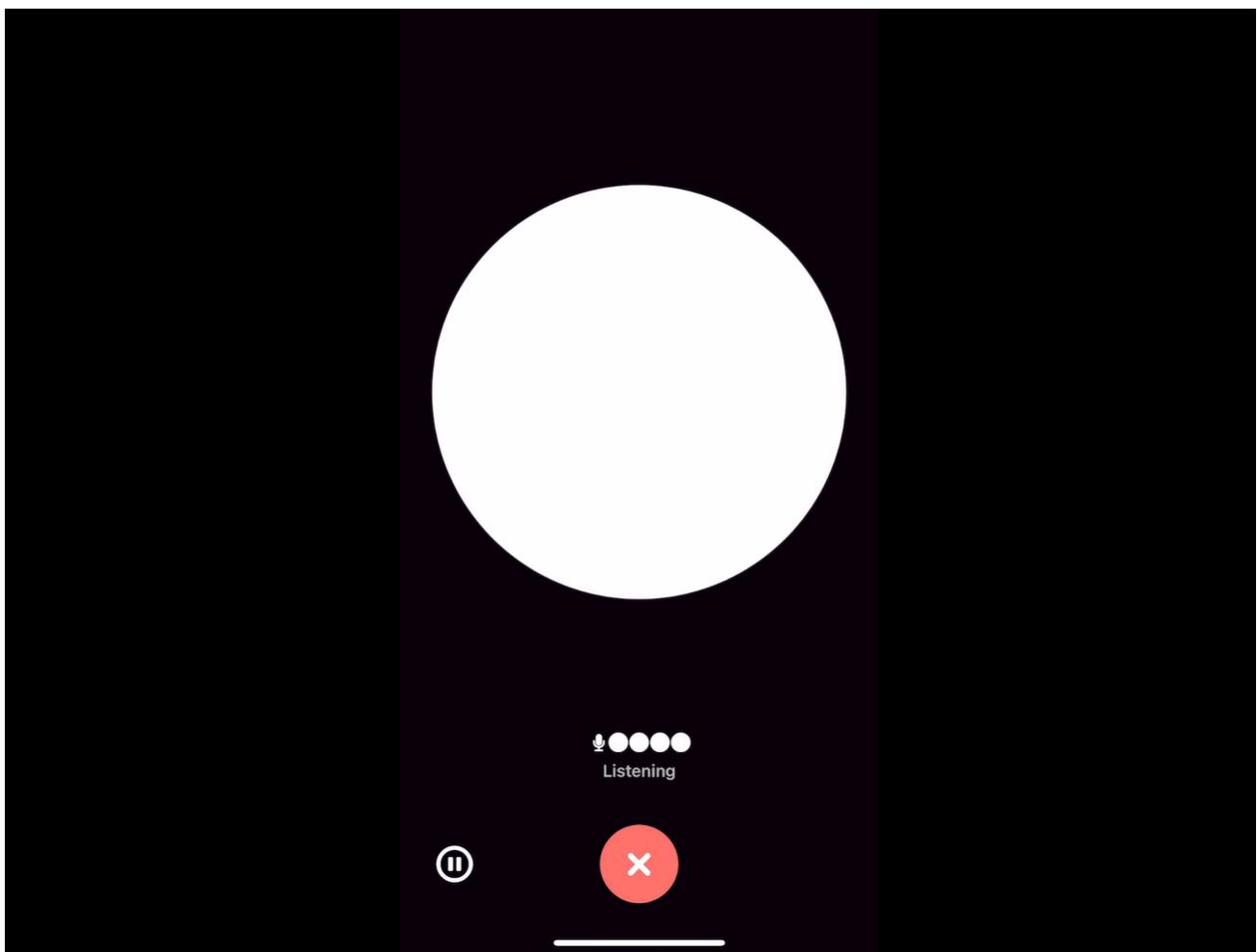


ChatGPT : le chatbot d'OpenAI devient plus humain

Les utilisateurs de ChatGPT peuvent désormais avoir une conversation vocale avec le chatbot ou encore lui montrer des photos pour étayer leur propos. Avec ces capacités supplémentaires, OpenAI cherche à proposer un échange qui ressemble de plus en plus à celui entre humains.



ChatGPT peut écouter quelqu'un et lui répondre. (Crédit : OpenAI)

« Prenez une photo d'un lieu lorsque vous voyagez et ayez une conversation en direct sur ce qui est intéressant à propos de ce point de repère. Lorsque vous êtes à la maison, prenez des photos de votre réfrigérateur et de vos placards pour savoir ce qu'il y a pour le dîner (et posez des questions

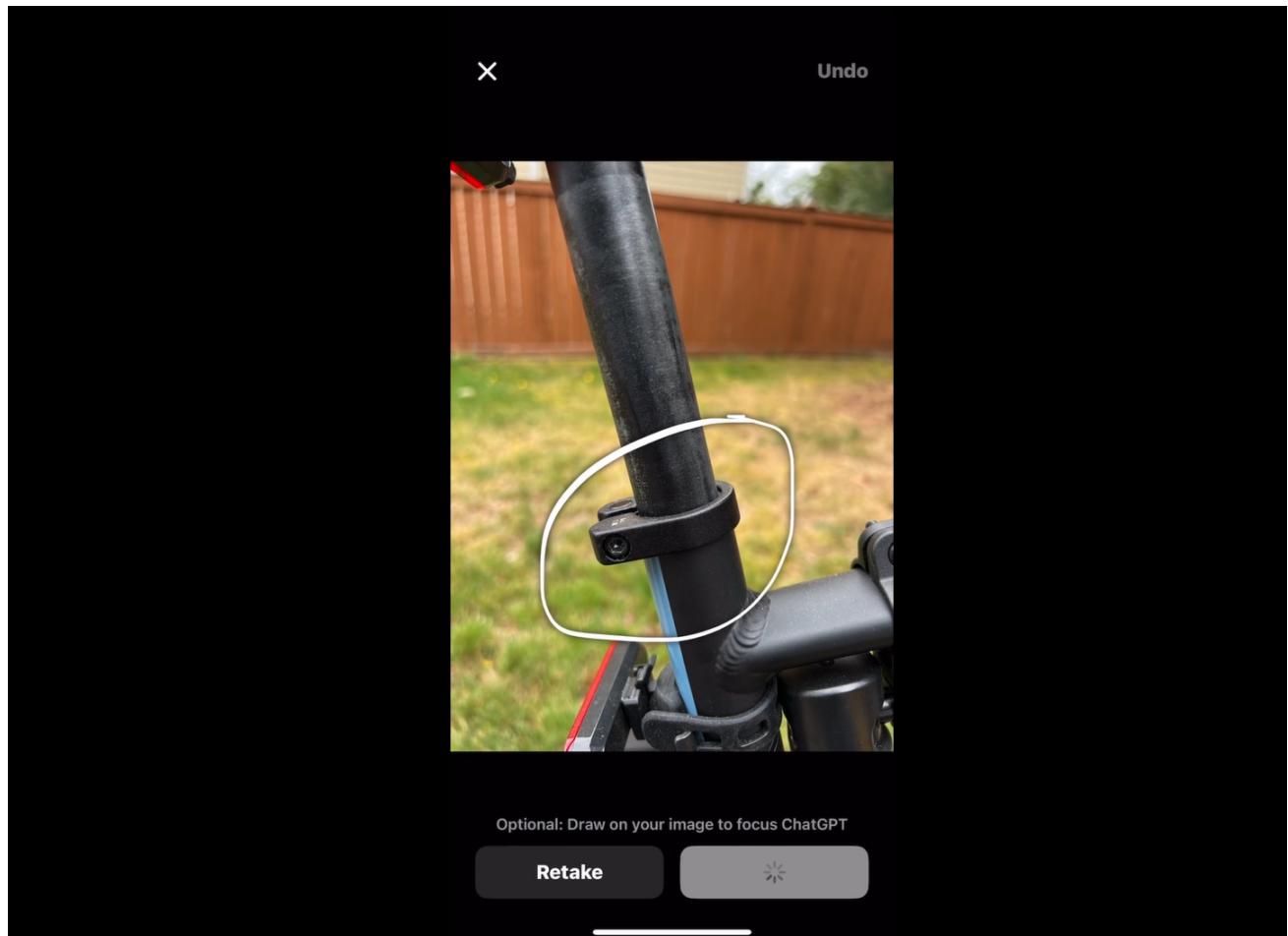
complémentaires pour obtenir une recette étape par étape). Après le dîner, aidez votre enfant à résoudre un problème de mathématiques en prenant une photo, en entourant le problème et en lui demandant de vous donner des conseils ». Voilà ce dont est désormais capable ChatGPT, le célèbre agent conversationnel d'OpenAI.

ChatGPT peut désormais voir, entendre et parler

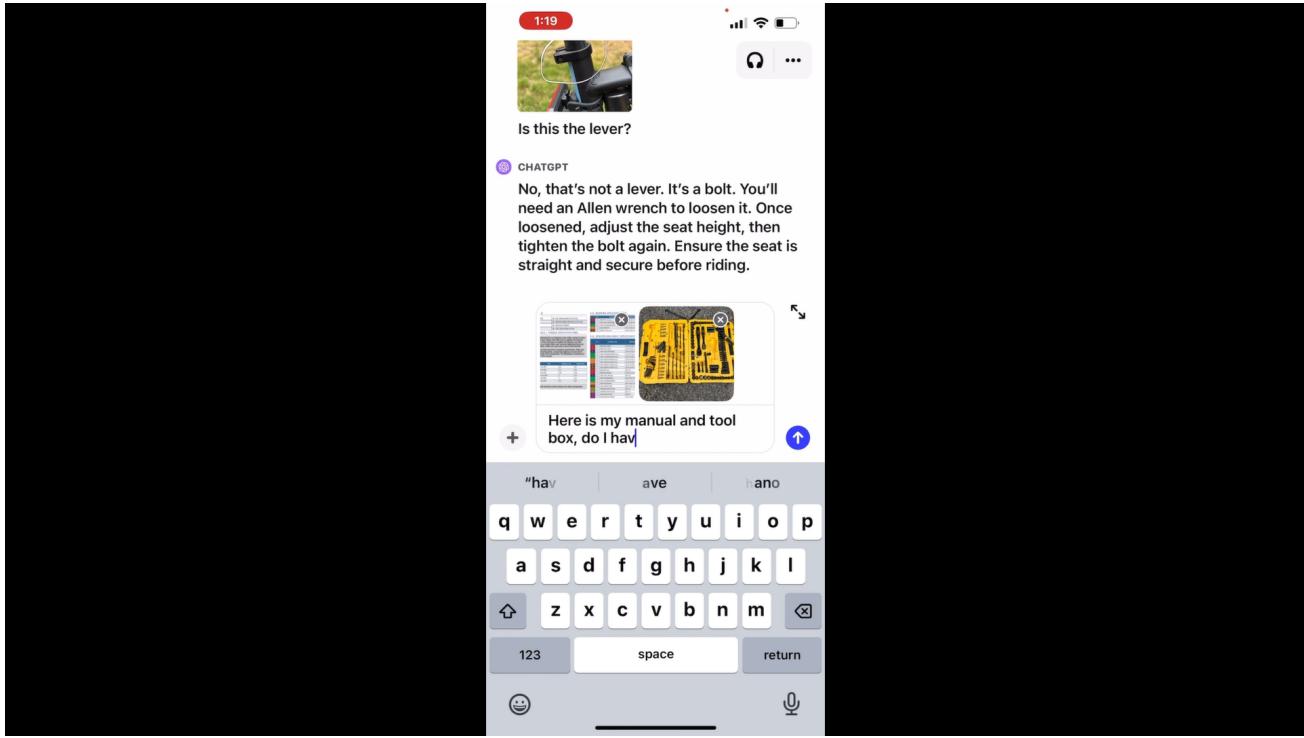
L'objectif est clair : faire de [ChatGPT](#) un outil inhérent à l'humain. La capacité vocale ajoutée à l'outil est alimentée par un modèle de synthèse vocale, capable de générer un son de type humain à partir d'un simple texte et de quelques secondes d'échantillon de discours réel. La firme précise avoir travaillé « avec des acteurs professionnels pour créer chacune des voix ». En complément, elle utilise également Whisper, son système de reconnaissance vocale open source, pour transcrire vos paroles en texte. Pour commencer à utiliser la voix, il suffit d'accéder à Paramètres → Nouvelles fonctionnalités sur l'application mobile et choisir les conversations vocales. L'utilisateur n'a plus qu'à appuyer sur le bouton du casque situé dans le coin supérieur droit de l'écran d'accueil et à choisir sa voix préférée parmi cinq voix différentes.

Autre changement de taille : il est possible de montrer une ou plusieurs images à ChatGPT. Prenant l'exemple du frigo, [OpenAI](#) raconte ceci : « Explorez le contenu de votre réfrigérateur pour préparer un repas » ou encore « analysez un graphique complexe pour des données liées au travail ». Il est même possible, pour se concentrer sur une partie spécifique de l'image, d'utiliser l'outil de dessin de l'application mobile. Pour l'utiliser, il suffit d'appuyer sur le bouton photo pour capturer ou choisir une image. Si vous êtes sur iOS ou Android, appuyez d'abord sur le bouton plus. La compréhension des images est assurée par les modèles GPT-3.5 et GPT-4 qui «

appliquent leurs compétences de raisonnement linguistique à un large éventail d'images, telles que des photographies, des captures d'écran et des documents contenant à la fois du texte et des images » est-il précisé.



Il est possible de poser une question à ChatGPT en ajoutant une photo pour mieux décrire le problème et entourer la zone sur laquelle doit se focaliser le chatbot. (Crédit : OpenAI)



Après avoir envoyé une photo et obtenu une réponse de ChatGPT, l'utilisateur peut renvoyer deux autres photos pour compléter son propos. (Crédit : OpenAI)

Des risques pointés du doigt

OpenAI tient toutefois à avertir ses utilisateurs que ces capacités présentent également de nouveaux risques, tels que la possibilité pour des acteurs malveillants d'usurper l'identité de personnalités publiques ou de commettre des fraudes. Elle précise que cette technologie est donc utilisée uniquement pour le chat vocal. Spotify s'appuie notamment sur cette capacité pour le projet pilote de sa fonction de traduction vocale, « qui aide les podcasteurs à étendre la portée de leurs récits en traduisant les podcasts dans d'autres langues avec la voix des podcasteurs eux-mêmes ».

En ce qui concerne les risques liés aux images, la firme précise que les modèles basés sur la vision présentent également des défis, « allant des hallucinations sur les personnes à la confiance dans l'interprétation d'images par le modèle dans des domaines à fort enjeu ». Le modèle a donc été

testé en amont par des membres de la fameuse « Red Team » d'OpenAI pour les risques dans des domaines tels que l'extrémisme et la pertinence scientifique, ainsi qu'avec un ensemble diversifié d'alpha testeurs. En complément, la société a également pris des mesures techniques pour limiter considérablement la capacité de ChatGPT à analyser et à faire des déclarations directes sur les personnes, « car ChatGPT n'est pas toujours précis et ces systèmes doivent respecter la vie privée des individus » affirme-t-elle.

Un déploiement progressif

Disponible uniquement auprès des utilisateurs Plus et Enterprise – au cours des deux prochaines semaines – cette mise à jour de l'outil devrait être accessible à d'autres types d'utilisateurs, y compris les développeurs, par la suite. OpenAI précise ainsi que la voix est disponible sur iOS et Android (opt-in dans vos paramètres) tandis que les images seront disponibles sur toutes les plateformes. Il est précisé que le modèle possède certaines limites, notamment linguistiques : « Le modèle est efficace pour transcrire des textes en anglais, mais ses performances sont médiocres dans d'autres langues, en particulier celles dont l'écriture n'est pas romaine. Nous déconseillons à nos utilisateurs non anglophones d'utiliser ChatGPT à cette fin ».

Source : www.lemondeinformatique.fr

→□ Accéder à CHAT GPT en cliquant
dessus