

Prendre en main la transcription

Studios recevant une version originale sans script

Transcrire un épisode sans script

Un programme arrive sans texte, ou avec un script qui ne correspond plus au montage.

5 minutes de réglage, puis l'attente

01 Donner la source

La transcription travaille sur **l'audio**, pas sur l'image. Collez l'adresse du média, ou choisissez-le dans **Médias**. Un fichier audio seul suffit et va plus vite qu'une vidéo entière.

URL AUDIO À TRANSCRIRE (MP3 RECOMMANDÉ — LÉGER)
 ou

Le champ d'adresse de l'audio à transcrire, et le bouton de fichier local.

02 Choisir la langue, ou ne pas choisir

Détection auto est le réglage recommandé, et il l'est vraiment : sur un audio qui mêle deux langues (du wolof et du français dans la même scène) il donne souvent un meilleur résultat que forcer une langue.

Forcer une langue n'est utile que si la détection se trompe, ce qui arrive sur les programmes très courts.

LANGUE

 **Détection auto (recommandé)**

Le menu Langue, ouvert sur la détection automatique.

03 Régler le niveau et les événements

Soignée est plus lente et plus juste ; **Standard** suffit pour un premier repérage. La case **Inclure les événements** ajoute les rires, la musique et les réactions, utile en doublage, où ils occupent une place réelle dans la bande.

MODÈLE

Soignée (recommandé)

 Pour un audio mêlant wolof et français, la détection auto donne souvent un meilleur résultat que de forcer une langue.

☒ Inclure les événements (pires, musique, réactions) dans la transcription

Le choix du modèle, Soignée ou Standard, et la case des événements.

04 Travailler sans réseau

Un **moteur hors ligne** s'installe depuis **Moteurs locaux**, en un clic. Le contenu ne quitte plus le poste et rien n'est facturé à la minute : c'est ce qu'il faut pour un programme sous embargo, une clause de confidentialité, ou une cabine sans connexion.

En contrepartie, **aucune séparation des locuteurs** : toutes les répliques arrivent sur un seul personnage, à répartir ensuite. Et c'est le processeur de la machine qui travaille — comptez plus long que la durée du programme.

Moteurs locaux

Deux fonctions tournent sans réseau une fois installées : la **transcription hors ligne** et la **séparation voix / M&E**. Elles ne sont pas dans l'installateur — ce sont des centaines de mégaoctets qui ne servent pas à tout le monde.

Transcription hors ligne ✓ prête

Le contenu ne sort pas de la machine et rien n'est facturé à la minute. En contrepartie, **aucune séparation des locuteurs** : toutes les répliques arrivent sur un seul personnage.

Grand modèle rapide (recommandé) · 547 Mo

Le meilleur rapport qualité/temps. Tient les langues africaines. C'est celui à prendre si vous hésitez.

installé

Retirer

Grand modèle complet · 1030 Mo

La meilleure reconnaissance, nettement plus lent. Pour un épisode difficile, pas pour une saison.

Installer

Petit modèle · 181 Mo

⚠ Faible sur les langues africaines et sur l'audio difficile. À réserver à un poste modeste, en connaissance de cause.

Installer

Séparation voix / M&E non installée

Sépare les dialogues d'origine de la musique et des effets. Le M&E obtenu est un vrai stem, livrable.

Environ 1 Go à télécharger, plusieurs minutes. Python 3.14 détecté.

Installé depuis le dépôt officiel de ses auteurs ; ses poids sont téléchargés par Demucs lui-même à la première séparation.

Installer

Fermer

La fenêtre Moteurs locaux : la transcription hors ligne et la séparation voix / M&E, à installer sur le poste.

Prendre en main la transcription · Axeho

2 / 3

05 Lancer, et laisser tourner

L'écran annonce la phase en cours et le temps écoulé, et estime la durée restante à partir de vos transcriptions précédentes. Vous pouvez interrompre à tout moment.

Transcription

Transcrits depuis une URL, un fichier hébergé, ou un fichier local (audio de tout type — mp3, wav, m4a, aac, ogg, flac, opus... — ou vidéo). Regarde la vidéo (MP4) pour corriger.

D'où vient le texte ?

Je n'ai pas de texte
Le faire écrire depuis l'audio ou la vidéo

J'ai un script
SRT, VTT, DFX, TTAL, TTML, TXT, Excel

J'ai déjà une transcription
Coller un résultat au format JSON

URL AUDIO À TRANSCRIRE (MP3 RECOMMANDÉ — LÉGER)

ou

URL VIDÉO POUR LE LECTEUR (MP4 — OPTIONNEL)

⚡ Connexion lente ? Télécharge la vidéo une fois, puis ouvre-la en « Vidéo locale » : lecture et navigation instantanées, sans réseau.

💡 Mets le MP3 dans le champ Audio (transcription rapide) et le MP4 dans le champ Vidéo (pour voir qui parle). Les deux sont synchronisés par les timecodes. Si tu n'as qu'une vidéo, mets-la dans le champ Audio.

MODÈLE

Soignée (recommandé)

LANGUE

Détection auto (recommandé)

💡 Pour un audio mêlant wolof et français, la détection auto donne souvent un meilleur résultat que de forcer une langue.

☒ Inclure les événements (rires, musique, réactions) dans la transcription

➤ Diagnostic et réponse brute

L'écran de transcription : source, langue et niveau de rendu avant de lancer le travail.

06 Récupérer le texte

Les répliques arrivent **timecodées** dans le script, prêtes à être adaptées. C'est le point de départ de tout le reste : la détection, la bande, les voix.

TC	PERSO	TEXTE	ACTIONS
▶ 0:00.40 0:05.40 5.00s	SAMARA	The night before my mating ceremony, my twin sister took my partner's bed.	↑ ↓ ↺ ↻ 🔍 🗑
+ Insérer ici			
▶ 0:06.10 0:08.20 2.10s	FINN	You knew. You always knew.	↑ ↓ ↺ ↻ 🔍 🗑
+ Insérer ici			
▶ 0:08.90 0:12.40 3.50s	SAMARA	Now she can marry that monster in my place.	↑ ↓ ↺ ↻ 🔍 🗑
+ Insérer ici			
▶ 0:13.20 0:16.00 2.80s	ALARIC	Nobody asked you anything.	↑ ↓ ↺ ↻ 🔍 🗑
+ Insérer ici			
▶ 0:17.00 0:20.60 3.60s	SIENNA	He looked broken, as if he had not slept.	↑ ↓ ↺ ↻ 🔍 🗑
+ Insérer ici			
▶ 0:21.40 0:24.00 2.60s	SAMARA	Maybe. I do not know any more.	↑ ↓ ↺ ↻ 🔍 🗑
+ Insérer ici			

Les répliques transcrites, timecodées et attribuées à leur personnage, dans l'éditeur de script.

Ce tutoriel en ligne : <https://axeho.studio/tutos-transcription.html>