TRANSKIT

Automatic **Rich Transcription** Toolkit

Technology for the automatic **rich transcription** of **audiovisual content**

The **Transkit** toolkit generates rich multimedia transcriptions automatically.

Features:

- **Supported languages:**
  Basque, Spanish, English, Catalan
- **Operation modes:**
  - **Live mode** generates transcriptions in real-time with low latency using an audio stream as input.
  - **Batch mode** works off-line and can transcribe previously recorded audio or video.
  - **Dictation mode** lets the user dictate documents through the microphone to obtain text transcriptions automatically in real-time.
- **Adaptation capabilities:**
  - **Different topics and domains**
  - **Particular acoustic conditions**
  - **Specific speaker**
- **Generable metadata:** audio language, audio transcription, background conditions, speaker changes, speakers, sentence units, proper nouns and acronyms.

Applications

- **Multimedia asset management**

Using rich audio transcriptions to generate metadata that can facilitate the search and management of multimedia material.

- **Transcription automation**

The toolkit supports the manual transcription process and make it more productive.

- **Live and pre-recorded subtitling**

Dictation and automatic transcription in live or batch mode are the first steps towards automatic intralingual subtitling.

- **Open-Source Intelligence**

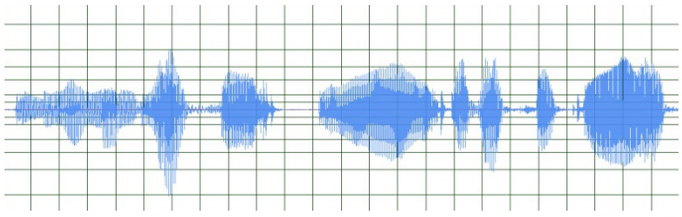
The intelligence collected from publicly available resources benefits from automatically processing the growing amounts of audiovisual material available in the Internet.

- **Speech analytics**

Automatic rich transcription helps analyse the information embedded in recorded customer care calls.

- **Human computer interaction**

The toolkit supports the introduction of speech in the interaction loop.



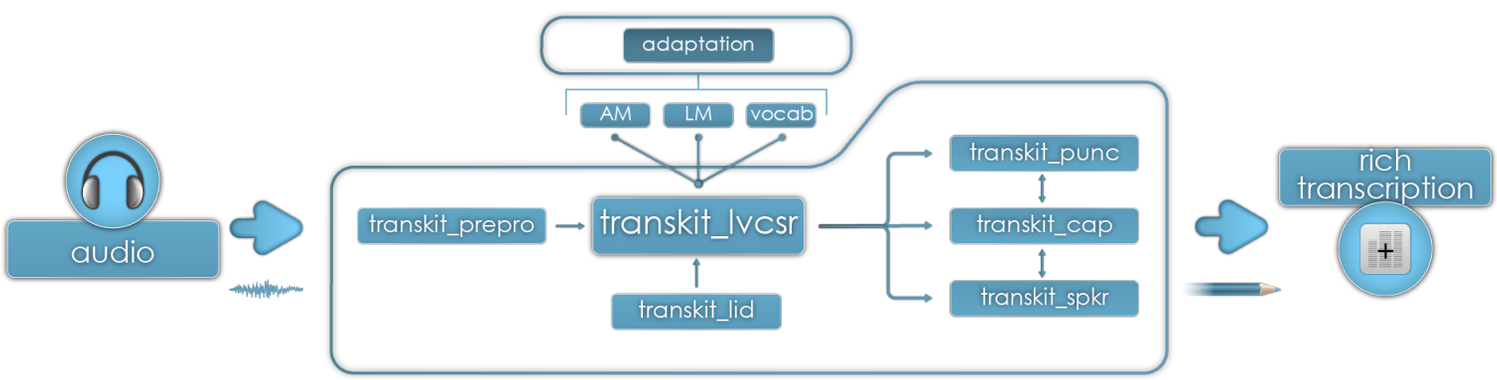
powered by: vicomtech
visual interaction & communication technologies

Transkit Modules and Architecture

Module	Description	Dependencies*	Language	Platforms
transkit_prepro	**Audio Pre-Processing** Normalization, non-speech segments detection and background classification	LIUM_SpkDiarization HTK	Java C	**Machine & Servers** Linux
transkit_lid	**Language Identification** Automatic identification of the language spoken in the audio	LIUM_SpkDiarization	Java	
transkit_lvcsr	**Large Vocabulary Continuous Speech Recognition** Automatic generation of the raw transcription from audio input	KALDI	C++	
transkit_punc	**Automatic Punctuation** Addition of punctuation marks to the raw transcription	Numpy Theano	Python	
transkit_cap	**Automatic Capitalization** Detection and capitalization of named entities in the raw transcription	OpenNLP Moses	Java C++, Perl	
Transkit_spkr	**Speaker Diarization** Automatic segmentation and clustering of the speakers in the audio	LIUM_SpkDiarization	Java	

The **Transkit** toolkit can be deployed on local machines or servers running GNU/Linux operating system

***All transfer processes in Vicomtech-IK4 are performed according to strict procedures that ensure legal control of the final software.**



vicomtech
visual interaction & communication technologies

Contact: **transkit@vicomtech.org**