

Open AI Whisper audio transcription

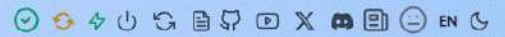
It looks like I could not install whisper because of this error:

cannot import name 'install_whisper' from 'lollms.services.stt.whisper.lollms_whisper' (c:\users\alino\downloads\lollms-webui\lollms_core\lollms\services\stt\whisper\lollms_whisper.py)

OK



One tool to rule them all



- Discussions
- Playground
- Personalities Zoo
- Apps Zoo
- Settings
- Help



New discussion



user

Bonjour, peux-tu me dire qui est Charles De GAULLE ?



lollms

just now



Processing Info

• Adding skills

Binding: gemini Model: gemini-pro

default

Send message...



Discuter avec des PDFs

Français



Télécharger des PDFs

Choisir des fichiers lolims_user_guide_FR.pdf

lolims_user_guide_FR.pdf



Discussion

AI

User

salut peux-tu me dire ce que je dois faire comme réglage exactement en lisant ce PDF ?

AI

Type your message...

Send





System status | 7.3 GB / 8.5 GB (85.8%) 144.3 GB / 254.0 GB (56.8%)

CPU Ram usage

Available ram: 1.2 GB

Ram usage: 7.3 GB / 8.5 GB (85.8%)



Disk usage:

Available disk space: 109.7 GB

Disk usage: 144.3 GB / 254.0 GB (56.8%)



> Smart routing configurations

> Main configurations

Smart routing configurations

Smart Routing Settings

Use Smart Routing:

Restore model after smart routing:

Router Model:

Models by Power:

Main configurations

General

Application logo:



Hardware mode:

CPU



Database path:

default

Add details to messages copied to clipboard:



Auto show browser:



Activate debug mode:



Activate showing the full prompt in console (for debug):



Show final full prompt in console:



Show chunks in console:



Debug file path:

Show news panel:



Auto update:



Auto title:



Model template

Start header id template:

Lollms communicatio ▾

Start header id template:

[INST]

End header id template:

:

Start user header id template:

[INST]

End user header id template:

:

End user message id template:

[/INST]

Start ai header id template:

[INST]

End ai header id template:

:

End ai message id template:

[/INST]

Separator template:

System template:

Full template:

[INST] Using this information: system message
[INST]user name: User prompt [/INST] [INST]ai personality: ai response[/INST]

Use continue message:

User

User name:

Use user description in discussion:

Use model name in discussion:

User avatar: 

Use User Name in discussions:

Maximum number of output tokens space (forces the model to have more space to speak):

Minimum number of output tokens space (forces the model to have more space to speak):



Security settings

turn on code execution:

turn on code validation (very recommended for security reasons):

turn on apply settings validation (very recommended for security reasons):

turn on open file/folder
validation:
turn on send file validation:

Knowledge database

Activate Skills library:

Skills library database name:

Latex

PDF LaTeX path:

Boost

Positive Boost:

Negative Boost:

Fun mode:

▼ Data management settings

Data Sources

Data Sources: [Add Data Source](#)

RAG Vectorizer: ▼

RAG Vectorizer model: ▼

Open AI key for open ai embedding method (if not provided I'll use)

RAG chunk size:



512

RAG overlap size:



0

RAG number of chunks:



4

Clean chunks:

Follow subfolders:

Check for new files at startup:

Preprocess chunks:

Activate multi hops RAG:

Use contextual summary instead of rag (consumes alot of tokens and may be very slow but efficient, useful for summary and global questions that RAG can't do):

Use all the document content

Data Vectorization

Save vectorized database:

show vectorized data:

Reformulate prompt before querying database (advised):

Force adding the first chunk of the file to the context:

Put Chunk Information Into Context:

Data vectorization method:

Data vectorization model (for Sentence Transformer):

Data visualization method:

Save the new files to the database (The database will always grow and continue to be the same over many sessions):

Data vectorization chunk size(tokens):

128

Number of chunks to use for each message:

2

Internet

Internet search

- Activate internet search:
- Activate internet pages judgement:
- Activate quick search:
- Activate search decision:

Internet vectorization chunk size:

512

Internet

Internet search

- Activate internet search:
- Activate internet pages judgement:
- Activate quick search:
- Activate search decision:

Internet vectorization chunk size:

Internet vectorization overlap size:

Internet vectorization number of chunks:

Internet number of search pages:

Services Zoo

Default services selection

Active TTS Service:	None	▼
Active STT Service:	None	▼
Active TTI Service:	None	▼
Active TTM Service:	None	▼
Active TTV Service:	None	▼

TTI settings

Use negative prompt:

Use AI generated negative prompt:

Negative prompt generation prompt:

Default negative prompt:

Full Audio settings

Listening threshold:

Silence duration (s):

Minimum sound percentage in recorded segment:

Volume amplification:

audio rate:

number of channels:

1

Buffer size:

10

Activate word detection:

Word detection wav file:

Audio devices settings

Audio Input device:

Audio Output device:

Lollms service

Host:

localhost

Enter access key

Add

Host: localhost

Access keys: Enter access key Add

Port: 9600

Activate headless server mode:

Activate lolms server:

Activate lolms RAG server:

Activate lolms TTS server:

Activate lolms STT server:

Activate lolms TTI server:

Activate lolms ITT server:

Activate lolms TTM server:

Activate ollama server emulator:


Activate openai server emulator:

Activate mistral ai server emulator:

Open AI Whisper audio transcription

openai whisper key:

Open Ai Whisper model:

whisper-1 

TTS services

Browser Audio TTS

Enable auto speak:

audio pitch:

