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

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

#### 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.

### 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.
### Transkit Modules and Architecture

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

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.*