

[Offrir un abonnement](#)[Mon compte](#)

Idées • Intelligence artificielle

IA générale, « super-intelligence » : l'avenir technologique est-il si inquiétant que ça ?

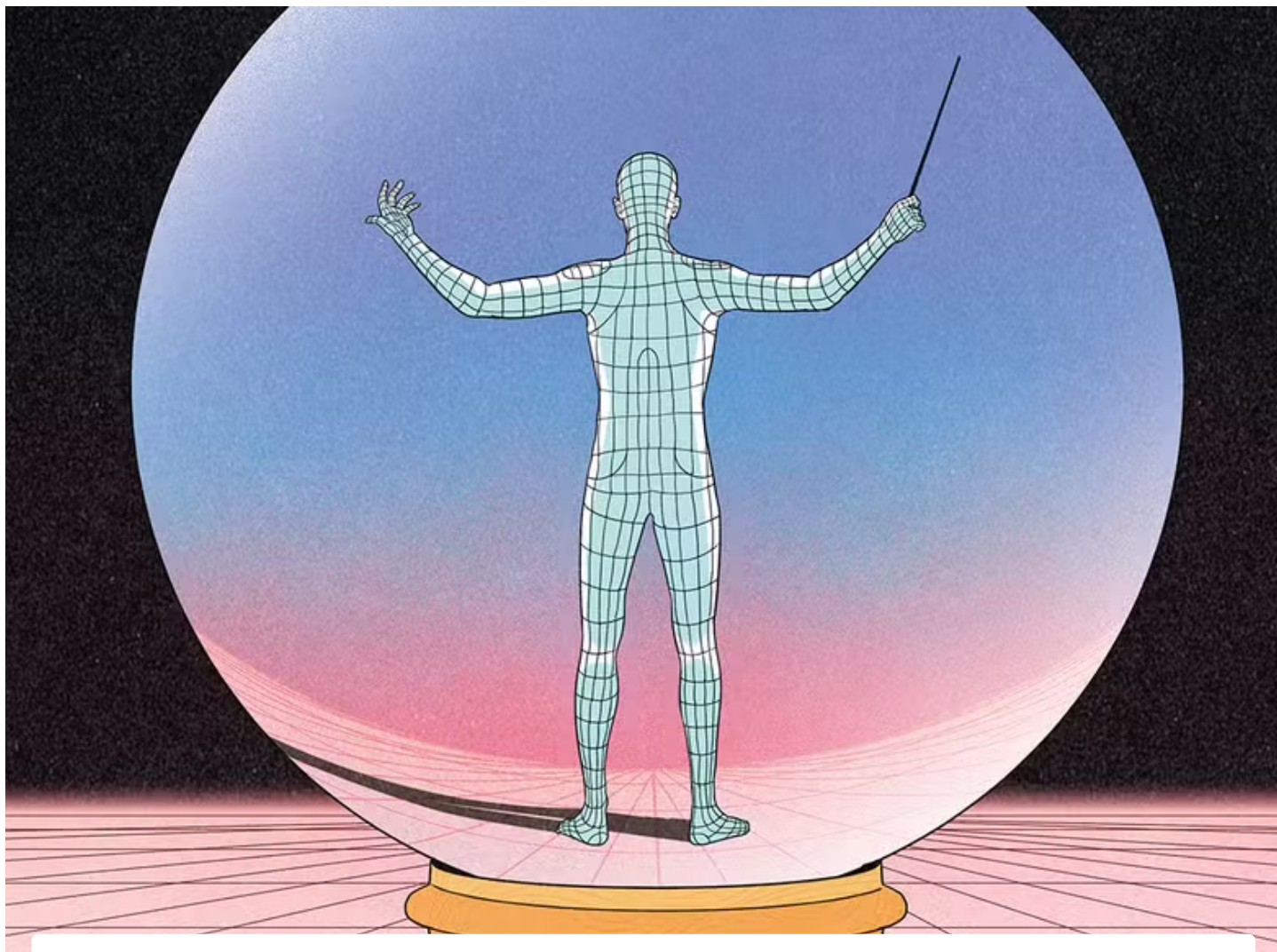
Décryptage Des machines aux capacités supérieures à celles des humains, c'est la prochaine rupture technologique annoncée. Elle effraie certains spécialistes qui publient une tribune alarmante. Mais derrière le terme très marketing, la définition est encore floue et les projets, balbutiants.

Par **Oscar Leroy** et **Xavier de La Porte**

Publié le 22 octobre 2025 à 17h00 , mis à jour le 23 octobre 2025 à 10h04

Lecture : 8 min.

Abonné

[Annuler](#)[Offrir cet article](#)[Commenter](#)

A peine prenons-nous la mesure de ce qui nous arrive avec les intelligences artificielles génératives (ChatGPT, Perplexity, Claude, etc.) qu'on nous annonce pour bientôt une nouvelle étape : l'intelligence artificielle générale. Sam Altman, patron d'OpenAI, l'entreprise à l'origine de ChatGPT, affirmait en janvier dernier : « *Nous sommes désormais convaincus de savoir comment construire l'IA générale.* » Huit mois plus tard, en septembre, il précisait : elle sera là d'ici à 2030. Demis Hassabis (Google DeepMind) et Dario Amodei (Anthropic, qui a créé le modèle Claude), font des prédictions du même ordre. On entend trois ans par-ci, dix ans par-là... Bref, on n'en serait pas loin. Mais de quoi s'agit-il ?

A lire aussi

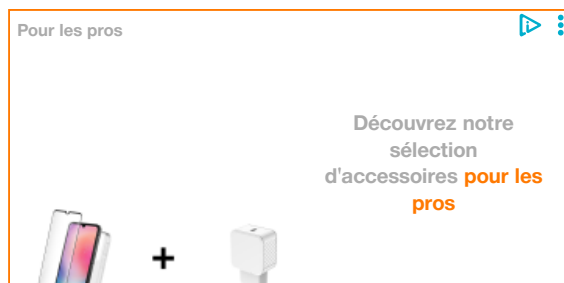


Analyse Post-vérité, contradictions et « vertige » : dans le cerveau de Sam Altman

[Abonné](#)

L'IA générale, c'est la réactualisation d'un rêve forgé par quelques chercheurs de génie lors de la célèbre conférence de Dartmouth (1956), qui a inauguré le champ de l'intelligence artificielle : fabriquer une machine qui puisse égaler, voire surpasser, le cerveau humain dans l'ensemble de ses aptitudes cognitives. Longtemps considérée comme un horizon inaccessible, cette ambition est devenue plus concrète avec les progrès de l'informatique depuis le début des années 2010. Puissance de calcul multipliée, accès à des données toujours plus nombreuses, nouvelles méthodes d'apprentissage... Tout cela a abouti à ce qu'on connaît aujourd'hui : des programmes avec lesquels on peut converser presque naturellement, qui comprennent ce qu'on leur demande, savent plein de choses, créent à la demande des textes et des images. C'est le monde des IA génératives actuelles. Mais ce n'est pas encore l'IA générale qu'on nous promet comme le Graal, telle une borne civilisationnelle.

Publicité



Annuler

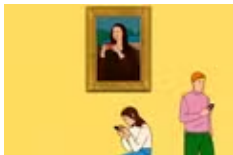
Mention l'

Quelques aptitudes – encore inaccessibles aux grands modèles de langage (*large language models*, LLM) comme ChatGPT – pourraient marquer le passage vers une IA générale : suivre des objectifs qu'elle s'est fixés elle-même, raisonner malgré les incertitudes, utiliser les compétences acquises dans un domaine pour les appliquer à un autre. Cela permettrait d'éviter les erreurs bêtes, de limiter la supervision humaine dans l'étape dite « d'apprentissage » et de réagir à des situations inconnues. « *Les modèles actuels sont entraînés sur des tâches très spécifiques, à partir d'immenses quantités de données*, explique Olivier Cappé, directeur de recherche au département d'informatique du CNRS. *Une IA qui sait créer une vidéo à partir d'un texte ne sait pas conduire une voiture.* » Un humain peut accomplir une tâche pour laquelle il n'a pas été entraîné. Si on lui donne un jeu de cartes, il peut avoir l'idée d'en faire un château, et saura intuitivement comment s'y prendre. Une machine qui n'aurait pas été entraînée dans ce but ne serait capable ni de l'un ni de l'autre.

« Risque existentiel »

Malheureusement, et c'est tout le problème, aucune définition solide de cette nouvelle IA ne fait consensus. Pour une raison simple : une intelligence égale ou supérieure à celle de l'humain, ça ne correspond à aucune notion informatique. « *Je vous mets au défi : essayez de trouver un test qui permette de déterminer si un modèle possède l'IA générale*, lance Michal Valko, qui a notamment été l'ingénieur en chef de Llama 3, l'IA générative de Meta. *Réussir la fusion nucléaire ou trouver le vaccin du Covid, ce sont des problèmes assez simples à définir. Mais l'IA générale...* » Selon le test qu'on choisit, on pourrait l'avoir déjà atteinte, ou ne jamais l'atteindre. Et c'est loin d'être anecdotique : les acteurs de l'IA peuvent employer ce terme pour décrire des choses très différentes. D'ailleurs, on parle aussi de « *superintelligence* ». Et quand on demande la différence à Patrick Pérez, chercheur en IA depuis trente ans et à la tête de Kyutai, laboratoire à but non lucratif, il hésite : « *A vrai dire, je ne sais pas très bien les distinguer.* »

A lire aussi



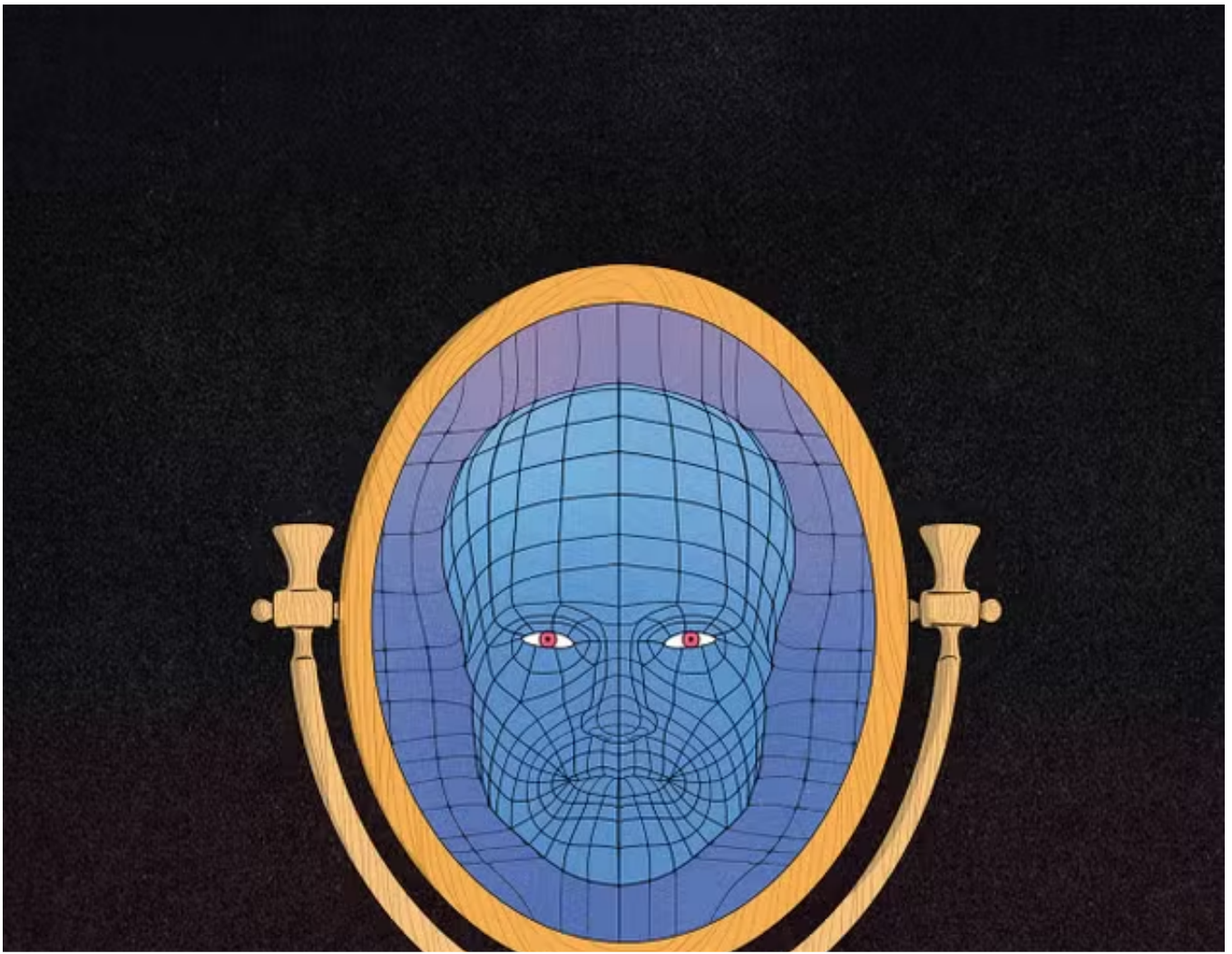
Décryptage Temps d'écrans, réseaux sociaux, IA... Sommes-nous devenus plus bêtes ?

Abonné

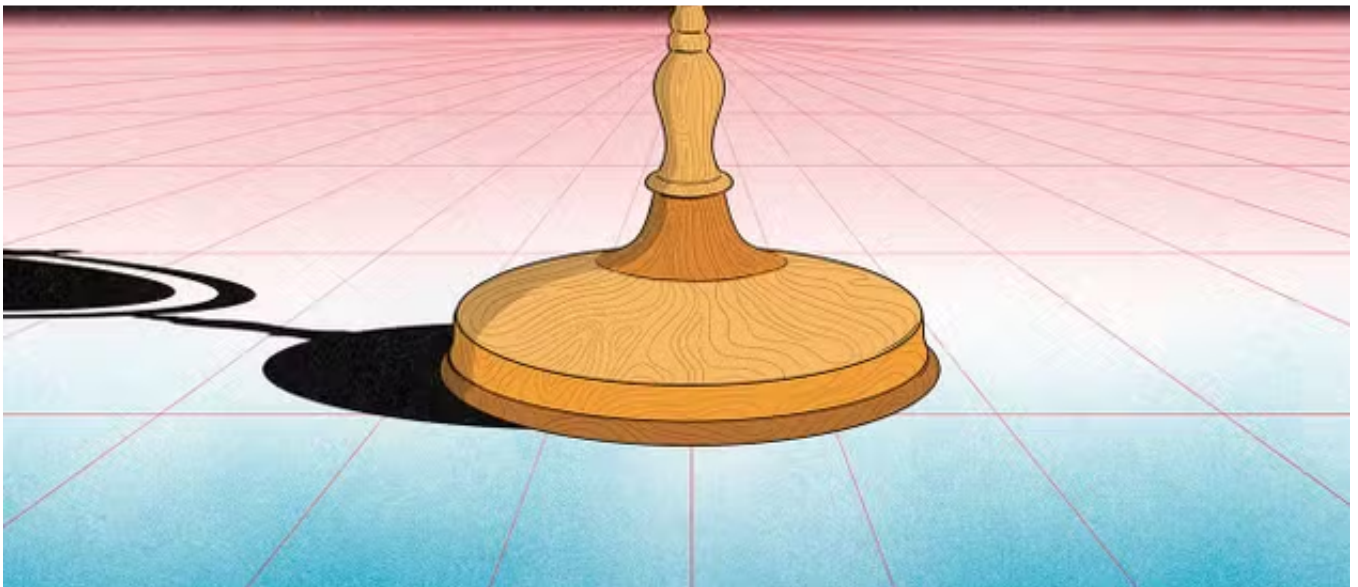
D'ailleurs, il est difficile de dire à quoi ressemblerait un monde avec l'IA générale. « *On peut imaginer des robots bipèdes capables de nous assister en chargeant un lave-vaisselle par exemple ou qui remplacent des métiers que les humains ne veulent pas faire comme l'aide aux*

[Annuler](#)

Dauphiné. Patrick Pérez, lui, envisage plutôt une intelligence numérique, qui agit de manière quasi autonome (on lui donne une tâche à effectuer, même très complexe, elle est capable de la décomposer en autant de sous-tâches que nécessaire). Sauf que, « plein de choses se font sans écran, nuance Olivier Cappé. Par exemple, une IA générale de ce type serait incapable de transporter des marchandises ».



Annuler



Bref, même les chercheurs les plus en pointe ne savent pas trop vers où on va. Ce qui n'empêche pas les fantasmes : se fixant ses propres buts, l'IA pourrait décider, par exemple, de nous éliminer... Dans une lettre ouverte [publiée ce mercredi](#), de grands noms de l'IA comme Geoffrey Hinton, Yoshua Bengio et Stuart Russell ont parlé à ce propos de risques qui pèseraient sur l'humanité, allant jusqu'à sa « *potentielle extinction* ». On n'en est pas encore là, mais « *dans un monde rationnel, on se demanderait d'abord à quoi l'IA générale*

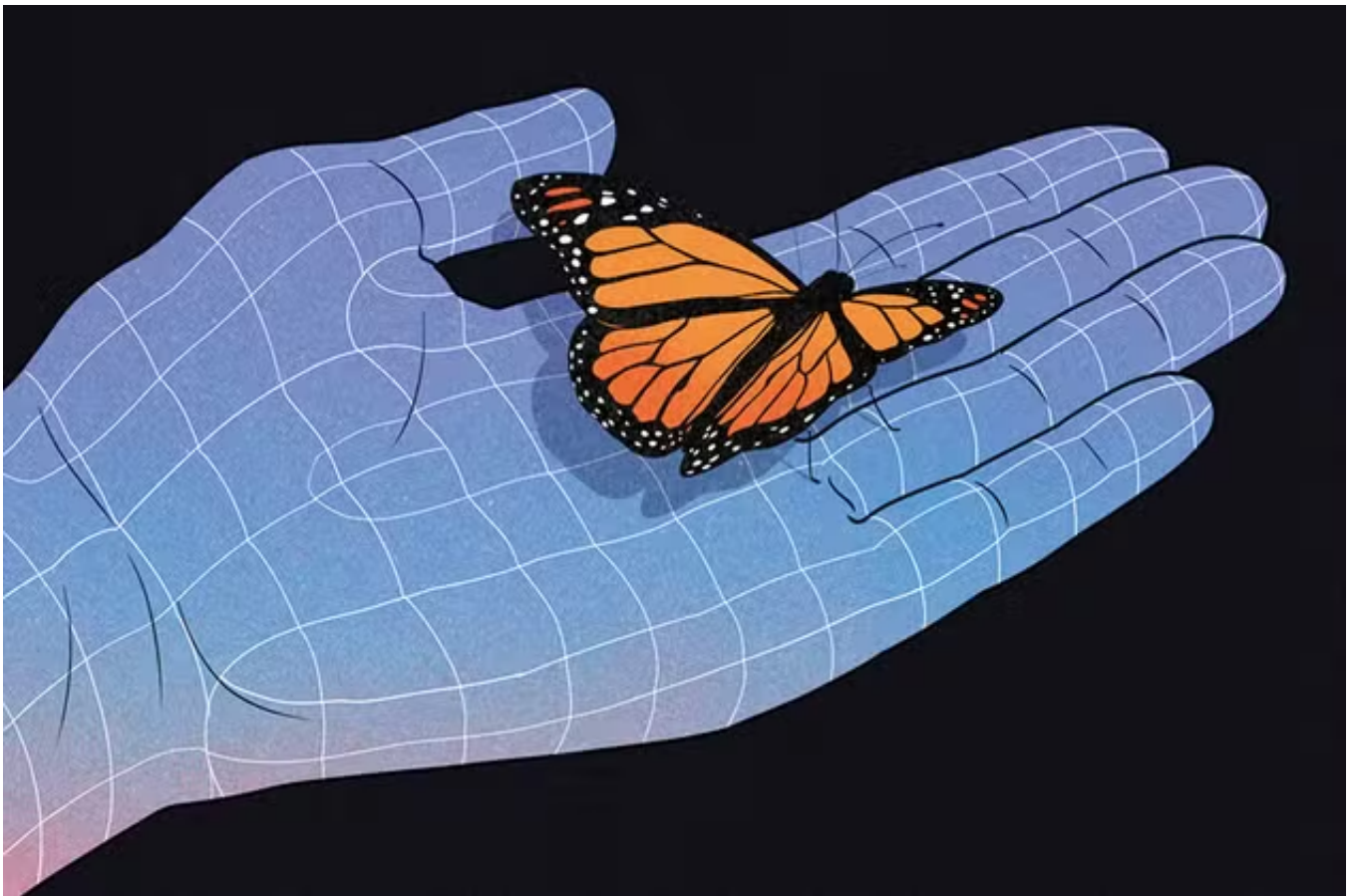
servirait avant d'essayer de la développer », s'agace Olivier Cappé. Hélas, nous ne sommes pas dans un monde rationnel...

Obtenir un modèle capable d'accomplir des tâches diverses

A ce propos, les laboratoires sont-ils vraiment en quête de l'IA générale, comme leurs dirigeants se plaisent à le proclamer dès qu'ils ont un micro sous le nez ? « *Le terme "IA générale" n'est pas employé au sein des grandes entreprises, explique Michal Valko, qui il y a un an encore travaillait chez Meta. Personne ne s'interroge : est-ce que ce modèle a atteint ce stade ? Les chercheurs se demandent plutôt si le modèle répond à tel critère clairement défini.* » Ce qui se rapproche le plus de ces recherches, ce sont, par exemple chez Google DeepMind, des ensembles de problèmes régulièrement proposés aux chercheurs : déplacer des cubes pour trouver un objet caché, sortir d'un labyrinthe, etc. Si ces tâches sont faciles à réaliser avec un algorithme spécialisé, le défi, ici, est d'avoir un seul modèle capable de toutes les résoudre. Lorsqu'il y parvient, la difficulté est augmentée. L'idée est d'obtenir un modèle de plus en plus fort pour accomplir des tâches diverses, et d'y aller pas à pas. « *Rappelons-nous ce qui s'est passé avec les échecs, explique Michal Valko. On y voyait le symbole de l'intelligence humaine, et on a fabriqué un programme bien meilleur que nous. L'étape d'après était le jeu de go, beaucoup plus difficile, et la machine a aussi battu l'humain. La quête de l'IA générale pourrait suivre ce schéma : on se donne des objectifs précis, on les atteint, on s'en fixe d'autres plus ambitieux et ainsi de suite. Mais est-ce parce qu'un modèle trouvera un remède universel contre le cancer qu'on va établir qu'il est plus intelligent que l'homme ? Ce n'est pas*

[Annuler](#)

du monde ».



Yann Le Cun, pionnier de l'IA, directeur de la recherche chez Meta, cherche à développer une IA capable de raisonner et d'interagir dans des environnements inconnus et sans

Annuler

A l'instar de Demis Hassabis, le PDG de Google DeepMind, il mise sur ces *world models*, dont la définition est volontairement très vague : des entités capables de faire une prédiction à partir de deux éléments, l'état et l'action, et d'en déduire une forme de représentation du monde fondée sur des lois causales. Exemple : si je lâche ma tasse de café, qu'est-ce qui se passe ? Et la prochaine fois, il se repassera la même chose. « *L'humain est un excellent "world model"* », explique Michal Valko. Un LLM, beaucoup moins, car si aucune donnée qu'on lui a fait ingérer pendant l'entraînement ne donne de réponse, il est incapable de la déduire. Des *world models* performants pourraient non seulement se rapprocher de nos capacités à anticiper les résultats d'une action, mais aussi prédire tout un tas de choses qui nous sont inaccessibles : par exemple, l'effet de molécules introduites dans le corps humain, ce qui accélérerait de façon phénoménale la recherche contre le cancer ; ou encore les conséquences sur le climat d'une nouvelle technologie. De tels modèles cocheraient de nombreuses cases de l'idée que l'on se fait d'une IA générale.

Une technologie qui surpasserait tout ce qu'on connaît aujourd'hui

Mais on en est très loin. D'autant que la tendance actuelle n'est pas à la rupture de paradigme. « *Les grandes entreprises s'en tiennent à de légères modifications de la méthode qui marche et est utilisée par tout le monde* », explique Michal Valko. Elles sont frileuses à l'idée

d'investir dans des techniques qui n'ont pas fait leurs preuves et qui n'assurent pas d'application d'envergure. Les moyens alloués par Meta au projet de Yann Le Cun – et c'est pareil dans les principaux labos – restent très limités par rapport à ce qui est investi dans les LLM. Pourquoi en parle-t-on tant alors ? Pourquoi ces annonces tonitruantes et récurrentes qu'on y est presque, que ça va arriver demain ?

A lire aussi



Yann Le Cun : « L'IA a un rôle important à jouer dans le monde d'après-Covid »

Abonné

« *L'IA générale fait partie du rêve vendu au public mais surtout aux investisseurs, explique Olivier Cappé. Les immenses avancées obtenues jusque-là n'auraient pas été possibles sans des financements faramineux. Les investisseurs doivent y croire.* » Sam Altman et ses pairs ont donc tout intérêt à garder l'idée vivante. Quitte à en faire un peu trop, comme en 2020 quand le patron d'OpenAI déclarait : « *Nous avons fait une promesse aux investisseurs : une fois que nous aurons mis au point ce système d'intelligence générale, nous lui demanderons de trouver un moyen de générer un retour sur investissement pour vous.* » Patrick Pérez tempère : « *Même s'il est absurde d'être aussi péremptoire et précis s'agissant de quelque chose qui n'est pas défini et dont on est aussi loin, on aurait tort d'en faire un simple fantasme, il s'agit d'un horizon qui motive, un vrai objectif pour ces boîtes.* » En effet, la première

Annuler

ne serait-ce que la possibilité de servir aussi bien le grand public que les chercheurs ultrapolyyvalent, qu'elle pourrait vendre aussi bien au grand public qu'aux chercheurs (dans la santé, l'environnement et partout ailleurs). « *Un tel modèle pourrait servir de base à tout un tas d'applications spécialisées* », explique Patrick Pérez. La promesse d'une position de monopole comme les Gafam les aiment...

Mais n'y a-t-il pas quelques risques à jouer de ce rêve d'une technologie qui surpasserait tout ce qu'on connaît aujourd'hui ? « *Ce terme d'"IA générale", comme celui d'"IA" avant lui, fait beaucoup de mal à la recherche, déplore Michal Valko. Il vend une perspective difficilement atteignable, qui n'est pas suivie de résultats à la hauteur. Ça a contribué aux booms et chutes successifs des investissements dans le secteur depuis soixante-dix ans.* » Autrement dit, les investisseurs, appâtés par les promesses à répétition, pourraient s'impatienter de ne pas voir la grande rupture, puis se lasser et couper les ressources, comme ce fut le cas pendant ce qu'on a appelé les « hivers de l'IA ».

« Cognition incarnée »

En effet, si les avancées sont inéluctables, l'horizon de l'IA générale pourrait ne jamais être atteint. Il demeure quelques questions fondamentales. Une intelligence qui égale – voire dépasse – les capacités cognitives de l'humain est-elle accessible sans corps ? C'est

une vieille question qui pourrait se résumer à : peut-on connaître le monde sans en avoir l'expérience ? On en comprend des aspects bien sûr, mais, comme le dit joliment Yann Le Cun, « *votre chat en sait plus sur la gravité que n'importe quelle IA* ». Autrement dit, il faudra peut-être que les machines s'incarnent pour prétendre à cette « intelligence générale ». Ce serait donc vers la robotique qu'il faudrait se tourner, en dotant celles-ci de capteurs. C'est ce que défend l'Australien Rodney Brooks, pionnier en la matière, avec sa notion de « *cognition incarnée* ». Et c'est la piste que privilégie Yann Chevalayre : « *Un robot qui expérimente son environnement, c'est la seule solution pour aller vers une forme d'intelligence générale. Un modèle auquel on ne fournit que des images sera nécessairement borné.* » Mais il faudrait sans doute que l'interaction de la machine avec son environnement dure longtemps avant que l'expérience de l'environnement soit satisfaisante. En effet, la question du temps est loin d'être triviale dans cette histoire, comme le souligne le philosophe suédois Nick Bostrom dans « Superintelligence » (Dunod, 2017). Pour lui, il a fallu pour arriver à ce que nous considérons comme notre « intelligence » un très long processus d'évolution de la vie, conjointement au développement de systèmes très complexes d'émotions, de contrôles et de valeurs, qui se sont fabriqués en interaction avec un environnement physique, biologique et social, lui-même très varié et en constante évolution. Bref, même si on trouve un jour la solution technique pour créer et incarner une IA générale, il faudra peut-être du temps à une machine pour agir intelligemment. Comme nous au fond...

Annuler



Offrir cet article



Commenter



Nos lecteurs ont lu ensuite



Entretien Marine Tondelier : « Pourquoi je suis candidate à l'élection présidentielle »

Abonné



Candidat à Paris, Emmanuel Grégoire se pose en anti-Bolloré

Abonné



Chronique Campagne de recrutement des profs, comme un grand écart entre les slogans et la réalité

Abonné



Entretien « Bruce Springsteen nourrit un rapport très ambigu avec la gloire et la célébrité » : on a parlé du Boss avec Scott Cooper, réalisateur de « Springsteen : Deliver Me from Nowhere »

Abonné

Annuler



Décryptage Quand l'IA est utilisée pour évaluer les risques d'arrêt cardiaque

Abonné



Chronique Le roman réaliste à lire pour mieux appréhender la bulle de l'IA (et l'inévitable correction boursière à venir)

Abonné

Dans la rubrique Intelligence artificielle



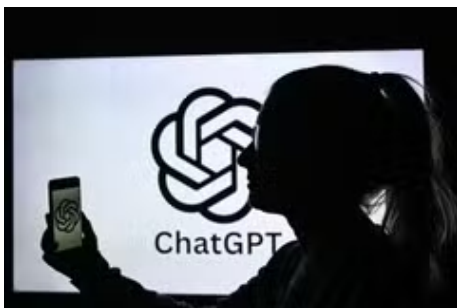
Décryptage OpenAI vs. Google : bienvenue dans la guerre des navigateurs web

Abonné



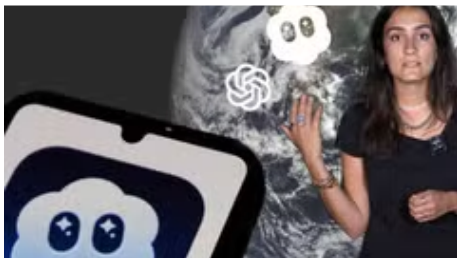
Chronique Le roman réaliste à lire pour mieux appréhender la bulle de l'IA (et l'inévitable correction boursière à venir)

Abonné

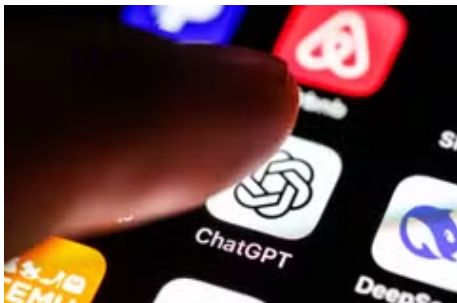


Informations obsolètes, détails inventés... Les IA comme ChatGPT ne sont pas fiables pour s'informer, alerte une étude européenne

Annuler



OpenAI : pourquoi Sora 2 est un gouffre énergétique



Atlas, le navigateur web d'OpenAI propulsé par ChatGPT pour concurrencer Google



Récit « Molière pouvait écrire une pièce en deux semaines. Avec l'IA, on a mis deux ans »

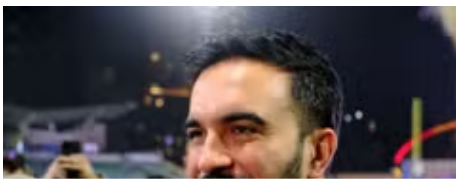
Abonné

Dans la rubrique Idées



Entretien Heide Goettner-Abendroth, anthropologue : « Personne ne s'est dit, je vais inventer le patriarcat ! »

Abonné



Carte blanche Mairie de New York : ce que dit l'ascension de Zohran Mamdani de la recomposition de la gauche américaine

Annuler



Interview Alice de Rochechouart, philosophe : « Le traitement réservé à Nicolas Sarkozy démontre que nous vivons dans une société de privilèges »

Abonné



Entretien Dans le nouvel « Astérix », le pirate noir n'a plus son « accent africain » : antiracisme ou excès de politiquement correct ?

Abonné



Rencontre Agnès Michaux a passé un an avec Huysmans. Elle en a tiré un petit chef-d'œuvre de biographie

Abonné



Carte blanche Dette française : « Il faut désormais qu'il y ait des adultes dans la pièce »

Sujets associés à l'article

Economie



BibliObs



Numérique



Annuler

Publicité

Cours de langue

Service partenaire

Cours d’italien

Test gratuit

Toutes nos langues

Cours d’allemand

Les univers

EcoloObs

BibliObs

Rue89

Annuler

Tendances

Culture

L’horoscope

Lab'O

Nouvelle ère

Dans l’actu

Budget

Gaza

Nicolas Sarkozy

Les services

Le magazine numérique

S’abonner / se désabonner

[Le Club Abonnés](#)

[Newsletters](#)

[Les mots croisés](#)

[La boutique](#)

[Les archives](#)

[Le plan du site](#)

[Bilan retraite](#)

[L’appli :](#)

[iOS | Android](#)

[Les sites du groupe](#)

[Le Monde](#)

[Télérama](#)

[Courrier international](#)

[Le HuffPost](#)

[Annuler](#)

[Le Monde diplomatique](#)

[Accès rapide](#)

[Programme TV gratuit et complet](#)

[La conjugaison](#)

[Mentions légales](#)

[|](#)

[CGV](#)

[|](#)

[Copyright](#)

[|](#)

[Publicité](#)

[|](#)

[Politique de confidentialité](#)

|

[Gestion des cookies](#)

|

[Aide / Contact](#)

© Le Nouvel Obs - 2024

Les marques ou contenus du site nouvelobs.com sont soumis à la protection de la propriété intellectuelle.

[Annuler](#)