. privatemode.ai
-
- 33 Kev, « Small Jev-like decision models you can train and run yourself », dépôt GitHub, consulté le 7 octobre 2026. github.com/jaredpalmer/kev

— PRESSE

-
- 34 Fernholz T., « A new kind of AI model from a ChatGPT inventor is thrilling developers », TechCrunch, 18 septembre 2026. techcrunch.com
-
- 35 Fernholz T., « OpenAI's Jev clone could help the frontier lab stop its swarming agents », TechCrunch, 30 septembre 2026. techcrunch.com
-
- 36 Fernholz T., « Amazon releases its own Jev clone as decision models flood the web », TechCrunch, 1^{er} octobre 2026. techcrunch.com
-
- 37 Brandom R., « How AI decision models could change content moderation », TechCrunch, 6 octobre 2026. techcrunch.com
-
- 38 Nani D., « TypeScript AI reportedly raising \$1B at \$10B valuation », Sovereign Magazine, 24 septembre 2026 (reprend The Information et Forbes, non lus). sovereignmagazine.com
-
- 39 Rao L., « Typesafe Raises \$14M From Shasta, Greylock, And Juniper To Commercialize Scala », TechCrunch, 22 août 2012. techcrunch.com