

ReadSpeaker® speechMaker Studio

ReadSpeaker speechMaker Studio (speechMaker Studio) est une application de synthèse vocale (Text-to-Speech, TTS) accessible via un navigateur, permettant de convertir des textes écrits en fichiers audio téléchargeables.

Description du produit

ReadSpeaker speechMaker Studio (speechMaker Studio) est une application SaaS (Software as a Service) accessible via un navigateur, permettant à plusieurs utilisateurs licenciés au sein d'un compte client de créer, gérer et convertir des textes saisis par l'utilisateur en fichiers audio téléchargeables à l'aide de la technologie de synthèse vocale (TTS) de ReadSpeaker.

L'accès à speechMaker Studio nécessite une authentification par identifiant et mot de passe. Un compte client comprend un nombre défini de licences utilisateur tel que spécifié dans la Confirmation de commande. Les utilisateurs disposant du rôle approprié (Propriétaire ou Administrateur) peuvent inviter des utilisateurs supplémentaires dans les limites des licences attribuées au compte.

Pour générer un fichier audio, l'utilisateur saisit ou colle du texte dans l'application, sélectionne la langue et la voix, choisit le format de sortie audio et ajuste les paramètres vocaux disponibles tels que le débit, la hauteur et le volume. L'utilisateur peut prévisualiser l'audio généré dans l'application. L'aperçu audio ne consomme pas de crédits. Les crédits sont déduits uniquement lors de l'export d'un fichier audio.

La génération audio est effectuée à l'aide de la technologie de synthèse vocale neuronale (TTS) de ReadSpeaker. Les fichiers audio exportés sont téléchargés sur l'appareil local de l'utilisateur.

Fonctionnalités

Tableau de bord

Après authentification, l'utilisateur est redirigé vers le tableau de bord. Le tableau de bord fournit une vue d'ensemble de :

- Activité récente au sein du compte client
- Notifications système
- Informations relatives au compte

Les informations affichées peuvent varier en fonction du rôle attribué à l'utilisateur.

Bibliothèque de contenu

Le contenu textuel peut être stocké dans la bibliothèque de contenu de speechMaker Studio.

Un fichier contenant du contenu textuel est appelé script. Un script peut contenir un ou plusieurs blocs de texte.

Les scripts peuvent être organisés en dossiers. Les dossiers et les scripts peuvent être regroupés en projets.

Tous les utilisateurs au sein d'un compte client peuvent accéder à l'ensemble du contenu stocké dans la bibliothèque de contenu. Les utilisateurs disposant du rôle Lecteur ont un accès en lecture seule et ne peuvent ni créer, ni modifier, ni supprimer du contenu.

Synthèse vocale

speechMaker Studio convertit les textes fournis par l'utilisateur en audio à l'aide de la technologie Neural Premium TTS de ReadSpeaker.

Les langues et voix disponibles dans l'application dépendent du pack vocal activé pour le compte client. La liste des langues et voix activées est affichée dans la [Galerie de voix](#) au sein de l'application.

Studio de contenu

L'ensemble du contenu est stocké dans le compte client et est accessible conformément aux rôles attribués aux utilisateurs.

La sélection d'un script depuis la bibliothèque de contenu ou le tableau de bord l'ouvre dans le Studio de contenu.

Chaque script peut contenir un ou plusieurs blocs de texte.

Les blocs de texte peuvent disposer de paramètres individuels pour :

- [Voix de lecture](#)
- [Débit \(vitesse\)](#)
- [Hauteur de voix](#)
- [Volume](#)

L'utilisateur peut ajuster la durée des pauses entre les blocs de texte ainsi que la durée du silence initial optionnel au début de l'audio, afin d'aligner les messages vocaux sur des fenêtres temporelles prédéfinies, par exemple.

Les outils suivants sont disponibles pour un contrôle détaillé du rendu vocal :

- [Éditeur de prononciation](#)
- [Éditeur d'intonation](#)
- [Outil de pause](#)

Voix de lecture ajustable

Une voix de lecture est sélectionnée parmi la liste des voix activées pour le compte client. Différentes voix peuvent être attribuées à des blocs de texte individuels au sein d'un même script. Un script peut ainsi contenir plusieurs locuteurs.

Débit réglable

Le débit peut être ajusté au moyen d'un curseur sur la plage prise en charge par la voix sélectionnée, d'un débit plus lent à un débit plus rapide par rapport au débit par défaut.

Hauteur de voix réglable

La hauteur peut être ajustée au moyen d'un curseur sur la plage prise en charge par la voix sélectionnée, d'une hauteur plus basse à une hauteur plus élevée par rapport à la hauteur par défaut. La hauteur peut être ajustée indépendamment pour chaque bloc de texte.

Volume réglable

Le volume peut être réglé au moyen d'un curseur sur la plage prise en charge par la voix sélectionnée, d'un volume plus faible à un volume plus élevé par rapport au volume par défaut. Le volume peut être réglé indépendamment pour chaque bloc de texte.

Aperçu : écouter l'audio avant export

L'audio généré peut être prévisualisé dans l'application avant export. L'aperçu est disponible pour :

- Le script complet
- Les blocs de texte individuels
- Les modifications de prononciation
- Les modifications d'intonation

L'aperçu audio ne consomme pas de crédits. Les crédits sont déduits uniquement lors de l'export de l'audio.

Éditeur de prononciation

L'éditeur de prononciation permet de modifier la prononciation d'un mot au sein d'un script. La prononciation peut être modifiée selon l'une des méthodes suivantes :

- Orthographe alternative - La prononciation d'un mot peut être remplacée en utilisant une orthographe différente lorsque la prononciation TTS par défaut ne correspond pas à la prononciation souhaitée.
- Transcription phonétique - La prononciation d'un mot peut être définie à l'aide de symboles phonétiques sélectionnés depuis un clavier IPA numérique.

Ces modifications de prononciation s'appliquent uniquement à l'occurrence spécifique du mot dans le script.

Les modifications de prononciation peuvent être prévisualisées avant application.

Éditeur d'intonation

L'intonation peut être modifiée au niveau de la phrase en ajustant la hauteur relative et/ou la durée des mots individuels au sein de la phrase.

L'éditeur d'intonation affiche la phrase sélectionnée sous forme de courbe, chaque mot étant représenté par un point pouvant être repositionné afin de modifier la hauteur. Une valeur de durée peut être attribuée à chaque mot afin de modifier sa longueur.

Les ajustements de hauteur et de durée sont relatifs au rendu par défaut de la voix sélectionnée. Des ajustements identiques peuvent produire des résultats différents selon les caractéristiques de la voix sélectionnée.

Les modifications d'intonation peuvent être prévisualisées avant application.

Outil de pause

Des pauses peuvent être insérées à tout endroit dans le texte. La durée des pauses est sélectionnée parmi des valeurs prédéfinies au sein de l'application. Les pauses insérées sont indiquées visuellement dans l'éditeur.

Export audio

Les fichiers audio sont téléchargés sur l'appareil de l'utilisateur lors de l'export.

L'audio peut être exporté à différents niveaux :

- **Script** - Export d'un script unique depuis le Studio de contenu ou la bibliothèque de contenu.
- **Projet** - Export d'un projet depuis la bibliothèque de contenu.

L'export d'un projet génère une archive ZIP reproduisant la structure des dossiers du projet dans la bibliothèque de contenu. Seuls les dossiers contenant des scripts sont inclus dans l'archive exportée.

Les crédits sont déduits lors de l'export.

Galerie de voix

La Galerie de voix affiche les langues et les voix activées pour le compte client. Les voix peuvent être filtrées à l'aide des fonctions de recherche et de filtrage disponibles. La liste des voix est classée par langue et par nom de voix.

Chaque entrée de voix comprend un avatar, un drapeau indiquant la langue, ainsi qu'un texte d'exemple pouvant être lu afin de prévisualiser la voix.

Les voix peuvent être marquées comme favorites. Les voix marquées comme favorites sont affichées en haut de la liste de sélection des voix dans le Studio de contenu ainsi que dans la Galerie de voix.

Des langues et voix supplémentaires peuvent être ajoutées à la Galerie de voix du compte client.

Crédits de production

speechMaker Studio utilise un modèle basé sur des crédits pour l'export audio.

Les crédits sont déduits uniquement lors de l'export de l'audio. L'aperçu audio dans l'application ne consomme pas de crédits.

Un crédit est facturé par caractère lors de l'export d'un fichier audio. Les caractères sont comptabilisés hors espaces. Les caractères introduits ou modifiés via des modifications de prononciation sont inclus dans le décompte des caractères.

Les comptes clients comprennent généralement une allocation initiale de crédits à des fins de test, telle que définie dans la Confirmation de commande applicable.

Des packs de crédits supplémentaires peuvent être achetés par les Administrateurs du compte. Les quantités de crédits et les périodes de validité sont définies dans la Confirmation de commande applicable. Les achats de crédits sont soumis à confirmation par ReadSpeaker avant que les crédits ne soient appliqués au compte client.

Consultation du solde restant

Le nombre de crédits restants, représentant la quantité de texte (mesurée en nombre de caractères hors espaces) pouvant encore être convertie en audio, est affiché dans la barre supérieure de l'interface speechMaker Studio.

Le solde restant est visible pour tous les rôles utilisateurs.

Des crédits supplémentaires peuvent être achetés à tout moment par les Administrateurs du compte.

Lorsque le solde de crédits atteint zéro, l'export audio n'est plus disponible tant que des crédits supplémentaires n'ont pas été appliqués au(x) compte(s) client(s).

Utilisateurs et rôles

Un compte client comprend cinq licences utilisateur. Des licences utilisateur supplémentaires peuvent être ajoutées.

Les utilisateurs au sein d'un compte speechMaker Studio peuvent se voir attribuer les rôles suivants :

- Propriétaire - Attribué au premier utilisateur créé pour le compte client. Le compte client doit disposer d'un Propriétaire. Le Propriétaire possède les mêmes autorisations qu'un Administrateur.
- Administrateur - Peut créer, modifier et supprimer l'ensemble du contenu dans la bibliothèque, inviter des utilisateurs, attribuer des rôles utilisateurs et acheter des crédits speechMaker Studio.
- Éditeur - Peut créer, modifier et supprimer l'ensemble du contenu dans la bibliothèque.
- Lecteur — Dispose d'un accès en lecture seule au contenu. Les Lecteurs ne peuvent ni créer, ni modifier, ni supprimer du contenu. Les Lecteurs peuvent prévisualiser l'audio dans le Studio de contenu.

Spécifications techniques

Compatibilité avec tous les appareils

speechMaker Studio est une application accessible via un navigateur, prise en charge par les versions actuelles des principaux navigateurs web.

Solution 100 % cloud; pas de téléchargement ni d'installation

speechMaker Studio est entièrement hébergé côté serveur. Aucun téléchargement ou installation de logiciel n'est requis pour son utilisation. L'application est hébergée par ReadSpeaker et accessible via un navigateur web pris en charge.

Hébergement du produit

speechMaker Studio est hébergé par ReadSpeaker dans un environnement hautement sécurisé, doté d'un niveau élevé de redondance et de basculement, et présentant les caractéristiques suivantes :

- Double équilibrage de charge et basculement redondants de niveau entreprise
- Réseau 100 % Gigabit (en interne et vers Internet)
- Matériel serveur IBM avec moteur de virtualisation
- Alimentation électrique redondante
- Architecture hautement évolutive et à haute disponibilité
- Infrastructure respectueuse de l'environnement
- Stockage de type Storage Area Network (SAN) via Gigabit iSCSI
- Alimentation de secours par onduleur (UPS) et groupe électrogène diesel
- Surveillance vidéo et audio 24 h/24, 7 j/7
- Sécurité de niveau 3
- Système automatique de lutte contre l'incendie à base d'Argonite
- Surveillance automatique des serveurs
- Sauvegarde automatique quotidienne

Formats d'entrée pris en charge

speechMaker Studio accepte une entrée en texte brut. Le texte doit être saisi directement dans l'Éditeur, soit en tapant, soit en copiant-collant depuis une source externe.

Limite de taille du texte

speechMaker Studio dispose d'une limite par défaut de 20 000 caractères, hors espaces, pour la conversion texte vers la parole par fichier audio exporté produit.

Formats audio pris en charge

Les fichiers audios exportés sont disponibles dans les formats suivants :

- MP3 (24, 48 et 128 Kbits/s, 22,1 KHz et 44,1 KHz ; 192 Kbit/s, 44,1 KHz)
- Ogg Vorbis (débit variable Q=0,2)
- PCM (8 bit/16 KHz, 16 bit/8 KHz, 16 bit/16 KHz, 16 bit/44,1 KHz)
- IMA-ADPCM (4 bit/8 KHz)
- WAV (U-law 16 bit/8 KHz, A-law 16 bit/8 KHz)
- WAV (8 bit/16 KHz, 16 bit/8 KHz, 16 bit/16 KHz, 16 bit/44,1 KHz, 16 bit/48 KHz)

Les formats disponibles pour un compte client sont définis dans la Confirmation de commande applicable.

Types de voix

speechMaker Studio prend actuellement en charge les packs vocaux Neural Premium de ReadSpeaker.

Éléments du produit

Application SaaS

Le composant principal de speechMaker Studio est l'application SaaS (Software as a Service) ReadSpeaker speechMaker Studio, hébergée par ReadSpeaker.

L'application est accessible aux utilisateurs licenciés via une connexion authentifiée.

Identifiants de connexion

L'accès à speechMaker Studio est fourni via des identifiants de connexion propres à l'utilisateur, associés à un compte client licencié. Aucun composant logiciel supplémentaire n'est requis pour les utilisateurs finaux.

Assistance et documentation

L'assistance technique est fournie par ReadSpeaker. Les coordonnées et les horaires d'assistance sont disponibles à l'adresse suivante :

<https://www.readspeaker.com/fr/support/>.

Options

Des fonctionnalités optionnelles peuvent être ajoutées à speechMaker Studio. Les informations tarifaires sont disponibles auprès de votre responsable de compte.

Taille maximale supérieure pour la saisie de texte

La limite par défaut de 20 000 caractères par fichier audio produit peut être augmentée sur une base spécifique à chaque client.

Exigences et limitations

Connexion Internet

Une connexion Internet active est requise pour utiliser speechMaker Studio.

Taille par défaut pour la saisie de texte

La taille maximale par défaut pour la conversion texte vers la parole par fichier audio produit est de 20 000 caractères hors espaces. Cette limite peut être augmentée sur une base spécifique à chaque client ; voir la section Options ci-dessus.

Export audio

Le client peut utiliser les fichiers audio exportés uniquement dans le cadre défini dans la Confirmation de commande.

Certaines voix peuvent être soumises à des restrictions d'utilisation. Les restrictions applicables sont affichées dans l'interface d'export lorsque des voix concernées sont incluses dans l'export. Le client est responsable de s'assurer que l'audio exporté est utilisé conformément aux restrictions d'utilisation applicables.

Types de voix

speechMaker Studio prend en charge les voix Neural Premium standard (« off-the-shelf ») de ReadSpeaker. Les voix Neural Standard et les voix personnalisées (Custom Voices) de ReadSpeaker seront prises en charge dans une version ultérieure.

Les types de voix activés pour un compte client dépendent du pack vocal configuré.

Il n'est pas possible d'utiliser des voix de types différents au sein d'un même script. Tous les blocs de texte d'un script unique doivent utiliser des voix du même type.

Prise en charge de HTML5 et CSS3

Un navigateur prenant en charge HTML5 et CSS3 est requis pour accéder à l'ensemble des fonctionnalités de speechMaker Studio.

Modèle tarifaire

speechMaker Studio est soumis à des frais d'installation ainsi qu'à des frais fixes mensuels ou annuels de service et de maintenance.

La production audio est facturée en fonction du nombre de caractères convertis, conformément aux tarifs spécifiés dans la Confirmation de commande applicable.

Utilisation prévue

speechMaker Studio peut être utilisé uniquement pour convertir des segments de texte en fichiers audio conformément à la présente spécification produit. Toute utilisation en dehors du périmètre défini dans cette spécification produit n'est pas considérée comme une utilisation prévue.

Clause de non-responsabilité

Les fonctionnalités décrites dans la présente spécification produit sont garanties uniquement si les exigences énoncées sont respectées. Les spécifications et fonctionnalités peuvent être modifiées par ReadSpeaker sans préavis.