

# EXPLORATION DE L'IA À TRAVERS CHATGPT : ENTRE SECRETS ET DÉFIS ÉTHIQUES

ÉDITION #11 — MARS 2024



# RÉSUMÉ D'ACTUALITÉ

Explorez les récentes avancées et découvertes dans le monde dynamique de ChatGPT et OpenAI. Des mises à jour passionnantes aux développements innovants, plongez dans les dernières actualités qui façonnent notre avenir numérique.

## Pourquoi Elon Musk a dû rendre Grok, sa réponse à ChatGPT, open source ?

OpenAI

Elon Musk

Grok

Elon Musk a annoncé qu'il rendrait son chatbot "Grok", destiné à concurrencer ChatGPT, open source, après avoir poursuivi OpenAI en justice pour avoir gardé sa technologie secrète. Grok, développé par la société d'IA xAI fondée par Musk, sera disponible gratuitement pour tous, contrairement à OpenAI, qui garde son code privé.

L'initiative d'Elon Musk intervient après avoir accusé ses cofondateurs d'OpenAI d'avoir manqué à leur promesse de rendre l'intelligence artificielle de la société publique. La décision de rendre Grok open source pourrait être interprétée comme un moyen pour Musk de montrer son engagement envers cette vision.

Grok avait été initialement réservé aux abonnés premium de la plateforme sociale X, mais sa libération permettra à n'importe qui de télécharger et d'utiliser le modèle. Cette démarche pourrait aider xAI à susciter l'intérêt des développeurs et à élargir sa base d'utilisateurs, fournissant ainsi des données pour améliorer sa technologie.

En rendant Grok open source, Musk emboîte le pas à Meta (anciennement Facebook), qui adopte une approche similaire avec ses modèles d'IA génératifs. Cela pourrait cependant le plonger dans le débat sur les avantages et les risques de donner accès à des modèles d'IA puissants à tout le monde.

**SOURCE :** Wired (11 Mars)

## Accord de partenariat entre Le Monde et OpenAI dans le domaine de l'intelligence artificielle

OpenAI

Le Monde

Partenariat

Le Monde a conclu un accord pluriannuel avec OpenAI, une société majeure dans le domaine de l'intelligence artificielle (IA), marquant ainsi le premier partenariat entre un média français et un acteur de cette industrie. Cet accord permettra à OpenAI d'utiliser le contenu du Monde pour entraîner ses modèles d'IA, notamment son outil ChatGPT, afin d'améliorer la pertinence des réponses fournies. En retour, Le Monde bénéficiera d'une source de revenus supplémentaires significative.

Ce partenariat est conçu comme un véritable partenariat à long terme, permettant aux deux parties de collaborer sur le développement de projets et de fonctionnalités utilisant l'IA. Il prévoit également que les références aux articles du Monde soient mises en évidence dans les réponses générées par ChatGPT, avec un lien hypertexte et le titre des articles utilisés comme références.

En plus de consolider le modèle économique du Monde, cet accord est une reconnaissance de la fiabilité du travail éditorial du journal. Il s'inscrit dans une série d'initiatives visant à protéger l'identité éditoriale du Monde tout en explorant les possibilités offertes par l'IA pour le journalisme.

**SOURCE :** Le Monde (13 Mars)

## Étude : ChatGPT confronté à 50 000 dilemmes du tramway

ChatGPT

Étude

Des chercheurs ont testé la capacité de ChatGPT à prendre des décisions morales en simulant plus de 50 000 scénarios du dilemme du tramway. Ces scénarios ont été générés par l'expérience de la Machine Morale, qui pose des questions complexes sur les choix que pourraient rencontrer les voitures autonomes. Les résultats montrent que, dans l'ensemble, les grands modèles de langage (LLM) comme ChatGPT prennent des décisions similaires à celles des humains, mais présentent parfois des différences subtiles.

Par exemple, certains modèles ont une préférence plus marquée pour la protection des piétons et des femmes, ce qui pourrait nécessiter un calibrage supplémentaire avant une utilisation réelle dans les voitures autonomes. De plus, il est souligné que les modèles d'IA sont souvent récompensés pour leur confiance dans leurs prédictions, ce qui pourrait influencer leurs décisions vers des choix plus intransigeants, et que la provenance des données d'entraînement peut introduire des biais, comme une préférence pour sauver les femmes.

Les chercheurs appellent à une meilleure compréhension des mécanismes de raisonnement des LLM pour garantir leur alignement avec les normes éthiques mondialement reconnues.

**SOURCE :** ARS Technica (13 Mars) — [Consulter l'étude](#)

## Fuite accidentelle : OpenAI dévoile-t-il GPT-4.5 Turbo ?

OpenAI

ChatGPT

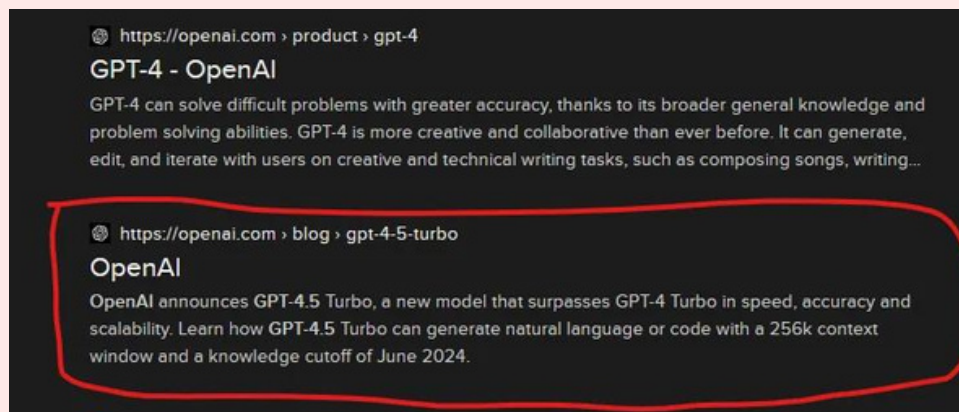
GPT-4.5 Turbo

Un article retiré du blog d'OpenAI a récemment fait allusion à un potentiel nouveau modèle d'intelligence artificielle, le GPT-4.5 Turbo. Bien que l'article ait été supprimé, des détails ont été indexés par les moteurs de recherche avant sa disparition, suscitant des spéculations parmi les utilisateurs de Reddit et Twitter.

Les informations disponibles suggèrent que le GPT-4.5 Turbo serait plus rapide, plus précis et plus évolutif que son prédécesseur, le GPT-4 Turbo. De plus, il semblerait que la nouvelle version dispose d'une base de connaissances jusqu'en juin 2024, indiquant une mise à jour importante par rapport au GPT-4 Turbo dont la base de connaissances s'arrête à avril 2023. Enfin, une fenêtre contextuelle plus grande de 256 000 jetons a été mentionnée, ce qui représente une augmentation par rapport aux 128 000 jetons du GPT-4 Turbo.

Cette fuite survient alors que le GPT-4 Turbo devient accessible à tous les utilisateurs gratuits de Microsoft Copilot, soulignant peut-être une transition imminente vers le GPT-4.5 Turbo.

**SOURCE :** O1Net (14 Mars)



## xAI publie le code de base du modèle Grok en open source, sans le code d'entraînement

xAI

Elon Musk

xAI, dirigé par Elon Musk, a décidé de rendre open source le code de base du modèle Grok AI, sans inclure toutefois le code d'entraînement. Ce modèle, comportant 314 milliards de paramètres, est désormais disponible sur GitHub sous licence Apache License 2.0, autorisant les cas d'utilisation commerciale. Bien que Grok-1 ait été entraîné sur une pile "personnalisée", les détails spécifiques n'ont pas été divulgués par l'entreprise.

Il convient de noter que la version open source du modèle ne comprend pas les connexions au réseau social X, contrairement à la version accessible aux utilisateurs Premium+ du réseau social. Plusieurs entreprises renommées, telles que Meta et Google, ont également ouvert une partie de leurs modèles d'IA. Certains développeurs d'outils alimentés par l'IA envisagent déjà d'intégrer Grok dans leurs solutions, comme l'a mentionné le PDG de Perplexity, Aravind Srinivas. Cette décision intervient alors que Musk est engagé dans une bataille judiciaire avec OpenAI, l'accusant de trahir l'objectif à but non lucratif de l'IA.

**SOURCE :** Techcrunch (14 Mars)

## Apple envisage un partenariat avec Google pour des fonctionnalités alimentées par Gemini sur les iPhones

Apple

Google

Gemini

Selon Bloomberg, Apple cherche à conclure un partenariat majeur avec Google pour exploiter le modèle d'IA Gemini sur les iPhones. Cette collaboration placerait Google dans une position dominante, car la société est déjà le fournisseur de moteur de recherche préféré sur les iPhones pour le navigateur Safari. Apple envisage de licencier la technologie d'IA de Google pour introduire des fonctionnalités alimentées par l'IA avec des mises à jour iOS plus tard cette année. En outre, la société a également eu des discussions avec OpenAI pour potentiellement utiliser des modèles GPT. Cette initiative survient alors qu'Apple ressent la pression de rattraper ses concurrents dans le domaine de l'IA, tels que OpenAI, Microsoft, Anthropic, et même Google. Bien que les propres modèles d'Apple pourraient alimenter certaines des fonctionnalités sur l'appareil dans la prochaine mise à jour logicielle iOS 18, la société explore également des partenariats avec des fournisseurs externes pour des cas d'utilisation d'IA générative.

**SOURCE :** Techcrunch (18 Mars)