

## Daily Log

### Monday October 28

Began to Go Through the CSV File of TED Talks and Downloaded and Transcribed Each Audio File. Had to Listen and Find the Length of Introduction to Ignore That Section in the Audio File.

### Tuesday October 29

Decided to Organize Transcripts in a Google Sheets File. Continued to Go Through the TED Talks and Repeat the Process. Started From Bottom of CSV File, to Make Sure Partner and I Did not Overlap TED Talks.

### Thursday October 31

Continued Going Through CSV File and Transcribing Different Talks. Removed a Few Talks if Speaker Did Not Enunciate Words Well or Talked too Quickly. Each Translation Took About 5 Minutes and The Whole Process for Each Took Around 9 Minutes.

## Timeline

Date	Goal	Met
Oct 18	Get Results from IBM Watson Speech to Text for all TED Talks	No, Had Trouble With the Size of Audio File Limit
Oct 25	Research Possible Text Summarization APIs and Methods	No, Got a Good Result From One TED Talk
Nov 1	Get Results from IBM Watson Speech to Text for Multiple TED Talks	Yes, Got Transcripts for Around 40 Different TED Talks
Nov 8	Research Possible Text Summarization APIs and Methods	
Nov 15	Continue to Research Possible Text Summarization APIs and Methods	

## Reflection

It was somewhat boring to repeat this process for many TED Talks. However, once I developed a system it became somewhat quicker. I decided to remove a few transcripts if the speaker had a thick accent, did not enunciate words well, or spoke too quickly. However, this was only a handful of the many TED Talks. We made good progress throughout class and over the weekend and we got around 40 excellent transcripts. We believe this is enough to begin looking into text summarization APIs and methods and will begin researching next week. While researching text summarization we still hope to continue updating our database of transcripts.