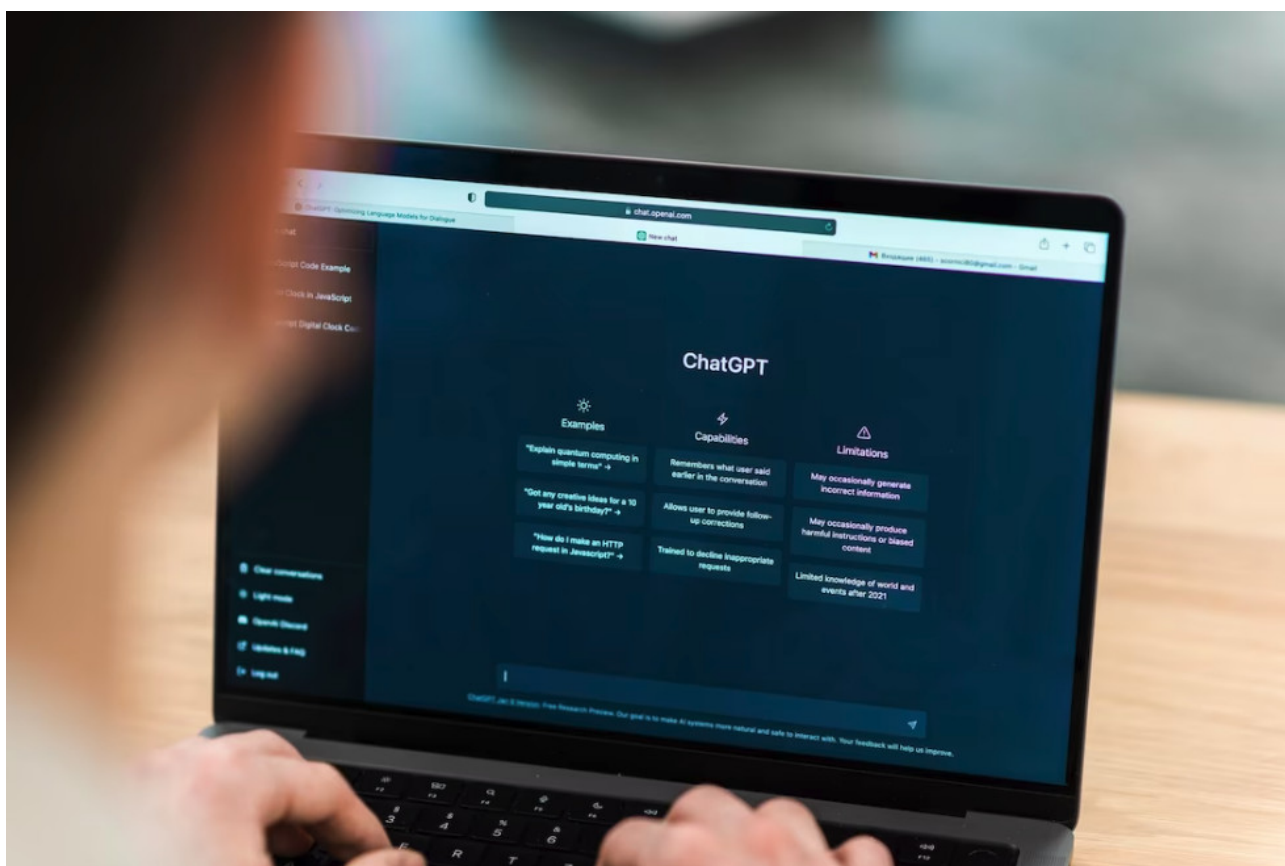


Valgo - Blog

<https://www.valgo.es/>

Nuevas funciones de ChatGPT: Ahora habla y comprende imágenes



OpenAI ha anunciado **mejoras en ChatGPT**, que incluyen nuevas **capacidades de voz e imagen**. Estas funciones revolucionarán la interacción con esta inteligencia artificial generativa y estarán disponibles para los usuarios de los planes Plus y Enterprise en las próximas semanas. OpenAI se ha centrado en la seguridad y utilidad al implementar estas características, reconociendo los desafíos y riesgos asociados con la voz y la imagen en la IA.

En lo que respecta a la voz, OpenAI ha enfocado sus esfuerzos en el **chat de voz**, creando capacidades de conversación por voz seguras y útiles. Esto permite a los usuarios solicitar información, escuchar historias o simplemente mantener una conversación con ChatGPT. Esta nueva interacción busca hacer la **experiencia más personal y accesible**. Para comenzar a utilizar esta función, los usuarios pueden dirigirse a la **Configuración en la aplicación móvil**, seleccionar "Nuevas funciones" y

optar por las "Conversaciones de voz". Luego, pueden elegir una de las 5 voces disponibles para su asistente. La tecnología subyacente se basa en un modelo de conversión de texto a voz que puede generar audio realista a partir de texto y una breve muestra de voz. Estas voces se han creado en colaboración con actores de voz profesionales. La tecnología Whisper de reconocimiento de voz de código abierto de OpenAI transcribe las palabras habladas en texto, permitiendo una experiencia de conversación más rica y natural con ChatGPT. Además, OpenAI está colaborando con organizaciones externas, como Spotify, para aplicar esta tecnología de voz en casos de uso adicionales, como la traducción de voz en podcasts.

La otra gran adición a ChatGPT es su **capacidad para comprender y discutir imágenes**. Los usuarios ahora pueden mostrar una o varias imágenes y obtener respuestas útiles basadas en lo que se ve en las imágenes. Esta capacidad tiene **aplicaciones diversas**, desde solucionar problemas técnicos hasta planificar comidas con los alimentos disponibles en el refrigerador o analizar gráficos complejos para obtener información relevante en el trabajo. La comprensión de imágenes se basa en los modelos GPT-3.5 y GPT-4 multimodales, que aplican su razonamiento lingüístico a imágenes de todo tipo, incluyendo fotografías, capturas de pantalla y documentos que combinan texto e imágenes. Esto promete hacer que las conversaciones sean más informativas y prácticas. OpenAI, ha priorizado la utilidad y seguridad al implementar esta función. Han tomado medidas técnicas para limitar la capacidad de ChatGPT para analizar y hacer afirmaciones sobre personas en las imágenes, respetando así la privacidad de los usuarios. Esto garantiza que la IA no realice análisis intrusivos o inapropiados de las personas en las imágenes compartidas, asegurando un uso seguro y ético de la tecnología de imágenes.

Estas mejoras implementadas a ChatGPT, con capacidades de voz que permiten conversaciones habladas y con la capacidad de comprender y discutir imágenes. Estas adiciones prometen **revolucionar la forma en que interactuamos con la IA** y están diseñadas con un fuerte enfoque en la seguridad y la utilidad. Los usuarios de los planes Plus y Enterprise podrán disfrutar de estas emocionantes funciones en las próximas semanas, lo que abrirá nuevas posibilidades en diversos campos.

Link to Original article: <https://www.valgo.es//w/valgo/34180/blog?elem=306774>