

Amberscript

NORDUnet 2021 Speech2Text Tender

Annex 1:

Detailed answers to the statement of Requirements

Structure of this document

| | |
|---|----|
| Structure of this document | 2 |
| 2.1.3 EC - Please provide a description of the proposed solution | 11 |
| About Amberscript: | 11 |
| Innovative: The most accurate service in Europe | 11 |
| Human: Amberscript is a fully European company | 11 |
| Reliable: We deliver all-in-one solutions: | 12 |
| Our core technology is available in the form of 4 customer-facing products which allows Amberscript to create a simple IT architecture for a all-in-one solution: | 12 |
| Customized Automatic Speech Recognition: | 13 |
| Integrations into video platforms: | 13 |
| Intuitive text-editing tools for highest productivity: | 13 |
| Europe's largest network of professional transcribers and captioners: | 14 |
| Amberscript is a trusted transcription and captioning partner of hundreds of universities across Europe: | 14 |
| A selection of our customers in Europe: | 14 |
| A selection of our clients in the Nordics: | 15 |
| Integrations with Kaltura, Mediasite and Panopto | 15 |
| Kaltura Integration: | 15 |
| Mediasite integration: | 15 |
| Panopto integration: | 16 |
| Next to Kaltura, Mediasite and Panopto, Amberscript integrates into the following media service platforms: | 16 |
| For transcription and subtitling of audio and video files, using a web-portal for upload and retrieval. | 16 |
| For transcription and subtitling of audio and video files using an API for integration with University legacy systems for both upload and retrieval of audio and video files. | 17 |

| | |
|--|-----------|
| Manual/Human captioning and transcription. | 18 |
| Future Outlook: Technologies that Amberscript is currently developing that can help NORDUnet to communicate in a fully accessible manner: | 19 |
| Live-captioning technology | 19 |
| Automatically translated captions: | 20 |
| 2.1.5 EC - Scope and purpose of data processing | 20 |
| 2.2.3 EC - Customer Institution support | 20 |
| 2.2.4 EC - Incidents | 22 |
| Structural incidents: | 22 |
| Amberscript has defined 4 different levels of communication | 22 |
| Amberscript's status page: | 22 |
| Uncritical updates: | 22 |
| Critical updates and incidents and mission-critical updates and incidents: | 22 |
| For customer-related incidents and customer support: | 23 |
| Response Times: | 23 |
| 2.3.2 EC - Describe the process of adding a new end Customer Institution to the service. | 25 |
| Webform to request access : | 25 |
| Onboarding flow for the web-portal or human transcription: | 25 |
| Onboarding flow for media services (Kaltura, Panopto, Mediasite) | 26 |
| Onboarding flow for API users | 26 |
| 2.3.3 EC - Monthly usage and utilisation reporting should be made available to each Customer, to allow for consumption based invoicing of each Customer Institution. | 27 |
| 3.1.2 EC - As some video and voice recordings contains personal sensitive information it is important to understand how the data is kept secure when handled by the service. | 28 |
| Technical and organizational security measures ("TOMs") | 30 |
| 1. Confidentiality | 30 |

| | |
|--|-----------|
| 2. Integrity | 31 |
| 3. Availability | 31 |
| 4. Resilience | 31 |
| 5. Timely restoration of access to Personal Data | 32 |
| 6. General Technical measures | 32 |
| 7. General Organisational measures | 33 |
| 3.1.3 EC - The solution offered should support a number of different input formats for audio and video, both compressed and uncompressed. | 33 |
| 3.1.4 EC - The solution offered should support a number of different output formats. | 34 |
| Captions: | 34 |
| Transcriptions: | 34 |
| Developer Formats & Integrations into QDA software: | 35 |
| 3.1.6 EC - Integration with SAML based identity federations (Web-portal) | 35 |
| Onboarding a new federation: | 36 |
| User flow and login: | 37 |
| 3.1.7 EC - Please describe your access control mechanism and any supported integration (Web-portal) | 37 |
| Access rights & user roles: | 37 |
| Principle of least privilege: | 38 |
| 3.1.9 EC - Describe the HTML user interface (Web-portal) | 39 |
| Login | 39 |
| The user dashboard | 40 |
| The upload flow | 40 |
| Upload via the web-portal: | 41 |
| How to order transcriptions: | 42 |
| How to order captions: | 44 |
| Upload via the smartphone App | 45 |

| | |
|---|----|
| Email notification of a delivered file: | 46 |
| Export of the file | 46 |
| 3.1.11 EC - Compliance with the EU Web Accessibility Directive (Web-portal) | 47 |
| Buttons | 48 |
| Links | 48 |
| Images | 48 |
| Headings | 48 |
| Forms | 48 |
| Modals | 49 |
| Tables | 49 |
| Toasts | 49 |
| Drawers/Panels | 49 |
| Skip Links | 49 |
| Touch Targets | 49 |
| Focus Indicators | 50 |
| Typography | 50 |
| Language | 50 |
| Layout | 50 |
| Audio | 50 |
| Video | 50 |
| Screen Readers | 50 |
| Keyboard-Only | 51 |
| 3.1.13 EC - Describe the provided captions text editor used for inspecting and/or adjusting of the automatic transcription (Web-portal) | 51 |
| Efficiency Potential for Transcription: | 52 |
| Efficiency Potential for Captioning: | 52 |
| Automatic realignment and AI-based subtitle segmentation: | 53 |

| | |
|---|-----------|
| Adding words to the custom dictionary | 59 |
| Spell-check and auto-correction functionalities | 59 |
| 3.1.15 EC - Describe the API | 60 |
| POST: Upload media files | 60 |
| GET: Get the status of the job | 61 |
| GET: Export files into a variety of formats (STL, SRT, VTT, Text, JSON) | 61 |
| DELETE: Delete a job with all related files | 61 |
| GET: List of previous jobs | 62 |
| 3.1.16 EC - Integration with Media Services API components | 62 |
| Integration with Kaltura REACH (in operation) | 62 |
| User rights and workflow: | 62 |
| Editing captions within Kaltura: | 64 |
| How to set up the REACH integration per customer institution: | 65 |
| Current status of the Kaltura REACH integration: | 66 |
| Integration with Mediasite (in operation) | 66 |
| User rights and workflow: | 66 |
| How to set up the Mediasite integration per customer institution: | 66 |
| Current status of the Mediasite integration: | 68 |
| Integration with Panopto (ready to deploy): | 68 |
| Current status of the Panopto integration: | 68 |
| User rights and workflow | 68 |
| How to set up the designated integration with Panopto | 69 |
| Integration into other media service providers: | 69 |
| 3.1.17 EC - Manual transcription or subtitling | 70 |
| Summary of the workflow for human transcription and captioning services | 70 |
| Detailed description of the workflow: | 70 |

| | |
|---|----|
| Step 1: Submit a file for human transcription and captioning services | 70 |
| Step 2: Automatic Speech Recognition (ASR) & Assigning of a transcriber | 71 |
| Step 3: Professional Transcription | 71 |
| Step 4: Quality Control | 72 |
| Step 5: Delivery | 72 |
| Step 6: Deletion | 73 |
| 1) Deletion in the Amberscript Dashboard. | 73 |
| 2) Optional: Automatic deletion after a safety period | 73 |
| Quality assurance at Amberscript | 73 |
| People: | 74 |
| Process: | 74 |
| Tools | 75 |
| Amberscript's Quality Management System: | 75 |
| Selection of transcribers at Amberscript: | 76 |
| How does Amberscript select and work with transcribers? | 77 |
| Screening | 77 |
| Onboarding & Initial Training of Transcribers | 77 |
| Continuous learning and development: | 77 |
| Feedback loop of data into ASR models: | 78 |
| 3.2.2 EC - Expected accuracy for automatic transcription | 78 |
| Measurement of accuracy: | 78 |
| Description of the dataset: | 79 |
| Description of the dataset: | 79 |
| Results Word-Accuracy measurement | 80 |
| 3.2.3 EC - Describe any additional language(s) or dialects supported as part of the automatic transcription services. | 80 |
| Languages supported in the NORDUnet environment: | 80 |

| | |
|---|----|
| Status of Icelandic: | 81 |
| Status of Norwegian Nynorsk: | 81 |
| 3.2.5 EC - The Nordic R&E community needs the solution offered to be able to adapt to R&E domain-specific language. | 82 |
| The Speech recognition (ASR) | 82 |
| Analysis of the performance of the current models | 82 |
| Data collection | 82 |
| Training of the ASR | 83 |
| Performance analysis | 83 |
| Implementation and ongoing training | 83 |
| Post-processing and text-formatting: | 83 |
| The acoustic and linguistic model | 84 |
| The development of domain-specific models | 85 |
| Continuous improvement and data collection of domain-specific data: | 86 |
| The Amberscript web-portal and transcription editor | 86 |
| Human transcription and captioning | 87 |
| Anonymized relevant data from other Amberscript customers | 87 |
| Public data sources | 87 |
| User-input | 88 |
| Feedback-loop with media services | 89 |
| Future Outlook: End-to-end ASR: | 90 |
| 3.2.6 EC - The Nordic R&E community needs the various speakers in a recording to be recognized as best as possible | 91 |
| Modules that are used to detect speaker changes: | 91 |
| Voice-activity detection: | 91 |
| X-vector diarisation: | 91 |
| Diarisation included in the ASR solutions and meta-data: | 92 |

| | |
|--|-----|
| 3.2.7 EC - Often multiple languages are being used in a single recording, in a teaching environment, i.e. Finnish and English. | 92 |
| Use of specific terminology and anglicisms in a non-English language | 92 |
| Inclusion in the proposed solution: | 93 |
| Detection of the language of an audio file if the language is unknown | 93 |
| Inclusion in the proposed solution: | 93 |
| Code-switching within a single recording (switching from one language to another within the same conversation) | 93 |
| Inclusion in the proposed solution: | 93 |
| The current research approach looks as follows: | 94 |
| 3.2.8 EC - Music and background noises | 95 |
| 3.2.9 EC - Punctuation | 95 |
| Grammatical factors | 96 |
| Speech patterns: | 96 |
| AI-based triggers: | 96 |
| 3.3.1 MR - As part of the evaluation NORDUnet wants to perform a test of the web portal functions and the accuracy of the automated transcription, using generic language for Danish, English, Finnish, Norwegian and Swedish. | 97 |
| Test Accounts & free credit for automatic and human transcription: | 97 |
| User Instructions: | 97 |
| Upload files by selecting 'Upload' in the dashboard | 97 |
| Select languages, number of speakers and type of project (transcription or subtitles) | 98 |
| Select 'automatic transcription' or 'manual / human transcription' | 98 |
| Optional: For human transcription select 'invoice' as payment option and fill in test-credentials | 99 |
| Customer Support: | 99 |
| 3.4.2 EC - Service Capacity | 100 |
| 3.4.3 EC - Automatic Transcription processing speed | 100 |

| | |
|--|-----|
| Future outlook: End-to-end ASR: | 102 |
| 3.4.4 EC - NORUnet wants as few limitations in the service as possible | 102 |

2.1.3 EC - Please provide a description of the proposed solution

Question: *In its solution description, the tenderer shall provide a good and comprehensive account of how the needs (ref.2.1.1) are met with the solution offered.*

The service supports the use cases for the service, as listed below:

- *For subtitling of video files and integration with the NORDUnet on-premise operated media services. Currently, MediaSite, Kaltura, and Panopto.*
- *For transcription and subtitling of audio and video files, using a web-portal for upload and retrieval.*
- *For transcription and subtitling of audio and video files using an API for integration with University legacy systems for both upload and retrieval of audio and video files.*
- *Manual/Human captioning and transcription.*

The description must include information about major technical service components and architecture, including hosting sites, sub-contractors and other provisions. We are looking for simplicity and robustness in the setup.

Answer:

About Amberscript:

Amberscript's outspoken mission is to 'make all audio accessible'. We believe that Europe will make a shift towards a more inclusive society for deaf and hard of hearing in the next few years. Therefore our vision is that, in a couple of years, every video will have very accurate subtitles. In order to democratize access to top-notch captioning and transcription services, Amberscript merges artificial and human intelligence into the best transcription and captioning solutions with the highest accuracy possible. At the core of everything that Amberscript does stand 3 values:

Innovative: The most accurate service in Europe

Amberscript uses the most innovative language technologies in order to find innovative solutions to easily transform audio into text. By combining highly accurate automatic speech recognition with human language experts, Amberscript creates a self-reinforcing cycle of improvement. Amberscript's AI modules learn from every correction that the language experts make.

Human: Amberscript is a fully European company

Headquartered in the center of Amsterdam, Amberscript's team consists of more than 16 nationalities that follow our mission to make all audio accessible. Amberscript has a network of more than 1000 language experts in over 12 different European languages that deliver human captions with the best quality possible.

Amberscript forms long-term partnerships with like-minded organizations in order to deliver customized solutions that achieve the highest level of accuracy. On top of that, Amberscript understands that the implementation of new technologies requires change management and dedicated customer service.

Reliable: We deliver all-in-one solutions:

Amberscript designed a solid infrastructure which is the basis for all required services: ASR, integrations into media services, text-editing tools and human captions. Having one single backbone for all these services, allows Amberscript to deliver reliable services from one solid infrastructure. At the same time, we are in full control of our customer's data and make sure that data is processed fully compliant to legislation around privacy and IT security.

Our core technology is available in the form of 4 customer-facing products which allows Amberscript to create a simple IT architecture for a all-in-one solution:

Amberscript's core technology is deployed in the cloud environment of Amazon Web Services located within the EU. Through a powerful API, the core technology of Amberscript is made available to clients through four products which match all use cases described in requirement 2.1.1.

Major technical components include:

- the Automatic-Speech-Recognition (ASR) workers
- the Amberscript API
- the transcription and captioning editor
- the platform to coordinate human transcription and captioning solutions

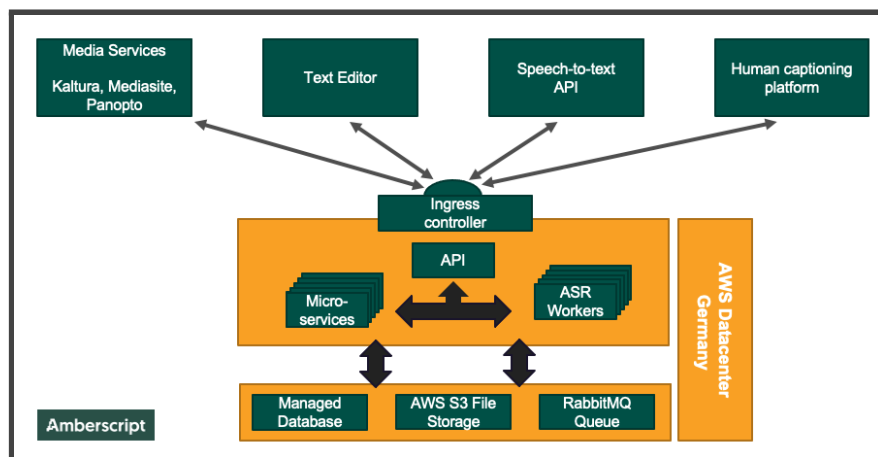


Image: Amberscript's main architectural setup

All of Amberscript's service components are hosted on data centers of Amazon Web Services within the European Union. No data will ever leave the European Economic Area.

For detailed explanations, see below:

Customized Automatic Speech Recognition:

Customized Automatic Speech Recognition (ASR) engines for European languages for highest accuracy. We are experts in developing speech recognition modules for European languages and specific domains within those languages. Amberscript has proven that generic speech recognition models can easily be outperformed by domain-specific ASR models as Amberscript has developed for the Dutch education sector, the Dutch political sector and various German cities and parliaments.

Integrations into video platforms:

Active integrations with video platforms such as Kaltura, Mediasite and Panopto and a powerful API. We believe that processes should be as automated as possible, therefore we directly integrate with media services that are used by our customers, but we also offer a stand-alone API so that our customers can use our services wherever necessary.

Intuitive text-editing tools for highest productivity:

Intuitive text-editing web-portal so that everyone can create transcripts and captions with the help of AI. Our editor is used by more than 1 million users from Europe to transform their audio and video more efficiently into text. Current users of this solution include more than 100 universities in Europe such as: University of Amsterdam, University Utrecht, University of Tilburg, University of Frankfurt, Humboldt University of Berlin, University of Barcelona, HEC Paris, and many others.

Europe's largest network of professional transcribers and captioners:

Europe's largest platform of professional transcribers and captioners to offer manual/human captions and transcriptions. More than 1000 transcribers from different language areas are working daily on our platform to deliver transcripts and captions of the highest quality. In the past years, Amberscript delivered human transcription and captioning projects to universities such as: University of Amsterdam, INSEEC Business School, Stockholm University, University of Malmö, University of Copenhagen, HAMK Hämen University and many others.

Amberscript is a trusted transcription and captioning partner of hundreds of universities across Europe:

A selection of our customers in Europe:



A selection of our clients in the Nordics:



Integrations with Kaltura, Mediasite and Panopto

Kaltura Integration:

Amberscript and Kaltura have a close partnership in which Amberscript supplies high-quality captioning services (automatic and human) to Kaltura's European clients. Amberscript one of the few (and the only European) companies with a native integration into Kaltura REACH, which makes it possible to seamlessly request captions (automatic, human and translations) from the Kaltura user interface. The integration is commercially deployed and currently running for NORDUnet.

- ***For a detailed description of the Kaltura REACH integration, please refer to 3.1.16***

Mediasite integration:

Amberscript is one of the few selected partner captioning providers that are integrated into Mediasite. As the only European 'partner captioning provider', Amberscript provides accurate ASR and human captioning solutions to Mediasite users within the EU.

The integration is commercially deployed from Mediasite version v.8.0.1 onwards.

- **For a detailed description of the Mediasite integration, please refer to 3.1.16**

Panopto integration:

Amberscript and Panopto have been setting up an integrated partnership in the past months. The workflows and IT infrastructure has been designed, and the integration is ready to be fully developed with a customer. In case of a contract award, Amberscript will collaborate with NORDUnet and Panopto to finalize the development of the integration. With the integration finalized, Amberscript will be the only European captioning provider within the Panopto ecosystem.

- **For a detailed description of the Panopto integration, please refer to 3.1.16**

Next to Kaltura, Mediasite and Panopto, Amberscript integrates into the following media service platforms:

- Opencast Video Management System
- Aventia
- Mediaflow
- Mjoll
- Streamio

Next to process automation for the creation of captions, Amberscript has integrations with most commonly used Qualitative-Data-Analysis software packages such as ATLAS.tio, MAXQDA and NVivo. These integrations allow researchers to easily import their Amberscript-transcriptions into QDA software packages for further analysis. These integrations will be available to NORDUnet free of charge.

For transcription and subtitling of audio and video files, using a web-portal for upload and retrieval.

Amberscript developed an intuitive web-portal which is used by more than 1 million users across Europe for transcription and captioning.

The web-portal allows users to:

- Securely upload audio and video files through their browser
- Transform these audio- and video files into text either automatically with the use of automatic speech recognition or manually with the combination of AI and professional language experts
- Edit the transcriptions and captions
- Export transcriptions into various text formats (such as .docs, .txt or JSON)
- Export captions into various subtitle formats (such as .srt, .vtt or .stl). Amberscript developed a unique algorithm that allows users to automatically transform transcripts into synchronized and well-formatted subtitles by applying different subtitling rules (e.g. BBC guidelines: Maximum 2 rows, maximum 32 characters per row).
- The web-portal is used by high-profile organizations such as the European Commission in Brussels and more than 200 universities across Europe.
- Edits that are made within the web-portal are anonymized and used to improve the speech-recognition capabilities of Amberscript over time. Users naturally have the option to opt-out of this process, but as many users want to contribute to a better speech recognition, the speech-corpus of Amberscript is steadily growing and thereby the speech recognition models become more and more accurate over time. This allows us to significantly outperform other speech recognition providers, with as much as 20 to 30%.
 - **for a detailed description, see 3.1.7**

For transcription and subtitling of audio and video files using an API for integration with University legacy systems for both upload and retrieval of audio and video files.

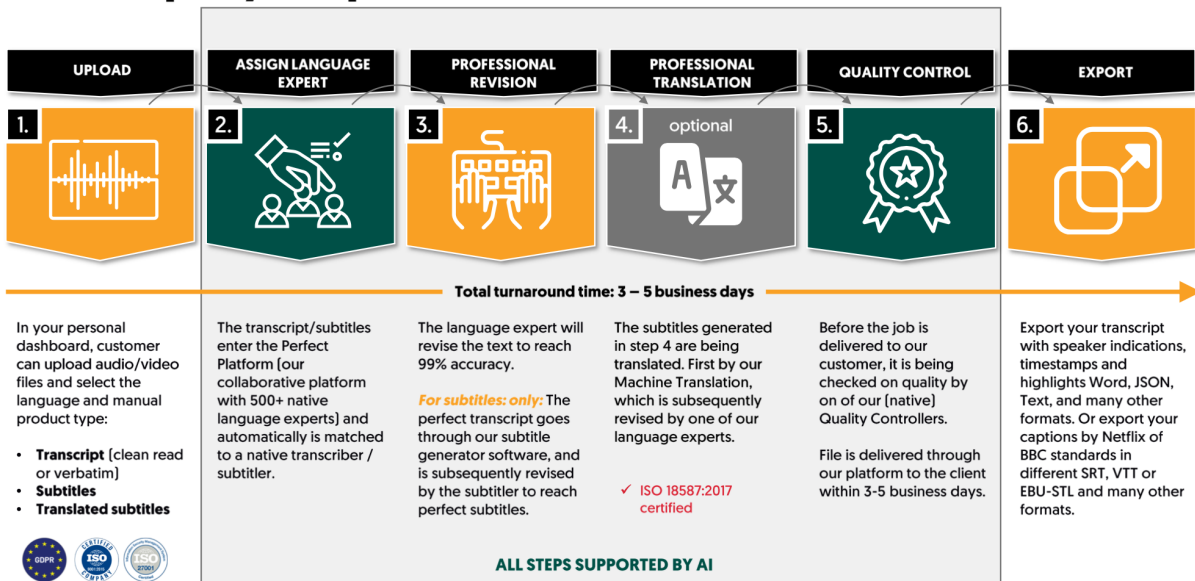
- Amberscript believes that automation is the only way to make audio and video accessible at a large scale. Therefore we have developed an easy-to-use API which allows users to use all of Amberscript's services directly via API.
- The API description can be found on github: <https://amberscript.github.io/api-docs/#introduction>
- The API allows users to order different transcription types such as 'Transcription' or 'captions'.
- For both transcription types, different jobtypes are available:
 - Direct: Automatic transcriptions / captions delivered by ASR
 - Perfect: Fully corrected transcriptions / captions delivered by a combination of ASR and professional language experts
- The API is developed based on REST standards
- Data is encrypted using the latest version of HLS (in motion) and with AES-256 encryption (at rest).
-
- ***For a detailed description, please refer to 3.1.15***

Manual/Human captioning and transcription.

- In the past years, Amberscript has become one of the largest and most reliable providers of professional transcription- and captioning services in Europe.
- We use a multitude of technologies to automate as much as possible in the transcription and captioning process so that the overhead costs are kept low. At the same time, these technologies allow us to offer the best quality at fastest delivery times.
- We combine technology with a community of highly qualified language experts who are native speakers in the languages they are transcribing or captioning in. Many of our language experts have an academic degree in linguistics or other relevant fields.

Our network of language experts is 24/7 available for transcripts, subtitles and translated subtitles

The step-by-step workflow for our manual services



- In order to ensure the highest quality, the files are processed through a highly accurate Automatic Speech Recognition (ASR) module, which creates a draft transcript in less than 20 minutes. The draft transcript, along with detailed work instructions is then assigned to the best-qualified transcriber on the Amberscript platform. The transcriber uses the AI-supported Amberscript-Editor to review the draft transcript and ensure it is accurate and free of error. Once the transcription is complete, the files are submitted for Quality Control (QC), where final checks are performed based on Amberscript's detailed quality control guidelines. The customer is notified via email when the transcripts are ready, and they can view the file by logging into their secured account on Amberscript. Users can re-listen to certain audio fragments and export files from the system
- As of may 2021, Amberscript transcribed more than 25.000 hours of video and audio files in different European languages.
- Currently, Amberscript's manual captioning solutions are used by organizations such as 'Hämeen University' (Finland), Sveriges Kommuner og Regioner (Sweden), Region Hovedstaden (Denmark), The Norwegian Ministry of Education (Norway), the City of Monheim (Germany), or Wageningen University (the Netherlands).
- **For a detailed description, please refer to 3.1.17**

Future Outlook: Technologies that Amberscript is currently developing that can help NORDUnet to communicate in a fully accessible manner:

Live-captioning technology

In 2020, Amberscript launched a domain-specific ASR model for the German national Parliament of Mecklenburg-Pomerania which currently creates accurate live-captions for all parliamentary hearings. The launch of this model was positively received from interest groups for deaf and hard of hearing in Germany as it supports the real-time accessibility of the parliamentary video stream.

Amberscript is currently working on integrating live-captioning capabilities into a variety of workflows for its customers to also make live-events digitally accessible.

The Challenge:

In light of the recent legislation around digital accessibility, the State Parliament of Mecklenburg-Pomerania wanted to ensure that citizens with auditory disabilities cannot only follow pre-recorded debates, but that they are equally enabled to follow live-debates in the state parliament. Next to that, the Parliament wanted to automate large parts of the internal process of parliamentary documentation. To achieve high accuracy of automated speech recognition for the local accents and the specific terminology in the parliament, the Parliament partnered up with Amberscript.

Mecklenburg
Vorpommern
Landtag



Our Solution:

As a first step, Amberscript collected hundreds of hours of domain-specific audio recordings, which were transcribed and turned into training data for the Parliament. Based on this dataset, a customer-specific ASR model was developed, which was specifically tailored to political terminology and local accents of Mecklenburg-Pomerania. After a successful integration into the systems of the Parliament for post-transcription, Amberscript innovated even further to create the first live captioning technology in Germany. All public debates are now streamed with easily readable captions via Amberscript's live captioning technology.

Link to a video sample:

<https://www.amberscript.com/en/news/Amberscript-launches-live-captions-technology>

Automatically translated captions:

Amberscript is currently investigating possibilities to integrate machine-translation capabilities into its services so that speech in one language can effortlessly be translated into another language. The first projects have yielded promising results and as the transcription and machine-translation accuracies increase, Amberscript expects to integrate machine-translation services into its services in the near future.

2.1.5 EC - Scope and purpose of data processing

Question: *Please provide your standard Data Processing Agreement or other/similar documentation describing the scope and purpose of data processing and the parties involved. Including any data retention policy etc.*

We are looking for completeness of documentation in relation to GDPR and information security.

Answer:

Please see the sample Data Processing Agreement '20210517 Sample Data Processing Agreement NORDUnet'

2.2.3 EC - Customer Institution support

Question: Each Customer Institution should have the ability to report problems directly to the vendor. Please describe the different ways to report problems to your support / Helpdesk, e.g. phone, mail, web, etc. Please also describe how progress of any reported issue can be monitored.

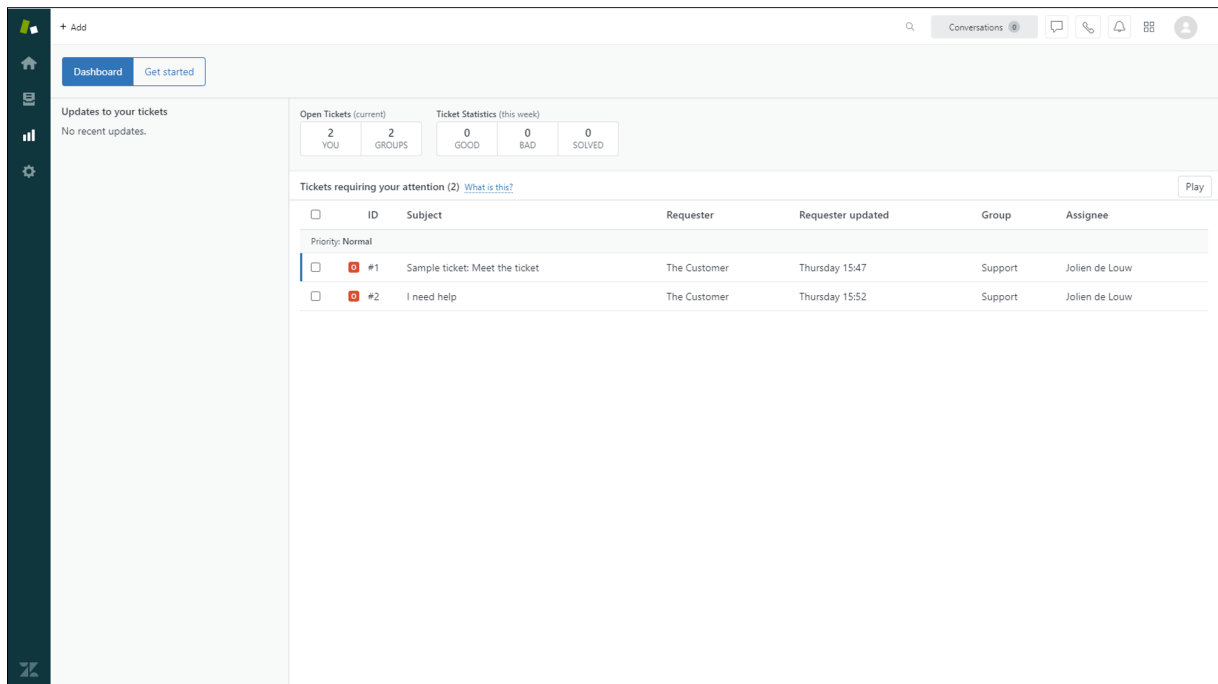
We are looking for ease of use and flexibility in options for communicating with the support / helpdesk.

Answer:

Amberscript has a customer support team that is fully dedicated to helping customers as fast as possible.

- The helpdesk is available via the following channels:
 - Chatbot with FAQs and instruction videos: 24/7
 - Live chat and email support: From 9:00 to 18:00 (CET) (In Swedish, Norwegian, Danish, Finnish or English)
 - Phone support: From 9:00 to 18:00 (CET) in English
- For any incoming issue through any of the channels, a ticket will be created by our helpdesk. The customer will be immediately informed about the creation of a new ticket via email.
- Using the ticketing system Zendesk, Amberscript can efficiently create and monitor the progress of customer-related issues. The customers who are involved in a ticket (e.g. the reporting customer or admins) have access to their tickets at all times.

- The average response time to incoming requests is 1 hour on business days (during working hours).
- The tickets are categorized and prioritized based on the severity of the ticket.
- Customers will automatically receive an email notification once the support ticket has been updated by our customer support or tech team
- NORDUnet will receive a quarterly report on the number of tickets created and the average time to resolve support tickets.



- In case of emergency, NORDUnet can reach out to their dedicated customer success manager who is available 24/07 for issues that require immediate assistance.

2.2.4 EC - Incidents

Question: Please describe the process and methods for communicating incidents to Customers.

Looking for **robustness in the process** and **flexibility in the options for consuming the incident information.**

Answer:

Amberscript believes in handling and resolving incidents as fast as possible so that our customers never need to delay their work because Amberscript isn't available.

To adhere to this ambition, Amberscript distinguishes between structural incidents and customer-related incidents.

Structural incidents:

Amberscript has defined 4 different levels of communication

- Amberscript status page: A website to track the status of the service
 - Uncritical updates (e.g. newsletters)
 - Critical updates (e.g. change of core functionalities)
 - Mission-critical updates (e.g. planned downtime, unavailability of the service)
-
- **Amberscript's status page:**
 - Amberscript offers a status page which shows the status and availability of services that are available to Nordunet.
 - Scheduled downtimes and maintenance windows will be announced on the status page.
 - **Uncritical updates:**
 - When creating an account with Amberscript, all users will be enrolled in the mailing list for uncritical updates. Users have the option to opt out in every email and directly at signup or from every email they receive
 - **Critical updates and incidents and mission-critical updates and incidents:**
 - Organization admins and group admins are automatically enrolled for critical- and mission critical updates. Every user has the option to opt out from these mailing lists in the emails or in their personal preferences
 - When necessary infrastructure upgrades require a short downtime, the time window for scheduled downtimes will be on Saturdays between 3:00pm - 4:00pm CET
 - Admin users will be informed at least 1 week in advance before scheduled downtimes for maintenance.
 - For other mission critical incidents (such as unscheduled downtime), admin users will be informed as soon as the incident is identified.

For customer-related incidents and customer support:

- Amberscript has implemented an Incident Management Processes that distinguishes between Incidents (service interruptions) and Service Requests (customer or user requests that do not represent a service disruption, such as a password reset). Service interruptions are handled through Incident Management, and Service Requests through Request Fulfilment.
- The Incident Management process can be triggered in various ways: A user or customer may report an issue, technical staff may notice a (potential or actual) failure, or an Incident may be raised automatically by an event monitoring system

- All Incidents are logged in our Incident Management System 'Zendesk' where their status can be tracked, and a complete historical record maintained. Initial categorization and prioritization of Incidents is a critical step for determining how the Incident will be handled and how much time is available for its resolution.
- Amberscript has implemented processes to identify incidents and performance issues even before customers do. Amberscript's infrastructure is constantly monitored for KPIs such as:
 - Server Load
 - Transcription time
 - Transcription queue
 - Infrastructure health
- Amberscript offers a range of tools and documentations that allow users to solve smaller incidents by themselves. These tools include a chatbot which is available 24/7 and a database of frequently asked questions (FAQs) with detailed solutions and tutorials.
- In case a customer reports an incident, Amberscript's first and second line support can help the customer solve the incident as fast as possible. Once the incident is solved, the customer receives a notification from the first- or second line support or via the ticketing system.

Uptime: The uptime of our software was 99,9%, measured in the past 12 months. This is due to our advanced technical system that scales automatically in parallel with the growth of the number of customers.

Response Times:

We have a clear incident management in place, and commit to the following bug response time:

| Priority | Response | Required | Remedy time | Final | Escalation |
|--|--|--|---|---------------------|---|
| P1 = Amberscript product or service is completely unavailable | Within 15 minutes during business hours. | 24/7 efforts until problem resolved | 6 hours if it's within Amberscript's control | <= 21 calendar days | CEO |
| P2 = Major feature(s) are impaired | 4 business hours | All diligent efforts considering the circumstances and nature of the problem to resolve as soon as reasonably possible | 12 hours if it's within Amberscript's control | <= 30 calendar days | After 8 business hours, escalate to P1 |
| P3 = Minor feature(s) are impaired | 8 business hours | Diligent efforts to resolve during business hours | Best effort | Best effort | After 24 business hours, escalate to P2 |
| P4 = Request a feature or raise a question about the use of the system. It has a little or no material effect on ability to use Amberscript products and services. | 48 business hours | Log any feature request and prioritize | Best effort | Best effort | N/A |

2.3.2 EC - Describe the process of adding a new end Customer Institution to the service.

Question: Please describe the steps and required information to on board a Customer Institution.

We are looking for **a simple process and clarity in the information that need to be provided, preferably a webform or similar.**

Answer:

Once the framework agreement with NORDUnet has been signed, customers will be able to make call-off agreements according to the service catalogue in the framework agreement. For an end customer institution to request access, Amberscript set up a simple process so that end customers will start their onboarding flow within 24 hours of submitting a request.

Webform to request access :

- 1) The end customer institution fills in a webform (see example on the right)
- 2) Amberscript's support team proceeds with the onboarding process according to the desired delivery format.
- 3) Amberscript assigns a dedicated onboarding manager that serves as the main point of contact to the end customer institution throughout the entire onboarding process. The onboarding manager will be responsible to enable customer institutions to work with Amberscript as smoothly as possible.
- 4) Amberscript's support team and onboarding manager will provide customers with the following onboarding information within 24 hours after submitting a request through the webform:

Onboarding flow for the web-portal or human transcription:

- Amberscript creates the requested accounts for organization admins and group admins.
- Amberscript sends out detailed instructions and onboarding tutorials (consisting of PDF instructions and a recorded webinar) to the admin users
- Admin users can then add and delete other users within their admin dashboard

The screenshot shows a webform titled 'Amberscript' with the NORDUnet logo. The form contains the following fields and options:

- Organisation Name:** Stockholm University
- Reference to call-off agreement:** SU12356
- Main contact person:** Thomas Oest (First Name: Thomas, Last Name: Oest)
- Email:** test@su.se
- Desired Product:**
 - Web-portal or human transcription
 - Integration into media services: (Kaltura REACH, Panopto, Mediasite)
 - API
- User information - Web-portal or human transcription:**
 - User 1:** Thomas Oest (First Name: Thomas, Last Name: Oest)
 - Email:** test@su.se
 - Type of user:**
 - Organisation Admin
 - Group Admin
 - Cost-center for this user (or any other invoice-related information):** SUS4321

- The Amberscript web-portal is designed to be as intuitive as possible. However, experience shows that in order to successfully onboard a large amount of users, a multi-channel onboarding programme can ensure the success of the implementation. Therefore Amberscript offers the following onboarding resources free of charge:
 - (1) **Monthly onboarding webinar with new admin users:** One of Amberscript's support team members explains step for step how new admin users can set up their Amberscript organizations in a short webinar of 15 minutes. This webinar is conducted once a month live with the option to ask questions at the end of the webinar. Next to that, the webinar is available as video-on-demand at all times with the option to ask questions directly to Amberscript support team via email or chat.
 - (2) **Onboarding webinar for regular users:** Amberscript is easy to use. However, there are some shortcuts and gimmicks that need some explanation to reach full productivity. In this webinar, one of Amberscript's transcription experts will explain the functionalities of the transcription and captioning tool and share best-practices on how to work even more efficiently.
 - (3) **Frequently Asked Questions:** In our FAQ, more than 100 questions are listed with detailed answers and detailed instructions on how to use the Amberscript platform and how to solve frequently asked questions
 - (4) **Automated onboarding email flow:** All users that will be added automatically receive an onboarding email flow with detailed instructions on how they can access the web-portal, how they can add new files, how they can order automatic or human transcriptions and captions and how they can edit files as fast as possible.
 - (5) **Dedicated customer-support helpdesk:** Our support team is available between 9:00 and 17:00 on weekdays. Please find a detailed description of our customer support services in 2.2.3.

Onboarding flow for media services (Kaltura, Panopto, Mediasite)

- Amberscript reaches out the person who requested access to the services with detailed instructions on how to access Amberscript's transcription services through media services.
- Amberscript will provide personalized onboarding to the person requesting access to ensure a smooth integration process.

Onboarding flow for API users

- Amberscript will send out detailed API description documents to the person who requested the API access.
- Next to that, Amberscript will share a document with best-practices and use cases from other customers with the person who requested access.

- The requester also will have access to Amberscript's technical support in case there are any technical questions about the implementation of the API.
- The onboarding manager will be the main point of contact for the customers to make sure that the API implementation works as smoothly as possible. The onboarding manager will coordinate with Amberscript's tech support and development team to enable the customer to integrate the API as smoothly as possible.

2.3.3 EC - Monthly usage and utilisation reporting should be made available to each Customer, to allow for consumption based invoicing of each Customer Institution.

Question: Please describe the process for making these monthly usage reports available and if possible provide a sample report.

We are looking for reports to be made available in a machine consumable form .csv, .xml or similar and flexibility in how they are made available.

Answer:

Amberscript tracks the usage based on 4 hierarchical user levels

- 1) NREN
- 2) End Customer Institution
- 3) Groups (e.g. Faculties, research teams)
- 4) Users

Based on experience with different universities across Europe, the ability to attribute usage on a group level (e.g. for faculties or research) times is crucial for customers in the education sector to allocate costs to the correct cost-centers.

Amberscript can therefore easily track the usage of each individual end-user and allocate them on any hierarchical level so that NORDUnet and it's member institutions can allocate them to the necessary cost-centers. Reports are available on the same hierarchical levels:

- NORDUnet will receive a monthly report via email
- NRENs can request a report at all times
- End-customer institutions can find a real-time usage report in their organization admin dashboard
- Group admins can access a real-time usage report in their group-admin dashboard
- Users can track their own usage in their individual dashboard

As the reports are exported directly from the backend of Amberscripts infrastructure, the reports can be provided in either .csv, .xml or .json format.

Please find a sample report in .xls, .csv and .pdf format **in attachment 'SAMPLE Usage Report 2021'**.

3.1.2 EC - As some video and voice recordings contains personal sensitive information it is important to understand how the data is kept secure when handled by the service.

Question: Please describe the level of encryption applied in both transit and at rest, as well as any other measures in place to ensure the privacy and integrity of the data.

We are looking for best/highest level of encryption and measures in place to ensure the data privacy.

Answer:

Amberscript processes data of the highest confidentiality on a daily basis. Therefore, it is crucial for Amberscript to structure all processes, systems and tools in such a way that the data of our clients is treated with high IT-security and privacy standards.

Amberscript takes a variety of measures to ensure the highest standards of IT security:

- Amberscript's services are protected by the latest version of HLS
- Amberscript's database and hard-disks are encoded.
- Encryption in transit: Traffic to the database is encrypted using SSL and all communication between frontend and backend is encrypted using the latest version of TLS.
- Encryption at rest: All data is encrypted using AES-256
- Amberscript accounts are protected by strong passwords (12 characters, combination of letters, digits and special characters, upper-case required)
- Two-factor authentication and/or external authentication via SAML2.0 is enabled for logins on the Amberscript platform.

- All files are stored and processed in ISO27001 and SOC2 certified data centers of Amazon Web Services, located in Frankfurt, Germany.
- Files will never leave the data centers of Amberscript
- Files will be stored and processed exclusively within the European Economic Area. No files will ever be processed or stored outside of the EEA.
- Amberscript defined Technical and Organizational Security Measures (TOMs), which are inspired by ISO27001 IT security standards.
- Amberscript is currently pursuing ISO27001 and ISO9001 certifications. See Annex 'Annex 3 - LoI ISO 27001 and ISO9001 certification' for a confirmation of the planned audit.
- Amberscript creates back-ups of data on an hourly basis. The main server and all back-up servers are synchronized continuously. Files that are added/deleted to the main server will be therefore automatically added/deleted to/from all back-up servers.
- Amberscript strictly follows the principle of least privilege:
 - the assigned resources (transcribers and quality checkers) have the bare minimum privilege to 'listen and edit only' while streaming the audio files and transcripts. This ensures that the transcribers can't download, copy, paste or export the files.
 - Audio and text files will never leave the secured Amberscript platform
 - Audio and text files cannot be downloaded or copied by Amberscript staff, transcribers or quality checkers
 - As few staff members as possible will have access to the clients of the customers. All access is limited to time so that staff has access only as long as it is strictly necessary.
- Amberscript's staff and transcribers are continuously trained on IT security and data privacy measures
- All transcribers must sign Non-disclosure agreement (NDA) and Data-processing agreement (DPA) as part of the transcriber selection process. Detailed training is provided to understand the content and penalties imposed by Amberscript in case of a breach.
- Amberscript directly employs all members of Quality Control. To qualify as a quality control member, resources must have at least five years of transcription experience and be native speakers of their assigned language.

- Amberscript’s environment is protected by an Application Security Mangement software which protects microservices, applications and APIs. The following tables shows measures taken to protect the infrastructure against different types of malicious attacks:

| Amberscript | | | | |
|--|---------------------------------------|--|-----------------|--|
| Amberscript's Application Security Management System mitigates vulnerabilities effectively | | | | |
| Vulnerability | In-App Web-Application Firewall (WAF) | Content-Security Policy and Security Headers | User Monitoring | Runtime Application Self Protection (RASP) |
| Account Takeover (ATO) | | | X | |
| lickjacking | | X | | |
| Cross-site Scripting (XSS), Reflected and Stored | X | X | | |
| Local File Inclusion (LFI) | X | | | X |
| MIME sniffing | | X | | |
| NoSQL Injectiong | X | | | X |
| Referrer Leakage | | X | | |
| Remote Code Execution (RCE) | X | | | |
| Security Scanners | X | | | |
| Server-side Request Forgery (SSRF) | X | | | X |
| Shell Injection | X | | | X |
| Shellshock | X | | | X |
| SQL Injection | X | | | X |

Technical and organizational security measures (“TOMs”)

Amberscript takes various Technical and Organizational Measures (TOMs) to ensure full security and full adherence to GDPR legislation. Please see the following list of TOMs which are also included in Amberscript’s standard DPA:.

1. Confidentiality

TOMs to ensure the confidentiality of the Personal Data processed by Amberscript.

1.1 Data is encrypted in transit: data transmitted between the user's browser and Amberscript is always encrypted over HTTPS using TLS protocols with minimum 128-bit keys and using SHA256 certificates. The Processor uses modern, strong ciphers for encryption. Known-weak ciphers are explicitly disabled with regular protocol reviews.

1.2 Data is encrypted at rest: data at rest, for example in backups or on the Processor's computers, is always encrypted using AES encryption with minimum 256-bit keys. This mitigates the risk of data falling into unauthorised hands, for example due to network exfiltration or stolen devices.

1.3 Data retention policies: the Processor maintains policies to ensure the minimal amount of Personal Data is retained and that Personal Data is not retained any longer than necessary. This mitigates the risk of accidental or deliberate disclosure of Personal Data.

1.4 The principle of least privilege access is embedded at all levels in the Processor, from staff down to operating system server processes. This ensures that only the data that is authorised to be processed may be accessed. This mitigates the risk of accidental or deliberate disclosure of Personal Data.

1.5 The data centres used by the Processor implement multiple physical access controls to prevent unauthorised people from physically accessing data processing equipment which processes or uses Personal Data.

1.6 The Processor only authorises specific staff to access the Service's production systems. This mitigates the risk of accidental or deliberate disclosure of Personal Data.

1.7 The Controller's Personal Data is kept logically separate from other Personal Data. This mitigates the risk of accidental disclosure of Personal Data.

2. Integrity

TOMs to ensure the integrity of the Personal Data processed by Amberscript.

2.1 All access of Personal Data, including use of the Service by the customer, is logged and an audit trail of changes is maintained. This mitigates the risk of Personal Data being altered or deleted either accidentally or deliberately and supports the data subjects' right to rectification.

2.2 The customer may view, update, and delete all their Personal Data held in our systems. This mitigates the risk of Personal Data becoming inaccurate or out of date and supports data subjects' right to rectification.

2.3 The principle of least privilege access is embedded at all levels in the Processor, from staff down to operating system server processes. This ensures that only the data that is authorised to be processed may be accessed. This mitigates the risk of accidental or deliberate alteration or destruction of Personal Data.

2.4 Amberscript only authorises specific staff to access the Service's production systems. This mitigates the risk of accidental or deliberate alteration or destruction of Personal Data.

2.5 Amberscript maintains separate development and production systems utilising different security tokens, passwords, and privileges. This mitigates the risk of accidental or deliberate alteration or destruction of Personal Data.

3. Availability

TOMs to ensure the availability of the Personal Data processed by Amberscript.

3.1 The availability of the Service is monitored continually. This supports the data protection principle of transparency.

3.2 Amberscript only authorises specific staff to access the Service's production systems. This mitigates the risk of accidental or deliberate interference with the Service which could affect availability.

3.3 If Personal Data is no longer required for the purposes for which it was processed, it is deleted promptly or when the customer commands to do so. It should be noted that with each deletion, the Personal Data is only locked in the first instance and is then deleted for good with a certain delay. This is done in order to prevent accidental deletions or possible intentional damage.

4. Resilience

TOMs to ensure the resilience of the Personal Data processed by Amberscript.

4.1 Data is backed up offsite frequently as per the data retention policy. This mitigates the risk of data loss, destruction or damage.

4.2 Data centres used by Amberscript utilise multiple redundant network connections to major internet exchanges. This provides resilience in the face of adverse network conditions.

4.3 Data centres used by Amberscript utilise redundant UPS power supplies supported by diesel generators for standby power. This mitigates the risk of power outages and provides resilience in the face of electrical supply problems.

4.4 Data centres used by Amberscript utilise modern fire systems for prevention, detection and response with direct connections to the local fire service.

4.5 Data centres used by Amberscript provide automatic protection against distributed denial of service (DDoS) attacks. This provides resilience in the face of network attacks whether directed against the Service or others on the network.

4.6 DNS services used by Amberscript are built on distributed, redundant architectures. This provides resilience in the face of adverse network conditions.

5. Timely restoration of access to Personal Data

TOMs to ensure the timely restoration of access to the Personal Data processed by Amberscript.

5.1 Amberscript uses modern devops practices, including but not limited to infrastructure-as-code, to enable new servers to be commissioned as necessary in a timely fashion. This mitigates the risk of loss of access to Personal Data in the face of problems with existing servers, including but not limited to web servers, application servers, and database servers.

5.2 Amberscript's backup and restoration processes enable the timely restoration of Personal Data from backups as per the data retention policy. This mitigates the risk of loss of access to Personal Data caused by problems with the database.

6. General Technical measures

Amberscript implements general technical measures, including but not limited to the following, to support the confidentiality, integrity, availability, and resilience of Personal Data.

6.1 Physical security

- Office premises protected by locks and alarms.
- Old computer equipment securely formatted before disposal.

6.2 Device security

- All computers and devices use full disk encryption.
- All backup media use full disk encryption.
- All computers and devices regularly updated and security-patched.
- All passwords generated by and stored in an industry-leading password manager.

6.3 Network security

- All networks protected by firewalls.
- Access to Personal Data over public wifi is prohibited unless a VPN is used.

6.4 System security

- All passwords and authentication keys rotated regularly.
- All servers are regularly updated and security-patched.

6.5 Website security

- All web traffic protected by HTTPS / TLS and appropriate security headers.
- Web server ciphers regularly reviewed and known weak ciphers disallowed.
- An industrial application security monitoring system is implemented to identify and protect outside threats (such as DDoS attacks etc)

6.6 Data centre security

- Manned 24hr/day all year.

- Entry controlled via electronic access control terminals.
- Continual high definition video surveillance.
- All personnel movements recorded and documented.

6.7 Data security

- Backups stored off-site
- Data deleted when no longer needed

6.8 Software development

- All developers are familiar with the OWASP Top Ten web application security risks.
- All software must pass automated tests before deployment.
- Data privacy is always a fundamental requirement for the Service's software.

7. General Organisational measures

7.1 Staff with access to Personal Data only process that data when instructed to do so and only within the scope of the instructions.

7.2 Staff are trained on:

- responsibilities under GDPR;
- staff responsibilities for Protecting Personal data, including the collection, processing and use of Personal Data only within the framework and for the purposes of their duties
- security policies.

7.3 Personal Data is only accessed as needed and only when approved by the customer, or by technical staff for necessary support and maintenance of the Service (principle of least privilege applies).

7.4 All staff members sign data-processing agreements as parts of their employment agreement.

7.5 Only designated staff can access software platforms and internal networks

7.6 Personal Data used for internal purposes only e.g. as part of the respective customer relationship, only in case a third party in form of a subcontractor is necessary for the service, personal data will be passed on. But solely under consideration of contractual arrangements and appropriate data protection regulatory requirements.

7.7 The transfer of Personal Data to a third party (sub-contractors) is only made if a corresponding contract exists, and only for the specific purposes. More information on subcontractors can be found in Article 10 of the Data-Processing Agreement.

3.1.3 EC - The solution offered should support a number of different input formats for audio and video, both compressed and uncompressed.

Question: Please list all video and audio input file formats currently supported by the Service. We are expecting that all major flavours like;

- Video: MKV (Matroska), MOV, MP4

- Audio: WAV, MP3, ... (AIFF, FLAC) are supported, but as many as possible to ensure maximum flexibility.

Answer:

Amberscript currently supports the following file formats:

- mp3, mp4, wav, m4a, m4v, mov, wma, aac, opus, flac, mpg

Additional formats at contract initiation:

Amberscript makes use of the ffmpeg open source library to accept and convert files. On request, Amberscript will make the following formats available as soon as the contract with NORUnet starts free of charge:

- audio formats: **aif, dss, snd, gsm, amr, ac3, aiff, flac, ogg, wav, aifc, 8svx, au, caf, mp2, wma, mp2, oga.**
- video formats: **avi, mpeg, mxf, vob, mkv, m2ts, mts, ogv, ts, 3gp, flv, rm, webm, wmv, mk3d,**

3.1.4 EC - The solution offered should support a number of different output formats.

Question: Please list all output formats currently supported for transcribed material by the service. We are expecting that all major flavours like WebVTT, SRT, to be supported, but as many as possible to ensure maximum flexibility.

Answer:

Amberscript currently supports all major output formats for captions and transcriptions such as:

Captions:

Currently available: WebVTT, SRT and EBU-STL

The following output formats will be added free of charge: DXFP / TTML / .sub / .ass

Transcriptions:

- .doc / .txt

Developer Formats & Integrations into QDA software:

Thousands of qualitative researchers across Europe use Amberscript's transcripts for qualitative analysis purposes. We have therefore developed research and software development-oriented output formats. Therefore, the following output formats are supported:

- .xml
- .json

As an example, ATLAS.ti and MAXQDA qualitative research software offer a feature to integrate Amberscript transcripts directly (e.g. in MAXQDA go to the import tab and choose Transcripts > From Amberscript (.json)).

These output formats can also be used to empower software built by our clients with our highly accurate speech recognition technology and services.

3.1.6 EC - Integration with SAML based identity federations (Web-portal)

Question: Each Nordic Country has their own SAML2 based Identity federation which is used by almost all Customer Institutions. As part of the onboarding and signing of a call-off with the respective Customer, the Bidder shall have a working integration.

The Identity federations are:

Denmark: WAYF <https://wayf.dk/en>

Finland: HAKA <https://www.csc.fi/en/haka>

Norway: Feide <https://www.feide.no/>

Sweden: SWAMID <https://www.sunet.se/services/identifiering/swamid>

Please describe any already working SAML2 integrations and any other methods for external authorisation in place. Describe your plan for ensuring support for the National federations.

Answer:

Amberscript is currently working with several organizations that require external authorisation with Active Directory or SAML environments such as the European Commission in Brussels (Active Directory) or INSEEC Business School in France (SAML 2.0).

For NORDUnet, Amberscript proposes the following approach to enable external authentication for the National Federations:

Onboarding a new federation:

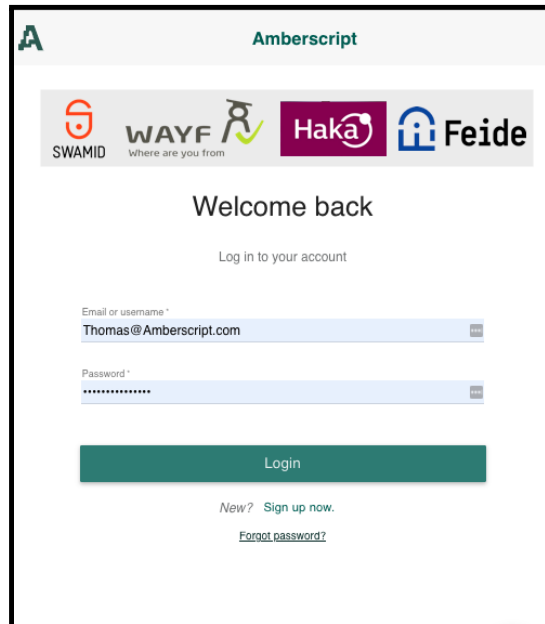
Amberscript uses an open-source technology called 'Keycloak' to enable external authentication for its enterprise customers. Building the integration with a new identity federation will take approximately 7-10 business days and will be completed by Amberscript free of charge as soon as a call-off is received by an end-customer from a specific identity federation.

As part of the onboarding process, Amberscript will include the following technical milestones to enable external authentication for the national federations of NORDUnet.

- 1) Define roles and permissions for NORDUnet users
- 2) Enable account creation via identity federation
- 3) Enable ability to log in and log out using SAML
- 4) Enable Session Management
- 5) Enable ability to delete an existing account
- 6) Connect to identity federations:
 - a) WAYF
 - b) HAKA
 - c) FEIDE
 - d) SWAMID

User flow and login:

Amberscript will create a separate login URL for NORDUnet (e.g. <https://nordunet.amberscript.com>) in which the login screen offers the ability for users to authenticate themselves using SAML as shown below:



In this login screen, the user can click on the relevant logo for SAML authentication (e.g. Haka). The authentication will take place in the background and if successful, the client will be redirected to his/her Amberscript dashboard where the user can start using Amberscript's services.

3.1.7 EC - Please describe your access control mechanism and any supported integration (Web-portal)

Question: List or describe all supported methods and integrations. We are looking for flexibility in the ways to manage the access control.

Answer:

Access rights & user roles:

For the web-portal, Amberscript designed a role-based access control mechanism with 3 different hierarchies and roles:

1) Organization Admin:

- a) Can create and delete new **Group Admins** within the organization with certain user-rights

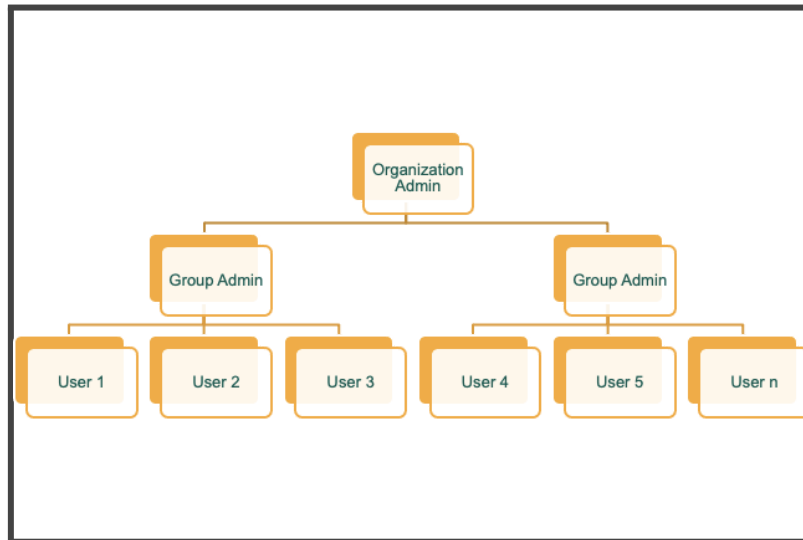
- b) Can create and delete new **Regular Users** within the group according to the assigned user-rights
- c) Can use Amberscript according to the assigned user-rights

2) Group Admin:

- a) Can create and delete new **Regular Users** within the group according to the assigned user-rights
- b) Can use Amberscript according to the assigned user-rights

3) Regular Users:

- a) Can use Amberscript according to the assigned user-rights



Principle of least privilege:

Within this environment, Amberscript applies the principle of least privilege in all functionalities. Regular users can track their own usage, but can't monitor the usage of the group or of the organization. Group Admins can track the usage of the group and if single users, but cannot see the usage of the entire organization.

User-rights can be defined as follows on all hierarchical levels (Organization, Group and Regular User);

- Automatic Transcription: Access (yes / no) / Unlimited use? (yes / no) / Limited use (hours per month / year)
- Automatic Captioning: Access (yes / no) / Unlimited use? (yes / no) / Limited use (hours per month / year)
- Human Transcription: Access (yes / no) / unlimited use? (yes / no)
- Human Captioning: Access (yes / no) / unlimited use? (yes / no)
- Human Translated Subtitles : Access (yes / no) / unlimited use? (yes / no)

3.1.9 EC - Describe the HTML user interface (Web-portal)

Question: The service shall have a responsive and accessible HTML-based user interface for upload and retrieval of audio and video files. Please describe all functions of the user interface. **We are looking for ease of use and simplicity in choosing options.**

Answer:

Amberscript's user interface and text-editor is used by more than 1 million users worldwide to easily order transcriptions and captions (both: automatic and human).

In order to submit files for transcription or captioning, users can follow 3 easy steps:

- 1) Login to the secured Amberscript environment and upload a file via the responsive web-interface.
- 2) Receive an email notification that a caption- or transcription-file was delivered
- 3) Download the final file

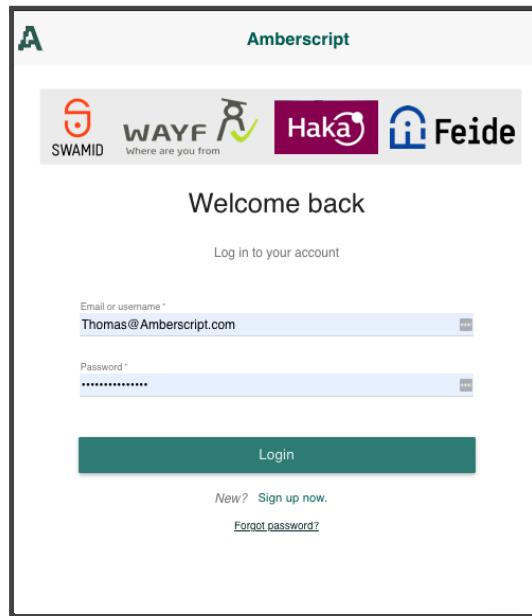
All steps in the process are conducted in the secured and responsive web-portal of Amberscript. Next to the fully responsive HTML interface, Amberscript launches a smartphone application (for iOS and android) in August 2021 which will be available to NORDUnet users free of charge.

Within the application, it will be possible to easily make audio recordings and to submit these recordings for automatic or human transcription into the secured web-portal of Amberscript.

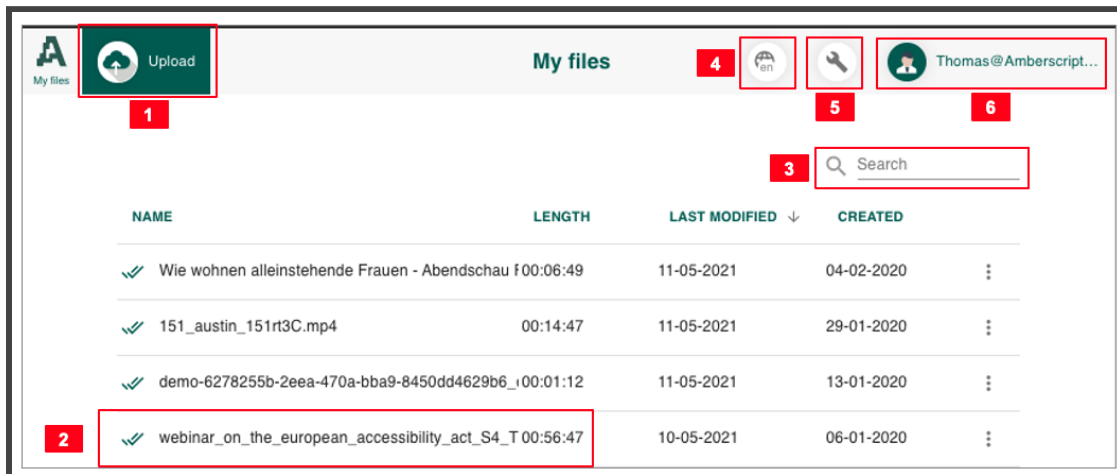
The steps are described in detail in the following paragraphs:

Login

NORDUnet users will be able to log in securely via a customized log in page for NORDUnet (e.g. <https://nordunet.amberscript.com>). In the log in screen, users are asked to authenticate themselves securely with a username, a strong password and with SAML 2.0 external authentication (see 3.1.6 for more details).



The user dashboard

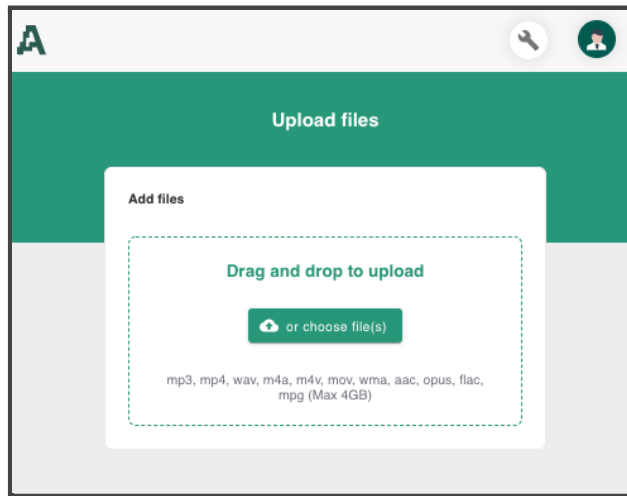


In the dashboard, the users has 6 buttons available:

- 1) The upload button - this button leads to the upload flow to add new media files
- 2) Transcribed files - this is an overview of previously uploaded transcription or captioning files
- 3) Search: This allows the user to search for the name of previously uploaded files
- 4) Language settings: The user interface is available in a variety of languages
- 5) Admin Settings: For Organization Admins and Group Admins only
- 6) Personal user settings: Users can manage their personal settings such as name, password, usage, invoices and logging out securely.

The upload flow

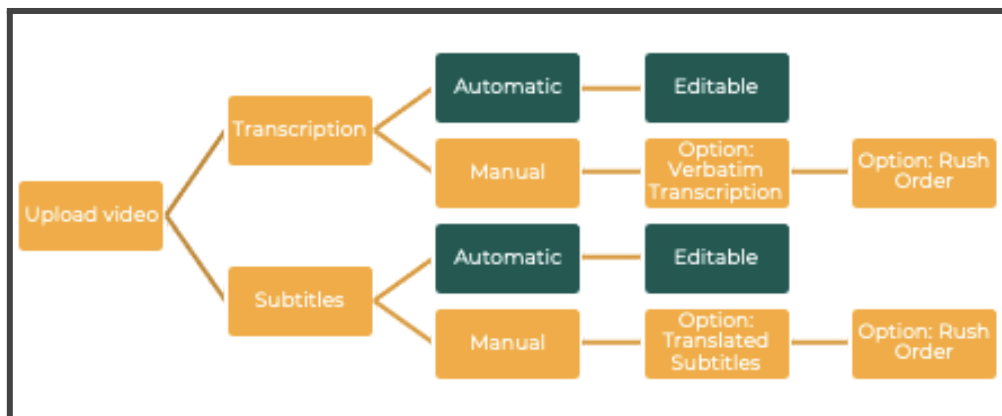
Upload via the web-portal:



In the upload window, files can be selected from multiple sources. Either the user can simply drag and drop the files into the uploading frame or the user can select files from the hard-drive from the device they are working on.

Within the next few months, Amberscript will add additional functionalities so that files can be uploaded directly from Youtube, Vimeo, Dropbox and other cloud environments. These developments will be available for NORDUnet users free of charge.

The ordering process follows a simple logic as described below in the following chart:



Within the intuitive order form, each user can make a simple choice of which services he wants to order. See detailed description below:

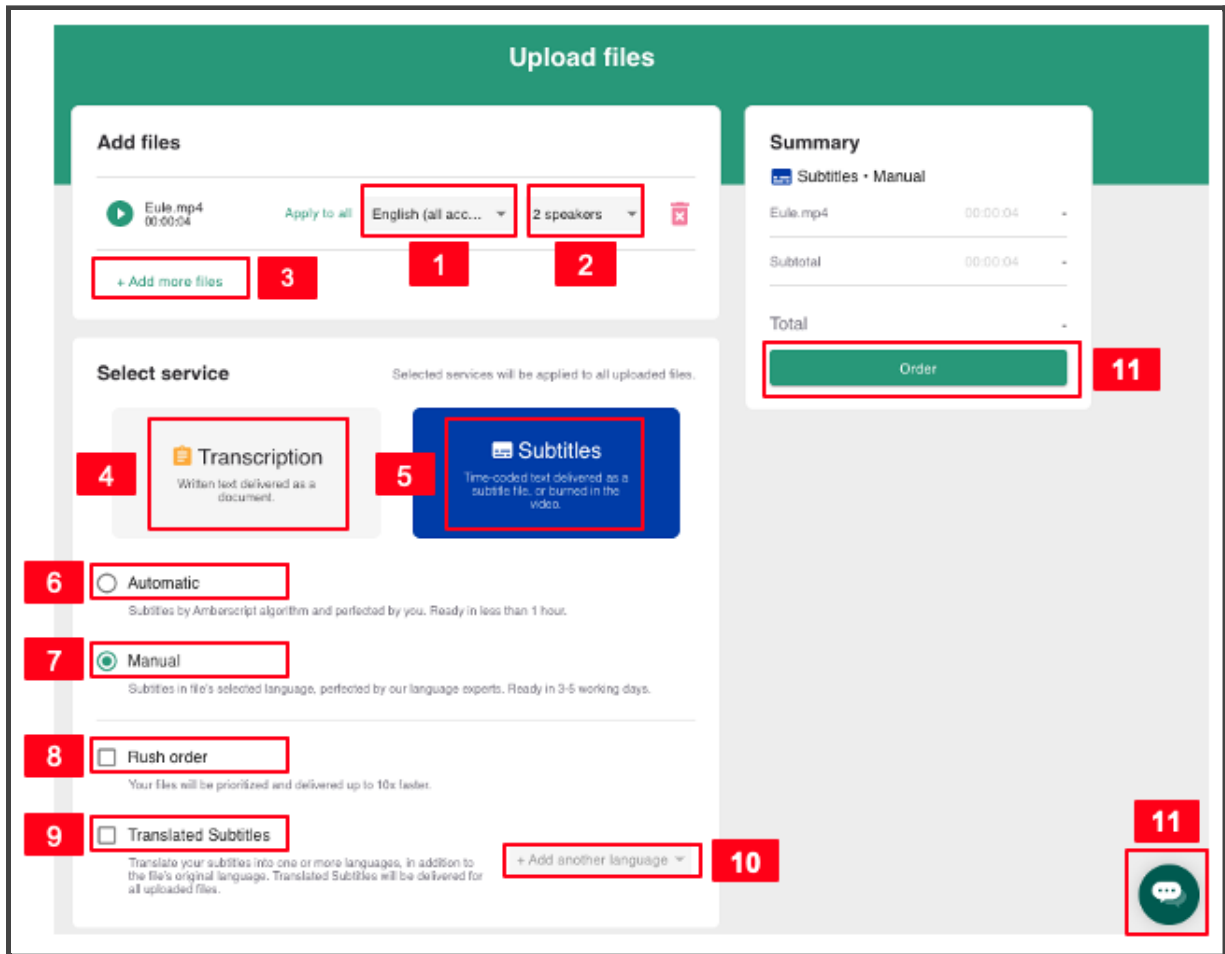
How to order transcriptions:

The screenshot shows the 'Upload files' interface. At the top, there's a green header with the text 'Upload files'. Below it, the 'Add files' section shows a file 'Eule.mp4' with a duration of '00:00:04'. To the right of the file name, there's a dropdown menu for language, currently set to 'English (all acc...)', and another dropdown for the number of speakers, set to '2 speakers'. A red box with the number '1' highlights the language dropdown, and a red box with the number '2' highlights the speakers dropdown. Below the file list, there's a '+ Add more files' button with a red box and the number '3'. The 'Select service' section has a sub-header 'Selected services will be applied to all uploaded files.' Below this, there are two service cards: 'Transcription' (with a red box and number '4') and 'Subtitles' (with a red box and number '5'). Underneath these are four radio button options: 'Automatic' (with a red box and number '6'), 'Manual' (with a red box and number '7'), 'Verbatim' (with a red box and number '8'), and 'Rush order' (with a red box and number '9'). On the right side, there's a 'Summary' panel showing 'Transcription · Manual' and a table with columns for file name and duration. The table lists 'Eule.mp4' with a duration of '00:00:04' and a 'Subtotal' of '00:00:04'. Below the table, there's a 'Total' row and a green 'Order' button with a red box and the number '10'. In the bottom right corner, there's a green chat icon with a red box and the number '11'.

1. Select the language of the media file that the user has uploaded. In the NORDUnet environment, the following languages will be available:
 - a. English (EU accents)
 - b. Swedish
 - c. Danish
 - d. Norwegian (Bokmal)
 - e. Finnish
 - f. Icelandic (beta)
2. Select the number of speakers in the media file. The indication of the number of speakers helps the 'speaker diarization module' to segment the text and to label the speakers accordingly.
3. Add an additional file: With this button you can simply add more files
4. Order Transcriptions: This switch allows you to order all available transcription services
5. Order Subtitles: This switch allows you to order all available subtitling services
6. Automatic Transcription: This service uses automatic speech recognition to create a draft-transcript of the media file within a few minutes. The draft subtitles can be corrected by the user in the transcription editor (see point 3.1.3)

7. Manual Transcription: This service allows the user to order professional transcriptions which are created by a combination of Automatic Speech Recognition and professional human transcribers (see more details in point 3.1.17)
8. Add-on for 'verbatim transcription': Amberscript's default transcription style is called 'clean-read'. In clean-read transcription, professional transcribers optimize the readability of the transcribed text while maintaining the integrity of the spoken language fully. Repetitions, stutters, hesitations and speech errors are omitted and slightly adjusted to increase readability. If the users orders 'verbatim transcription', repetitions, stutters, hesitations and other sound events are fully included in the transcription. Verbatim are usually used for scientific research where the way of speaking is equally important as the content of the speech
9. Add-on for 'Rush orders': Amberscript's standard turnaround time for human transcription is 3-5 business days. When selecting a rush order, the file will be prioritized and delivered up to 10 times faster (usually within 24 hours).
10. Order button: Once the user selected all options for their order, the order can be confirmed with the 'order' button. After clicking on this button, the order cannot be changed anymore and the process of creating the transcription will start right away.
11. Chatbot and live-chat: The chatbot is designed to give the user easy access to more than 100 frequently asked questions. The user also has the option to contact our customer support team straight from the chatbot.

How to order captions:



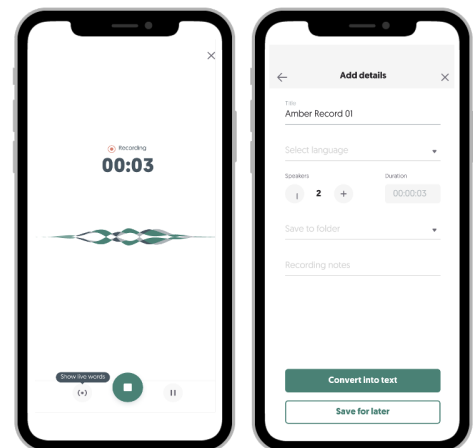
1. Select the language of the media file that the user has uploaded. In the NORDUnet environment, the following languages will be available:
 - a. English (EU accents)
 - b. Swedish
 - c. Danish
 - d. Norwegian (Bokmal)
 - e. Finnish
 - f. Icelandic (beta)
2. Select the number of speakers in the media file. The indication of the number of speakers helps the 'speaker diarization module' to segment the text and to label the speakers accordingly.
3. Add an additional file: With this button you can simply add more files
4. Order Transcriptions: This switch allows you to order all available transcription services
5. Order Subtitles: This switch allows you to order all available subtitling services
6. Automatic Subtitles: This service uses automatic speech recognition to create a draft-transcript of the media file within a few minutes. The draft subtitles can be corrected by the user in the captioning editor (see point 3.1.3)

7. Manual Subtitles: This service allows the user to order professional subtitles which are created by a combination of Automatic Speech Recognition and professional human captioners (see more details in point 3.1.17).
8. Add-on for 'Rush orders': Amberscript's standard turnaround time for human transcription is 3-5 business days. When selecting a rush order, the file will be prioritized and delivered up to 10 times faster (usually within 24 hours).
9. Add-on for 'translated subtitles': Via this button, the user can order translated subtitles in multiple languages. The translated subtitles are created and quality-checked by 2 certified translators before delivery to the client.
10. Add on to add multiple languages: Via this button, the user can add multiple languages in which the video shall be subtitled. This service is oftentimes used for promotional videos and public content such as 'MooCs'.
11. Order button: Once the user selected all options for their order, the order can be confirmed with the 'order' button. After clicking on this button, the order cannot be changed anymore and the process of creating the transcription will start right away.
12. Chatbot and live-chat: The chatbot is designed to give the user easy access to more than 100 frequently asked questions. The user also has the option to contact our customer support team straight from the chatbot.

Upload via the smartphone App

Amberscript's smartphone application for android and iOS will be available to NORDUnet users free of charge. The smartphone application allows users to easily and intuitively make audio recordings straight from their smartphone. Common use-cases include:

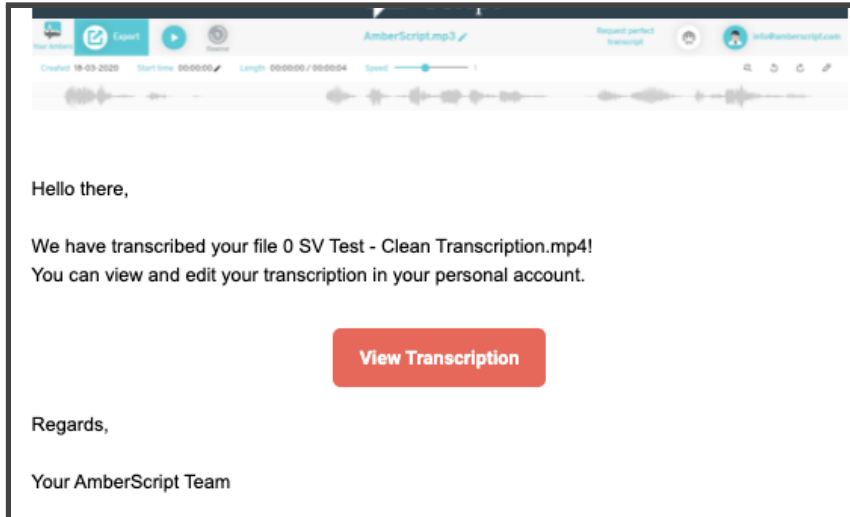
- Researchers using their smartphone to record interviews for qualitative research purposes
- Users record online meetings (e.g. MS Teams) by putting their smartphones next to the laptop speakers to record meetings and conversations
- Students record interviews for their research projects from their smartphones



After recording, users will have the opportunity to directly submit their recording to Amberscript's web-portal for automatic and human transcription. All data that is submitted via the smartphone application will be synchronized with the web-application so that users can easily correct their automatically generated transcriptions. Also human transcription services can be ordered through the smartphone application.

Email notification of a delivered file:

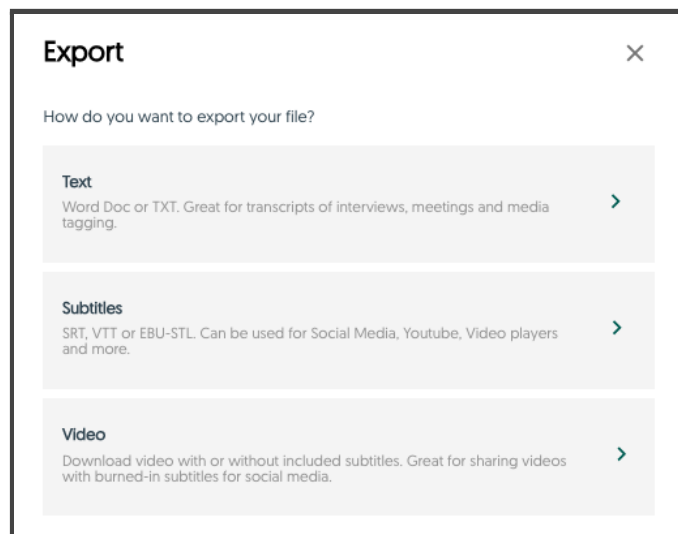
When the file is finished, the user receives an email notification such as this:



When clicking on the button 'View Transcription' the user is redirected directly to the finalized transcription or caption file. From there, the user can export the file into different formats.

Export of the file

When clicking on the 'Export' button in the transcription or captioning editor, the user has the option to export the text file into various formats. Depending on the format of the file, it is possible to export the transcript as text (in .docx or .txt format), as subtitles (in .srt, .vtt or .stl format) or even as video with burnt-in captions.



3.1.11 EC - Compliance with the EU Web Accessibility Directive (Web-portal)

Question: Please describe how the user interface ensure compliance with the EU Web Accessibility Directive and WCAG standards and the guidelines for universal design.

We are looking for compliance that is better than the minimum requirement i.e. WCAG AAA

Answer:

With the outspoken company vision to make all audio accessible, accessibility is at the very core of Amberscript's company identity. Amberscript works with a company specialized on digital accessibility (<https://technobility.nl/>) to design all user-interfaces of our products in the most accessible way possible.

Amberscript also collaborates with associations of people with auditory impairments throughout Europe. Europe's deaf and hard of hearing are well organized in national and international interest groups. Amberscript collaborates with the organizations to map its member's needs, introduce and improve new technologies and jointly advocate rights together. Amberscript closely collaborates with foundations and interest group such as:

- Foundation Hoormij (<https://www.stichtinghoormij.nl/>) in the Netherlands
- Høreforeningen (<https://hoeforeningen.dk>) in Denmark
- Hörselskadades Riksförbund (<https://hrf.se/>) in Sweden



The collaboration includes giving free access to members with a hearing loss of our partner associations to the Amberscript transcription platform. These members use

Amberscript to transcribe lectures, interviews or conversations they recorded when they don't have easy access to sign-interpreters.

In order to design our applications in an accessible way, Amberscript created an internal accessibility checklist for all of Amberscript's applications and web-interfaces:

Buttons

- Buttons are used for triggering actions, not for navigating.
- All buttons have clear labels explaining their purpose.
- Icon-only buttons have an aria-label attribute that provides additional information for screen reader users.
- Buttons have contextual labels that provide information for screen reader users about which item this action is being performed on. (For example, if you have a list of 10 items in a to-do list, and each one has a Delete button, you need to provide a contextual label in the form of an aria-label on each button so that the screen reader will see something like "Delete item: walk the dog".)
- div and span elements with click handlers are not used in place of button elements.

Links

- Links are used for navigating, not for triggering actions.
- Links have contextual labels that provide more information for screen reader users. (For example, if your text is simply "Read more", either change the visible text to "Read more about web accessibility" or else leave the text as "Read more" and add an aria-label that says "Read more about web accessibility".)

Images

- Text is not used in images.
- Each image has an alt attribute. (Background images or other images that you want the screen reader to ignore can have an empty alt attribute like alt="".)
- Alt text describes the intent of the image, not describes the image literally.

Headings

- The page has one and only one h1 element.
- The first heading on the page is an h1 element. (Usually. If you have heading elements inside the navigation, you may make an exception.)
- Heading levels are not skipped. (For example, an h2 element must not be followed by an h4 element without an h3 element between them).

Forms

- All inputs have associated labels.
- Placeholder text is used for example input and not as a label.
- Formatting expectations for user input are displayed.
- Required fields are easily identified.
- Color is not the only indicator for feedback. (You should use a combination of color, icons, and text.)
- Error messages are helpful and appear close to the input.
- Every element is reachable by the keyboard.
- Inputs and buttons have focus indicators.
- Tab order makes sense.
- fieldset and legend elements are used to group inputs.

Modals

- When closing a menu or modal, focus is returned to where it was previously (for instance, back to the button that the user clicked to open the modal).
- When a modal is open, hitting the Escape key closes the modal.
- When a modal opens, focus starts on the Close button or on the first element inside the modal.
- When a modal is open, keyboard-only users and screen reader users are restricted to only interacting with the content inside of the modal. (They should not be able to focus on, read, or interact with content outside of the modal in any way).

Tables

- Tables are not used for layout purposes. Tables are used to present data.
- Tables have clear headings on every column or row (use the scope attribute).

Toasts

- Users have enough time to read and respond to toasts or other messages on the screen. (For example, toasts generally shouldn't disappear automatically after a few seconds, especially if there are buttons in the toast that a user may need to interact with. It's a much better practice to allow the user to dismiss a toast when they are ready.)

Drawers/Panels

- Content that expands and collapses allows the screen reader to read the current state for whether the content is expanded or collapsed.

Skip Links

- The very first element on the page is a visually hidden button that appears when focused on and allows the user to skip past the navigation to the main page content.

Touch Targets

- Touch targets are at least the size of an average person's finger (at least 44x44px).
- There is adequate space between touch targets.

Focus Indicators

- It is easy to see what element is currently active or focused.
- Focus indicators are not color-only.

Typography

- The base font size is around 16–20px.
- em or rem is used instead of px to create a more responsive experience when changing the font size.
- Text that is in all caps is cased normally in the HTML and transformed to all caps using CSS.
- Content is not aligned “justified”.

Language

- The lang attribute is present on the html tag.
- Content that is in a different language than the rest of the page is wrapped in its own lang attribute.
- Abbreviations are inside abbr tags and also have a title attribute included which contains the full name (for instance, the full name of an organization if only the acronym is shown on the page).

Layout

- The user can increase the font size of the text on the page by up to 200% without loss of content or functionality.
- The page layout doesn't require the user to scroll both horizontally and vertically. (Preferably, only vertical scrolling is needed.)

Audio

- Audio has a transcript provided either inline or as a link to a text file.
- Audio is not autoplayed.
- Audio longer than three seconds can be paused or stopped by the user.

Video

- Video is captioned or has a transcript provided either inline or as a link to a text file.
- Captions can be toggled on and off.
- Non-visual cues are narrated/captioned (for example, when a character laughs).
- Video is not autoplayed.
- Video longer than three seconds can be paused or stopped by the user.

Screen Readers

- All visible text can be read by a screen reader
- All interactions can be completed using only a screen reader.
- The screen reader makes an announcement if the page content changes significantly (for instance, new search results have loaded or page content has been filtered by some criteria).
- All error messages are announced by the screen reader.
- When content is deleted or removed from the page, focus is moved backwards to something previously seen rather than forwards.
- Screen readers can focus and read disabled elements (for example, a disabled button).
- Content that is inherently inaccessible, like drag and drop functionality that requires a mouse, has an alternative implementation that screen reader users can use.

Keyboard-Only

- All interactions can be completed using only the keyboard.
- There are no keyboard traps (places that you can get into via the keyboard but can't get out of via the keyboard).
- There are visual focus indicators for every focusable element on the page.
- The tab order is consistent, and focus moves through the page in a way that makes sense.
- Tabbing with the keyboard does not focus on disabled elements (for example, a disabled button).
- Content that is inherently inaccessible, like drag and drop functionality that requires a mouse, has an alternative implementation that keyboard-only users can use.

3.1.13 EC - Describe the provided captions text editor used for inspecting and/or adjusting of the automatic transcription (Web-portal)

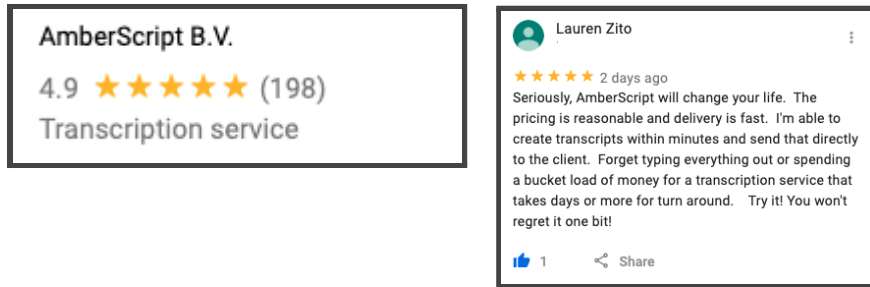
Question: Please describe the text editor including interfacing, i.e. is it a stand-alone application or integrated in the service HTML interface.

In the description of the captions editor, the Bidder shall provide a comprehensive description on how the process of correcting transcriptions are working effectively, including descriptions of links/synchronisation between written text and audio, timing adjustments of text blocks, autosave/save routines and character counter and i.e. wind-rewind, shortcuts. It's Possible also add screenshots to support the description.

The more functionality that is available in the editor the better.

Answer:

Amberscripts transcription editor is used by thousands of users daily across more than 100 countries. In total, more than 1 million users signed up to use the editor in the past 3 years. On review platforms such as Google, Amberscript's text editor is reviewed with a 4.9/5 by people from various fields and expertises.



The text editor was designed to efficiently transcribe audio into text. After a new video or audio file is submitted via the HTML interface (as described in 3.1.9), the file is automatically transcribed by automatic speech recognition so that the user can get to work immediately. The transcript for a file of 1 hour duration is typically readily transcribed within 15 minutes.

Below a comprehensive description of the editor follows:

Efficiency Potential for Transcription:

Without any professional tools, it takes the average person between 5 and 6 hours of work to transcribe an interview of 60 minutes.

The Amberscript transcription editor was designed in a way that combines professional features that transcription-professionals use with an intuitive user interface so that also people without a background in transcription can drastically improve their efficiency. With the help of automatic speech recognition, synchronization of audio and text and other professional tools such as spell-check, the average user spends 1-2 hours of editing time on an interview of 60 minutes. This represents an improved efficiency by 73-80% compared to transcription without tools.

Efficiency Potential for Captioning:

In order to create captions for a video, typically the following steps have to be taken:

- 1) Transcription of the speech into text
- 2) Synchronisation of the audio and the text
- 3) Segmentation into subtitles that appear at the right moment, but that take into account factors like reading speed, number of rows, number of characters per rows, line-breaks and others

Without any prior knowledge in captioning, without knowledge in best-practices for subtitling and without any professional tools, it is very difficult for teachers, students and learning instructors to create captions.

If they manage to create captions at all, typically the process takes 10-12 hours of manual work for a video of 60 minutes.

Amberscript's captioning editor makes use of advanced AI-based algorithms so that the user only has to correct the transcription (see step 1 above). Once the transcription is finalized, the AI will take care of the synchronization with the text (step 2 above) and the segmentation into common subtitling formats such as BBC or Netflix (see step 3 above).

Automatic realignment and AI-based subtitle segmentation:

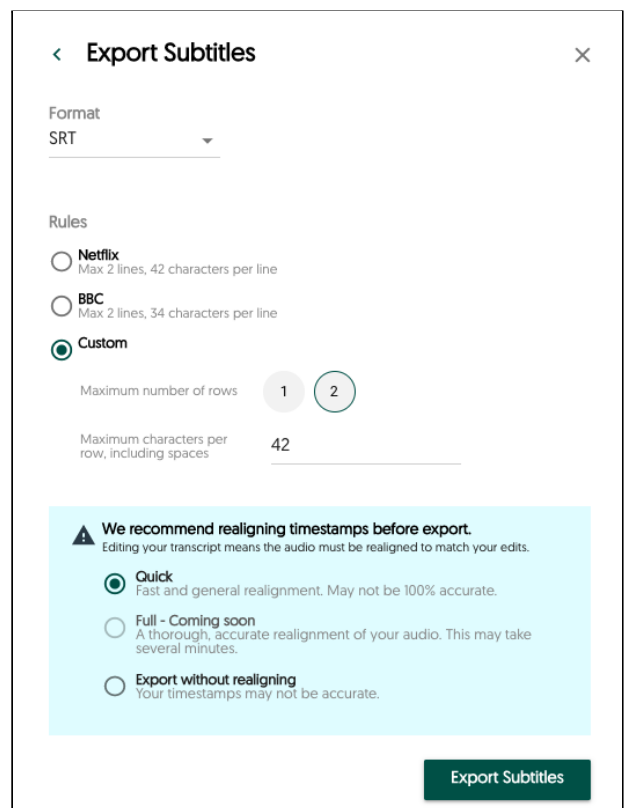
One of the biggest advantages of ASR is that every recognized word receives time-codes which can be used to synchronize the audio to the text. However, if the user deletes or adds words, these time-codes will be lost and the text will become asynchronous to the audio. Without the time-codes, the text cannot be transformed into synchronized subtitles.

That's why Amberscript developed two AI-based modules called 'automatic realignment' and 'AI-based subtitle segmentation'. Any captioning editor that does not offer 'automatic realignment' will not be able to deliver synchronized captions reliably.

When creating captions with Amberscript's captioning editor, the 'automatic realignment' function will be triggered automatically when the user wants to export the transcript as a 'subtitle file' (see screenshot).

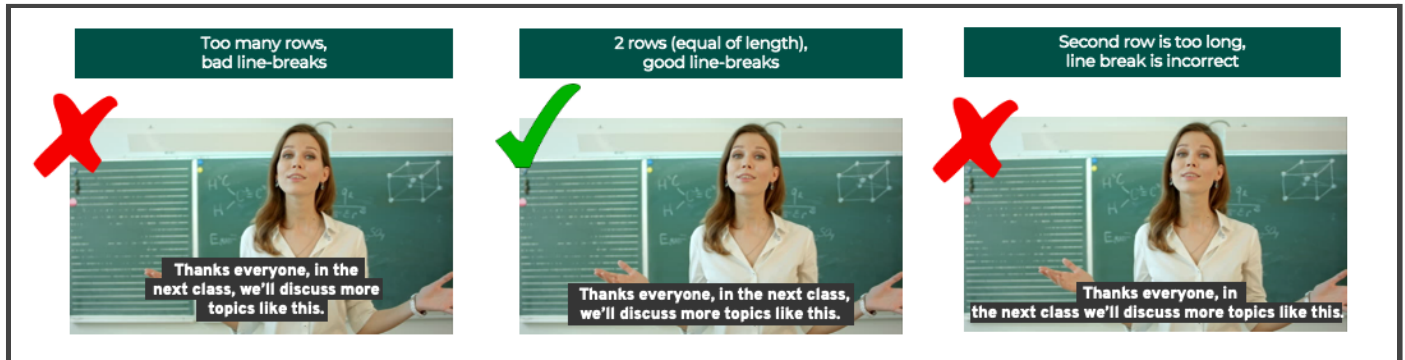
Next to that, the user has the option to choose between different subtitle segmentation rules. The standard rules include 'BBC-standards' (2 rows, 32 characters per row) and 'Netflix-standards' (2 rows, 42 characters per row). The user can also use a 'custom setting', where the user can determine the number of rows and number of characters per row.

The AI-based segmentation algorithm is trained on thousands of hours of manually formatted subtitles and captions in order to determine the right positions for line-breaks in the subtitles.



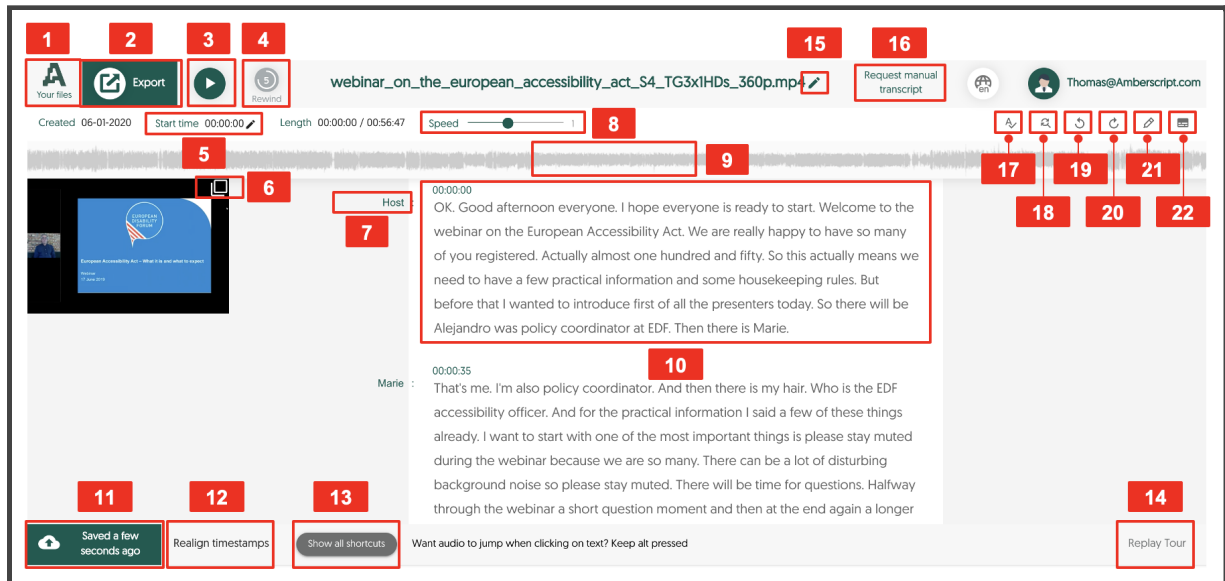
Amberscript

See below an example of the ideal segmentation of subtitles with correct line-breaks (in the middle) and 2 examples of typical mistakes that captioning editors without a specifically trained segmentation algorithm.



Combining these 3 technologies, Amberscript users typically spend 1,5 - 2 hours manual work to create captions for a video of 60 minutes, representing an efficiency increase of 80-90%.

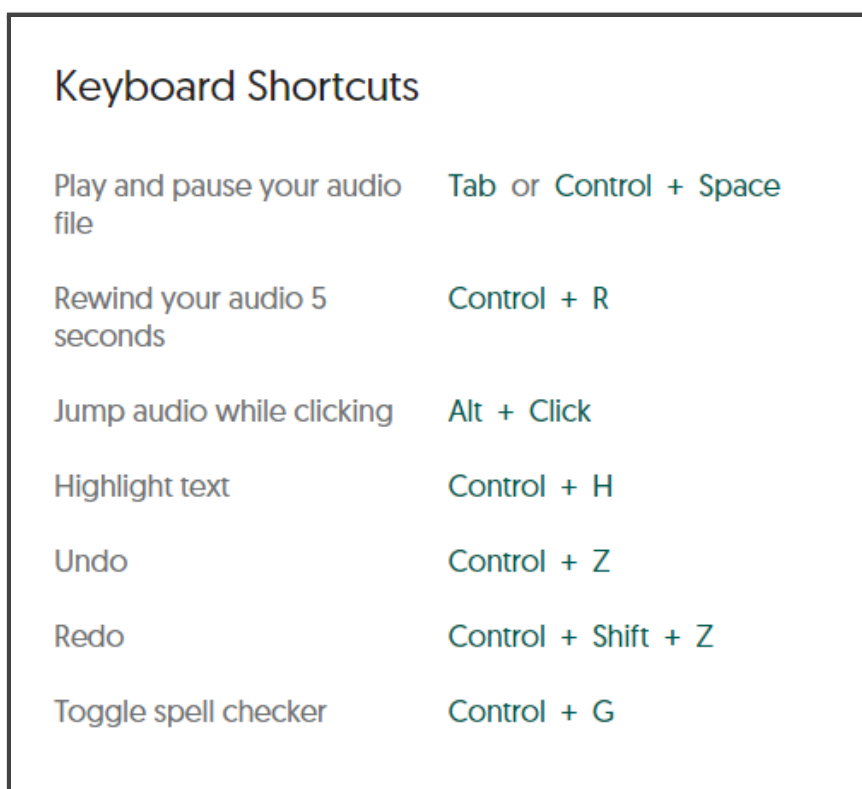
The Transcription Editor:



Description of text-editing functionalities in detail:

1. The Amberscript logo in the top-left corner brings the user back to the central dashboard
2. The 'Export' button allows the user to export the transcript or caption file in various formats. The full export workflow will be explained in detail below.
3. The play/pause button allows the user to play or pause the audio or video that is being played. This can also be done with the shortcut "Tab or Control + Spacebar".
4. The rewind button allows the user to rewind the audio or video by 5 seconds while making edits. This is especially convenient when the user is scanning through the document and notices a mistake to listen again to what is being said exactly before making adjustments. The shortcut for this button is "Control + R"
5. The start time button makes it possible to change the timestamps throughout the transcript. When the audio starts at the first minute, you can set this timer to 00:01:00 instead of 00:00:00. All time stamps in the document will then be adjusted for this change.
6. With this button the user can make the video slightly larger or smaller, depending on which size the video is currently set.
7. By clicking on the speaker the user can create new speaker names or use already saved names to rename the speaker which was clicked on.
8. With the speed slider, it is possible to adjust the speed of the audio. Normal speed is at 1. It can be slowed down to 0.1 and accelerated up to 1.9.
9. The audio waveform can be used to quickly skip parts where no audio is recorded or to skip to a specific section of the audio file. The transcript will then automatically follow to where the user has clicked. When the user highlights pieces of text, these highlights will also be shown in this waveform.

10. The user will be able to edit the text and add highlights with the shortcut “Control + F”. It is also possible to scroll to the text and use shortcut “Alt + Click” to sync the audio to the text part that has been clicked on.
11. The save button allows the user to save the adjustments that have been made in the transcript manually. Here the user can also see when the last automatic or manual save took place.
12. When the user clicks on the button “Realign timestamps” the timestamps that are shown, will be quickly and automatically realigned to the audio. This is especially relevant when a user deleted or added words to the automatic transcription, but wants to use the text output for subtitles or captions. Without realignment of the time codes, the subtitles will not be synchronous to the audio.
13. With this button you will see a cheat-sheet, where all possible shortcuts are shown. The cheat-sheet looks like this:



| Keyboard Shortcuts | |
|--------------------------------|------------------------|
| Play and pause your audio file | Tab or Control + Space |
| Rewind your audio 5 seconds | Control + R |
| Jump audio while clicking | Alt + Click |
| Highlight text | Control + H |
| Undo | Control + Z |
| Redo | Control + Shift + Z |
| Toggle spell checker | Control + G |

14. When a user enters the editor for the first time, a tour will be shown with all the features of the transcription editor. In case the user would like to follow this short tour again, it can be done by clicking on this button.
15. With the pencil button that is shown next to the file name, the user will be able to change the name of the file. This file name is automatically set to the same filename of the file that has been uploaded.
16. When a user is working on correcting the automatic transcription and feels like it is too much work. It is possible to request a manual transcription made by one of Amberscript’s language experts. In this case, Amberscript assigns a skilled transcriber to bring the text to 100% accuracy and the user will be notified as soon as the transcript has passed Amberscript’s internal quality check.

17. This button allows the user to toggle the spell checker, which can also be done with the shortcut “Control + G”. The Spell-check allows users to receive feedback on misspelled words directly within the transcription editor.
18. The replace button can be used to replace a word throughout the full text with another word. This is useful when the user notices that, for example, a name is often (incorrectly) transcribed in a certain way. It is then possible to change this for the full document. It could also be useful for words that have different correct ways of spelling, depending on region or dialect. An example of this is colour, which can be written as colour (UK English) and color (US English).
19. The undo button, which can also be triggered with the shortcut “Control + Z”, is used to undo the latest adjustments to the transcript.
20. The redo button, which can be triggered with the shortcut “Control + Shift + Z”, has the opposite effect of the undo button.
21. This is the highlight button, which allows the user to highlight pieces of text. This can also be done by selecting the words or sentences and using the shortcut “Control + H”. Once a highlight have been made, it is also shown in the waveform above the transcript. When exporting the file, it is possible to export the text with highlights or to even export ‘only highlights’ to create a summary of only the most important segments of the text.
22. This button allows the user to switch between the transcription editor (which is optimized for transcription) and the captioning editor (which is optimized for the creation of captions and subtitles).

The Export Workflow

The export workflow has a few steps. First the user needs to select the button that is mentioned above as number 2. The user will then see the following pop-up.

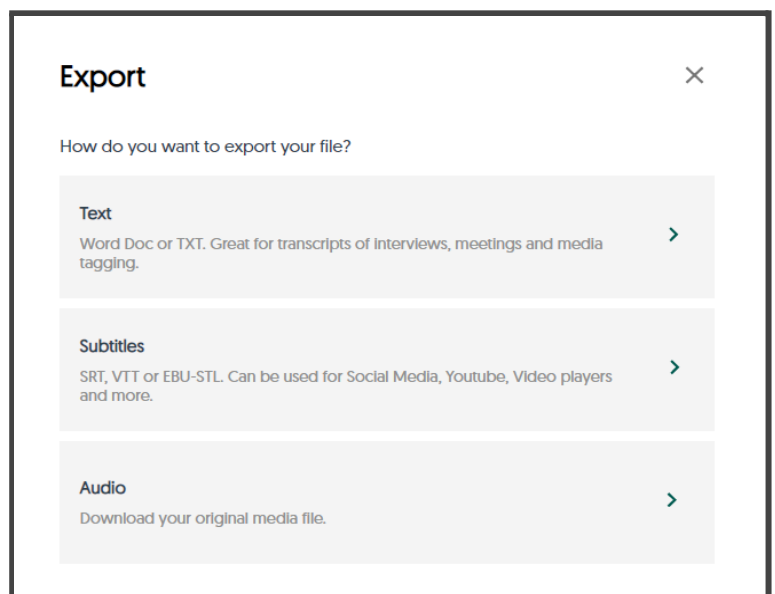
As can be seen in the screenshot above, it is possible to export the transcript as text or captions, or to download the original audio/video file. Below there is an oversight of the supported formats.

Transcriptions

- .doc
- .txt
- .xml
- .json

Captions

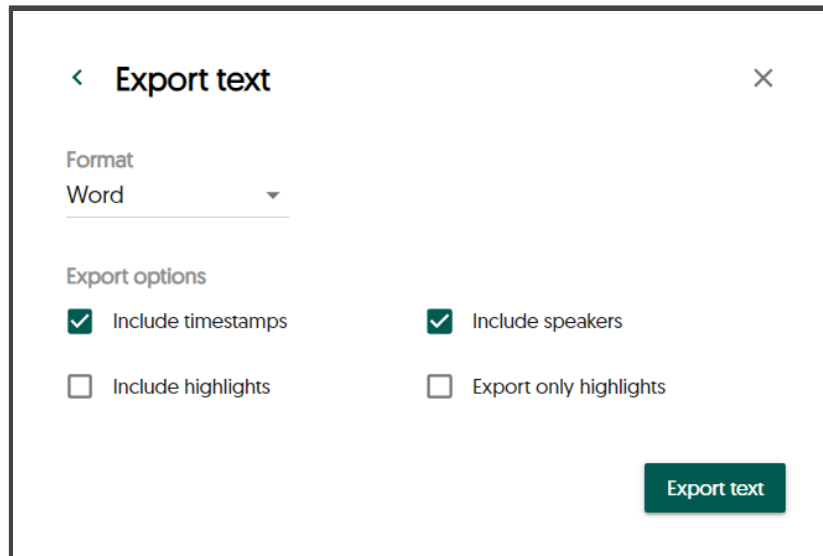
- WebVTT
- SRT
- EBU-STL



Caption formats that are not part of our standard delivery, but can be requested free of charge.

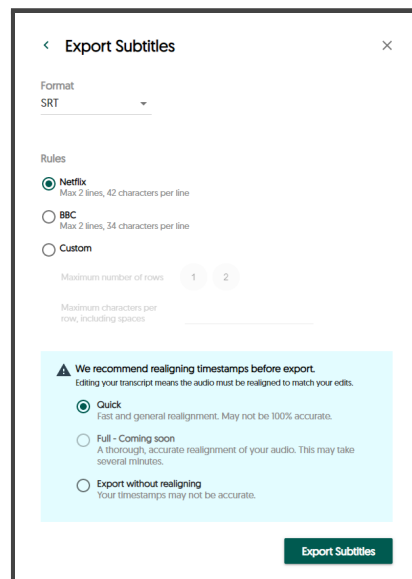
- DXFP
- TTML
- .sub
- .ass

When a user selects to export text they will see the following menu on their screen.



The user will have the possibility to select a desired format and to choose for the options that are shown in the screenshot.

When the user selects to export subtitles they will see the pop-up below.



Here, the user is able to select the desired captioning rules or create a custom rule. The user will have the possibility to realign the transcript to make sure that file that will be exported has a more accurate timecoding.

Adding words to the custom dictionary

It is possible to add custom words to the customer's dictionary in order to improve the accuracy of the automatic speech recognition. These words can be submitted to Amberscript through a webform (see point 3.2.5) and will be added to the Lexicon subsequently.

The feature of adding words allows the customer to achieve a higher speech recognition accuracy for topics, keywords and special keywords that are in the current media (e.g. "COVID-19") without making large software adjustments. New frequently used words can simply be added to the lexicon of the Automatic Speech Recognition and the words will be recognized in the next files that will be submitted.

Spell-check and auto-correction functionalities

In the transcription tool, state of the art spell-check in English, Swedish, Norwegian Bokmal, Danish and Finnish will be available to avoid typographical errors.

Typographical errors will automatically be red-lined in the editor. The user will then see improvement proposals of the correct spelling that the referent can choose from.

3.1.15 EC - Describe the API

Question: Please describe in detail all aspects of the API, or provide a copy of you API documentation. We are looking for as many function as possible to be supported by the API, and for an easy to use and navigate documentation

Answer:

All of Amberscript's services (ASR, transcription, captioning, human services) can be ordered through the API so that customer institutions can integrate the full range of accessibility services into their existing workflows. Integrations with media services such as Mediasite are build upon Amberscript's standard API and therefore, the API offers extensive professional functionalities.

The API documentation is available on github with the following link: <https://amberscript.github.io/api-docs/#introduction>

The API is developed influenced by REST standards and is very intuitive to integrate into existing workflows.

In essence, the API allows the user to to the following API calls:

1) POST: Upload media files

File requirements:

- Media with the following formats can be uploaded: wav, mp3, m4a, aac, wma, mov, m4v, mp4, opus, ogg, flac
- Media files can have a maximum file size of 8GB

Query parameters:

- **Language:** Allows the user to indicate the language of the speech the file contains
- **transcriptionType:** Allows the user to select 'Transcription' or 'Captions'
- **jobType:** Jobtype 'direct' refers to automatic speech recognition. Jobtype 'perfect' refers to human transcription and captioning services.
- **numberOfSpeakers:** Allows the user to indicate how many speakers are involved. This indication improves the segmentation of the text into various speaker-segments.
- **callbackURL (optional):** Allows the user to indicate callback URL to receive status updates
- **transcriptionStyle (optional):** Allows the user to differentiate between clean-read and 'verbatim' transcription style (not relevant for captions)
- **turnaroundTime (optional):** Allows the user to order different delivery times via the API.

2) GET: Get the status of the job

This simple endpoint allows users to request the current status of their job. By submitting a job-ID, the users receives one of the following responses:

- 1) Open: The job is still open and not finished yet
- 2) Done: The job is done and the output file is ready for download
- 3) Error: There was an error (e.g. the file does not contain speech or the file is corrupted) and the job cannot be finished.

3) GET: Export files into a variety of formats (STL, SRT, VTT, Text, JSON)

An output file can be downloaded multiple times and in multiple formats such as: STL, SRT, VTT, Text or JSON.

In order to influence the segmentation of the captions that are downloaded (for subtitle files), the user can indicate 3 parameters:

- 1) maxCharsPerRow: Maximum number of characters per row, Default is 42 according to 'Netflix subtitling standards'
- 2) maxNumberOfRows: Maximum number of rows per subtitles (default is '2' according to 'Netflix subtitling standards')
- 3) maxScreentimePerRowSeconds: Sets the maximum screen time of each subtitle multiplied by the number of rows (default is 2 according to 'Netflix subtitling standards').

The parameters and file formats provide the maximum flexibility of generating output from Amberscript's services while recommending an industry standard with the default settings ('Netflix subtitling standards').

4) DELETE: Delete a job with all related files

In order to give the user full control of the data, it is possible to delete the jobs and all related files (media file and all corresponding text or caption files) by using the DELETE endpoint.

The user indicates a job-ID and all files related to this job will be deleted from all servers (and back-up servers) immediately.

5) GET: List of previous jobs

The GET JOBS endpoint allows the user to generate a list of jobs with different filter criteria such as:

- Jobtype
- Job status
- Transcription type

This overview allows users to get a simple overview for several use cases such as:

- Identifying all 'done' jobs and delete all related files
- Determine the usage of the past period
- Identify all 'errored' jobs and submit files again

3.1.16 EC - Integration with Media Services API components

Question: Please describe in detail how the integration with the media services is done, e.g. integration with the Kaltura REACH module. It should also describe the support for subtitling/captioning and transcription. Information about existing implementations is preferred.

Answer:

Integration with Kaltura REACH (in operation)

In the past years, Amberscript and Kaltura developed a close partnership. Amberscript is deeply integrated into the Kaltura REACH framework. Kaltura REACH is the Kaltura service for providing captions and enrichment services. Kaltura REACH supports captioning, transcription, translation, audio description, chaptering, in-video and cross-library search and discovery, deep-linking capabilities, and metadata and keyword extraction.

The configuration of Amberscript subtitling is configured directly within the Kaltura configuration.

A detailed guide on the configuration can be found here:

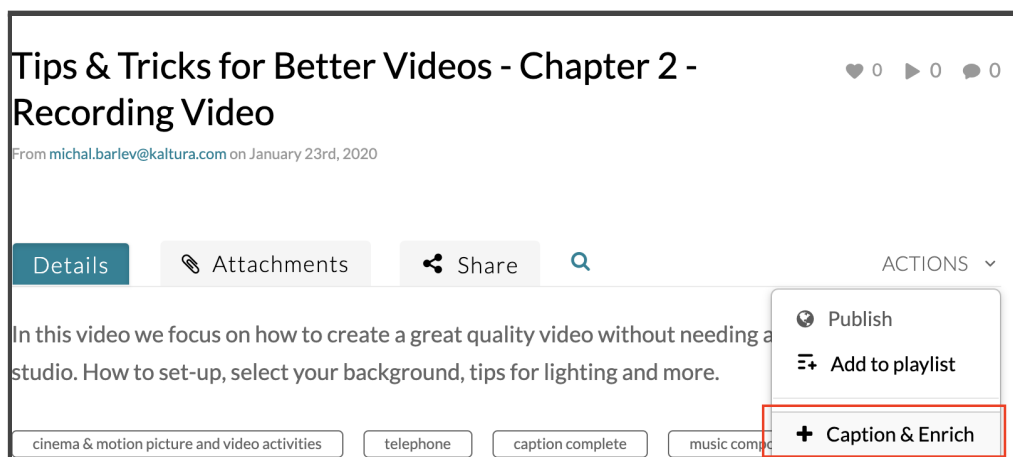
<https://knowledge.kaltura.com/help/kaltura-reach-v2-kmskaf-administration-guide>

User rights and workflow:

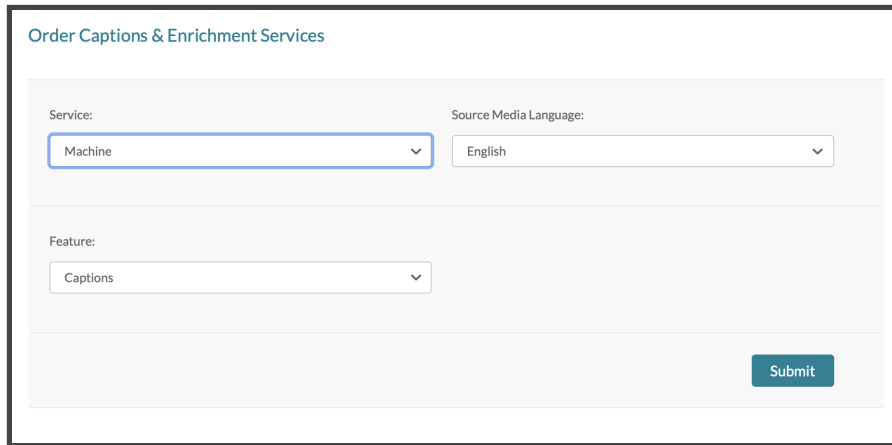
The user rights can be configured flexibly so that every teacher / every user / every admin user can order automatic captions, human captions or even translations with Amberscript directly from the Kaltura user interface.

In the Kaltura environment, captions can be easily ordered with the following steps:

1. Select My Media from the User drop down menu and click on a media entry.
2. Select Caption & Enrich from the Actions drop down menu. The Captions and Enrichment Services Requests screen is displayed.



3. In the following pop-up, simply select the type of captions (automatic or human) and the input language:



Order Captions & Enrichment Services

Service: Machine

Source Media Language: English

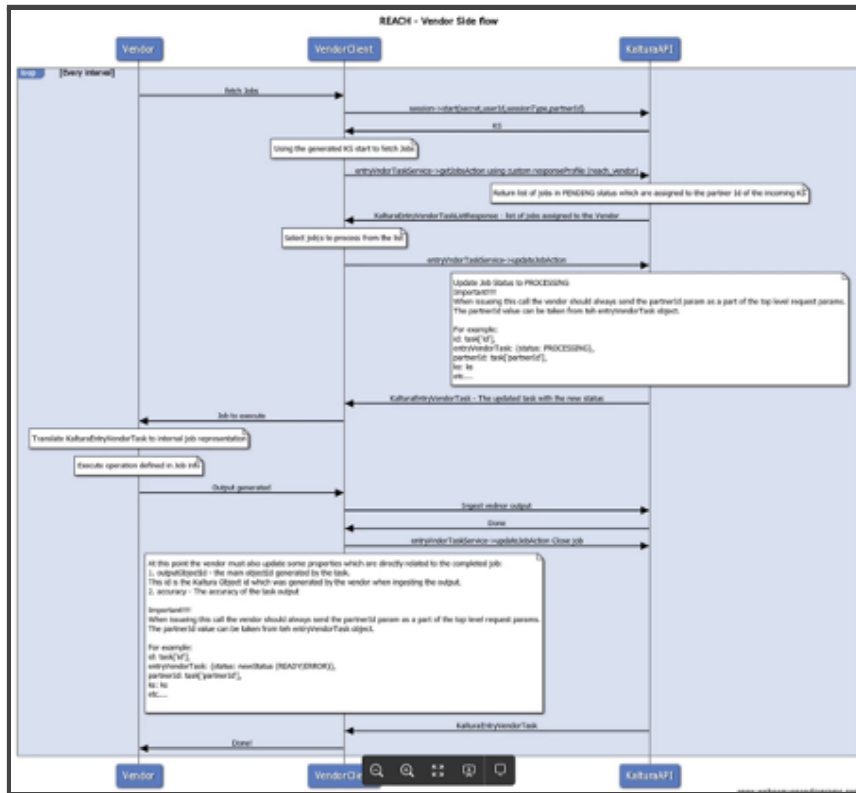
Feature: Captions

Submit

4. If the captions' requests require approval – the request status will be “Pending approval” until the requests are approved. After captions' requests are approved, the requests are sent automatically.
5. The captions will be uploaded to the media file when completed. You will receive a message that your request has been received.

The integration is based on workflows to fetch jobs/tasks from Kaltura. The job/task info will contain all relevant information for Amberscript to perform the task at hand such as the jobtype (transcription, captions), the language, the transcription type (automatic or human), the language combinations (for translations) and others. The job will contain account settings as well which are highly configurable within the Kaltura environment.

The detailed flow API flow is described below.



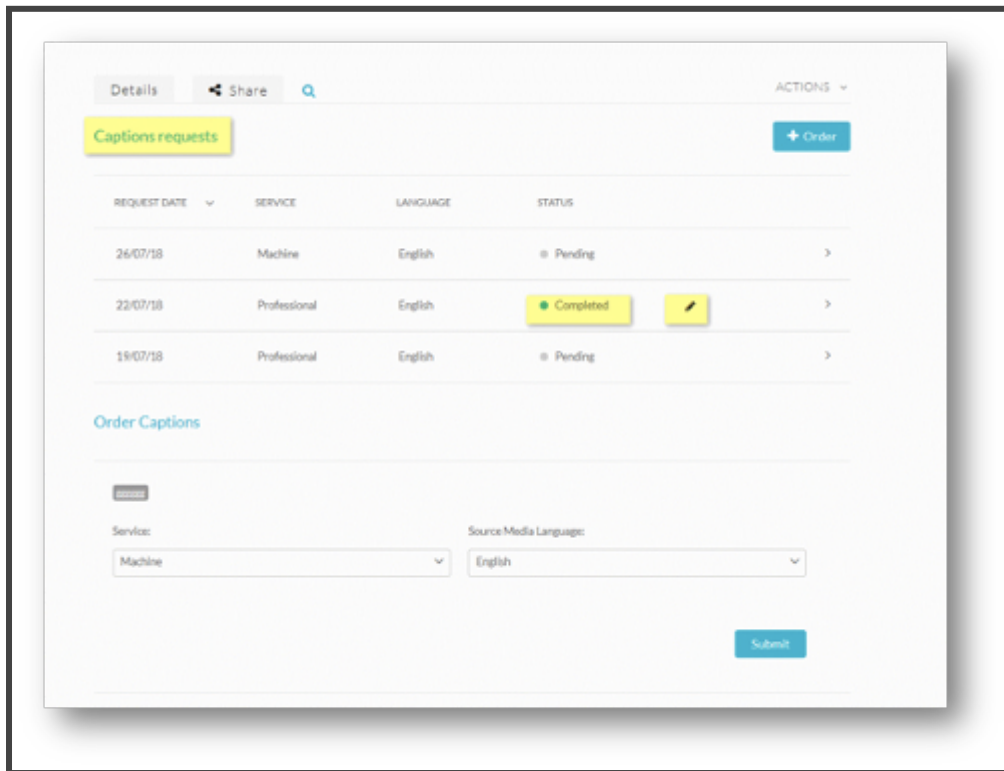
To execute API calls against Kaltura a valid KS (KalturaSession) is generated with permissions and user roles. The KS is generated using the dedicated account info.

Editing captions within Kaltura:

Editing of Captions/Transcripts can be done with the native Kaltura Caption Editor. There are two ways to access the Captions Editor from the media page:

- Actions menu - Select Caption & Enrich.
- Actions menu - Select Edit.

Below image shows the caption editor. It is possible to change the caption, edit timestamps, search, replace and add speaker ID.



See further details about the Caption Editor here

<https://knowledge.kaltura.com/help/editing-captions-reach-v2>

How to set up the REACH integration per customer institution:

NORDUnet can independently configure the Kaltura platform in a way so that Amberscript's services will be available to customer institutions and NORDUnet can create new REACH-instances for each end-user.

A scoping document will be shared in advance and should be jointly filled in by the Institution and the NORDUnet.

The scoping document will list questions about the initial setup e.g. which languages, which services, turnaround time and initial setup requirements. Details about the options can be found here

<https://knowledge.kaltura.com/help/kaltura-reach-v2>

Current status of the Kaltura REACH integration:

The integration is commercially deployed and is in use by several NORDUnet member organizations and by universities such as the University of Amsterdam, Saxion College or the university of Leiden in the Netherlands.

Integration with Mediasite (in operation)

User rights and workflow:

Amberscript has been working with Mediasite in the past few months to develop an integration with Mediasite v8.0.1 and Amberscript's services for automatic and human captioning. In the Mediasite environment, Amberscript will receive the status of 'partner captioning provider' from July 2021 onwards.

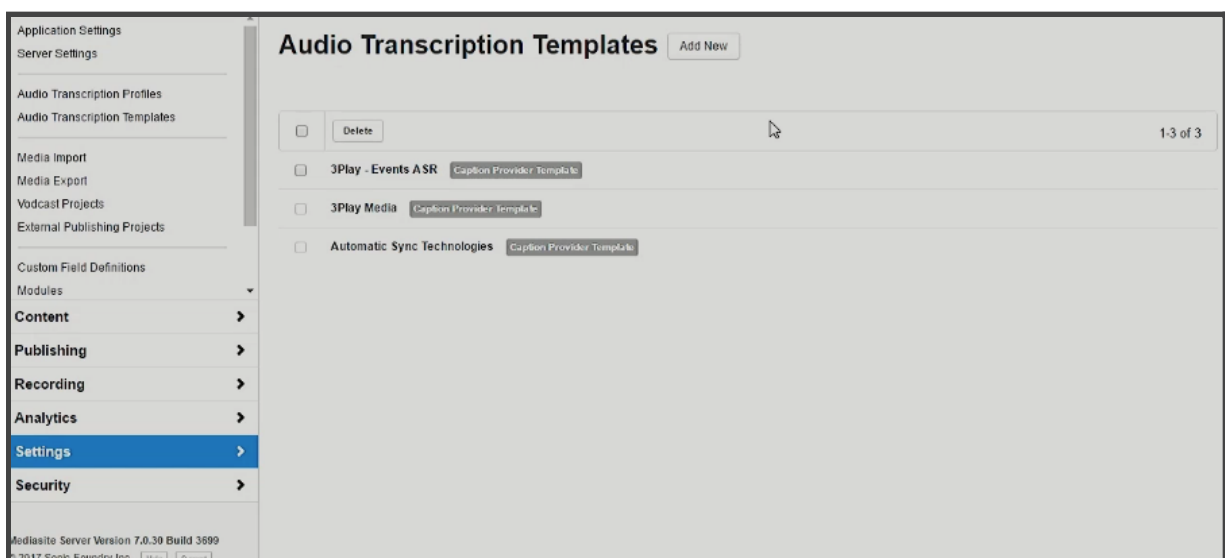
Mediasite admins can assign who in the organization (teachers, users, admins) can submit Mediasite objects (folders, presentations, templates or schedules) for captioning. Also individual files can be submitted for automatic and human captioning directly from the Mediasite user interface.

Mediasite then uploads media files along with a manifest describing the captioning job (automatic captioning or human captioning, delivery times) to Amberscript. Once the job is completed, Amberscript posts an SRT file back to Mediasite

How to set up the Mediasite integration per customer institution:

When configuring the Mediasite instance, NORDUnet admins and Mediasite admins can configure caption providers within Mediasite as explained below:

- Set up a captioning template for Amberscript



- Choose a name, the server URL that your provider gives you, and the Tracking URL Method. Amberscript will also share an Authentication Method that admins can select from the dropdown menu.

- Choose the formats for captions under Supported Caption Formats.
- Choose the kinds of media to be sent to Amberscript under Supported Media Formats.

Add Caption Provider Template Save Cancel

Name

Server URL

Tracking URL Format

Authentication Method
Password or Encryption Key

Supported Caption Formats

- DFXP
- SRT
- SMI

Supported Media Formats

- WMV
- MP3
- MP4
- WMA

- Set up a profile. The profile is your means of connection with the provider.
- In the main menu, go to Audio Transcription Profiles. Enter username and password (to be received from Amberscript) in the spaces provided.
- Decide if you want to use the Request Rush Processing feature, which will let Amberscript know you want rushed processing.
- Click Save to have Mediasite automatically upload the media file. This gets sent to Amberscript for captioning.

Current status of the Mediasite integration:

The integration is set up, thoroughly tested and will be deployed commercially in July 2021 with version 8.0.1. of Mediasite. The integration will automatically be available to NORDunet's on-premise environment (v.8.0.1).

Integration with Panopto (ready to deploy):

Current status of the Panopto integration:

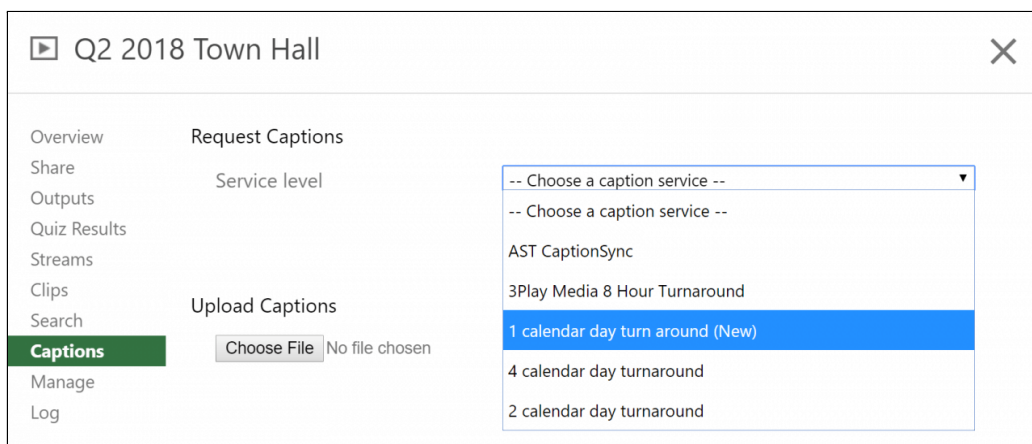
Amberscript and Panopto have been setting up an integrated partnership in the past months. The workflows and IT infrastructure has been setup, and the

integration is ready to go live with a customer. In case of a contract award, Amberscript will collaborate with NORDUnet and Panopto to finalize the setup of the integration. With the integration finalized, Amberscript will be the only European captioning provider within the Panopto ecosystem. The workflow is designed to look as described below:

User rights and workflow

Panopto admins can determine which users have access rights to Panopto's captioning services. Panopto users can then determine which videos (our course, or libraries) they want to submit for captioning. Within the platform, users have both options: Automatic captioning and human captioning.

The user interface will look as follows:

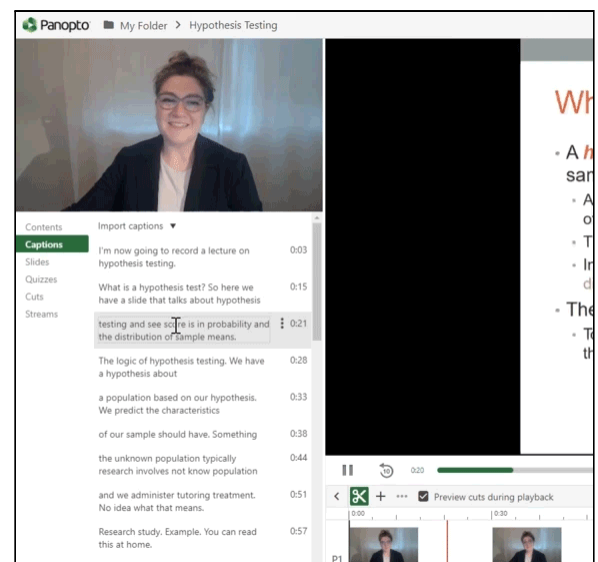


Captions will automatically be added to the video file as soon as the captions are delivered.

Automatically generated captions can be corrected and edited in a captions-editor within Panopto:

How to set up the designated integration with Panopto

After deployment of the integration, Panopto admins will be able to configure their Panopto environment in a way that Amberscript will be the standard captioning provider (for human and automatic captions). The administrators can also define the ordering rights of each individual user.



Integration into other media service providers:

Next to the integrations with the media service providers, Amberscript's API is used by many other media service providers to integrate automatic and human transcription and captioning services into their workflow. Amberscript has deployed integrations running with the following media service providers:

- Streamio.com (Sweden) - <http://streamio.com/>
- Aventura.no (Norway) - <https://www.ventura.no/corporatetv/>
- Mimir (Norway) - <https://mjoll.no/mimir/>
- Company Webcast - <https://www.companywebcast.com/>
- Planet e-stream - <https://www.planetestream.co.uk/>
- Opencast - <https://opencast.org/>
- Mediaflow (Sweden) - <https://www.mediaflow.com/>

Further media service providers can use Amberscript's API to integrate all of Amberscript's services seamlessly into their workflow.

3.1.17 EC - Manual transcription or subtitling

Question: Please describe in detail the process and workflow for manual transcriptions and subtitling.

It is preferable if the manual transcription and subtitling can be triggered and managed directly from the web portal or through the API.

Answer:

Summary of the workflow for human transcription and captioning services

Submitting audio files for transcription is possible via all channels:

- via the amberscript web-portal
- via Amberscript's API
- via media-services such as Kaltura, Panopto and Mediasite

After a maximum of 3-5 business days, the user receives an email notification that the transcription is completed. The user can then review and download the transcription.

For the web-portal, each user of the customer will receive login details to their personal Amberscript account. The user logs into their secured Amberscript account and uploads the audio files for transcription. Files are stored in ISO27001 and SOC2 certified data centers. The files are processed through a highly accurate Automatic Speech Recognition (ASR) module, which creates a draft transcript in less than 30 minutes. The draft transcript, along with detailed work instructions is then assigned to the best-qualified transcriber on the Amberscript platform. The transcriber uses the AI-supported Amberscript-Editor to review the draft transcript

and ensure it is accurate and free of error. Once the transcription is complete, the files are submitted for Quality Control (QC), where final checks are performed based on Amberscript's detailed quality control guidelines. The customer is notified via email when the transcripts are ready, and they can view the file by logging into their secured account on Amberscript. Users can re-listen to certain audio fragments and export files from the system. The files are automatically deleted from the system after a deletion period that is agreed upon with the customer.

Detailed description of the workflow:

Step 1: Submit a file for human transcription and captioning services

To begin the process, a file will be submitted to Amberscript via the secured web-portal, via the Amberscript application, via an integration into media services or via API.

Within the Amberscript editor, it is also possible to upgrade a file that has been submitted for automatic transcription to human transcription.

In the web-portal, the customer can easily upload a file by clicking on the Upload button. It is also possible to upload the files in a batch of up to 10 files at a time. The customer then selects the required language, dialect and the number of speakers for each file before uploading it.

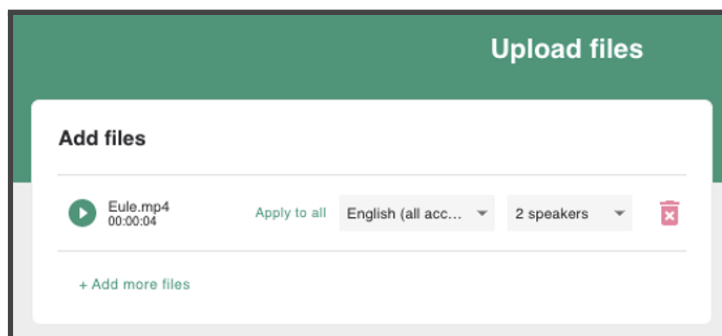


Image: Select Language and number of speakers

All audio files are encrypted and uploaded into ISO27001 and SOC2 certified data centers located within the European Union. Amberscript follows the principle of least privilege where the assigned transcriber and quality checker have the bare minimum privileges (listen and edit only) while accessing the files.

After uploading the file, no further action is required from the user. Amberscript's project management team will receive the file in their dashboard and they will take care of the process until the final transcript is delivered.

Step 2: Automatic Speech Recognition (ASR) & Assigning of a transcriber

A specifically trained Automatic Speech Recognition (ASR) module processes the files to deliver a draft transcript in less than 30 minutes. The data is processed and stored strictly within secured data centers located in Germany.

Amberscript's project managers then assign the draft transcript to the best-qualified transcriber. Transcribers are matched based on factors such as customer timeline and availability of the transcriber, qualifications, capabilities to transcribe certain dialects, customer specifications and topic of the audio.

Step 3: Professional Transcription

The assigned transcriber receives detailed work instructions along with a customer-specific glossary and temporary 'listen-and-edit-only' access to the audio file.

The detailed work instructions contain crucial information such as:

- Transcription guidelines of the client for this file (clean-read / full-verbatim / captions / SDH captions)
- A customer-specific glossary which contains crucial information such as names, specific terminology and other words that are helpful to deliver a fully accurate transcription
- Information on the number of speakers

Next to the customer-specific guidelines, each transcriber is trained on Amberscript's transcription and captioning guidelines.

Transcribers access their secured Amberscript account to listen to the audio file and to make edits to the draft transcript. This closed system ensures that transcribers do not save or download any customer files, and they use the advanced Amberscript Editor to edit and correct the draft transcript. To ensure an even higher level of IT security, all transcribers, project managers and quality checkers employed by Amberscript adhere to strict privacy and security standards and undergo regular training in IT security and privacy.

The transcribers are instructed to review and edit the entire transcript draft from beginning to end to ensure that it is delivered with a 99,5% accuracy rate. Once the transcription is completed, the transcriber submits the files for Quality Control and immediately loses all access to the files.

Step 4: Quality Control

The assigned quality checker receives temporary access to the files and reviews the submitted work according to the Quality Control guidelines. Quality Checkers are highly skilled transcribers or captioners which have multiple years of experience in the field of transcription. They are instructed to review the entire audio file (from beginning to end) to ensure thorough quality control.

In case the Quality Checker rejects the transcript from the transcriber, the following workflow is triggered:

- If the file is rejected by the Quality checker, the file is submitted for review with detailed feedback to the transcriber. The transcriber receives a new deadline of 12 hours to resubmit the transcript after making the necessary edits.
- Once again, it goes through the Quality Control process, and if approved, the files are submitted for customer review.

In case the file is accepted by the Quality checker, the file will automatically be delivered to the customer.

Step 5: Delivery

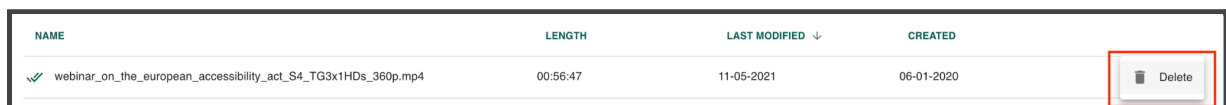
Customers are notified via email about the delivery, and they log into their secured Amberscript account to review the transcript. The customer can either approve or reject the submission and provide the necessary feedback. If the submission meets the customer's expectations, the customer can export the file in any of the required format that includes .doc, .txt, .srt, .vtt or .stl.

Step 6: Deletion

To guarantee a secure transcription process, Amberscript offers 2 workflows to delete files from all servers:

1) Deletion in the Amberscript Dashboard.

The customer can easily delete all their audio and text files from the dashboard with a single click, and the files are immediately erased from all the Amberscript servers. This option puts the user in full control of the data that is shared with Amberscript. When the file is deleted from the Amberscript dashboard, all files (including audio and transcriptions) will be deleted immediately.




| NAME | LENGTH | LAST MODIFIED ↓ | CREATED | |
|--|----------|-----------------|------------|--|
| ✓ webinar_on_the_european_accessibility_act_S4_TG3x1HDs_360p.mp4 | 00:56:47 | 11-05-2021 | 06-01-2020 |  Delete |

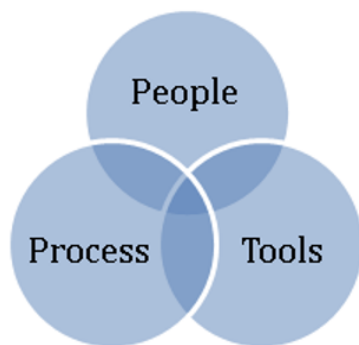
Image 6: Easily delete files from the Amberscript dashboard:

2) Optional: Automatic deletion after a safety period

The customer also has the option to determine a deletion policy with or without a safety period. Files will then automatically be deleted 5, 10 or 30 days after delivery. The user will receive a notification for every file of which the 'deletion date' approaches.

Quality assurance at Amberscript

To maximize the output and improve the quality of our transcription platform, Amberscript actively applies principles of Lean Six Sigma and ISO9001 under the leadership of the Head of Operations, who is a qualified Lean Six Sigma practitioner. Amberscript is also currently pursuing ISO9001 and ISO27001 certifications.



The three elements of Amberscript's quality control are People, Process and Tools.

People:

- At Amberscript, rigorous processes are set in place to select and onboard qualified transcribers.
- High profile and critical jobs, such as files from NORDUnet, are assigned to only level 5 transcribers who meet various conditions such as
 - o having an academic degree
 - o completion of 100 hours of transcription with First Acceptance Rate (FAR) 95% or higher
 - o adherence to client deadlines in 100% of cases.
- Amberscript directly employs all members of the quality control department. To qualify as a quality control member, resources must complete at least five years of transcription experience and be native speakers of their assigned language.

Process:

- Each submitted file goes through the following process:
 - Automatic transcription

- Manual review and editing (from beginning to end) by a professional transcriber
- Manual review and editing (from beginning to end) by a professional Quality checker
- Submission to the client
- Amberscript has defined thorough transcription guidelines to ensure that transcribers deliver quality work according to customer specifications and according to transcription guidelines for 'clean-read' or 'full verbatim' transcription standards
- For every project, a customer-specific glossary is created, including specific details such as frequently used abbreviations, names of companies or organizations, names of speakers, specific terms of products, and other relevant information.
- The entire process is secured and monitored in the Amberscript platform.

Tools

- Amberscript uses technology to automate as much of the transcription process as possible. Our 'transcription factory' uses a unique combination of assistive technologies and professional transcribers.
- The Automatic Speech Recognition (ASR) tool delivers a draft transcript in less than 30 mins with 85% accuracy. The draft transcript formats the text into speaker segments with a clear distinction between the speakers involved and provides the timecodes per word. It also identifies the topic of the file that enables matching to qualified transcribers and indicates possible errors that the transcriber will verify.
- Amberscript's intelligent algorithm together with experienced project managers match the best transcriber for the job from our pool of highly qualified transcribers after considering various factors like timeline, customer requirement, accent and topic of the audio.
- The Amberscript Editor has several handy tools such as slow/fast player, rewind/fast forward, spell check, and highlight inaudible segments that experts review.
- All time-consuming overhead activities such as downloading or uploading files, putting files into a transcription program, receiving guidelines via email etc. are automated.
- Also, advanced features such as spell check, grammar correction, shortcuts to play, pause, fast-forward, slow-down and rewind the audio, assist transcribers to perform at a high quality standard easily and efficiently.

Amberscript's Quality Management System:

Amberscript is constantly collecting data points to improve and streamline processes according to the Lean Six Sigma mentality. Within the platform, each process step is measured and timed. Bottlenecks are constantly being identified and processes are continuously improved.

On top of that, Amberscript's Quality Management System measures a variety of KPIs.

Transcribers are measured on:

- First Acceptance Rate (FAR) from Quality Control
- Second Acceptance Rate (SA) from Quality Control
- On Time Delivery (OTD)

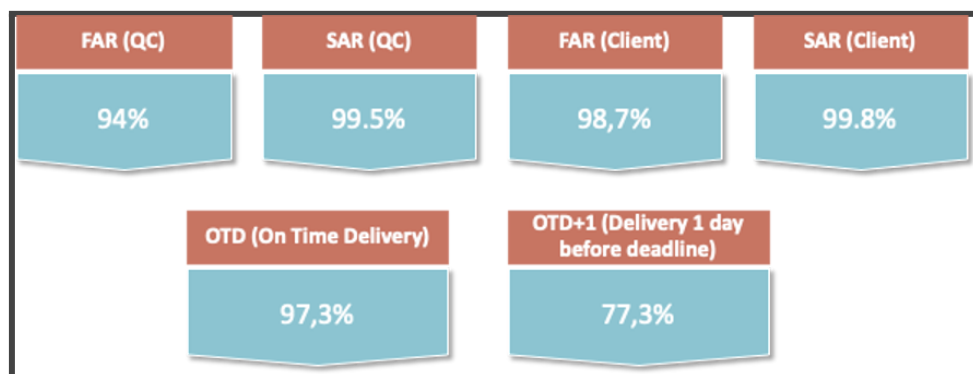
Quality Checkers are measured on:

- First Acceptance Rate (FAR) from the client
- Second Acceptance Rate (SA) from the client
- On Time Delivery (OTD)

Our Project Managers are measured on:

- Time-to-transcription (TTT) - the time it takes before a transcriber accepted the job and starts to transcribe
- On Time Delivery (OTD)
- Delivery before the customer deadline

From the past 3 months, Amberscript achieved the following scores:



Amberscript is proud to share these key performance indicators as they prove that:

- 97,3% of files are delivered to our customers on-time
- 77,3% of files are delivered to our customers 1 business day ahead of the customer deadline
- 98,7% of files are accepted by our customers right away
- 99,8% of the files are accepted by our customers after initial feedback

- For 0,2% of the files, Amberscript delivered the final transcript with a 100% discount to solidify the trust of our customers.
- Currently, more than 1800 hours of audio are transcribed per month by Amberscript in 10 different languages.

Selection of transcribers at Amberscript:

Amberscript guarantees a high quality of transcription services by selecting and training the best transcribers for the job. Transcribers are segmented into different groups based on dialects and domain knowledge and strictly follow the transcription guidelines for each job.

How does Amberscript select and work with transcribers?

Screening

- Amberscript enforces a rigorous screening process that ensures that only the most qualified transcribers can work for Amberscript.
- Each new transcriber who signs up on the Amberscript platform must show prior experience and training and must pass the capability test as part of onboarding.
- Before granting transcriber access to the platform, Amberscript conducts a background check to screen transcribers for:
 - o Their age (minimum age is 18)
 - o A background check is performed by requesting an official Certificate of Good Conduct from the local authorities
 - o Relevant experience in transcription, linguistics or other relevant fields.
- Once the background check is performed and the outcome is positive, the transcriber needs to pass the 'Amberscript Transcription Assignment' where 2 audio files need to be transcribed according to specific guidelines with an accuracy rate of 99,5%.
- After passing the transcription test, selected transcribers sign Non-disclosure agreements (NDA) and Data-processing agreements (DPA) before they can work for Amberscript.

Onboarding & Initial Training of Transcribers

Once transcribers are selected, Amberscript ensures that they are well trained on the following aspects to continuously meet Amberscript's quality criteria and to improve their skills:

- Amberscript Academy: Transcribers receive access to various free resources that include how-to guides, transcription guidelines, practice sessions and webinars
- Tool training: Detailed training is provided to the newly onboarded transcribers to use the Amberscript platform and Editor
- Privacy training: Transcribers are trained to handle data with the utmost care and respect the terms of the Non-disclosure agreement (NDA) and Data-processing agreement (DPA)

Continuous learning and development:

- Amberscript ensures that every transcriber in the community constantly improves themselves through regular feedback and continuous learning.
- Regular training and webinars are conducted to ensure all the resources are updated with the latest information and industry best practices.
- A strict rating system is put in place to evaluate each transcriber. This ensures that customers can rely on the best talent in the industry.
- Over time, transcribers can specialize in particular domains (e.g. legal transcription, academic transcription, political transcription) or progress into the Quality Checker role.

Feedback loop of data into ASR models:

The data that is created in the process of human transcription and captioning is the perfect data to improve the accuracy of domain-specific ASR models. Amberscript has used thousands of hours of this kind of data in order to train customer-specific models for a variety of customers.

Please see point 3.2.5 for more details.

3.2.2 EC - Expected accuracy for automatic transcription

Question: Please provide a % for the accuracy for automatic transcription for each of the 5 languages, for regular language as spoken in the capital areas.

- Danish
- Finnish
- Norwegian Bokmål
- Swedish
- English

An accuracy of 80-85% is expected but a higher accuracy is desirable. Please also describe how the results were obtained; Include a description of the data set, as well as the how the accuracy was measured.

Answer:

Measurement of accuracy:

The most frequently used measurement of ASR output is the word accuracy rate (WACC). The WACC is calculated as follows:

$$WAcc = 1 - WER = \frac{N - S - D - I}{N}$$

WAcc = Word accuracy rate

WER = Word-error-rate

N = number of words in the reference text

S = the number of substitutions

D = the number of deletions

I = the number of incorrect words.

The WER can be calculated by a tool which Amberscript provides free of charge:

<https://wer.amberscript.com/>

Next to the pure accuracy, other factors play an important role to create accurate transcriptions and captions such as:

- punctuation
- capitalization
- number denormalisation
- speaker diarization
- For captions: Synchronization
- For captions: Segmentation of output subtitles

For each language, Amberscript measured the accuracy and other important factors as shown below.

Description of the dataset:

Amberscript regularly keeps track of the performance of different ASR models. For this measurement, Amberscript created datasets that resemble real-world circumstances as much as possible.

Description of the dataset:

Each dataset in each language consists of 20 videos of equal length taken from different domains with speech such as:

- Television interviews

- The weather channel
- Daily news shows
- Political debates & speeches
- E-lectures and instructional videos
- 'Talking Head'-videos
- Screen-capture recordings

Results Word-Accuracy measurement

| Language | Current engine (Generic) | Expected accuracy (Domain-specific engine) |
|-----------|--------------------------|--|
| English | 92% | 94% |
| Swedish | 88% | 92% |
| Norwegian | 82% | 90% |
| Danish | 83% | 90% |
| Finnish | 85% | 90% |

By developing domain-specific speech recognition models, the accuracy is expected to increase significantly especially for domain-specific speech such as specific vocabulary, anglicisms and academic terminology.

3.2.3 EC - Describe any additional language(s) or dialects supported as part of the automatic transcription services.

Question: Please list all supported languages included in the service, support for Icelandic, Norwegian Nynorsk, Sámi languages or any major dialects within the Nordics, will score higher than other languages.

Answer:

Languages supported in the NORDUnet environment:

Via the Amberscript Editor and the API, the following languages will be supported:

- Icelandic (see current status below)
- Norwegian Nynorsk (see current status below)
- Northern Sami (see current status below)
- German
- Dutch
- French

- Italian
- Spanish
- Portuguese
- Arabic
- Chinese (Mandarin)
- Croatian
- Czech
- Danish
- Dutch
- English (all accents)
- Estonian
- Farsi (Iran)
- Filipino
- Finnish
- French
- German
- Greek
- Hindi
- Hungarian
- Icelandic
- Italian
- Japanese
- Latvian
- Lithuanian
- Norwegian
- Polish
- Portuguese (Brazil)
- Portuguese
- Romanian
- Russian
- Spanish
- Swedish
- Thai
- Turkish
- Ukrainian

Status of Icelandic:

In recent months, Amberscript started developing a speech recognition model for Icelandic language which is still under development. Apart from that, Amberscript partners with Tiro (<https://tiro.is/>) who developed an Icelandic Speech Recognition model. The Tiro model for Icelandic can be integrated into current services by August 2021.

Status of Norwegian Nynorsk:

Amberscript develops its own custom speech-to-text engines, and partners up with the best players in the market where Amberscript itself doesn't have the data

or capabilities to create its own engine. For Norwegian Nynorsk, Amberscript currently does not have the data to create an engine itself. However, Amberscript has a cooperation with a partner who can integrate Norwegian Nynorsk into the Amberscript platform. Therefore, Amberscript can make an engine of Norwegian Nynorsk available at the end of 2021 based on the demand of NORDUnet.

Status of Northern Sami

Amberscript investigated the possibility of launching an ASR model for Northern Samí in 2022. It seems Amberscript can build on top of the research described by Juho Leinonen and co-authors in his paper 'New Baseline in Automatic Speech Recognition for Northern Sámi' and that Amberscript can use the Northern Samí portion of the Giellagas Corpus of Spoken Saami Languages as a starting point when demand is there.

3.2.5 EC - The Nordic R&E community needs the solution offered to be able to adapt to R&E domain-specific language.

Question: The Nordic R&E community needs the solution offered to be able to adapt to the domain-specific language used in; healthcare, physics, mathematics, bioscience or other domains within the R&E community. This includes vocabulary ability to recognize special terms and abbreviations etc.

Describe ALL mechanisms for using either existing transcribed material to train the solution/service, or directly using i.e. a web portal text editor to achieve this.

Also describe possibilities for opting in/out of providing material. The more and flexible ways for training the solution the better.

Answer:

The Speech recognition (ASR)

Speech recognition models are a complex system of multiple modules such as (1) the acoustic and linguistic model, (2) punctuation models (see 3.2.9), (3) post-processing and text formatting and (4) speaker diarization (see 3.2.6). In the following, the sub-modules will be described separately:

As Amberscript mainly works with customers in the educational sector, all of the models that Amberscript creates align very well with the use case of NORDUnet. Amberscript's current models are already optimized for speech recognition in academic settings and will be optimized further with the input from NORDUnet from the channels described below.

In order to train a domain-specific model for NORDUnet, Amberscript follows a proven 5-step process.

1) Analysis of the performance of the current models

Analysis of the acoustic environment and the linguistic characteristics of the speech that should be transformed into text. The analysis is performed on a representative dataset and the word-error-rate is measured.

2) Data collection

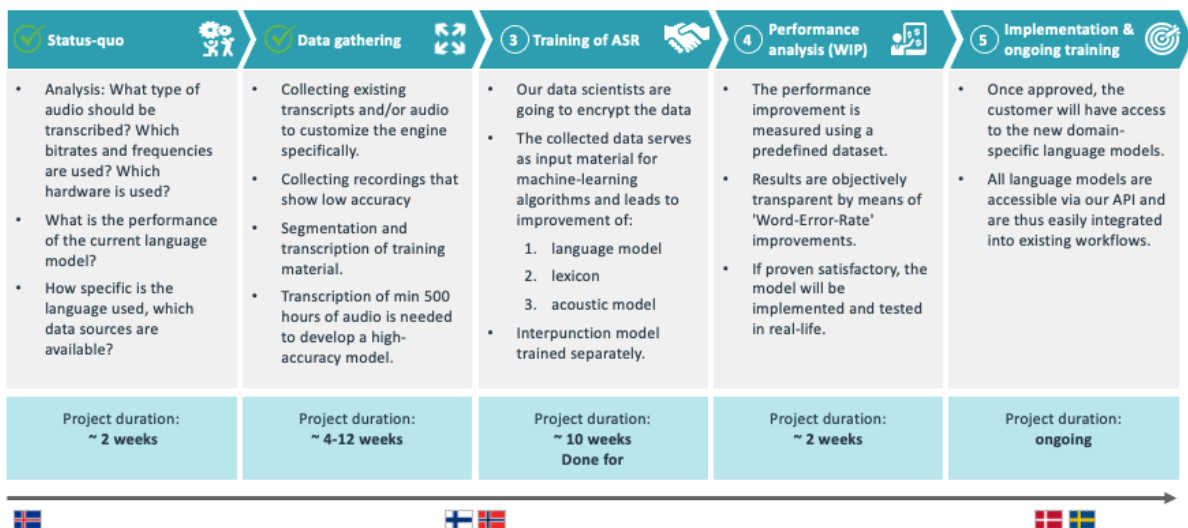
Data (audio & transcripts that are timely aligned) will be collected from public sources, from data that NORDUnet might be able to provide, but also from Amberscript's internal platform

3) Training of the ASR

In the next step, Amberscripts data scientists develop several iterations of linguistic and acoustic models. The lexicon will be optimized and formatting of the output can be specified.

4) Performance analysis

Based on Word-error-rate measurements on relevant audio files, the accuracy will be measured. Next to purely numerical 'accuracy rates', also other factors such as readability, punctuation, speaker diarization and capitalization will be taken into account.



5) Implementation and ongoing training

The key to achieve the highest accuracy possible is to continue training and improving the speech recognition models for domain-specific use cases over

multiple iterations. Amberscript uses multiple channels and technologies to collect the most relevant training data as described below.

Post-processing and text-formatting:

The output of a speech recognition model is usually a word-by-word transcription of the speech. Spoken language is oftentimes not easy to read. Therefore, post-processing and text-formatting can help to optimize spoken text slightly for readability. The post-processing module can be used for, amongst other things, the following linguistic characteristics:

- Spoken disfluencies such as 'uhm' and 'ah' can be automatically deleted in the text output
- Spoken repetitions and hesitations can be automatically summarized
- Numbers can be displayed in digits or text ('four hundred twenty' or '420')
- Dates can be formatted in certain way ('4th of February 2020' or '04.02.2020')
- The capitalization model determines, which words should be written with a capital letter (such as nouns, names or abbreviations of for example parties)

Post-processing can save a lot of time for users as it automates tasks that are usually done manually. In collaboration with the client, the post-processing module can be customized in order to meet the transcription guidelines of the Starting.

The acoustic and linguistic model

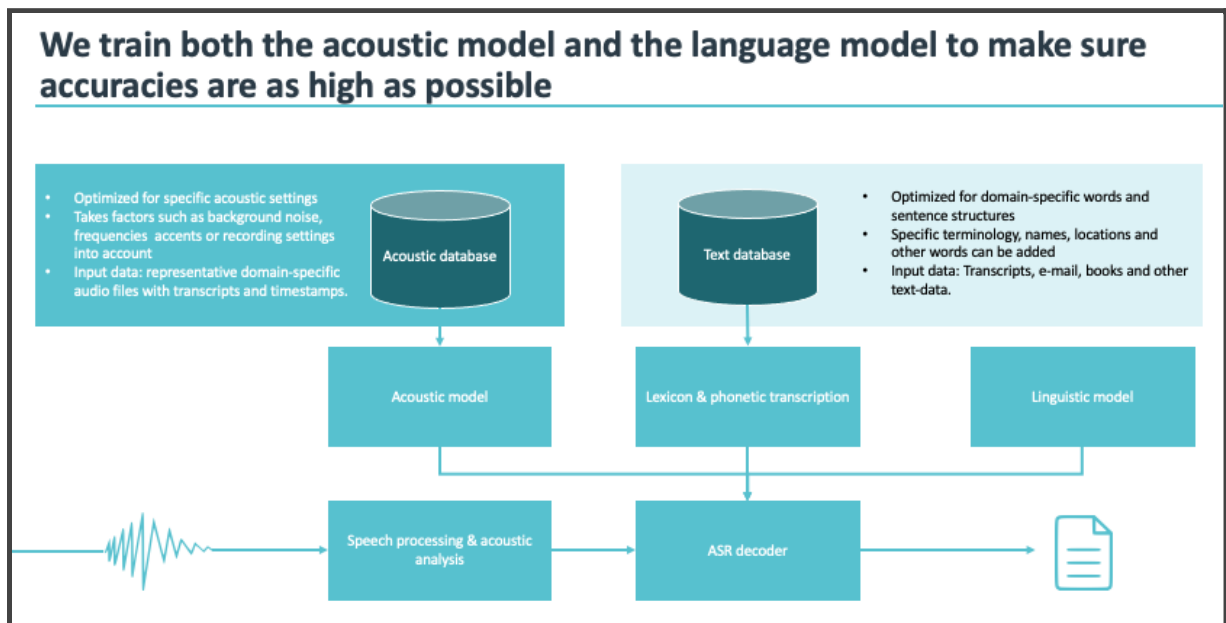
Currently, Amberscript uses the best generic ASR engines in the market and develops domain-specific speech recognition models based on the open-source framework Kaldi, which is the state-of-the-art speech-recognition framework that is used by most companies that are operating at the forefront of innovation. Kaldi is the fundamental framework for training, evaluating and running the ASR models in different languages.

Next to Kaldi, another technology called 'end-to-end ASR' is slowly gaining traction. Amberscript recently developed its first end-to-end ASR model for English language which showed promising results. In case the performance of end-to-end models starts to outperform the accuracy of Kaldi-based models, Amberscript will switch towards this technology in the future.

Amberscript's science team has perfected the process of training and deploying speech recognition models which are specifically trained for specific customers. As of today, Amberscript operates several ASR models which are continuously updated and tweaked for our main user group: Academic researchers and video content from education within Europe.

Many suppliers of ASR solutions offer customization capabilities which focus mainly on the linguistic model of the speech recognition. Words can be added to the lexicon and text can be submitted to be used for training purposes. This approach however neglects the acoustic model which is crucial to improve the recognition of speech with accents, with difficult acoustic characteristics (e.g. recorded and compressed online meetings) and with frequent use of anglicisms and specific terminology. What makes Amberscript unique is that we train both components of an ASR Model in order to achieve highest accuracies for educational content: The acoustic model as well as the linguistic models.

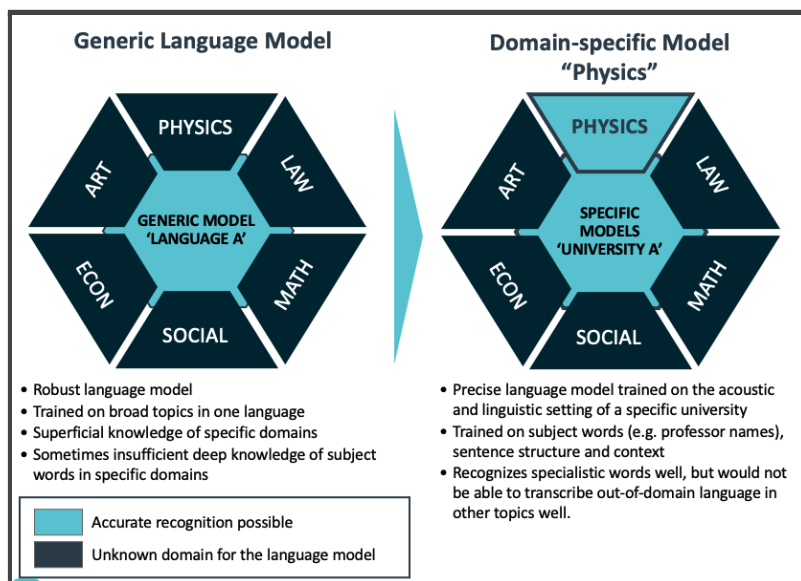
In the coming months and years, Amberscript will continue to improve domain-specific speech recognition models for the academic sector by adding vocabulary and acoustic data from various domains and acoustic settings.



The development of domain-specific models

In speech recognition, it is a best practice to develop ASR models that are tailored to the domain of speech which shall be transcribed. While a generic model is trained on a broad dataset of a variety of acoustic and linguistic speech, a domain-specific model is developed to achieve the highest accuracy possible while focussing in depth on one or more domains (see below 1).

Speech in educational videos and academic research is the perfect use-case to create a domain-specific model as the acoustic and linguistic characteristics are very specific and very constant.



Continuous improvement and data collection of domain-specific data:

The key in achieving high-accuracy speech recognition models is continuous training with human feedback loops. Amberscript automated a large part of the process to improve speech recognition models by implementing a direct feedback loop. With machine learning techniques and semi-supervised learning, the models are continuously improved with the help of linguists and data scientists who are native speakers in the Nordic languages.

There are multiple sources that Amberscript uses to collect relevant training data:

1) The Amberscript web-portal and transcription editor

Data that NORDUnet users transcribe with the Amberscript text-editor can be used to create NORDUnet-specific engines in the course of the partnership. Audio- and video data that is uploaded by users in combination with human edits that are performed by the users are regularly selected by an algorithm, anonymized and used to improve the engine over time.

In this way, Amberscript can select the most relevant data possible for NORDUnet as the future ASR models will be trained more and more by real input data from the NORDUnet community.

Opt-out: Per default, all users are opted in to submit their manual corrections to improve the engines over time. However, can simply opt-out of this training by sending an email to info@amberscript.com or by switching their account settings in the user overview.

2) Human transcription and captioning

Data that NORDUnet users submit for human captioning is the perfect training data to improve the ASR models that Amberscript will develop for the NORDUnet community because of the following reasons:

- The media files resemble the exact content that the ASR be optimized for in terms of acoustic- and linguistic characteristics
- Professional captioners and additional quality checks guarantee a high accuracy of the humanely-corrected text
- Each hour of video that is transcribed helps to improve the quality of the ASR engines significantly

Amberscript trained domain-specific models based on this kind of data for several customers such as these:

- Creating a domain-specific model for political debates for the state-parliament of Mecklenburg-West Pomerania in Germany. After using around 1.000 hours of manually transcribed videos in order to train the ASR model, the accuracy of the model increased by more than 9%
- Creating a political speech recognition for more than 100 municipalities and province-parliaments in the Netherlands. In the past 2,5 years, Amberscript transcribed hundreds of hours of parliamentary debates with human transcribers. In regular feedback loops, this data has been fed back into the ASR model which is trained on the political domain. With each iteration, the accuracy increased significantly leading to an overall accuracy of up to 93% by now (May 2021).

Opt-out: Per default, all users are opted in to submit their manual corrections to improve the engines over time. However, can simply opt-out of this training by sending an email to info@amberscript.com or by switching their account settings in the user overview.

3) Anonymized relevant data from other Amberscript customers

As mentioned previously, Amberscript's transcription software was used by more than 1 millions users who deliver audio-, video and text-data that Amberscript uses with their consent to improve ASR models.

NORDUnet will benefit from this huge speech corpus as Amberscript will select acoustically- and linguistically relevant data to improve the domain-specific engines of NORDUnet.

4) Public data sources

Amberscript is deeply connected in the ASR research community in Europe. Frequently, new speech corpusses are published by different stakeholders. Also in the network of NORDUnet, there are some universities doing groundbreaking research in the ASR field such as

Naturally, Amberscript screens these corpusses to determine whether any of the current engines can be improved by adding data from a public data source.

Some of the data sources that might be used for NORDUnet's engines include:

- Mozilla Common Voice Project
- The BigBrother Corpus, offered by Oslo University
- Other resources provided by the CLARINO in Norway project
- other publicly available corpusses that might be provided by the NORDUnet network

5) User-input

Amberscript can also use input data that is specifically uploaded by the users for training purposes to improve the engines. For this, Amberscript created a simple web-form which will be integrated into the user-interface of the web-portal by the end of 2021.

Form to collect data:

Amberscript created an easy-to-use form that can be used to upload training data such as human captions or transcriptions that have been created in the past:

The screenshot shows a web form titled "Submit Training Data to Amberscript". The form is divided into several sections:

- Header:** Amberscript logo with the tagline "We make audio accessible".
- Language Selection:** A section titled "For which language do you want to submit additional words?" with radio buttons for "English", "French", and "Other: _____".
- Instructions:** A section titled "Submit Training Data to Amberscript" explaining the form's purpose: "This form was designed to submit training data for Amberscript's speech recognition. Amberscript will use the data to improve the automatic speech recognition for the customer. Please submit manually corrected caption files, transcripts or single words. The name and photo associated with your Google account will be recorded when you upload files and submit this form. Not thomas@amberscript.com? [Switch account](#)".
- Keywords:** A section titled "Please submit keywords that you want to add to the speech recognition separated by a comma (e.g. USA, White House, ...):" with a text input field labeled "Your answer".
- File Upload:** A section titled "Please upload the corrected caption file (VTT or SRT):" with a button labeled "Add file".
- Validation:** A section titled "For validation purposes, please submit your company email address: *" with a text input field labeled "Your answer".
- Submit:** A large orange "Submit" button at the bottom right.

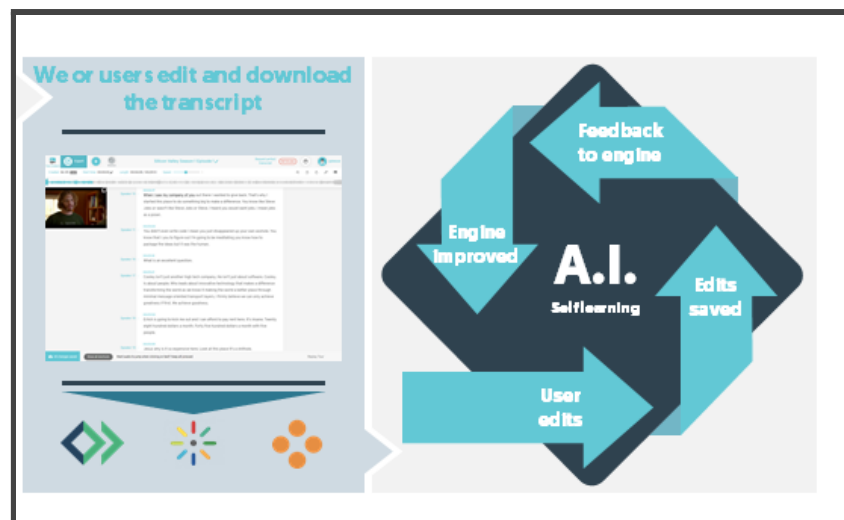
Submitting data via this form also allows users to add a maximum of 1000 words to their 'private dictionary' for automatic or manual transcription. These words will be added to the lexicon of the speech recognition immediately in order to increase the chance of recognizing specific terminology.

6) Feedback-loop with media services

It is expected that a lot of the usage from NORDUnet will come through media services (such as Kaltura, Mediasite and Panopto). Amberscript is currently

partnering with these media services to develop a feedback-loop mechanism. In the Kaltura-REACH environment, this feedback loop will look as follows:

- Users submit videos for automatic captioning through Kaltura to Amberscript
- Users receive automatic captions in the Kaltura platform
- Users make corrections to the automatically generated captions within the Kaltura environment
- These user-edits will be sent back to Amberscript on a monthly basis so that Amberscript can use these data points to improve the accuracy of the speech recognition with the data.
- This feedback loop will only be activated for users that request the service (opt-in) with their Kaltura contact person



Future Outlook: End-to-end ASR:

In the past months, Amberscript's machine-learning team has created the first ASR models based on a new technology called 'end-to-end' speech recognition. This machine-learning approach has the advantage that it can be trained on large corpora of audio data without requiring labelled text. Once a base model is developed, it is possible to create a domain-specific ASR engine with a relatively small amount of data (30 - 100 hours).

Amberscript has launched its first English end-to-end ASR in the last weeks and the results are very promising to continue developing this technology. While tests in Scandinavian languages are still lacking, Amberscript has all the machine-learning capabilities in house to create end-to-end ASR models for NORDUnet in the future based on data from all the sources described above. Especially in low resource languages, such as Icelandic, Amberscript will use end-to-end models to create high accuracy engines.

3.2.6 EC - The Nordic R&E community needs the various speakers in a recording to be recognized as best as possible

Question: Please describe how the solution effectively achieves this i.e. through the solution being able to distinguish different voices in one and the same sound stream. **Include information about how it will be possible to link metadata to the text based on this information.**

Answer:

Modules that are used to detect speaker changes:

In order to segment the output text into accurate speaker segments, Amberscript makes use of two AI-based algorithms that cluster the text based on speaker-labels: Voice-activity detection and x-vector diarisation.

Voice-activity detection:

- Voice activity detection is a neural network that splits the entire audio file into segments
- Segments are then evaluated and classified as 'speech' or 'non-speech'
- The classification is made based on a neural network that analyzes MFCC-features
- MFCC-features are a representation of the frequencies that are recorded on the audio file. The classifier distinguishes common features that are represented in human speech from 'background noise'

Once the output text is split into 'speech' and 'non-speech' segments, the x-vector diarisation module is used to add speaker labels.

X-vector diarisation:

- the x-vector diarisation module splits the 'speech segments' into various 'speaker segments' with labels for each speaker
- When submitting an audio-file to Amberscript, the user can indicate the number of speakers that are expected in the audio. Submitting this information, improves the accuracy of the speaker diarization significantly.
- the x-vector module analyses MFCC features of human speech to identify different voice patterns and to assign speaker labels. Amongst others, the following factors are taken into account:
 - pitch of voice (high-pitched voice vs. low pitched voices)
 - pace of speech (does a speaker speak fast or slow?)
 - patterns of speech (such as accents and other speech patterns)
 - volume of speech (loud vs. quiet)
 - pauses between speech
 - Speech events such as interruptions

- other voice patterns that are visible in the MFCC-features of the sounds recording

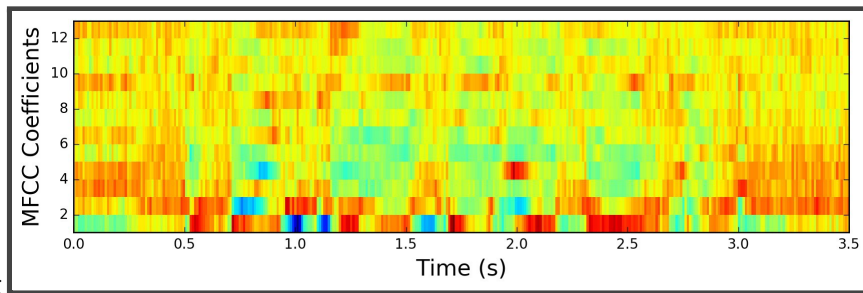


Image: An example of MFCC features that are analyzed by x-vector modules.

Diarisation included in the ASR solutions and meta-data:

Both modules are included in all of Amber Scripts ASR solutions (Text-editor, integration into media services, API).

When exporting a transcript or captions in JSON format, the recognized number of speakers is indicated in the header of the JSON so that it can be used as metadata.

This metadata can also be imported from Amberscript's API by mediaservices.

3.2.7 EC - Often multiple languages are being used in a single recording, in a teaching environment, i.e. Finnish and English.

Question: Please describe how the service supports automatic language detection and how it supports multiple languages within a single recording.

Answer:

In order to solve the challenge of language detection, there are 3 use cases which typically apply at universities:

1) Use of specific terminology and anglicisms in a non-English language

Amberscript creates language models in each language with a large number of training data. A very large part of the training material is usually from academia and research as this is Amberscript's main customer group.

As part of the training data, anglicisms are frequently included and added to the lexicon of the speech recognition system in order to improve the accuracy of recognition for specific terminology and anglicisms.

Inclusion in the proposed solution:

In the proposed solution, Amberscript will continue to train the domain-specific ASR models to recognize frequently used anglicisms and other specific terminology.

2) Detection of the language of an audio file if the language is unknown

In some instances, the person that submits a media file for automatic transcription, does not know in which language the media file is recorded. This sometimes happens at universities where lectures are held in multiple languages and where professors at time simply decide to have an entire lecture that was scheduled for 'English' in the local language.

Amberscript in the past used an algorithm which compares the 'average confidence score per language model' of a short sample of the submitted audio file to determine the language of a media file. The algorithm worked well if the amount of languages could be limited to 2. (e.g. the file is either 'English' or 'Dutch'). Classification worked less accurate if the language was completely unknown (e.g. 'the file could be any of these 10 languages').

Inclusion in the proposed solution:

In case of a contract award, Amberscript would like to investigate the concrete use cases together with NORDUnet to see where the algorithm can be deployed most effectively.

3) Code-switching within a single recording (switching from one language to another within the same conversation)

Code switching (which describes the event when a speaker switches from one language to another and back multiple times) is a problem that isn't yet broadly solved in the field of Automatic Speech Recognition.

In programme funded by the Dutch government, Amberscript is currently working on ASR models that support frequent code switching and is trained on two languages.

Inclusion in the proposed solution:

In case the programme is successful, Amberscript can include the first multilingual model for NORDUnet approximately in the first quarter of 2022. The language combination is to be determined together with NORDUnet.

The current research approach looks as follows:

- We will use the Kaldi toolkit (open source Apache License v2.0). The approach will be based on previous work done by TaipeiTech Lab and the Slovak Research and Development Agency, which already have experience in implementing and testing Deep Neural Network bilingual ASR. Simultaneously, research is done whether Wav2Vec end-to-end modelling would work with this approach.
- A large vocabulary continuous speech recognition module, which is already available in Kaldi and deployed by Amberscript in more than 10 engines, will be combined with a new language identification module which runs in parallel and is integrated with a gate mechanism.
- To train a bilingual speech recognizer, databases with sentences in both languages are required. Material of Dutch natives speaking English has to be included. The following datasets will be used:
 - ELRA; CAREGIVER Corpus (4.000 recordings English, 20 Dutch), Collins Multilingual database (2.000 recordings in both English and Dutch), MIST database (740 recordings in both English and Dutch). Corpus Gesproken Nederlands (CGN) (900 hours Dutch), LibriSpeech (1000 hours English), VoxForge (120 hours English, 10 Dutch) and Mozilla CommonVoice (1.469 hours English, 42 Dutch) datasets.
- A bilingual large vocabulary continuous speech recognition module will be trained using Kaldi. For the share-hidden-layer Neural Networks, the input and hidden layers will be shared for Dutch and English, while the output layer will be grouped. For better discrimination of similar phonemes all neurons in the output layer will be tuned at the same time. An LID-based gating mechanism will be built to control the LVCSR output and discriminate different languages. The output layer is not presented this way in a monolingual setup and will be of a large contribution to the community. The output scores of the model are multiplied by Language Identification modules scores.
- We will experiment with different setups. 43 MFCC feature vector together with 100 dimensional i-vector is the starting point, with 6 layers and 850 neurons for each layer following the AiShell/nnet3 recipe (425 for language identification training).
- The project builds on the following work (journals and names emitted for consistency):
 - Crosslanguage knowledge transfer using multilingual deep neural network with shared hidden layers, 2013
 - Automatic recognition of Slovak-English bilingual speech, 2019
 - Multi-task recurrent model for true multilingual speech recognition, 2016
 - Towards language-universal end-to-end speech recognition, 2018
 - Multilingual speech recognition training and adaptation with language-specific gate units, 2019
 - The Kaldi speech recognition toolkit, In Proceedings of IEEE 2013
 - Workshop on Automatic Speech Recognition and Understanding, ASRU 2011
 - Gated Module Neural Network for Multilingual Speech Recognition, 2018
 - End-to-end speech recognition in English and Mandarin, 2016

3.2.8 EC - Music and background noises

Question: Please describe the service ability to distinguish music/background noise from speech and its ability to label it accordingly.

Answer:

The labelling of music and background noises is not yet integrated into the current ASR engines. However, Amberscript researched and tested several technologies and found out a promising approach towards labeling of music and other background noises. Currently, Amberscript uses a voice-activity detection module which analyzes MFCC (Mel-frequency cepstral coefficients) features of the recorded audio in order to identify speech segments.

This module makes use of neural networks that are trained to label every segment in the audio with 'speech' or 'non-speech'.

This model can be repurposed to classify each segment of the audio into other clusters such as: [music], [applause], [laughter], [sneezing], [coughing] and others.

While Amberscript has a running model to label music and background noise in a staging environment, more labelled data is required to create a reliable model.

In case NORDUnet wishes to integrate these kinds of features into the transcription solution, Amberscript and NORDUnet can cooperate to create the specifications of sound labels and to create a corpus of labelled data subsequently.

3.2.9 EC - Punctuation

Question: Please describe the service ability to add punctuation to the transcribed text. More ways of detecting and adding punctuation is preferred. i.e. grammatic and pauses.

Answer:

Amberscript's punctuation models are based on a technology called 'Transformer'] and are specifically trained for each language as part of the 'post-processing' modules of the ASR solution. In order to create accurate punctuation models, more than 500 thousand training-sentences are required which ideally resemble natural speech of the domain that shall be transcribed. For this use case, Amberscript continuously collects input data from real user data in the academic domain. The data set is then cleaned and accurately labelled by linguists and language experts to ensure that the training data for the punctuation model is 100% accurate.

The current punctuation models in all languages (Swedish, Finnish, Danish, Norwegian and English) support signs such as dots (.), kommas (,), dashes (-), colons (;) and question marks (?) will be supported.

The punctuation models uses multiple indications to determine where to insert punctuation signs such as grammatical factors and speech patterns:

Grammatical factors

- using commas before any coordinating conjunctions that link two independent clauses
- using commas after dependent clauses that start sentences
- using commas to separate items in a series
- using commas after introductory phrases
- using commas before quotes
- using commas when the first word of a sentence is a freestanding 'yes' or 'no'
- using a dot to end a sentence
- inserting a dash for abrupt break in thought
- using colons for introduction of a list of items
- inserting colons to introduce indirect speech
- ending question-type sentences with a question mark

Speech patterns:

Next to grammatical rules, a second punctuation model also takes speech patterns into account such as:

- High-rising terminal indicates a question mark: High-rising terminal describes the fact that questions are often indicated by the speaker ending the sentence with a higher-pitched voice.
- Pauses in speech often indicate punctuations such as dots (.), colons (:) or dashes (-)
- repetition of the same voice pattern often indicates multiple items on a list ('I put the car, the home and the dog on my bucket-list')

AI-based triggers:

Typically, the punctuation models are trained on thousands of input sentences. Next to grammatical factors and speech patterns, the AI-based models also take other factors into account that might not be obvious to the linguists developing the AI modules.

3.3.1 MR - As part of the evaluation NORDUnet wants to perform a test of the web portal functions and the accuracy of the automated transcription, using generic language for Danish, English, Finnish, Norwegian and Swedish.

Question: The Bidder must provide five (5) test accounts for the web portal service, for NORDUnet to evaluate the web portal functions and accuracy in the automatic transcription. Each account must have a minimum of ten (10) hours of subscription time associated.

Please provide the access information for the five (5) accounts, and a contact person and contact information, should NORDUnet encounter any issues with the access to the web portal.

The accounts shall be available at the time of submitting the bid.

Answer:

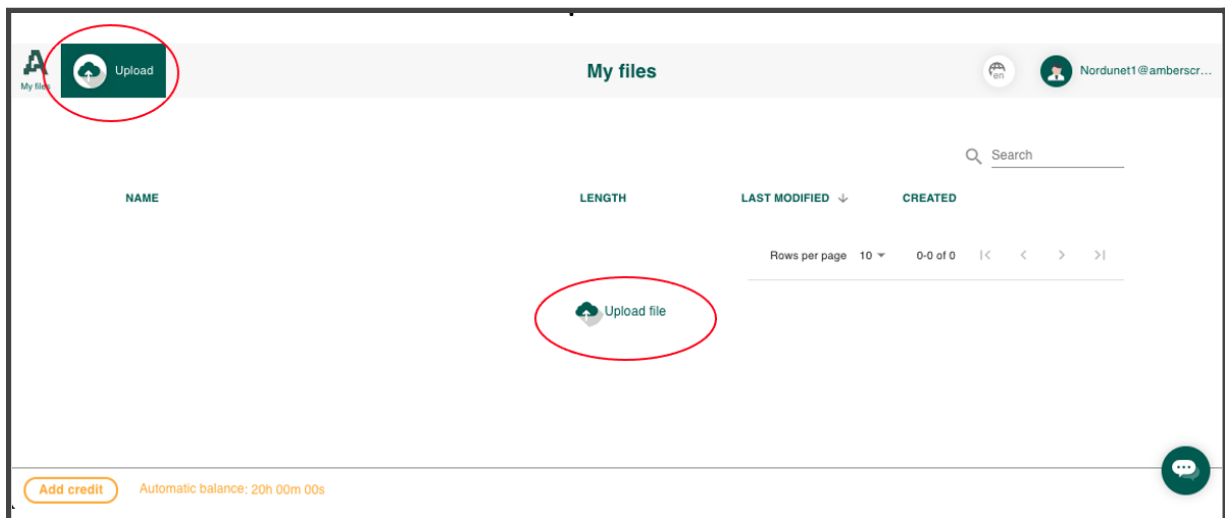
Test Accounts & free credit for automatic and human transcription:

Amberscript has created 5 test accounts for NORDUnet with 20 hours of credit for the automatic transcription feature.

Amberscript would like to offer NORDUnet to also test the human transcription or human captioning feature in **English** or **Swedish** for a maximum of 2 hours free of charge. In order to order manual services, please follow the instructions below.

User Instructions:

1) Upload files by selecting 'Upload' in the dashboard



2) Select languages, number of speakers and type of project (transcription or subtitles)

Upload files

Add files

0 SV Test - Clean Transcription.mp4 00:02:06 Apply to all English (all acc...) 2 speakers

+ Add more files

Select service Selected services will be applied to all uploaded files.

Transcription
Written text delivered as a document.

Subtitles
Time-coded text delivered as a subtitle file, or burned in the video.

Summary

0 SV Test - Clean Transcription.mp4 00:02:06

| | |
|----------|-------------------------------|
| Subtotal | Select service to see pricing |
| Total | Select service to see pricing |

Order

3) Select 'automatic transcription' or 'manual / human transcription'

Select service Selected services will be applied to all uploaded files.

Transcription
Written text delivered as a document.

Subtitles
Time-coded text delivered as a subtitle file, or burned in the video.

Automatic - Pay with your balance (20:00:00)
Transcript by Amberscript algorithm and perfected by you. Ready in less than 1 hour.

Manual
Transcript in file's selected language, perfected by our language experts. Ready in 3-5 working days.

4) Optional: For human transcription select 'invoice' as payment option and fill in test-credentials

Choose payment method

Credit card PayPal SEPA Invoice

Billing Details

First name* Last name*
Test Nordunet

Address* City*
Test street Test City

Country Postal code*
Denmark 1111
You can change your country in your settings

Company VAT number

Summary

Transcription - Manual

0 SV Test - Clean Trans... 00:02:06 -

Subtotal 00:02:06 -

Total -

Order

By continuing you agree to our Terms of Service

Nordunet can use the following credentials to log into the web-portal with the url:

<https://app.amberscript.com/>

| Username | Password |
|---------------------------|------------------|
| Nordunet1@amberscript.com | Amber@nordunet1! |
| Nordunet2@amberscript.com | Amber@nordunet1! |
| Nordunet3@amberscript.com | Amber@nordunet1! |
| Nordunet4@amberscript.com | Amber@nordunet1! |
| Nordunet5@amberscript.com | Amber@nordunet1! |

Customer Support:

During the test, NORDUnet can reach out to Thomas Dieste, Chief Commercial Officer of Amberscript for any support questions.

Thomas Dieste

thomas@amberscript.com

+31 623120978

3.4.2 EC - Service Capacity

Question: The NREN Customers and NORDUnet are expected to add a significant amount of load to the service. The usage of the service is expected to ramp up over time and potentially peak at 400,000 – 800,000 hours of transcriptions in year-3.

Please describe the current capacity in the service and how it will effectively scale to accommodate the volumes described.

Answer:

Amberscript's services (both: Speech-to-text, but also human services) are built for scalability. The ASR services exist within containers orchestrated by Kubernetes and deployed in cloud datacenters. This setup makes it highly scalable while reducing waste capacity as described below.

As the solution is cloud-based, Amberscript can scale the capacity of the solution nearly instantaneously and to a theoretically infinite scale while matching demand and supply of server capacity. Our system functions by reacting to changes on demand and scaling to appropriately match that demand. By polling the number of incoming job requests every 60 seconds, we can quickly react to changes in demand, scaling up Kubernetes deployments (transcription servers) and server nodes to increase our capacity to a theoretical infinite level within seconds

Our system also reacts appropriately in periods of low demand, scaling services down to reduce waste and therefore cost. Our scaling solution has been proven to work with the increase in demand over the past few months, and the system is already poised to handle continuations in that growth

A monitoring system is in place that alerts us when server/capacity errors are happening, so Amberscript's tech team can react to it

Amberscript's current transcription volume exceeds 800.000 hours per year by far and our infrastructure is set up to handle volume increases using the auto-scaling solution described above. In the past 12 months, Amberscript's infrastructure has been available with an uptime of more than 99,98%, while an uptime of 99,9% is guaranteed for our enterprise customers.

3.4.3 EC - Automatic Transcription processing speed

Question:

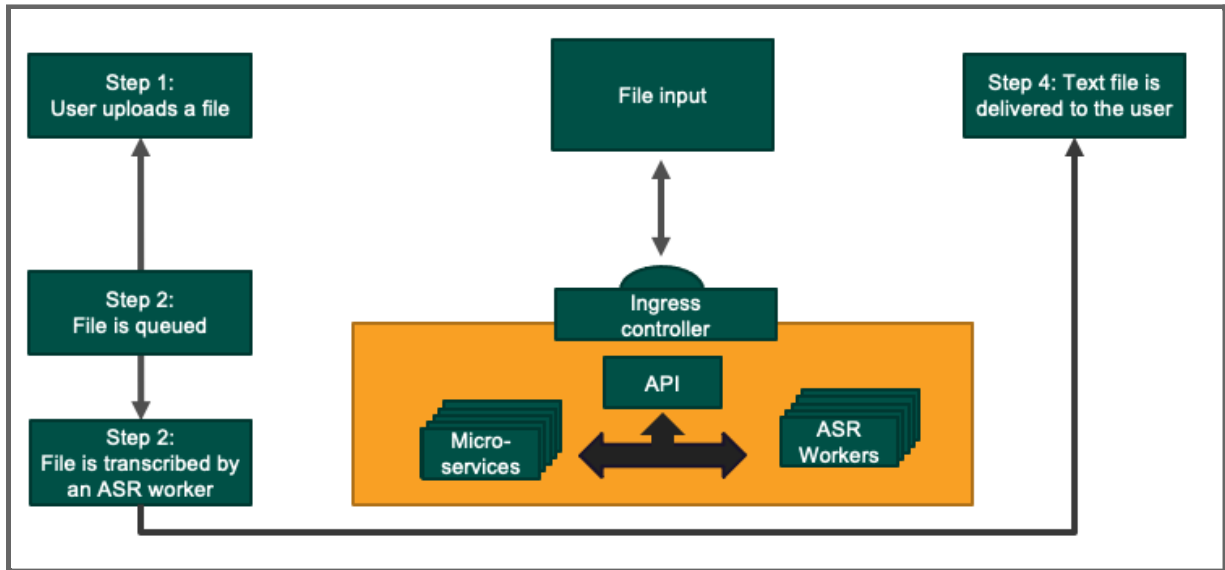
NORDUnet wants an automatic transcription process that is fast and effective. Please describe how the service support this, please provide examples and i.e. workflows for the API and Web portal scenarios. Including typical processing times.

Simple workflows and fast turnaround times are preferred.

Answer:

The automatic transcription process consists of 4 steps:

- 1) The user uploads a file
- 2) The file is queued
- 3) A free ASR-worker transcribes the file
- 4) The output file is delivered to the user



The same logic applies for all delivery methods and workflows such as the submission of files via the web-portal, via API or via media-services. Auto-scaling in Amberscript's cloud-environment allows Amberscript to scale ASR workers up and down based on real-time demand. Therefore, queuing times will be kept short at all times automatically.

The transcription time depends on the length of the audio file. The table below gives a good indication of the transcription time:

| Length of the media file | Typical processing time |
|--------------------------|-------------------------|
| 5 minutes | 2 minutes |
| 30 minutes | 10 minutes |
| 60 minutes | 15 minutes |
| 180 minutes | 45 minutes |

Future outlook: End-to-end ASR:

In the past months, Amberscript developed a speech recognition model which is based on end-to-end speech recognition (instead of the Kaldi framework). One significant advantage that end-to-end technology offers is that the transcription speed can be significantly increased as shown below:

| Length of the media file | Typical processing time |
|--------------------------|-------------------------|
| 5 minutes | 1 minutes |
| 60 minutes | 2 minutes |

Amberscript will keep on researching end-to-end technology and implement it for NORDUnet in case the accuracy of end-to-end models starts to exceed the performance of Kaldi-based models.

3.4.4 EC - NORDUnet wants as few limitations in the service as possible

Question: Please list known limitations in the service i.e. maximum length or input file size of a recording etc.

Limitations that inhibit an effective use of the service will affect the score.

Answer:

There are only 2 limitations that influence the size of files that can be uploaded.

Maximum length of video or audio files: **12 hours**

Maximum size of video or audio files: **8GB**