

Whisper Transcription App i UCloud

Whisper Transcription er en app udviklet til UCloud på AAU, den benytter OpenAIs Whisper Large Language Model. Alt omkring Whisper Transcription på UCloud foregår i den lukkede UCloud sky på SDUs eScience Center.

Alle der logger på UCloud har DKK 1000 i *credits* til rådighed, som omregnes til regnetid på UCloud; én transskribering afregnes efter hvor meget regnetid der benyttes. Man kan få flere *credits* ved at ansøge om mere regnetid - der skal altså ikke rigtige penge på bordet. Der kan ansøges om flere credits inde i UCloud. DKK 1000 rækker til rigtig mange transskriberinger.

Du kan læse mere om UCloud her: <https://docs.cloud.sdu.dk/>

Login i UCloud

Gå til <https://cloud.sdu.dk/app/login> og tryk på WAYF Login boksen:



Skriv Aarhus i søgefeltet og klik på Aarhus Universitet og log ind med din AU-konto:



- Always use the login provider that I choose now. At my.wayf.dk I can reset this preference or use a different login provider.

Showing 5 of 5:

[Aarhus School of Architecture](#)

[Aarhus School of Marine and Technical Engineering](#)

[Aarhus University](#)



Dashboard i UCloud

Når du er logget ind i UCloud lander du på dit dashboard, her kan du bl.a. finde dine filer i menuen til venstre og se dine seneste kørsler.

1: Søgefelt - find Apps og filer.

2: Genvej til Filer, favorit Apps, kørsler.

The screenshot shows the UCloud dashboard interface. At the top, there is a blue header with the UCloud logo, 'My Workspace', and a search bar labeled 'Search files and applications...'. A red box labeled '1' highlights this search bar. On the left side, there is a dark sidebar with menu items: Files, Shares, Resources, Apps, and Runs. A red arrow labeled '2' points to the 'Files' menu item. The main content area is divided into several sections: 'News' (UCloud 2023.3.0), 'Favorites' (No favorites), 'Recent runs' (a list of runs with status icons and timestamps), 'Recent notifications' (No notifications), 'Resource allocations' (a table of resource usage), and 'Resource usage' (a summary of resource usage). The 'Recent runs' section contains the following data:

Run Name	Status	Time
test_price_3h	Success	about 23 hours
test3_19092023	Success	3 days
test_19092023	Success	3 days
test2_19092023	Success	3 days
test19092023	Success	3 days

The 'Resource allocations' section shows the following data:

Resource Allocation	Usage
DeiC Interactive HPC (SDU) / u1-standard	997,74 DKK
DeiC Interactive HPC (SDU) / u1-	47,GB

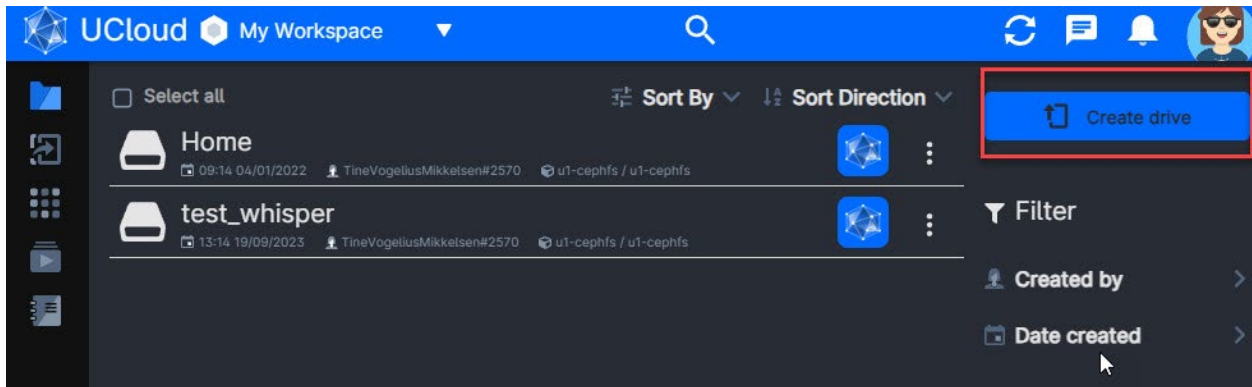
The 'Resource usage' section shows the following data:

Resource Usage	Usage
Compute	2,25 DKK

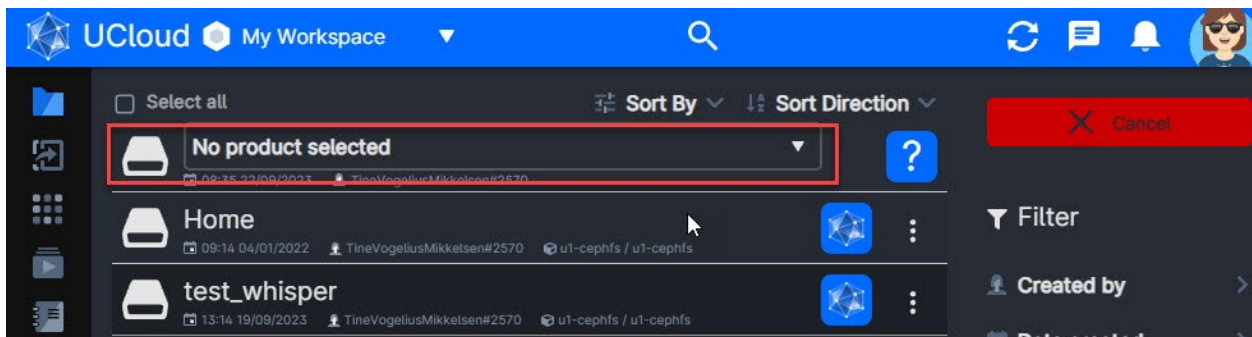
Upload fil til transkribering

Gå til Files i menuen til venstre.

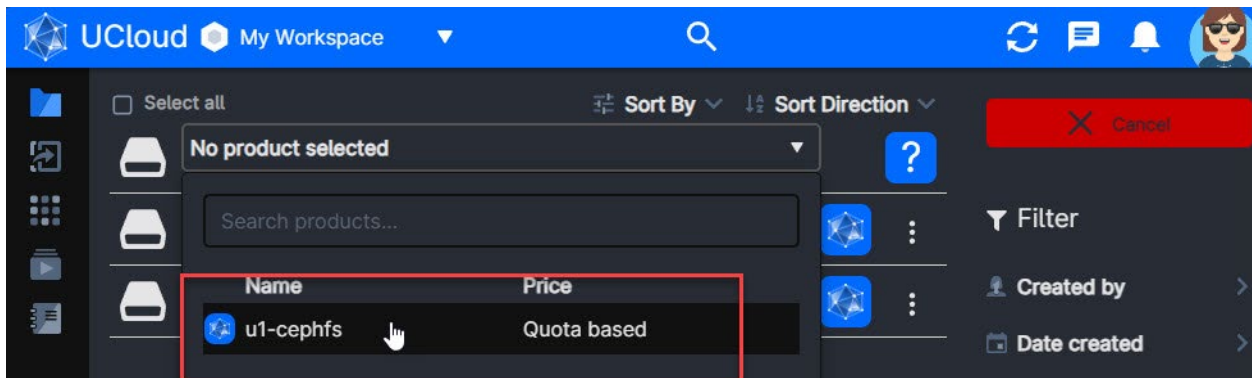
Opret en mappe til filen, klik på "Create drive" øverst til højre:



Klik i feltet "No product selected":

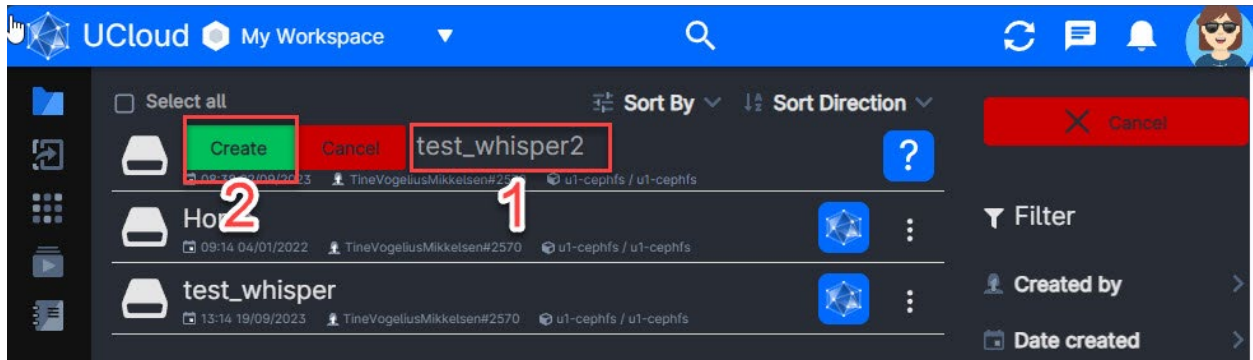


Vælg "u1-cephfs" i drop-down boksen:

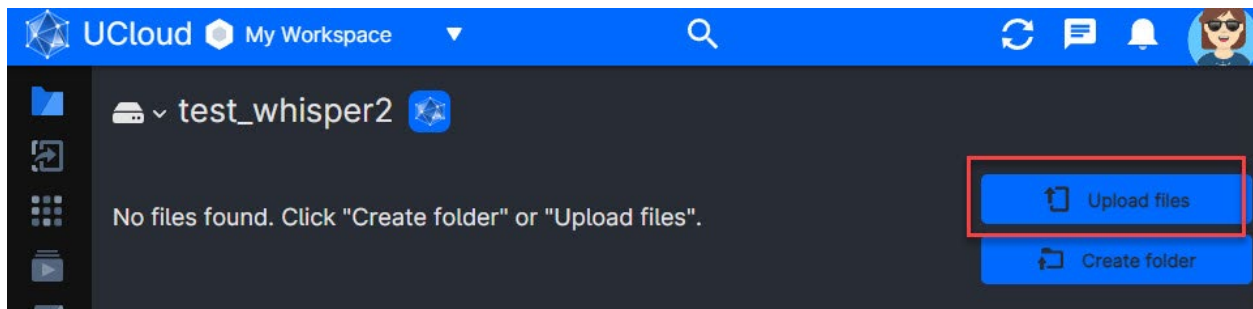


1: Navngiv din mappe – der må ikke være mellemrum, men gerne – bindestreg og *_ underscore*.

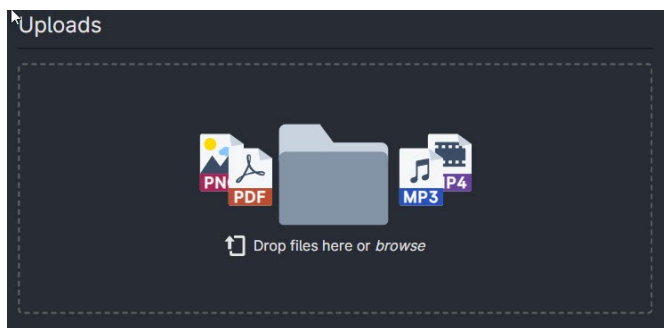
2: Klik på Create for at oprette.



Klik på mappen du har netop har oprettet, du kan evt. oprette flere undermapper her, og vælg "Upload files" i menuen til højre:



Drag & Drop filen du vil uploade eller klik i feltet for at *Browse* efter filen:



Når filen er uploadet får du et grønt flueben og kan klikke væk fra upload boksen.

Filformater og omdøbning af filnavne

Du kan benytte følgende filformater i Whisper Transcription app'en:

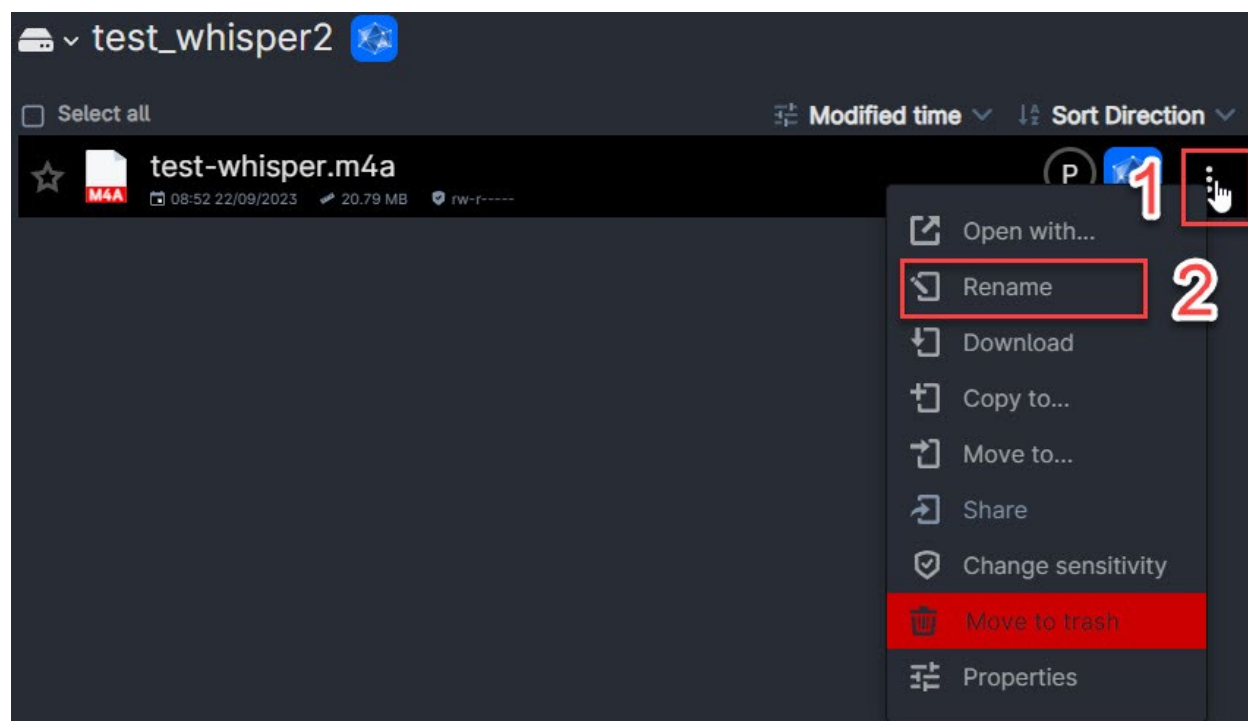
.mp3, .mp4, .m4a, .wav og .mpg

Eksempelvis optager en iPhone *voice memos* i .m4a formatet.

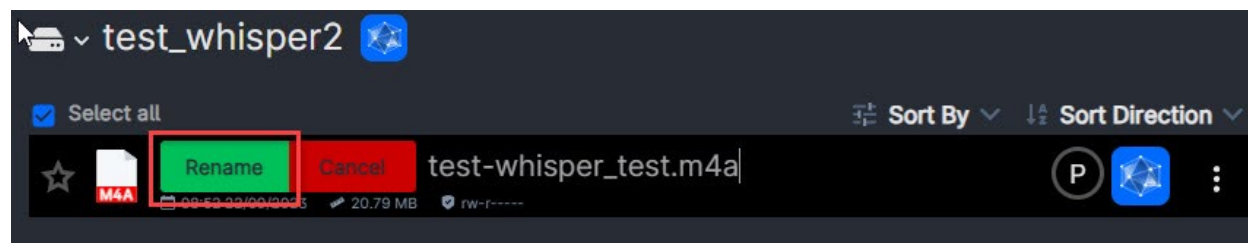
Når filerne skal transskriberes må der ikke være mellemrum i filnavnet men gerne – bindestreg og *underscore*. Du kan omdøbe filen via Files fra menuen til venstre og browse dig frem til den; når du har fundet filen:

1: Klik på de tre prikker.

2: Vælg "Rename".



Skriv det nye filnavn i feltet og klik på den grønne "Rename" knap:



Transskribering med Whisper Transcription

Find Whisper Transcription

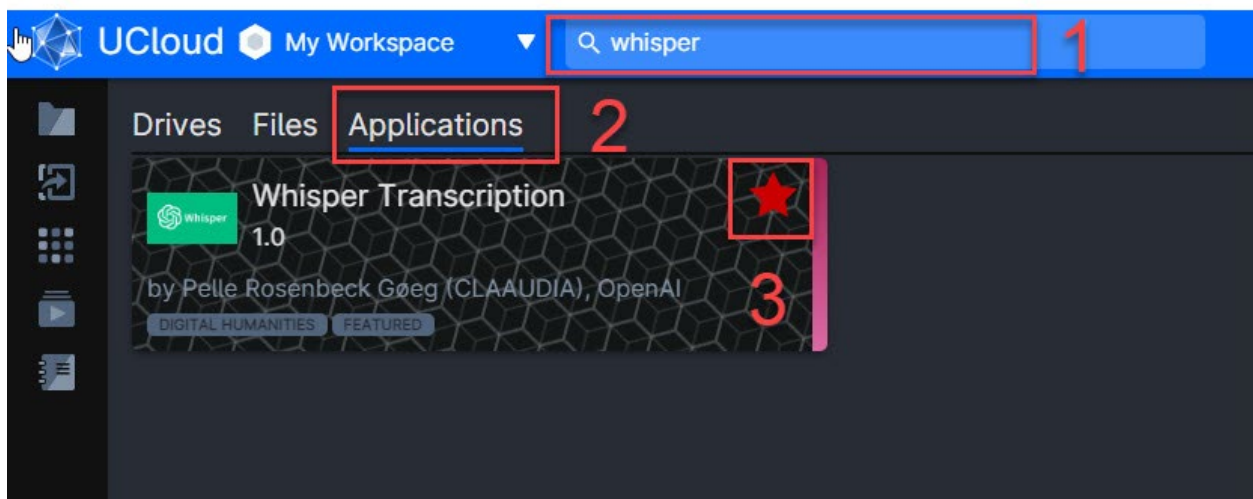
Find Whisper Transcription ved at søge på *Whisper* i søgefeltet øverst.

1: Skriv Whisper og tryk på enter.

2: Vælg "Applications".

3: Føj Whisper til favorit Apps ved at stjernemarkere den, så kan du hurtigere finde den næste gang via Apps i menuen til venstre på Dashboardet.

Klik på App'en for at åbne den.



Start transskribering

I det følgende skal du navngive dit "job", altså kørsel, og vælge hvor lang tid jobbet skal vare, og hvor meget regnekraft der skal benyttes til det og sætte jobbet i gang. Whisper Transcription finder selv ud af hvilket sprog der tales på optagelsen.

1: Navngiv dit job - det er vigtigt at der ikke er mellemrum og andre specialtegn i navnet. Bindestreg - og *underscore* _ er OK.

2: Et interview på omkring 40 minutter kan tage mellem 20 og 60 minutter om at blive transskriberet. Du kan sagtens sætte tiden til 2 timer eller mere hvis du er i tvivl om hvor lang tid det vil tage. Der trækkes kun *credits* for den faktiske regnetid jobbet varer.

3: Klik i feltet for at vælge regnekraft – en menu folder sig ud.

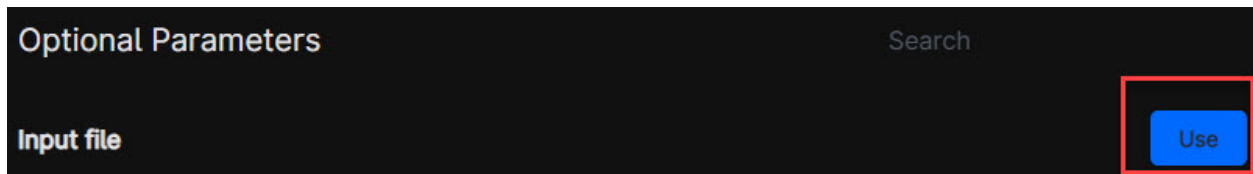
4: Det anbefales at vælge u1-standard-16 til transkriberingsjob.

The screenshot shows the Whisper Transcription interface with the following elements:

- Job name:** A text input field containing "whisper_test_03" with a red box and the number "1" next to it.
- Hours:** A text input field containing "2" with a red box and the number "2" next to it.
- Machine type:** A dropdown menu showing "No machine type selected" with a red box and the number "3" next to it.
- Machine selection table:** A table listing various machine types under the heading "DeiC Interactive HPC (SDU): u1-standard". The row for "u1-standard-16" is highlighted with a red box and the number "4" next to it.

Name	vCPU	Memory (GB)	GPU	Price
DeiC Interactive HPC (SDU): u1-standard				
u1-standard-1	1 (Intel Xeon Gold 6130)	6	None	0,07 DKK/hour
u1-standard-2	2 (Intel Xeon Gold 6130)	12	None	0,16 DKK/hour
u1-standard-4	4 (Intel Xeon Gold 6130)	24	None	0,33 DKK/hour
u1-standard-8	8 (Intel Xeon Gold 6130)	48	None	0,67 DKK/hour
u1-standard-16	16 (Intel Xeon Gold 6130)	96	None	1,36 DKK/hour

Under Optional Parameters skal du klikke Use ud for på Input file:

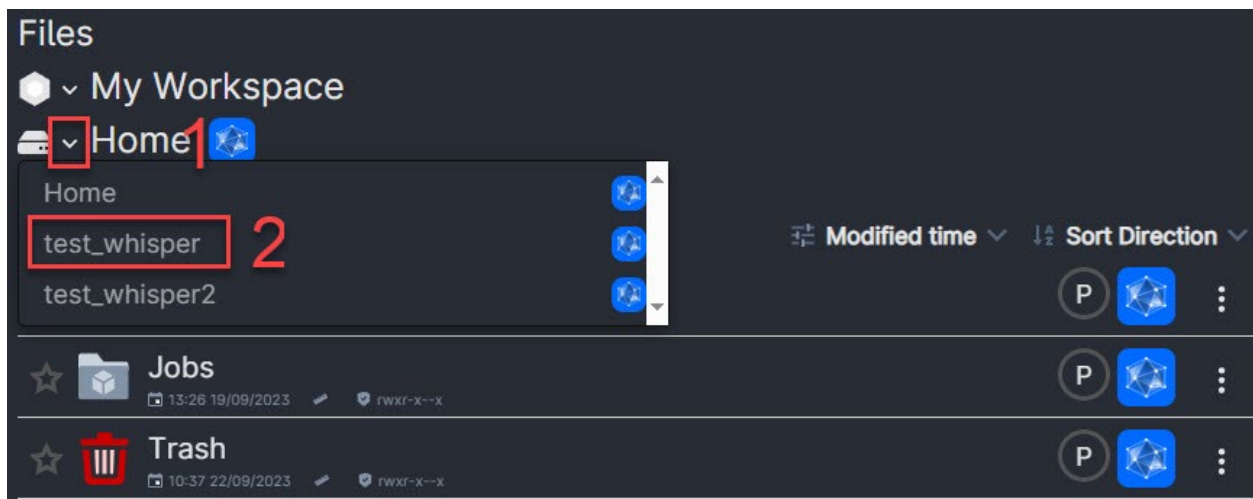


Klik i feltet der aktiveres.

Et nyt vindue åbner.

1: Klik på \vee ud for Home.

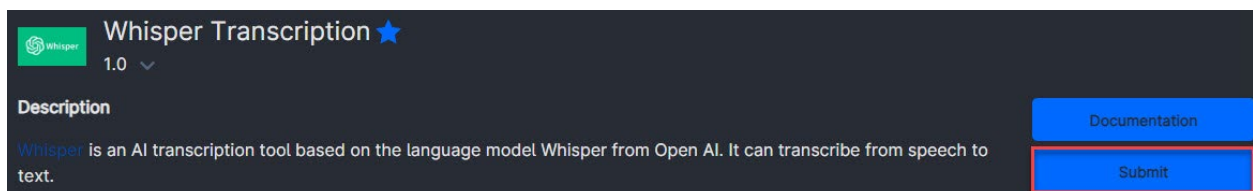
2: Vælg den mappe filen er i.



Når du har fundet filen skal du klikke på Use:



Tryk på Submit øverst til højre og jobbet går i gang:



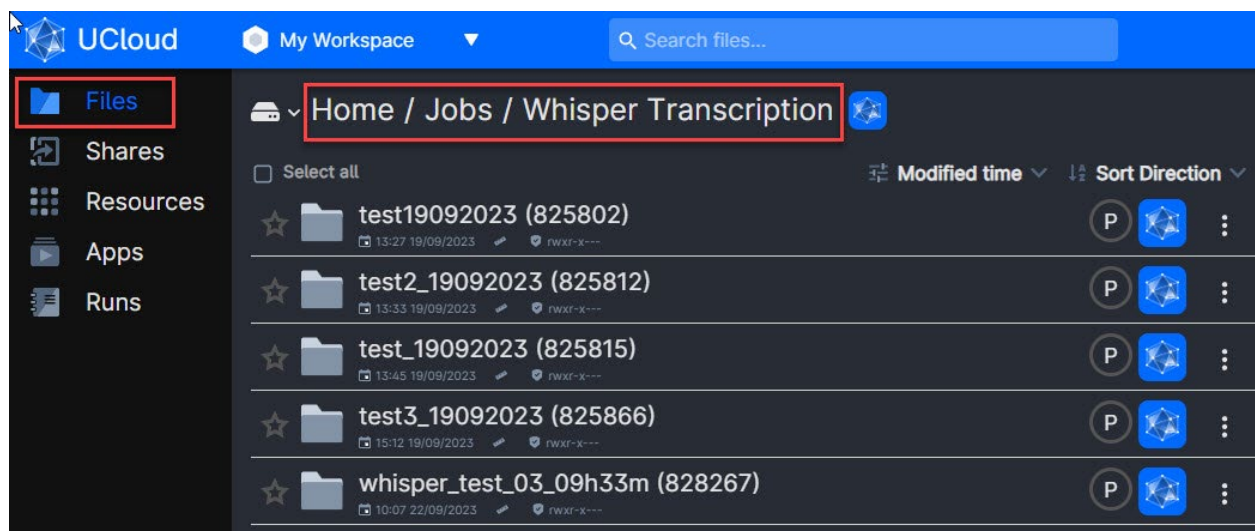
Mens jobbet kører kan du browse væk fra siden, fx til endnu en UCloud side for at sætte endnu et job i gang, lukke browservinduet eller helt slukke computeren. Når du vender tilbage til UCloud efter noget tid vil din transskribering være klar.

Find transskribering

Hvis du ikke har browsset væk fra UCloud jobbet kan du finde transskriberingen under "Job Completed" beskeden i "out" folderen.

Du kan også finde transskriberingen ved at klikke på Files i menuen til venstre, vælge "Home" og herunder "Jobs" og mappen "Whisper Transcription". Alle dine kørsler er placeret her som standard.

For hvert job vil der være en "Out" mappe, og det er her du finder transskriberingen.



Som standard genereres der syv formater af transskriberingen, der kan åbnes i forskelligt software, fx Notepad/TextEdit.

.txt – ren tekstfil.

.vtt – viser tidskoder optimeret til web.

.srt – undertekstformat med tidskoder.

.tsv – tab-separeret format med tidskoder.

.json – åben standard filtype, som kan benyttes i mange forskellige andre apps.

.csv – alle parametre fra Whisper LLM'en er indeholdt i denne filtype.

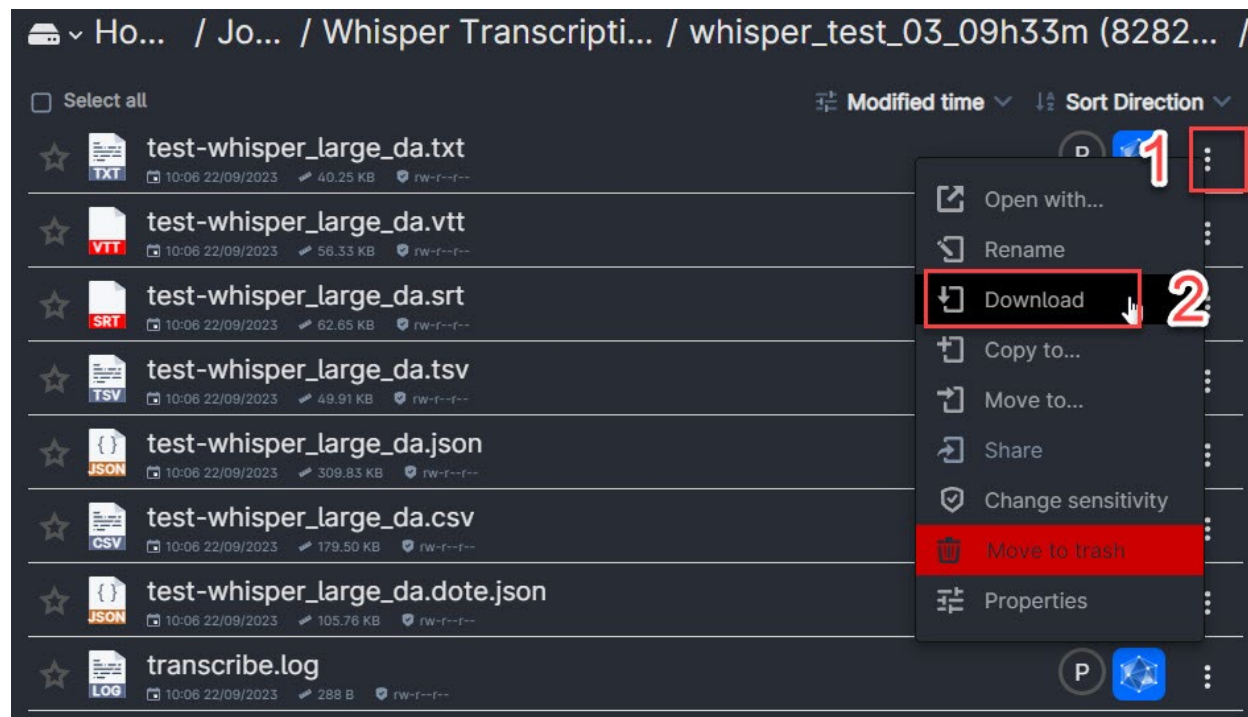
.dote.json – filtype der benyttes i DOTE transkriberingssoftware udviklet på AAU.

Download transskribering

Klik på den folder med det job du vil downloade og klik på "Out" folderen. Du ser nu alle filerne der er genereret for dit job.

1: Klik på de tre prikker ud for den fil du vil downloade.

2: Vælg download.



Flere funktioner

Der er mange flere avancerede funktioner i Whisper Transcription, fx mulighed for at fortælle hvilket sprog der skal transskriberes til (hvis der fx benyttes flere sprog i filen), igangsætte transskribering og download af flere filer samtidig.

Du kan se en demo på engelsk med de mange funktioner her (videoen varer 48 minutter):

<https://www.youtube.com/watch?v=M0IGco--8LQ>

Note omkring computergenereret transskribering

Vær opmærksom på at nøjagtigheden i outputtet fra en computergenereret transskribering ikke altid er på samme niveau som en manuel transskribering. Automatiserede transskriberinger indeholder ofte fejl, som skal rettes manuelt. Derudover fortolker automatisk talegenkendelse ikke altid tone og følelser korrekt, hvilket kan føre til unøjagtige transskriberinger. Automatisk transskriberingssoftware kan være mindre pålidelig end menneskelige transskribenter, hvis lyd kvaliteten er dårlig, eller der er meget baggrundsstøj. Endelig foretager transskriberingssoftware en oprydning i outputtet, så fyldord fjernes.