

## Transcription service

For automatic transcription of audio and video files, you can use Amberscript, ITS' transcription service. This tool is also suitable for files with sensitive data.

**Note:** This is a pilot. If you want to use Amberscript, we ask you to provide one-time feedback on your experience with this new service. The pilot project will run from 17 October 2022 to 30 September 2023.

For this pilot, 1,000 audio hours have been purchased. In the pilot phase, transcription costs will not be charged. We pursue a fair use policy. Please do not request more hours than you actually need.

If the total hours budget runs out before the end of the pilot and new budget is needed, we may decide to charge for new transcription requests. We will inform you in time.

Contact person Utrecht University: Bertil Schaart, [b.m.s.schaart@uu.nl](mailto:b.m.s.schaart@uu.nl)

Administrator Utrecht University: [transcriptie@uu.nl](mailto:transcriptie@uu.nl)

Responsible organisation: [ITS Research and Data Management Services](#)

## Create an account

Do you want to transcribe sensitive data? Please read Utrecht University's privacy policy before creating an account. You can find the policy here:

[https://tools.uu.nl/privacy\\_statement/transcriptie-tool-amberscript/](https://tools.uu.nl/privacy_statement/transcriptie-tool-amberscript/)

To create an Amberscript account, follow these steps:

Are you a student at Utrecht University and do you need this service for your research? Then your tutor should follow the procedure below.

1. Go to <https://app.amberscript.com/login> and choose *Login via SSO*. Then choose *Utrecht University* and log in with your Solis-id.

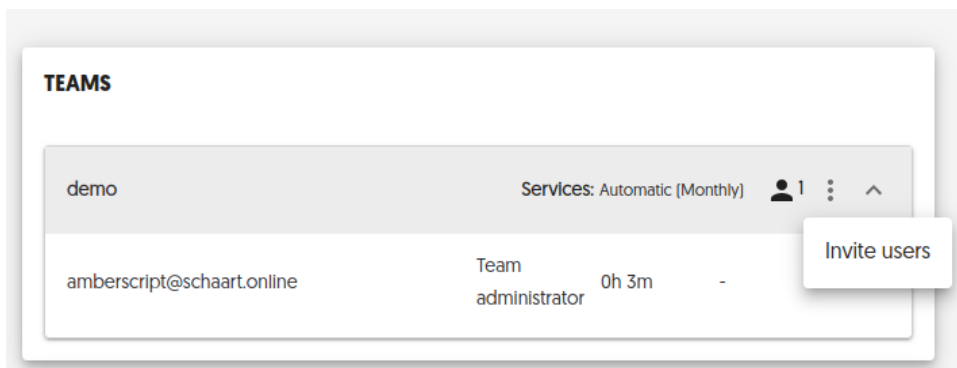
**Note:** have you previously created an account directly with Amberscript outside this pilot and have you used your @uu.nl e-mail address? Then you will first have to get all your files out of your Amberscript account. Then email [operations@amberscript.com](mailto:operations@amberscript.com) requesting them to delete your account. After that, you can re-register via <https://app.amberscript.com/login>. Choose *Login via SSO* and then select *Utrecht University*. Then you can log in with your Solis-id.

2. Your account has now been created, but is not yet linked to a faculty within Utrecht University. Therefore, please send an e-mail to the administrator of the transcription service at [Transcriptie@uu.nl](mailto:Transcriptie@uu.nl). In your request, write the following information:
  - the title of the research project you are working on
  - how long the audio recording(s) are that you want transcribed for your research
  - in which language/languages
  - do the recordings contain GDPR-sensitive information?
  - at which faculty you work

- how you will further edit the automatic transcription {not, self, working student, professional, ...}
- whether you are willing to provide feedback on your experiences with the transcription service
- whether you are willing to measure how much time manual post-processing takes up
- when you want to start the transcription
- when you think you have finished transcribing (including post-editing)

The administrator will then link you to the right faculty and assign you a time budget. Once you have been notified by the administrator, you can start using Amberscript. You can read how this transcription tool works in the manual.

Is a student going to work out the automatic transcription? Then invite them to your Amberscript Team: click on your account at the top right, choose <Manage Organisation>, click on the three bullets next to your team and choose <Invite users>. You can then invite them and give them credits.

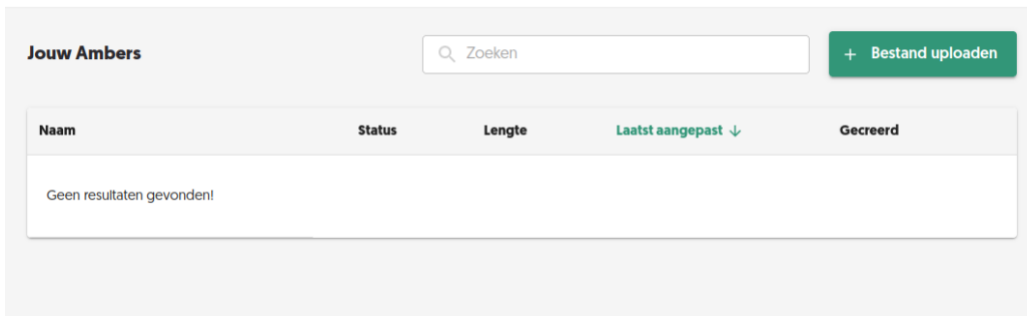


# Transcribing with Amberscript

1. Go to <https://app.amberscript.com/login> and choose Login via SSO. Next, enter the name Utrecht University and then log in with your Solis-id.
2. Choose [Upload]. Watch this instructional video for more details: <https://youtu.be/KZofgSc-k1s>

## Amberscript

Tegoeid 00h 09m 35s b.m.s.schaart@uu.nl



3. Select the language and number of speakers. Then choose whether it is [Transcription] or [Subtitles].
4. [OPTIONAL] If you have a lot of specific terms or proper names in your recordings, you can turn on the glossary. This greatly improves transcription. Click <Use glossary>, <Create new collection> and add the words.

### 4 Other details

Use glossary

Make use of a glossary to help our speech recognition recognise certain words

### Glossary

×

#### Collection name

Demo

#### Names

Can include nouns, such as a company name or spoken names. Hit enter per name.

Bertil × Schaart × Heidelberglaan × Uithof × Amberscript ×

1 Upload bestanden

58.mp3  
00:00:27 • Upload voltooid

Overal toepassen Duits 2 Sprekers

Meer bestanden toevoegen

2 Selecteer diensten

Geselecteerde diensten zullen worden toegepast op alle geuploade bestanden.

Transcriptie  
Krijg een transcript van je audio, bewerk en exporteer hem in meerdere formaten.

Ondertiteling  
Maak ondertitels voor je video en exporteer ze in meerdere formaten.

Samenvatting

Transcriptie • Automatisch

58.mp3 00:00:27

Totale bestand duur 00:00:27  
Uw tegoeid 00:09:35

Subtotaal Valt binnen tegoeid

Totaal Valt binnen tegoeid

Afrekenen

Veilig en versleuteld

5. Click on [Order]. The transcription process will now start automatically.

Amberscript Tegoeid 00h 09m 35s b.m.s.schaart@uu.nl

Jouw Ambers Zoeken + Bestand uploaden

Naam	Status	Lengte	Laatst aangepast	Gecreerd
58.mp3 Transcriptie • Automatisch • German	Wordt verwerkt	00:00:27	05-08-2022	05-08-2022

10 Rijen per pagina

6. You will receive an e-mail when the transcription is ready. In the meantime, you can upload the next audio file for transcription.

7. When the transcription is ready, you can edit it manually. You can see how to do this in this video: <https://youtu.be/Ay2mFabfiEc>

Amberscript Exporteren 58.mp3 Ontvang 100% accurate tekst b.m.s.schaart@uu.nl

Gecreerd 05-08-2022 Starttijd 00:00:00 Lengte 00:00:00 / 00:00:27 Speed 1

Speaker : 00:00:00 Sie interessieren sich für das Wetter in Südbayern und hören den Wetterbericht im Radio.

Speaker : 00:00:07 Und nun der Wetterbericht. Das Wetter Amtes München für Montag, den 28 Oktober Südbayern. Anfangs noch vereinzelt Regen, später zunehmend sonnig, mit höchst Temperaturen bis 18 Grad. Nachts geringer Bodenfrost. Nordbayern wechselhaft mit Schauern, dabei mild und sehr windig. Temperaturen nachts bis -2 Grad.

Opgelagen een paar seconden geleden Tijdsmerkingen aanpassen Laat alle shortcuts zien Wil je de audio automatisch laten verspringen wanneer je klikt? Houdt alt ingedrukt tijdens het klikken Replay Tour

8. After editing, click [Export] to download the text file. This video explains how to do it: <https://youtu.be/319ja9UH0UQ>

9. When you have finished editing and downloaded the files, you must delete the text files and audio files from Amberscript because of privacy rules. When doing so, make sure you save the downloaded files in the right place and remove them from the download folder.

10. For privacy reasons, you should not wait too long before starting manual editing. This prevents audio files containing sensitive information from being unnecessarily long on Amberscript's server.

## Advice

- The speed of automatic transcription depends on how busy the Amberscript server is. In your planning, allow for one hour of processing time per audio file. In practice, it will usually be faster.
- Automatic transcription has an accuracy rate of about 85%-90%. To improve the transcribed text, edit it manually. This obviously takes time, depending on the desired result. You can use the following rule of thumb, based on Amberscript's experience: manually editing the automatic transcription is about 25% of the audio duration. So an eight-hour recording takes about two hours of manual post-processing. We would love to hear about your experiences. Please send them to: [transcriptie@uu.nl](mailto:transcriptie@uu.nl).
- The length of the audio file naturally determines the processing time required, but other factors also come into play; accents, speed of speech, how clearly someone talks, the use of terminology not recognised by the programme, etc.