aiconix

Live
Onboarding

07/2023

Onboarding

After you have sent us your request, we will contact you as soon as possible to discuss the technical requirements of your live stream. Based on a requirements catalog (next page), we will determine with you the technical framework for a smooth process.

Testphase

A pre-test allows, taking into account the real circumstances, to evaluate the automated subtitling and translation.

Support by aiconix for questions / problems: support@aiconix.ai

Steps

1. We will set up an endpoint for you to send your live stream to aiconix

2. Give us (if necessary) a list of names and special terms ,which we will add to your personal digital dictionary or edit the names yourself in the admin area

3. As soon as the live stream starts, aiconix will automatically transcribe the audio track and convert it into subtitles

4. Either we wrap the subtitles back into the video or into a text-stream

5. Via URL, the subtitled video stream or text stream comes back to you:
   a. either you pick it up via URL
   b. or we push for you e.g. to YouTube, Twitch or other channels

6. If desired, you can use the live editor during the transcription to make any necessary corrections yourself before the transcription appears online

Please answer the following administrative questions:
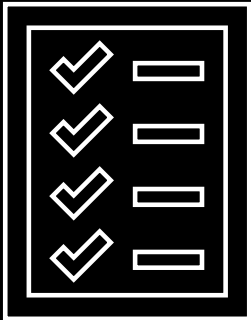
- On which days and at which times should the streaming take place?

- To what extent will the live streaming take place: Event day, conference week, permanent streaming?

- In which language should subtitles be provided?

- How many streams are to be subtitled?

- Which language will be spoken in the stream? Is there a language change?

- Should a manual correction of the subtitles be done additionally before the stream goes online?

- Is a translation of the live stream desired? If yes, in which language?

- Should a dictionary be included with technical terms and proper names?

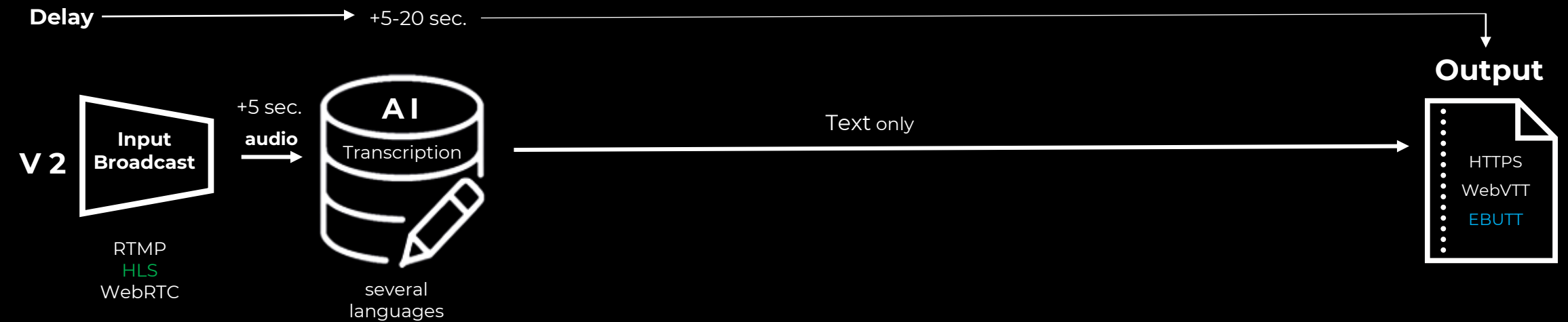**Please answer the following technical questions:**
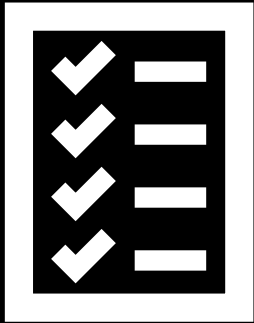
- Do you want the subtitles as burned-in subtitles or embedded closed captions (CC)?

- Is rtmp streaming possible for you?

- Parallel streaming or sequential?

- What happens to the transcribed stream: should it be pulled or pushed? If so, where to?

- What FPS, bitrate, resolution will be sent?

- Where does the stream come that will be send to aiconix?

- Is an SRT file needed afterwards?

Please also provide the details of the **contacts for <u>administrative</u> and <u>technical</u> support** for the livestream.

Live - technical process flow overview

Under implementation
In the Backlog

Ⓐ

Delay ————→ +5-20 sec. ————————————————→ +5 sec. ————

+65 sec. → +5 sec. → +65 sec. →

V 1

Input Live Stream

RTMP
HLS
WebRTC

+5 sec.
audio →

A I
Transcription

several languages

Editor
optional

Translation

optional

Post Editor
optional

video

Burnedin
Closed Captions

Pull per Stream-URL
Push optional

Output

RTMP
HLS
SRT

Delay ————→ +5-20 sec. ————————————————→

V 2

Input Broadcast

RTMP
HLS
WebRTC

+5 sec.
audio →

A I
Transcription

several languages

Text only

Output

HTTPS
WebVTT
EBUTT

Live – In- & Output Options

OPTION	Ingress	Egress	Video Codec	Audio Codec	Subtitle Format		
RTMP	✓	✗	H.264	AAC	Burnedin	Closed Captions	
WebRTC	✓	✗	H.264	OPUS	✗		
HLS	✗	✓	H.264	AAC	Burnedin	Closed Captions	WebVTT
ATT	✗	✓	✗	✗	JSON via HTTPS		
EBU-TT	✗	✓	✗	✗	XML via Websocket		

Egress – Traffic exiting across the aiconix network boundary.

Ingress – Traffic entering across the aiconix network boundary.

Ingest – Ingest describes the consumption of incoming traffic for example via an encoder.

Codec – Codec refers to the encoding of a media resource like a video or audio stream.

Transcription – Transcription is the creation of a text that describes what happens in a media resource – like the text of what is being said in a video or audio stream.

Push – In streaming context push refers to the starting of a process from the streaming origin – eg the video encoder. The receiving system is passively waiting.

Pull – In streaming context pull refers to the starting of a process by listening to the streaming origin. The receiving system is actively checking if new input is available.

Closed Caption Subtitles – Subtitles provided as CC are enclosed in the video stream or file as separate data stream.

Burned-In Subtitles – Burned-in subtitles are subtitles that are made part of the video frames. They are "burned" into the video frames.

FPS – FPS means "frames per second" and describes how many frames are provided in a second of video.

Bitrate – The bitrate of a data stream refers to the amount of data provided per second, 3 Mbit/s for example.

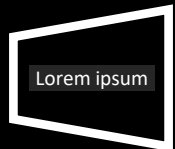**Resolution –** The resolution of an image or videoframe describes how many pixels make up the image.

Source

Live Stream

- The better the audio quality and level the better the result
- Clear sound
- Little to no background noise

Stream format →

- RTMP
- Push to Ingress-address
- Stream-key
- Ingoing resolution = outgoing resolution (25FPS/30FPS/50FPS/60FPS)

aiconix

Live-Transcription

1. Split of Video & Audio
2. Recognition and transcription
3. Re-Merge of Video & Audio

Dictionary

- List of words to improve recognition
- Names, Places
- Needed before stream start

Output

Endpoint

- Pull per Stream-URL
- Push (RTMP) optional
- Total delay 20s + 60s
- Pull per HLS-URL

Post-Stream

- Supply of SRT-file

Stream format ←

- RTMP or HLS
- Pull (Egress-address)
- Stream-key
- Burned in subtitles or
- Embedded closed captions (CC)
  - Subtitles in „608 over 708"
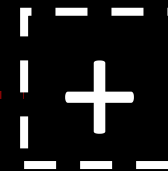 - Embedded in Stream-"Header" SEI user-data

Editor (optional)

- Online Corrections
- + 60s delay

Translations (optional)

- Translation to target languages
- + 5s delay

Options

- Font type
- Font size
- Line height
- Border width
- Max line count
- Background Color
- Outline Color

- Regular Color
- Highlight Color
- Subtitle burn-in (yes/no)
- Closed Captions (yes/no)
- Cc type (Roll-up/Pop-on)
- Max Cc Line Count (0, 1, …)

Example taken from:
https://www.ecb.europa.eu/press/pressconf/html/index.en.html

Limits

FPS:

- 60 interlaced
- 30 progressive

- Resolution: 1920x1080
- Bandwidth: 10Mbit/s
- Transcription languages:
  - English (en)
  - German (de)
  - Spanish (es)
  - French (fr)
  - Dutch (nl)
  - Italian (it)
  - Polish (pl)
  - Turkish (tr)
  - Russian (ru)
  - Swedish (sv)
  - Latvian (lv)

- Translations: 1 language
- Translation languages:
  - Bulgarian
  - Czech
  - Danish
  - German
  - Greek
  - English (default)
  - English (British)
  - English (American)
  - Spanish
  - Estonian
  - Finnish
  - French
  - Hungarian
  - Italian
  - Japanese
  - Lithuanian
  - Latvian
  - Dutch
  - Polish
  - Portuguese (Brazilian)
  - Portuguese (PT)
  - Romanian
  - Russian
  - Slovak
  - Slovenian
  - Swedish
  - Chinese

Live - examples

Example URLs

Endpoint: **endpoint**
Language identifier: **lang**
Steam-Key (unique/UUID):   **streamkey**
Base URL: **url**
Ingress-RTMP-Stream-Address: rtmp://**url**/**endpoint**-**lang**/**streamkey**
Ingress-WebRTC-Stream-Address: webrtc://**url**/**endpoint**-**lang**/**streamkey**
Egress-RTMP-Stream-Address: rtmp://**url**/**endpoint**-**egress**/**streamkey**
Editor-Address: (to be generated from frontend)
SRT-Stream-Address: https://**url**/export/subtitles/**endpoint**-**streamkey**.srt
SRT-Stream-Address for translations: https://**url**/export/subtitles/**endpoint**-**streamkey**-**lang(Translated)**.srt
Egress-HLS-Stream-Address: https://**url**/**endpoint**-**egress**/**streamkey**.m3u8

Example Editor Use

All „608-over-708" compatible players

- Youtube
- Twitch
- Facebook
- Vimeo
- VLC-player

List updated September 2022

CEA-708/EIA-608 embedded in the video elementary stream as described in ATSC A/72 (SEI user_data).
This format is common among television broadcast encoders.

CEA-608 standard characteristics

- Appearance: White text on black background
- Text size: Fixed
- Font: Fixed
- Languages: Primarily Western European languages.

NOT compatible

- iOS, Android, Safari video players (not compatible with RTMP)

Endpoint

URLS:
EndpointName = [abc0-9]
StreamKey = [abc0-9]
LanguageLocale = [de, en, …]

RTMP-Ingest: rtmp://streaming.aiconix.cloud/EndpointName-LanguageLocale/StreamKey
HLS-URL: https://live.v2.aiconix.cloud/EndpointName-egress/StreamKey.m3u8
Subtitle-URL: https://live.v2.aiconix.cloud/EndpointName-egress/StreamKey-sub-en.m3u8
https://live.v2.aiconix.cloud/EndpointName-egress/StreamKey-sub-de-13590.webvtt

WebVTT

WEBVTT

HLS-Manifest

#EXTM3U

#EXT-X-VERSION:3

#EXT-X-MEDIA:TYPE=SUBTITLES,GROUP-ID="subs",NAME="ru",DEFAULT=NO,FORCED=NO,URI="StreamKey-sub-en.m3u8",LANGUAGE="ru"

#EXT-X-MEDIA:TYPE=SUBTITLES,GROUP-ID="subs",NAME="de",DEFAULT=NO,FORCED=NO,URI="StreamKey-sub-de.m3u8",LANGUAGE="de"

#EXT-X-STREAM-INF:BANDWIDTH=2048000,SUBTITLES="subs"

/hls-video/EndpointName-egress/StreamKey.m3u8?hls_ctx=EndpointName-egress

Subtitle-Manifest

#EXTM3U

#EXT-X-VERSION:3

#EXT-X-MEDIA-SEQUENCE:13588

#EXT-X-TARGETDURATION:2.000000

#EXTINF:2.000000, no desc

StreamKey-sub-de-13590.webvtt

….

11

00:46:23.660 --> 00:46:27.349

Lorem Ipsum

12

00:46:27.349 --> 00:46:29.299

Hallo und guten Tag.