AI-powered Closed Captioning For Media & Entertainment

Industry Challenges

Content accessibility norms are guidelines laid by FCC that the M&E broadcasters must follow. They need to ensure their captions are high quality, consistent, and meet the needs of deaf and hard of hearing audiences. In the US, the FCC stipulates that captions generated should be of high accuracy, synchronous, complete, and have the correct placement on the screen.

However, manual closed captioning operations have multiple challenges.

- **Captioning takes time & effort**: Usually, manual captioning operations could take 10x more duration than the length of its audio files. As a result, captions for recurring TV shows could take up to a week.

- **Requires specialized talent**: Live captioning requires specialized and expensive skills similar to a court stenographer to transcribe with pace and accuracy.

- **Variable costs**: The cost of transcription depends on turnaround time and the volume of content.

Challenges with standard speech-to-text algorithm

- **Standard speech-to-text algorithms have a higher Word Error Rate (WER)**

- **Unable to adapt to business jargons & industry terms**

- **Detection of change of speaker and labelling of individual voices in the video is challenging (Speaker Diarization)**

- **Cannot accurately identify background noise & music**

Our Solution

Quantiphi’s closed-captioning engine is trained on media content for maximum relevance. The model uses active-learning architecture to generalize itself for different kinds of media content - news, recurring TV Shows etc.

Here, near-perfect accuracy is generated within captions utilizing ensemble models, Natural Language Processing and human-in-the-loop curation.
Features

- Reduced word error rate (WER)
- Auto-punctuation & capitalization
- Speaker Diarization
- ‘Human-in-the-loop’ architecture
- Direct transcription of non-native media audio (German, Italian, Japanese, Mandarin)

Business Impact

- Fast, accurate and scalable closed captioning operations.
- Content accessibility compliance (FCC)
- Improved working experiences for captioning teams & transcribers.
- Fraction of existing captioning cost

Case Study: Live captioning for a US News Broadcaster

Customer

US Broadcasting Major with eleven news broadcasting studios

Challenge

The customer wanted real-time transcriptions on their live broadcasted news feeds on the web and pre-recorded video news segments.

Live streaming still relied on old technologies, such as teleprompters, encoders, and other hardwired devices, prone to glitches and transcription errors.

Our Solution

Quantiphi developed a scalable web-based user interface (UI) which consisted of closed captions obtained using speech-to-Text API synced with the video feed. This enabled near real-time captioning and viewing of those captions on live news broadcasts.

Business Impact

- Complete transcriptions for recorded and live stream videos on their web-based portal.
- Over 85% transcription accuracy with no glitches or breaks in transcription.