

Mini-Project Proposal

Transcript Magic: Improving Transcription for Journalists (and Others)

Business Context

In journalism both accuracy and speed are essential. Journalists generally work under tight deadlines, sometimes competing to break stories first and often juggling the competing demands of busy editorial calendars. Despite these sometimes intense time pressures, however, journalists must uphold the highest standards of accuracy. To help prevent errors, many journalists will record their interviews or refer to previously recorded video and audio sources when researching and writing stories. Audio and video recordings help ensure accurate quotations, aid the process of writing, and document interviews for future reference, as needed.

Problem Statement

Audio and video recordings can help journalists write accurate stories but the most useful format for a journalist writing an article or book is text, especially if the interview is longer than a few minutes. It is much easier for a journalist to quickly scan a transcript of an interview to find a quote or specific topic rather than fast forward, skip around, or replay a recording to find what they seek. Therefore, journalists will frequently transcribe recordings. This is a time-consuming process for the journalist, or expensive, if they use a transcription service. Therefore, journalists commonly turn to machine transcription, using automatically generated transcripts from Zoom calls or from dedicated machine transcription services. Depending on the method of transcription, these transcripts are error-prone and poorly formatted, sometimes requiring manual review and correction.

Proposed Solution

I propose using the OpenAI API to correct transcripts or produce new transcripts that are more accurate and less expensive than other solutions. The tool would send text to one of OpenAI's models and return corrected text, or send an audio file and return a transcription (note, this is more experimental—the speech-to-text Whisper model is listed as “beta”—and I might find that it is not be feasible for the final product). A comparison of potential costs is provided below and shows that an OpenAI tool could provide superior cost efficiency.

- Transcription Services Recommended by the New York Times Wirecutter (<https://www.nytimes.com/wirecutter/reviews/best-transcription-services/>)
 - o Temi machine transcription: \$0.25 per minute
 - o GoTranscript human transcription: \$0.90 per minute
- OpenAI Services
 - o Whisper (beta): \$0.006 per minute
 - o text-davinci-edit-001 (alpha): currently free
 - o gpt-3.5-turbo: \$0.002 per 1K tokens (~750 words)

Note, that this tool would also have uses beyond journalism. Other organizations would also find it useful to more easily correct or produce transcripts of recordings, especially given the

increasing prevalence of recorded Zoom meetings. For example, nonprofits could use the tool for internal information sharing or to facilitate compliance with IRS regulations and open meeting laws by more easily creating records of Board meetings. The tool could also potentially be enhanced to provide useful summaries of recorded meetings and interviews.

i've had the privilege of knowing ting chang since 2000 when i became her colleague at mcgill university and it's
 2:05
 wonderful to have been friends since then she teaches now at the university of nottingham
 2:10
 she's published extensively on gustav kobe and on collecting
 2:16
 the latter is the focus of her wonderful book travel collecting and museums of
 2:21
 asian art in 19th century paris her talk today presents research from a
 2:26
 chapter that will be coming out in the making world's global invention in the early modern period
 2:33
 a volume that's the result of the core program making worlds that was held three years ago at the clark library
 2:40
 the title of her talk today is de monde games maps and world conquest in early

YouTube-generated transcript.
 (<https://www.youtube.com/watch?v=MyoLo4nz0t4>)

I've had the privilege of knowing Ting Cheng since 2000 when I became her colleague at McGill University, and it's wonderful to have been friends since then. She teaches now at the University of Nottingham. She's published extensively on Gustav Klimt and on collecting. The latter is the focus of her wonderful book, Travel, Collecting, and Museums of Asian Art in 19th Century Paris. Her talk today presents research from a chapter that will be coming out in the Making Worlds: Global Invention in the Early Modern Period, a volume that's the result of the core program Making Worlds that was held three years ago at the Clark Library. The title of her talk today is "De Monde, Games, Maps, and World Conquest in Early Modern France." So welcome, Ting.

Transcript corrected by ChatGPT

1:46 Introduces each speaker briefly.
 2:05 Introduces Ting Chang as a colleague since 2000 when she became her colleague at McGill University. Notes Ting teaches now at the University of Nottingham and has published extensively on Gustav Klimt and on collecting. Also mentions Ting's book "Travel Collecting and Museums of Asian Art in 19th Century Paris."
 2:48 Introduces the title of Ting's talk, "De Monde Games, Maps, and World Conquest in Early Modern France."

Transcript summarized by ChatGPT

Outputs

1. At a minimum, a Python program that uses the OpenAI API to correct transcripts.
2. Possible enhanced features include:
 - a. A speech-to-text feature that provides users with the ability to upload an audio recording to an OpenAI model and receive a transcript.
 - b. A summarization feature that produces a useful, short summary of an uploaded transcript.

Risks and Challenges

The first (and perhaps obvious) risk is that the OpenAI models will likely still produce transcripts with errors. As can be seen in the above sample images from a test with ChatGPT, the corrected transcript is not perfect, erroneously ‘correcting’ “gustav kobe” to “Gustav Klimt,” although it should be noted that *neither* are correct; the person referenced was the artist Gustave Courbet. A related risk is that the model will not only make incorrect edits but will introduce new errors or information. The intention, however, is not to produce a perfect transcript but to produce a better transcript with superior time and cost efficiency. A more serious risk is information security and privacy. Documentation of the tool will warn journalists (and others) that it should not be used for any sensitive materials as OpenAI not only stores inputs for a period of time but also uses inputs to refine its models.