

Découvrez VALL-E : une IA capable de reproduire votre voix en seulement 3 secondes.

Titre HTML : Microsoft développe un modèle d'intelligence artificielle capable de reproduire votre voix

Microsoft a annoncé la création de VALL-E, un nouveau modèle d'intelligence artificielle qui est capable de générer un message audio avec votre voix. La technologie fonctionne en utilisant deux entrées : un texte écrit et l'enregistrement d'une voix disant n'importe quel texte. Avec seulement trois secondes d'enregistrement, VALL-E peut reproduire votre voix sur n'importe quel texte.

Titre HTML : Comment fonctionne VALL-E

VALL-E est un modèle de génération de parole à partir de texte (text-to-speech synthesis ou TTS, en anglais). Le modèle fonctionne en convertissant le texte écrit et le texte audio sous un format mathématique adapté. Grâce à son entraînement, VALL-E sait comment passer du message audio introduit à celui qu'on veut lui faire dire avec la voix du locuteur de départ. Le modèle doit finalement rendre un fichier audio que le public peut comprendre, un message audio.

Titre HTML : Les performances de VALL-E

Les chercheurs de Microsoft annoncent que VALL-E dépasse les performances actuelles en termes de réalisme et de ressemblance avec la voix de la personne. Les performances ont été comparées à celles de l'un des derniers modèles de génération de parole YourTTS. Sur le naturel ou le taux d'erreur par rapport au script, VALL-E a produit de meilleurs

résultats.

Titre HTML : VALL-E ne fonctionne qu'en anglais

Bien que VALL-E ait été entraîné sur un corpus de 60 000 heures d'enregistrement en anglais avec 7 000 personnes différentes, il n'est pas encore prêt pour une utilisation dans une autre langue. Le modèle est également moins performant avec des accents qui n'étaient pas représentés dans la base de données d'entraînement.

Titre HTML : Les dangers de deepfakes vocaux

Ces technologies, qui permettent la génération automatisée de contenus audio et visuels, font craindre la propagation de deepfakes sur les réseaux sociaux. Il est facile de corriger un message où un mot a été écorché ou de créer automatiquement la traduction d'un discours. Cependant, cela pourrait se faire au détriment de la qualité.

Source : [numerama.com](https://www.numerama.com)

→ **Accéder à [CHAT GPT](#) en cliquant dessus**