Whisper Transcription App in UCloud

Whisper Transcription is an app developed for UCloud at AAU using OpenAI's Whisper Large Language Model. Everything about Whisper Transcription on UCloud takes place in the closed UCloud at SDU's eScience Center.

Everyone who logs in to UCloud has DKK 1000 in credits available, which are converted into computing time on UCloud; one transcript is charged according to how much computing time is used. You can get more credits by applying for more computing time - no real money is required. You can apply for more credits from within UCloud. DKK 1000 is enough for many transcriptions.

You can read more about UCloud here: https://docs.cloud.sdu.dk/

Login to UCloud
Go to https://cloud.sdu.dk/app/login and select the WAYF login box:

WAYF Login

Write Aarhus in the search box and select Aarhus University and log in with your AU-account:

☐ Always use the login provider that I choose now. At my.wayf.dk I can reset this pre use a different login provider.

aarhus

Showing 5 of 5:

Aarhus School of Architecture
Aarhus School of Marine and Technical Engineering
Aarhus University

Dashboard in UCloud

When you are logged in to UCloud, you will land on your dashboard, where you can find your files in the menu on the left and see your recent runs.

1: Search bar - find Apps and files.

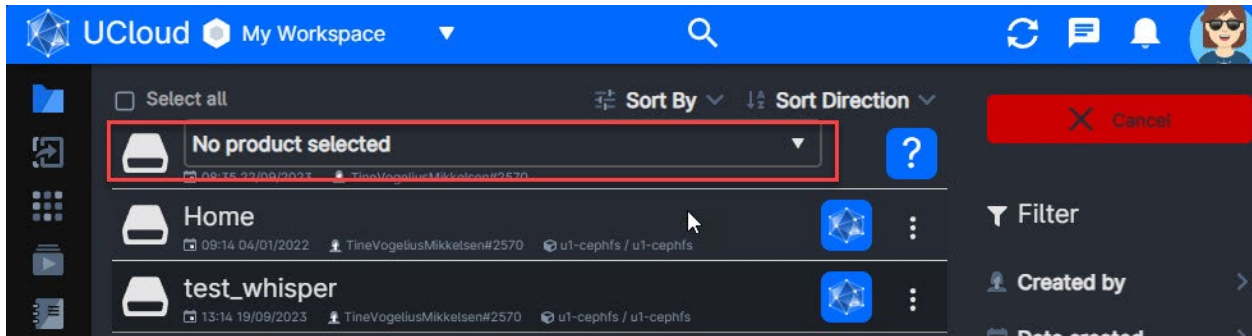2: Shortcut to Files, Favorite Apps, Runs.

Upload file for transcription
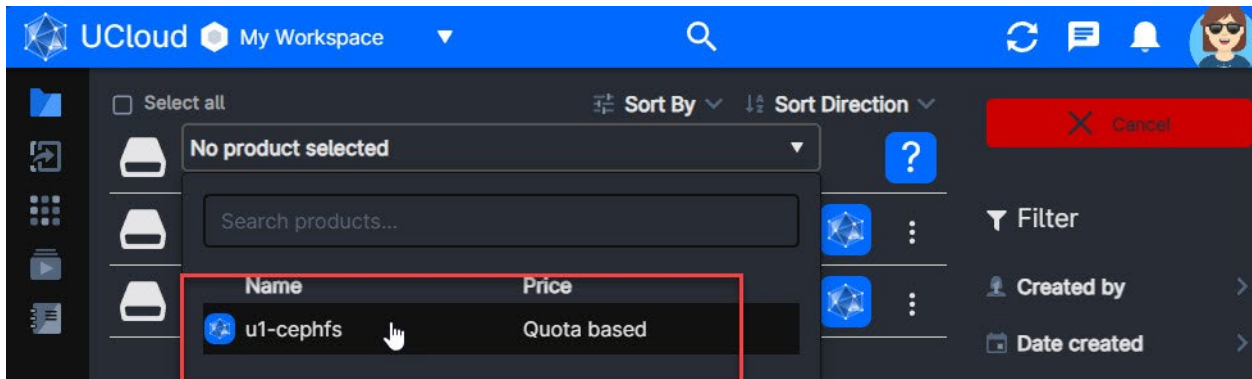
Go to Files in the menu on the left.

Create a folder for the file, click "Create drive" in the top right corner:
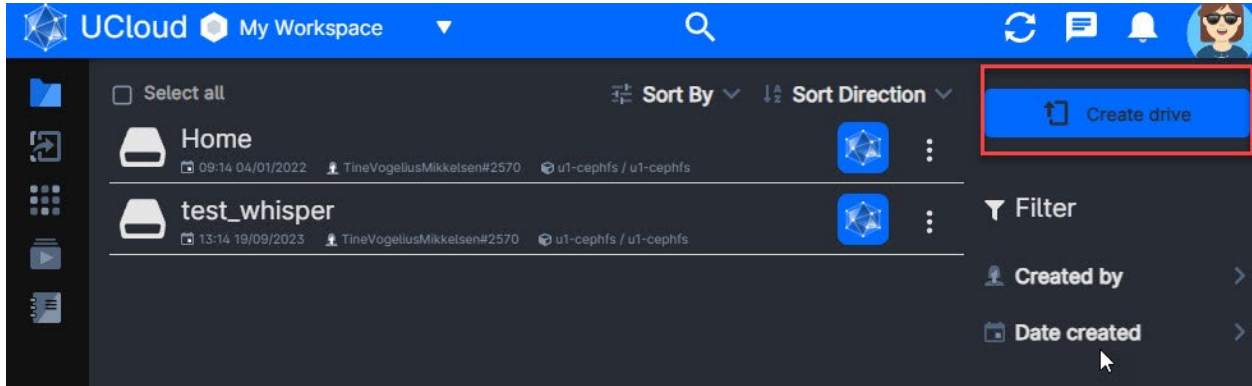
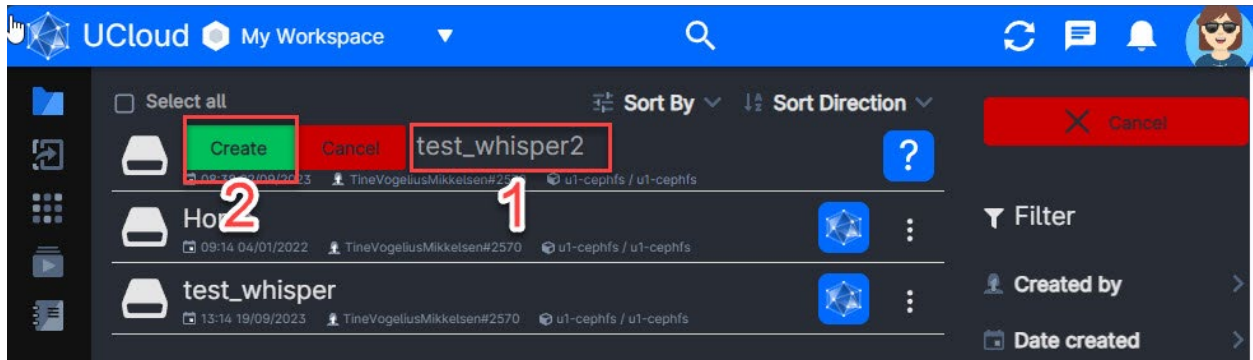Click in the "No product selected" field:

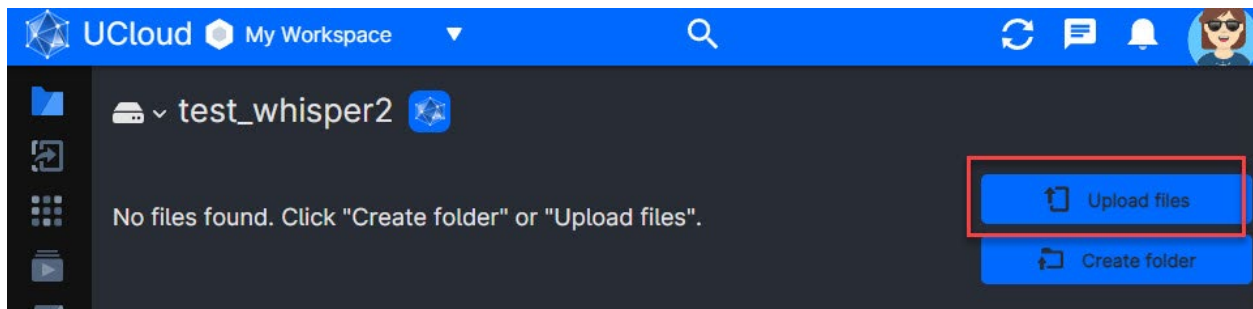Select "u1-cephfs" in the drop-down box:

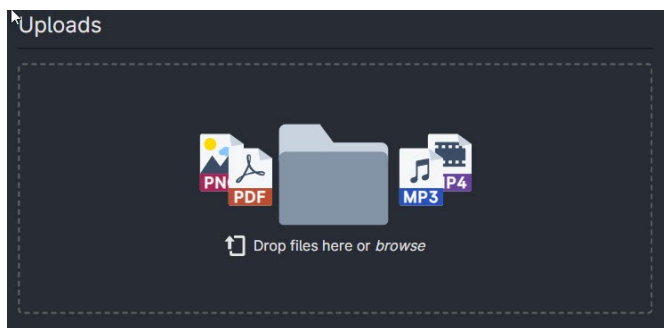1: Name your folder - no spaces, but - hyphen and _ underscore are acceptable.

2: Click Create to create.

Click on the folder you have just created, you can create multiple subfolders here, and select "Upload files" in the menu on the right:

Drag & Drop the file you want to upload or click in the box to *Browse* for the file:

Once the file is uploaded, you will see a green checkmark and can click away from the upload box.

File formats and renaming file names

You can use the following file formats in the Whisper Transcription app:
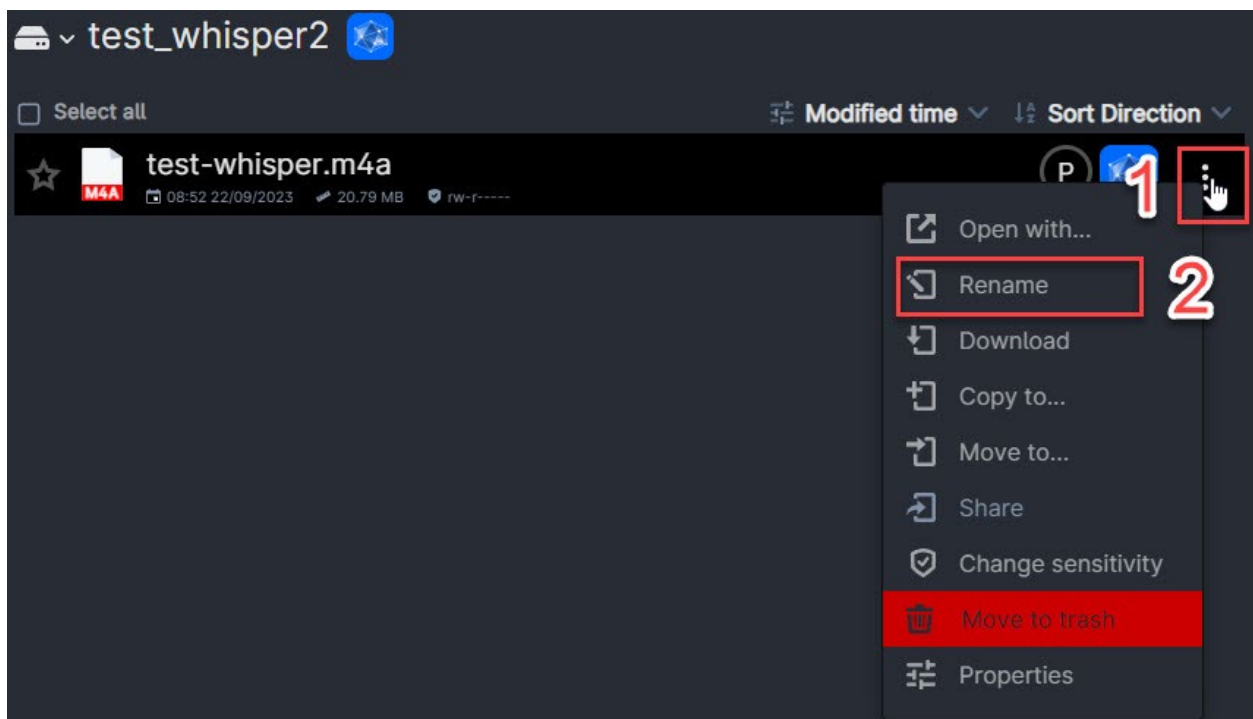
.mp3, .mp4, .m4a, .wav and .mpg

For example, an iPhone records voice memos in .m4a format.

When transcribing the files, there should be **no spaces in the file name**, but preferably - hyphen and _ underscore. You can rename the file via Files from the menu on the left and browse to it once you have found the file:

1: Click on the three dots.

2: Select "Rename".

Type the new file name in the field and click the green "Rename" button:
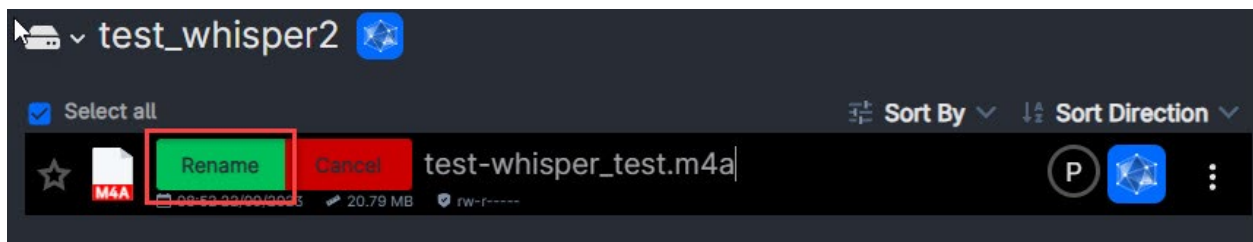
Transcribing with Whisper Transcription

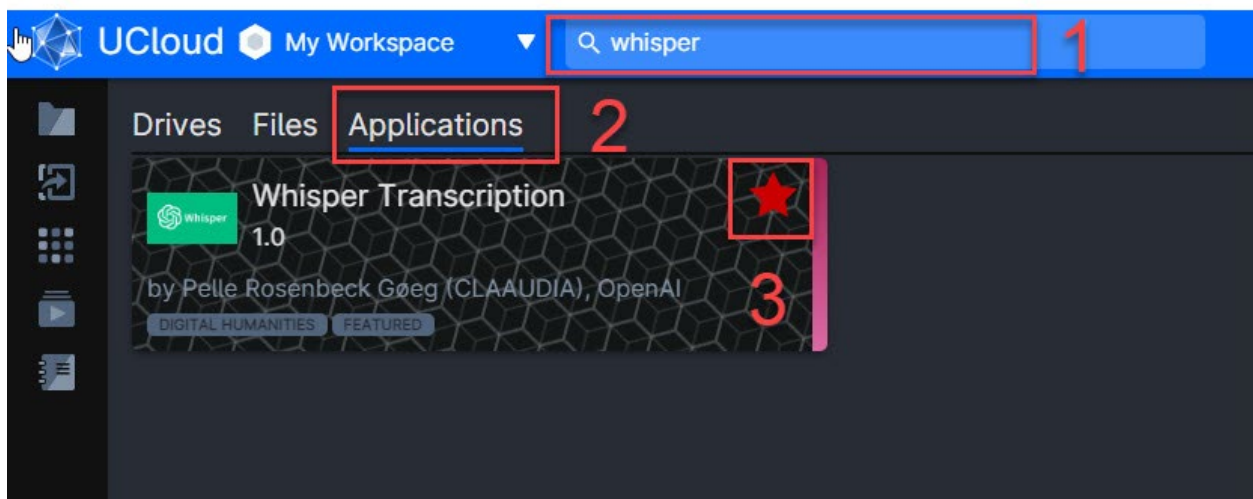## Find Whisper Transcription

Find Whisper Transcription by searching for *Whisper* in the search bar at the top.

1: Type Whisper and press enter.

2: Select "Applications".

3: Add Whisper to favorite Apps by star marking it, so you can find it faster next time via Apps in the left menu on the Dashboard.

Click on the App to open it.
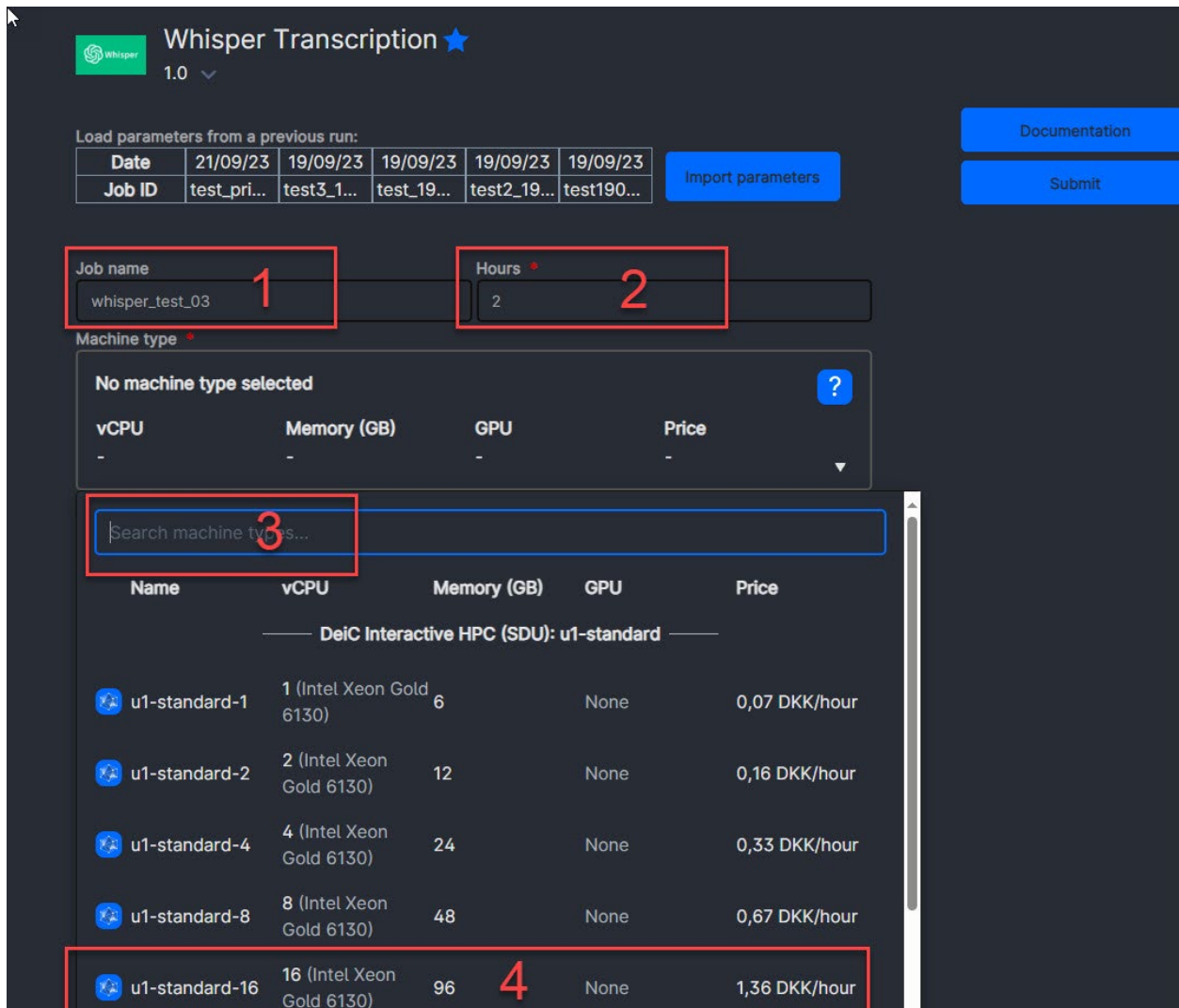
Start transcribing

Next, you need to name your "job", and choose how long the job should last and how much computing power should be used to run the job. Whisper Transcription will detect the language spoken on the recording.

1: Name your job - it's important that there are no spaces or other special characters in the name. Hyphens - and underscores _ are OK.

2: An interview of around 40 minutes can take between 20 and 60 minutes to transcribe. You can easily set the time to 2 hours or more if you are unsure how long it will take. Credits are only deducted for the actual running time of the job.

3: Click in the box to select computational power - a menu will unfold.

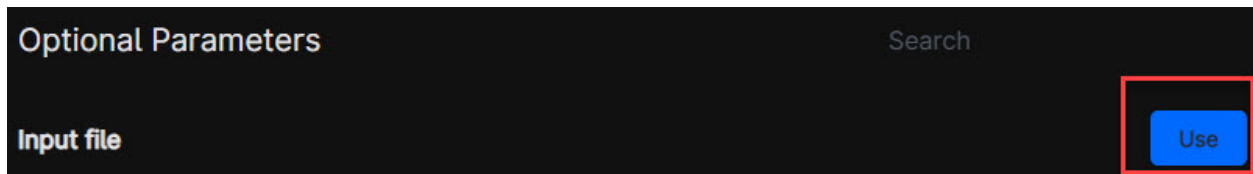4: It is recommended to select u1-standard-16 for transcription jobs.

Under Optional Parameters, click Use next to Input file:
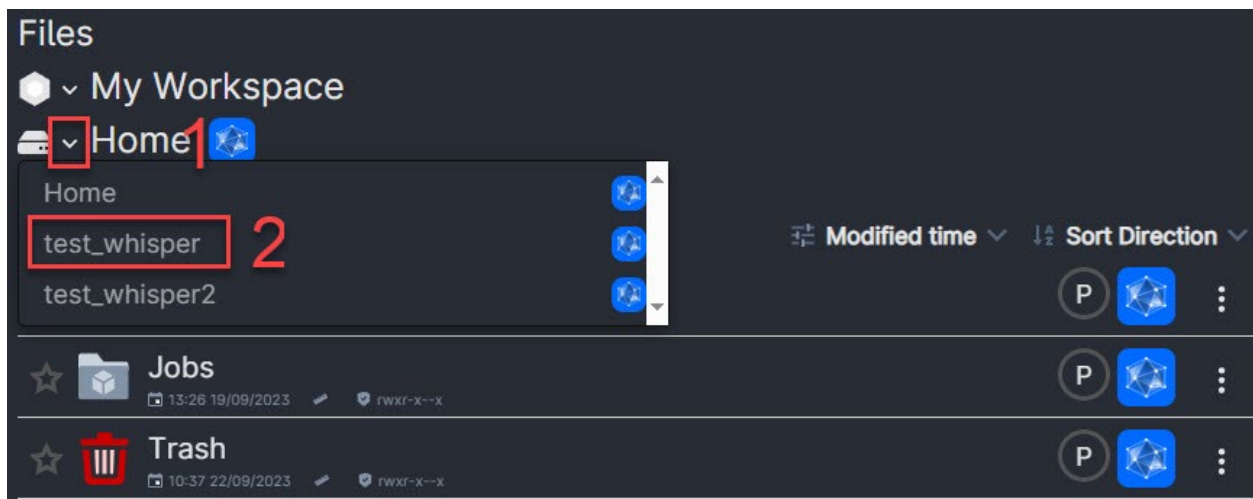

Click in the activated field.
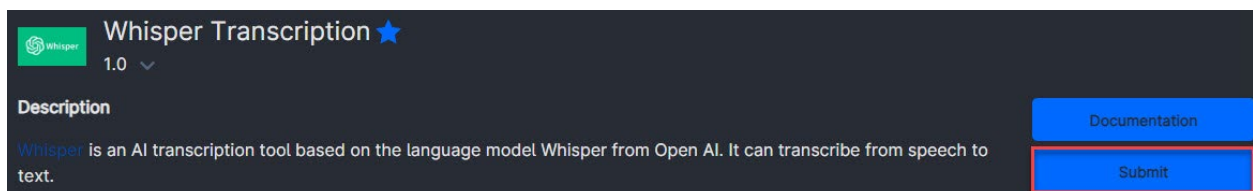
A new window opens.

1: Click ∨ next to Home.

2: Select the folder the file is in.

Once you have found the file, click Use:

Tap Submit in the top right corner and the job starts:

While the job is running, you can browse away from the page, e.g. to another UCloud page to start another job, close the browser window or turn off your computer completely. When you return to UCloud after some time, your transcription will be ready.

Find transcript

If you haven't browsed away from the UCloud job, you can find the transcription under the "Job Completed" message in the "out" folder.

You can also find the transcription by clicking on Files in the left menu, selecting "Home" and then "Jobs" and the folder "Whisper Transcription". All your runs are located here by default.

For each job, there will be an "Out" folder, and this is where you will find the transcription.
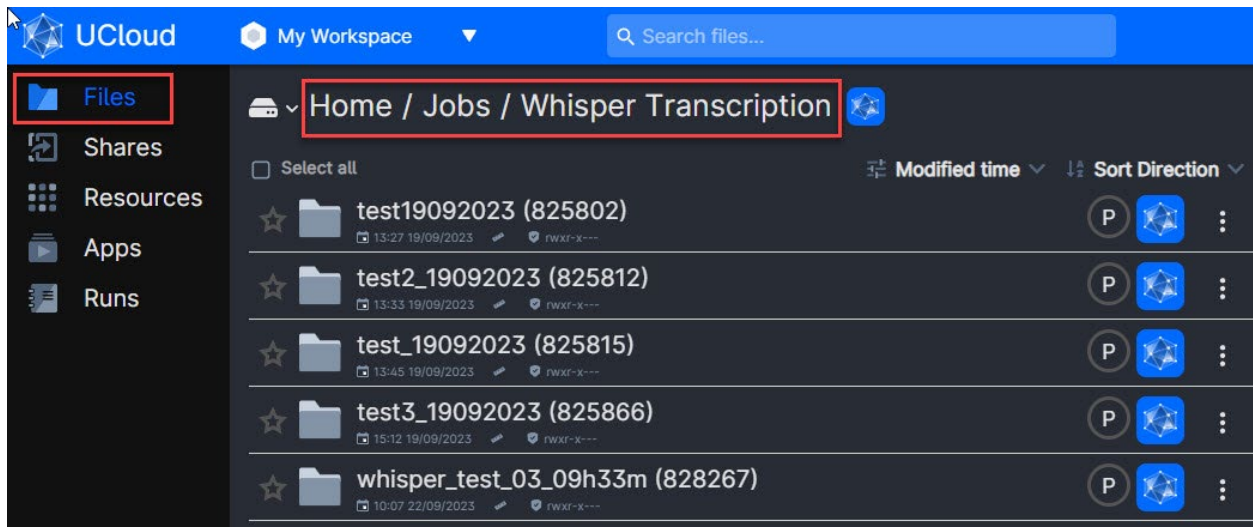
By default, seven formats are generated from the transcript that can be opened in different software, such as Notepad/TextEdit.

.txt - plain text file.

.vtt - displays timecodes optimized for the web.

.srt - subtitle format with timecodes.

.tsv - tab-separated format with timecodes.

.json - open standard file extension that can be used in many other apps.

.csv - all parameters from the Whisper LLM are contained in this file extension.

.dote.json - file extension used in DOTE transcription software developed at AAU.

Download transcript

Click on the folder with the job you want to download and click on the "Out" folder. You will now see all the files generated for your job.

1: Click the three dots next to the file you want to download.

2: Select download.

More features

There are many more advanced features in Whisper Transcription, such as the ability to tell which language to transcribe to (e.g. if multiple languages are used in the file), initiate transcription and download multiple files simultaneously.

You can watch a demo of the many features here (the video is 48 minutes long): https://www.youtube.com/watch?v=M0IGco--8LQ

Note about computer-generated transcription

Be aware that the accuracy of the output from a computer-generated transcription is not always at the same level as a manual transcription. Automated transcripts often contain errors that need to be corrected manually. Additionally, automatic speech recognition doesn't always interpret tone and emotion correctly, which can lead to inaccurate transcripts. Automatic transcription software can be less reliable than human transcribers if the audio quality is poor or there is a lot of background noise. Finally, transcription software cleans up the output to remove filler words.