

+ Présentation de l'outil Whisper + AI

15 janvier 2024

Géoteca, Olympes de Gouges

Brian Plaisant et Mélody Dumont

Présentation de l'outil whisper



Qu'est-ce que
l'outil Whisper ?

Quels sont les
prérequis ?

Comment installer
Whisper ?

Comment utiliser
Whisper et
l'optimiser ?

Retour utilisation –
Le cas d'entretiens
japonais

Qu'est-ce que l'outil Whisper AI ?



- C'est un outil de **retranscription automatique** de pistes audio développé par OpenAI.
 - En entrée → **une piste audio**.
 - En sortie → **fichiers textes** avec ou sans horodatage.
 - Modèles entraînés sur des centaines de milliers d'heures de pistes audio sous-titrées.
 - Différents modèles existent, adaptés en fonction des besoins et attentes de l'utilisateur.
- La qualité varie selon :
 - La **représentation des langues** spécifiques lors de l'entraînement des modèles;
 - **Le modèle** choisi;
 - La **qualité de l'audio**.
 - L'outil est **gratuit**, relativement **facile à installer**.
 - **L'outil n'est pas parfait**. Il nécessitera toujours une **supervision humaine**.



Éléments de base pour ce tutoriel

- **FFMPEG** : Bibliothèque de logiciels libres permettant à l'ordinateur de traiter n'importe quel flux audio ou vidéo (<https://ffmpeg.org/>).
- **Anaconda ou miniconda** : permet d'installer l'application dans un environnement définie (<https://docs.conda.io/projects/miniconda/>).

En plus

- **Pytube** : Si vous voulez retranscrire ou tester sur des vidéos Youtube.

Les prérequis : installation de FFMPEG



Sur Ubuntu et distribution similaire

- directement dans un terminal :

```
sudo apt-get install ffmpeg
```

- Allez dans le terminal et tapez

```
ffmpeg --help
```

Si l'aide fonctionne, c'est bon!

Sur Windows

- Téléchargez [ffmpeg](#)
- Dézippez, renommez le dossier "ffmpeg", déplacez le dans un dossier non-temporaire.
- Copiez le chemin d'accès au sous-dossier « bin » (ex : C:\chemin\ffmpeg\bin)
- Aller dans démarrer, tapez variable d'environnement, dans la fenêtre, cliquez sur « variables d'environnement... »
- Dans variables d'environnement système, cliquez sur Path, modifier et ajouter le chemin d'accès.
- Allez dans l'invite de commande et tapez

```
ffmpeg --help
```

Si l'aide fonctionne, c'est bon!

Les prérequis : installation de Miniconda



Sur Windows

- [Lien](#) pour télécharger le fichier d'installation
- L'installation est réussie si l'application "Anaconda prompt" est disponible dans le menu démarrer.

Sur Ubuntu et distributions similaires

- Sur Ubuntu et distributions similaires, entrez les lignes suivantes dans le terminal:

```
mkdir -p ~/miniconda3
wget https://repo.anaconda.com/miniconda/Miniconda3-latest-Linux-x86_64.sh -O ~/miniconda3/miniconda.sh
bash ~/miniconda3/miniconda.sh -b -u -p ~/miniconda3
rm -rf ~/miniconda3/miniconda.sh

~/miniconda3/bin/conda init bash
~/miniconda3/bin/conda init zsh
```

- L'installation est réussie si en redémarrant le terminal, " (base) " s'affiche.

Les prérequis : création d'un environnement conda



1.

Windows : Ouvrir "Anaconda prompt" en mode administrateur

Ubuntu : ouvrir un terminal

2. Créer un environnement au nom de votre choix et activez le.

```
conda create -n nom_de_lenvironnement  
conda activate nom_de_lenvironnement
```

3. Installez python 3.9 :

```
conda install python=3.9
```

4. Installez git (et pytube si nécessaire) :

```
conda install git  
(pip install pytube)
```

5. Vous pouvez aussi tout faire d'un coup en écrivant :

```
conda create -n nom_de_lenvironnement python=3.9 git  
conda activate nom_de_lenvironnement
```

Bravo, vous avez un environnement avec python d'installé!

Installation de Whisper



- Toujours dans le même terminal, collez :

```
pip install git+https://github.com/openai/whisper.git
```

- Patientez...
- Si vous avez la fibre, passez à l'étape suivante, sinon patientez encore.

- Pour vérifier si Whisper a bien été installé, tapez :

```
whisper --help
```

Ou

```
whisper -h
```


Comment utiliser Whisper ?



- **Utilisation de base de Whisper** (modèle *small* par défaut)

- Quittez le mode admin. Allez dans le dossier de vos audio et dans la barre du chemin d'accès, tapez cmd
- Pour sélectionner la piste audio :

```
whisper nomdufichier.mp3
```

```
whisper C:\Users\Utilisateur\nomdufichier.mp3
```

- Si un espace ou un accent se trouve dans le chemin d'accès, rajoutez des guillemets :

```
whisper "C:\Users\À ne pas faire\Musique\nomdufichier.mp3"
```

Comment utiliser Whisper ?



Options dans Whisper

- Utilisation avec un modèle spécifique (tiny, small, medium, large, large-v2, large-v3) :

```
whisper nom_du_fichier.mp3 --model <model>
```

- Si vous disposez d'une carte graphique Nvidia avec une quantité de mémoire vidéo suffisante :

```
whisper nom_du_fichier.mp3 --device cuda
```

- Si vous avez besoin de préciser la langue dans le cas où elle n'est pas automatiquement détectée :

```
whisper fichier.mp3 --language Japanese
```

- Si vous voulez la traduction de l'audio vers l'anglais :

```
whisper nom_du_fichier.mp3 --task translate
```

- Les commandes sont cumulatives :

```
whisper "au d.mp3" --model tiny --device cuda --language Latin --task translate
```

Comment utiliser Whisper ?



- Utilisez le modèle adapté à votre configuration et à vos attentes :

Size	Parameters	English-only model	Multilingual model	Required VRAM	Relative speed
tiny	39 M	<code>tiny.en</code>	<code>tiny</code>	~1 GB	~32x
base	74 M	<code>base.en</code>	<code>base</code>	~1 GB	~16x
small	244 M	<code>small.en</code>	<code>small</code>	~2 GB	~6x
medium	769 M	<code>medium.en</code>	<code>medium</code>	~5 GB	~2x
large	1550 M	N/A	<code>large</code>	~10 GB	1x

Optimisation de l'outil Whisper AI

+

•

○

	Interview Aurore Bergé (1139 mots)	Micro-trottoir poubelles à Paris (693 mots)
large-v2	1.4%	3.5%
medium	2.3%	9.8%
small	5.8%	8.1%
base	10.5%	14.3%
tiny	19.1%	25.1%

Tableau 1 : Taux d'erreur (WER) selon la qualité de la prise de son (en studio et hors-studio) et le modèle Whisper utilisé.

Retour d'utilisation



La qualité de votre retranscription dépendra aussi de la qualité de votre audio.

Points positifs de l'outil :

- Gain de temps.
- Retranscription efficace.
- Outils libre et gratuit.
- Simple d'installation et d'utilisation.

Limites de l'outil :

- Pas de différenciation entre les locuteurs.
- Mise à la ligne ne correspond pas toujours à un changement de locuteur (pause dans le discours).
- Éléments non retranscrits lorsque plusieurs locuteurs parlent en même temps.
- Oubli de certains éléments (adverbes, prépositions...)

+
• ◦

Merci pour votre attention

+
• ◦
