



## Say What? AI-Driven Voice-to-Text Transcription

Who would have predicted at the beginning of 2020 that the phrase “Zoom fatigue” would be a trending search on Google? As we have all become proficient in various videoconferencing platforms, we have moved many in-person conversations and presentations to the small—sometimes very small—screen. With that transition has come an expectation that virtual meetings and events will be recorded for later viewing. And yet, be honest: When is the last time you actually watched the entire recording of an event you couldn’t attend live?

All that changed with the advent of powerful, affordable, AI-driven voice dictation services. Zoom and Google Slides both offer free live auto-transcription, and, having watched both in action during recent virtual conferences, I think they are surprisingly good. Not only do they make presentations accessible to those with hearing loss or anyone needing to keep their audio turned off, they also generate a decent transcript. Other free and inexpensive audio-to-text automatic transcription services such as Otter.ai, Temi (temi.com), and Rev (rev.com) are expanding access to this technology to anyone who wants an audio file converted to text.

Combine affordable speech-to-text and the explosion of recorded audio and video content, and info pros finally have a way of making non-print material searchable, accessible, and sharable. When I am conducting research on an organization or an individual, I often find recordings of presentations. Until recently, I ignored these files, as I knew my clients would not have the time or patience to watch a 30-minute presentation, nor would they want to pay me to watch it. Even if there was useful information buried in that MP4 file, it was not cost-effectively accessible; audio and video content is difficult to skim or navigate through.

The implications of this new ability to transform audio into text are wide ranging. Now, when I find a potentially useful presentation or other recording, I have it auto-transcribed—Otter.ai is my current fave for accuracy and cost—and, in a fraction of the time it would take to listen to the audio, I can skim the transcript for useful nuggets of information. I include the snippet of transcription in my report to my client, along with a pointer to the recording and where in the timeline the relevant information appears. (I will pause here to note that I am not a lawyer, much less an intellectual property lawyer. Whether these transcriptions fall within fair use is an unresolved question, so check with your IP compliance folks.)

Realizing how easy it was for me to take advantage of this newly accessible technology, I started thinking of ways that info pros can leverage automated transcription to enhance

information discoverability and information services. Here are just a few ideas that come to mind:

- Info pros can identify podcasts that focus on new developments relevant to their industry. As each new issue drops, the audio file can be sent for auto-transcription and the transcript distributed to interested groups along with a link to the podcast itself.
- Libraries with spoken-word collections, such as oral histories, speeches, or interviews with people of note, can now create searchable transcripts of these resources, making them far more discoverable and accessible, both to researchers and those with hearing challenges.
- Libraries are creating more short video tutorials, knowing that many users default to learning something by finding a YouTube video on the topic. By adding auto-generated closed captioning to the videos, the value of the tutorials is enhanced, making them easier to understand and viewable without the audio playing.
- Reference librarians sometimes struggle to catch all the nuances of a research request, particularly in an unfamiliar area. If the reference interview is via videoconference and the patron agrees, recording the conversation and then transcribing it provides an easy way for the researcher to go back and see if there were any additional subtleties to consider.

What may be most transformative about AI-supported transcription is that it will now be possible for aggregators to provide access to spoken material. Imagine if news broadcasts or government proceedings could be cost-effectively auto-transcribed and quickly made searchable. Newspapers that provide multimedia stories can finally make that content findable by providing a transcript of the audio. Speeches and presentations from marginalized populations, which may not get coverage from traditional press outlets, can be uploaded to the web for wider discoverability. (Note that this is a double-edged sword, as it also can make inflammatory rhetoric and misinformation more accessible when the content is more easily findable.)

Info pros often bemoan the challenges of searching grey literature—all those resources outside traditional publishing and distribution channels. AI-supported transcription may finally make some of the multimedia grey literature more accessible.

---

*Mary Ellen Bates (mbates@BatesInfo.com, Reluctant-Entrepreneur.com) is seeking an auto-transcription from dog to human.*

*Comments? Email the editor-in-chief (marydee@xmission.com).*