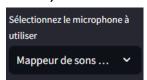
## **Documentation TalkMate**

## Page Menu

→ Configuration du Microphone à utiliser dans l'application durant la session (permet de parler à votre assistant par la suite)



→ Configuration de la voix synthétique souhaité, si vous n'avez pas toutes les voix du narrateur disponibles, veuillez regarder le ReadMe dans la partie "Configuration Synthetic Voices", et suivez les étapes à faire pour ajouter d'autres voix du narrateur windows dans votre application

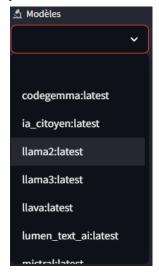


→ Dans l'onglet "Pour développeur uniquement", le bouton Paramètres Modèles permet de modifier l'instruction système de vos modèles Ollama en local, le bouton Quantize Modèles permet de quantizer (donc de réduire, compressé) en utilisant Docker un modèle et enfin le bouton Publier Modèles permet via l'API d'Ollama de publier votre modèle sur Ollama, pour chaque bouton, vous devez suivre les étapes dans le terminal

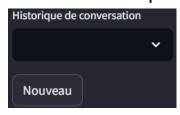


## Page Assistant

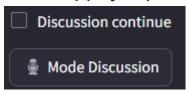
→ Vous pouvez choisir un de vos modèles que vous avez au préalable téléchargé sur Ollama et ou modifié en local



→ Ensuite vous avez votre historique de conversation qui permet de choisir une conversation et ou d'en créer une nouvelle en cliquant sur le bouton *Nouveau* 



→ Il y a 2 modes de discussions possibles ici, soit le mode Discussion continue qui permet de discuter continuellement avec votre modèle sans avoir à réappuyer sur un bouton et le Mode Discussion qui permet de parler une seule fois et d'avoir la réponse de la voix synthétique une fois et à chaque fois il faut réappuyer pour reparler avec le modèle



#### Page Summarize

→ Vous pouvez Upload un fichier pour le résumer



Ca va extraire le texte et ensuite utiliser un modèle de résumé universel qui va être utilisé pour résumer le mieux possible votre document



- → Vous pouvez personnaliser votre résumé via les *Propriétés Avancées* où vous retrouverez une liste de propriété à modifier selon vos besoins :
  - Choisir la longueur maximale du résumé
  - Choisir la longueur minimale du résumé
  - La taille du lot
  - Le top\_k et le top\_p qui sont expliqués via le bouton Info juste en dessous



→ Vous avez votre *Mode Audio* qui permet de parler via la voix synthétique pour vous lire le texte qui vous a été résumé



## Page QA Document

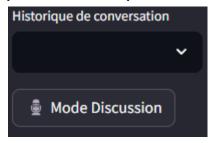
→ Vous pouvez Upload un fichier qui permet de parler par la suite via un LLM sur votre document upload



→ Vous pouvez choisir le modèle en local que vous souhaitez utiliser et le modèle d'embedding (pour la vectorisation) que vous souhaitez

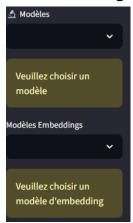


→ Vous avez ensuite votre historique de conversation comme pour la page Assistant mais ici vous allez retrouver vos discussions seulement de votre page QA Document & Vous avez votre mode discussion qui vous permet après avoir mis un document de parler une fois en cliquant sur le bouton pour poser votre question sur le document fourni



## **Page Question**

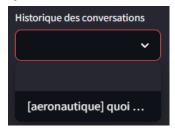
→ Vous devez choisir votre modèle LLM et votre modèle d'embedding en local comme pour la page QA Document



→ Si vous n'avez pas créé de profil vous devez en créer un en entrant le nom du nouveau profil et en appuyant sur le bouton créer



→ Vous allez avoir votre historique des conversations qui va afficher seulement vos conversations dans cette page et qui va mettre le nom de votre profil entre [] pour vous indiquer le profil que vous avez utilisé durant la conversation



# Information supplémentaire

→ Chaque historique de conversation, si vous sélectionnez un fichier cela va automatiquement vous proposer d'autres paramètres comme le fait de renommer le nom de la conversation sauvegardée, télécharger en CSV et enfin supprimer la conversation

