

Les outils embarqués dans Sharedocs – Whisper (SpeechToText)

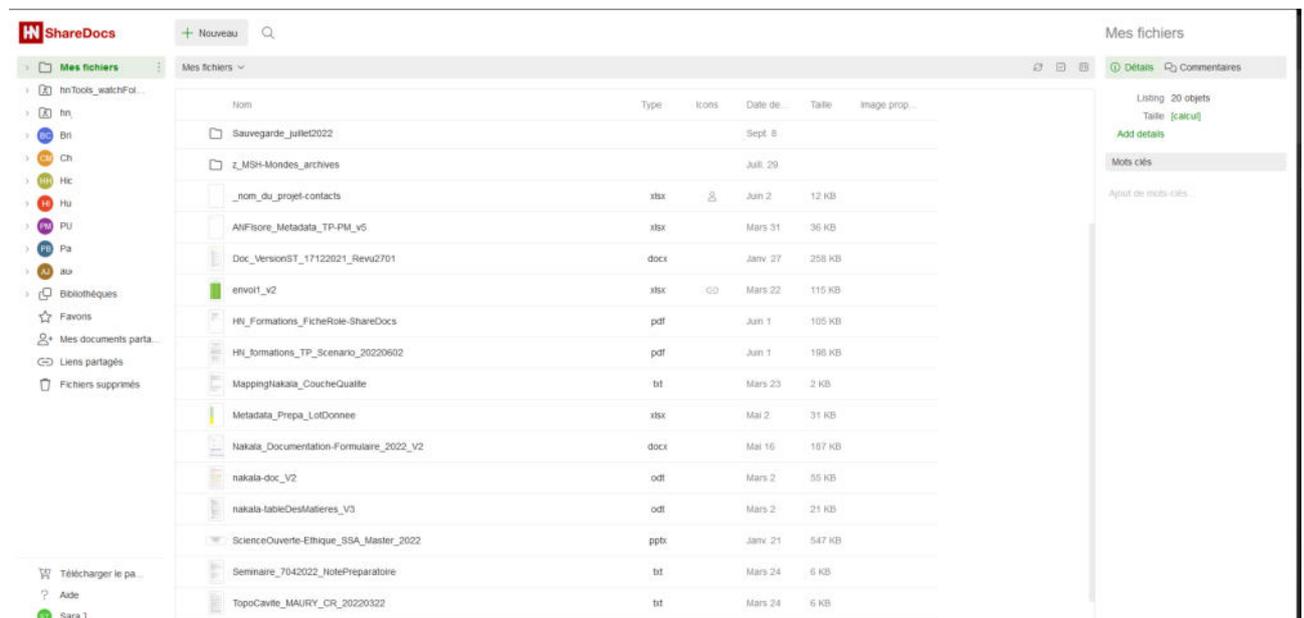
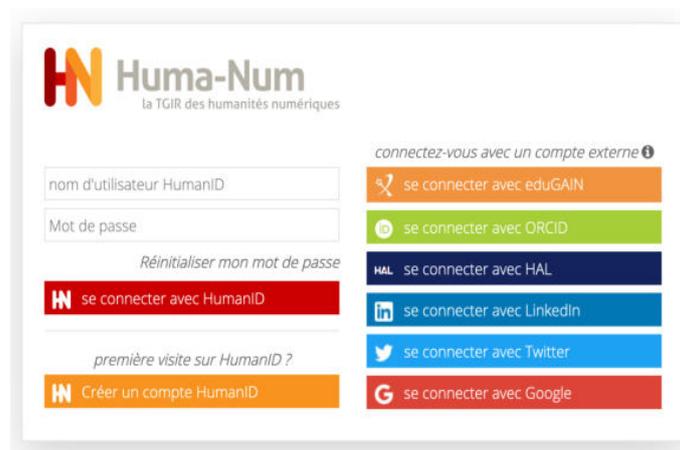
Petit guide ([CC-BY](#))

Aline Benvegnú dos Santos

MSHB Rennes

2024-11-20

1. Connectez-vous à votre compte Sharedocs

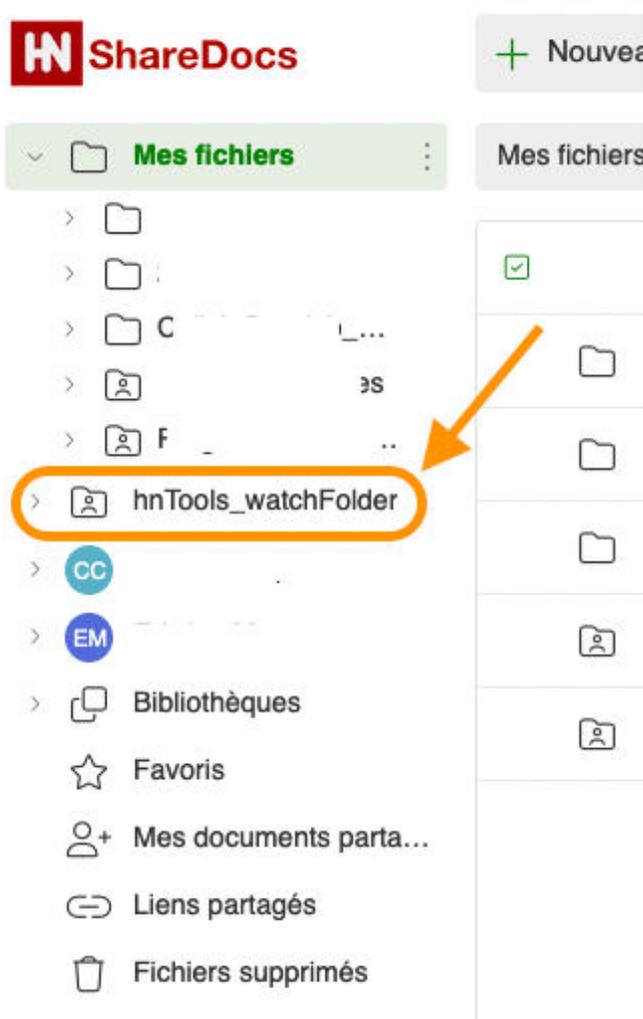


2. Le répertoire **hnTools_watchFolder**

Les outils de traitement embarqués dans Sharedocs se trouvent dans un répertoire qui fonctionne sur le principe du *Watch Folder* : ces outils n'ont pas d'interface graphique dédiée, il suffit de réaliser un glisser-déposer du fichier à être traité dans le bon dossier, et le travail se lance.

Ces outils sont à retrouver dans le **menu latéral gauche**, sous le répertoire **hnTools_watchFolder**.

(La documentation de Sharedocs explique le fonctionnement de ce répertoire : https://documentation.huma-num.fr/sharedocs-traitement/#fonctionnement-du-dossier-hntools_watchfolder)

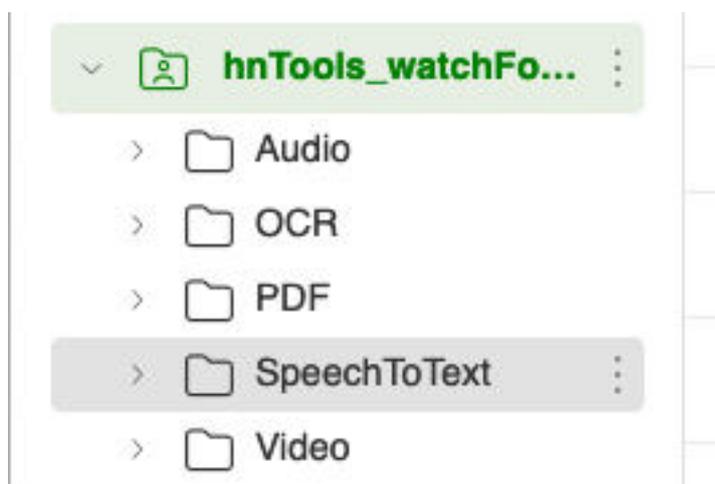


3. Accédez à Whisper via le répertoire SpeechToText

Whisper est intégré au dossier SpeechToText.

Après avoir cliqué sur "hnTools_watchFolder", une liste des dossiers des outils disponibles s'ouvre. Vous devez cliquer sur le dossier "SpeechToText".

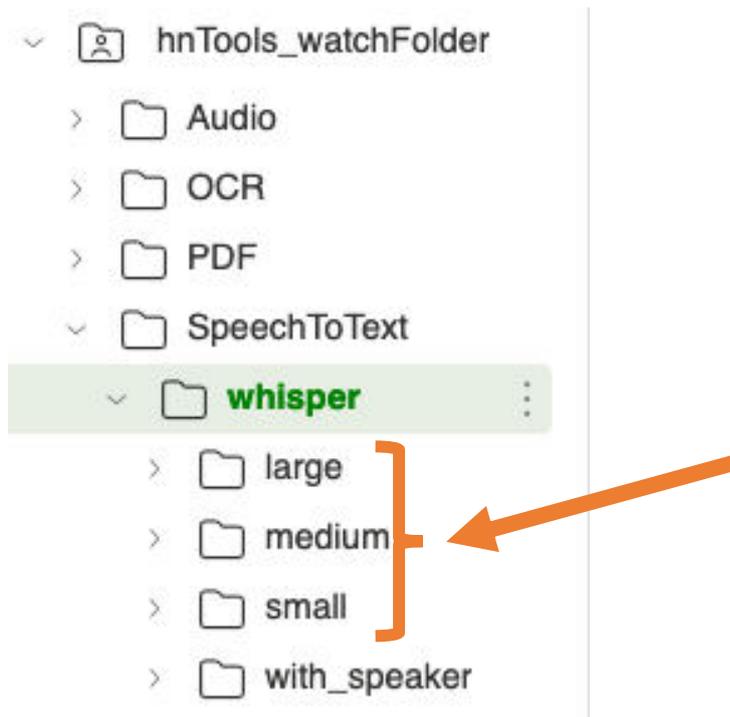
L'outil Whisper a été très récemment intégré à Sharedocs, pour cette raison il ne se trouve pas encore dans la documentation dédiée.



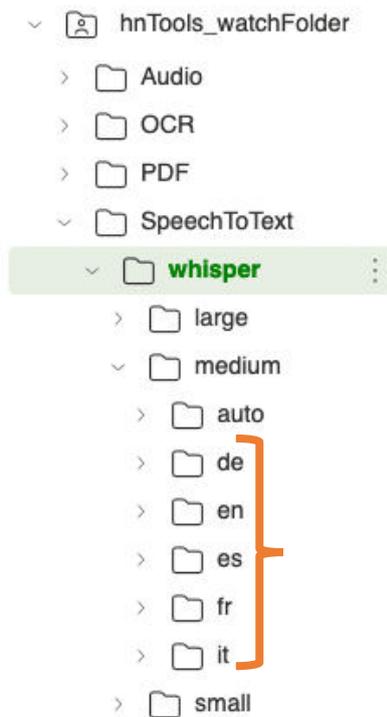
Puis, vous devez ouvrir le dossier "**whisper**".

Ensuite, il faut choisir l'un des trois **modèles de langage** : *large, medium, small*. Le modèle de langage désigne la taille du réseau de neurones utilisé par OpenAI pour faire de l'inférence. Ils correspondent, grosso modo, à un modèle plus robuste (large), avec un vocabulaire plus enrichi, mais qui va être beaucoup plus lent à la transcription, un modèle médian, et un modèle plus léger (small), pour un vocabulaire plus « simple », mais plus rapide à transcrire. (Il existe un dernier dossier, "with_speaker", qui sert à faire la diarisation du discours – il permet la reconnaissance des locuteurs.)

L'enjeu ici est de trouver le bon compromis entre la qualité et le temps de transcription (en fonction de la qualité et de la durée de l'enregistrement).



Une fois que vous avez choisi le modèle de langage, vous cliquez sur le dossier de la **langue** du discours (actuellement, sont disponibles à la transcription : l'allemand, l'anglais, l'espagnol, le français et l'italien).



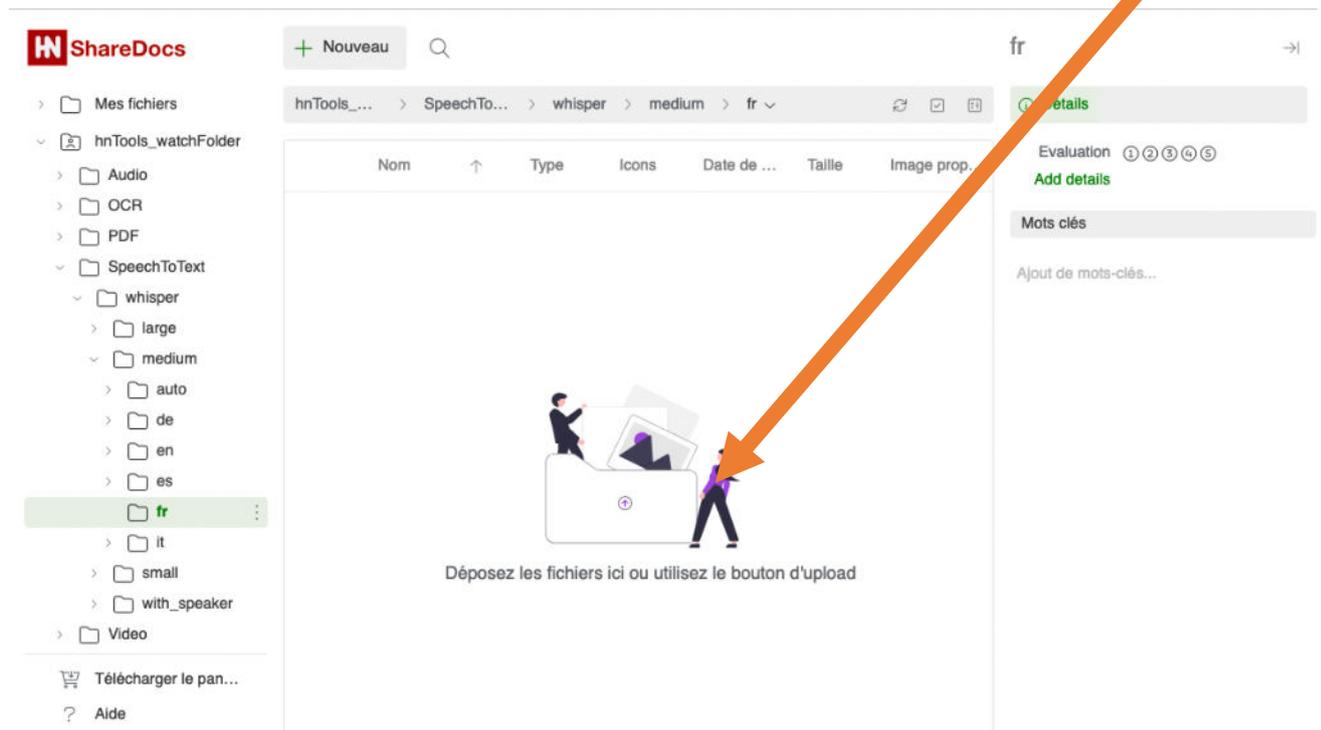
4. Déposez le fichier à transcrire

Une fois que la langue est définie, vous pouvez déposer dans le dossier correspondant le fichier audio à transcrire.

La transcription se lance automatiquement et vous devez attendre.

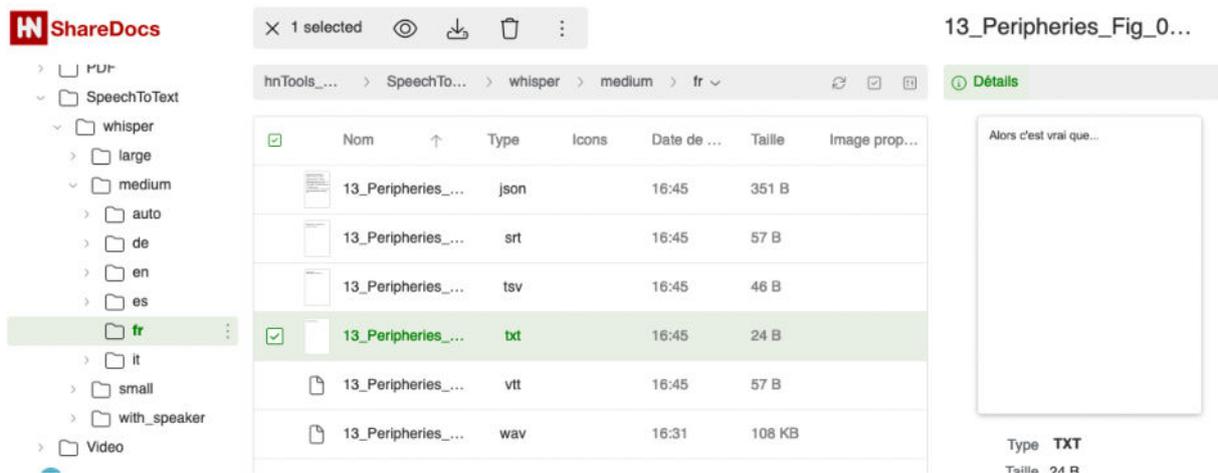
À la fin du traitement, un email vous est automatiquement envoyé pour vous indiquer que le document produit est prêt. Vous pouvez donc le récupérer.

Il est demandé de bien le supprimer du Watch Folder après utilisation.



À titre d'exemple, j'ai déposé un fichier audio en français (format .waw, durée 2 secondes, taille 108 kb) très court et léger, à être transcrit avec le modèle de langage medium :

Je retourne dans mon dossier Sharedocs, je vois que plusieurs fichiers de sortie ont été générés, dont du texte brut (.txt).



ShareDocs interface showing a file list in the 'fr' folder. The selected file is '13_Peripheries_...' with a size of 24 B.

Norm	Type	Date de ...	Taille
13_Peripheries_...	json	16:45	351 B
13_Peripheries_...	srt	16:45	57 B
13_Peripheries_...	tsv	16:45	46 B
13_Peripheries_...	txt	16:45	24 B
13_Peripheries_...	vtt	16:45	57 B
13_Peripheries_...	wav	16:31	108 KB

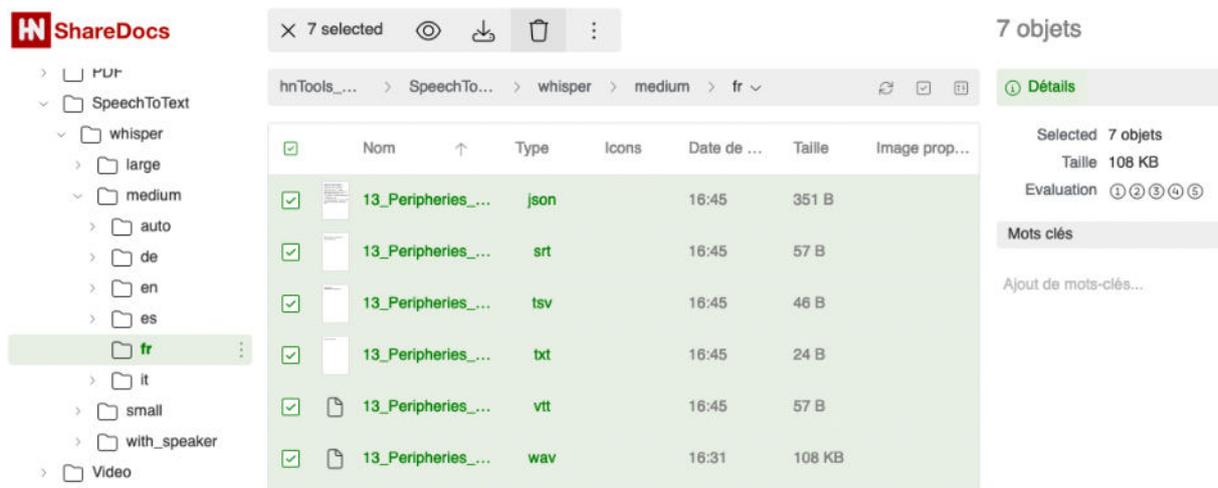
13_Peripheries_Fig_0...

Détails

Alors c'est vrai que...

Type **TXT**
Taille 24 B

Je télécharge les fichiers générés et je n'oublie pas de tout effacer de ce dossier de travail Sharedocs.



ShareDocs interface showing a file list in the 'fr' folder. All 7 files are selected. The total size is 108 KB.

Norm	Type	Date de ...	Taille
13_Peripheries_...	json	16:45	351 B
13_Peripheries_...	srt	16:45	57 B
13_Peripheries_...	tsv	16:45	46 B
13_Peripheries_...	txt	16:45	24 B
13_Peripheries_...	vtt	16:45	57 B
13_Peripheries_...	wav	16:31	108 KB

7 objets

Détails

Selected 7 objets
Taille 108 KB
Evaluation ① ② ③ ④ ⑤

Mots clés

Ajout de mots-clés...

Ce répertoire doit rester vide, sinon le travail risque de se relancer et de surcharger les serveurs.

- > PDF
- > SpeechToText
 - > whisper
 - > large
 - > medium
 - > auto
 - > de
 - > en
 - > es
 - fr
 - > it
 - > small
 - > with_speaker
 - > Video
- > CC
- > EM
- > Bibliothèques

hnTools_... > SpeechTo... > whisper > medium > fr

<input checked="" type="checkbox"/>	Nom	↑	Type	Icons	Date de ...	Taille	Image prop...
 <p>Déposez les fichiers ici ou utilisez le bouton d'upload</p>							

E
A
Mo
Ajo