Metadata Works From Home

How Student Workers Continued to Work While Increasing Collection Accessibility

Emily Vinson evinson@uh.edu
Andrew Weidner ajweidner@uh.edu
University of Houston Libraries

Project Overview

Project Goals

- Keep students employed and engaged in Library work
- Improve accessibility to AV collections for deaf and hard of hearing users

Project Background

Transcription Accuracy and Cost Study

- Internal grant awarded late 2018
- Examined accuracy of various transcription methods

Transcription Accuracy and Cost Study

Methods

- Professional transcriptionist services
- Online Artificial Intelligence-based services
- Free "hacks"
- Untrained study participants

Transcription Accuracy and Cost Study

Sample Videos

- Single speaker, clear audio, scripted speech

- Crosstalk and off-mic speech

- Multiple speakers, off-mic and unclear speech

- Bilingual scripted speech

- Background music, repetitive language

Transcription Accuracy and Cost Study

Average Error Rate and Cost
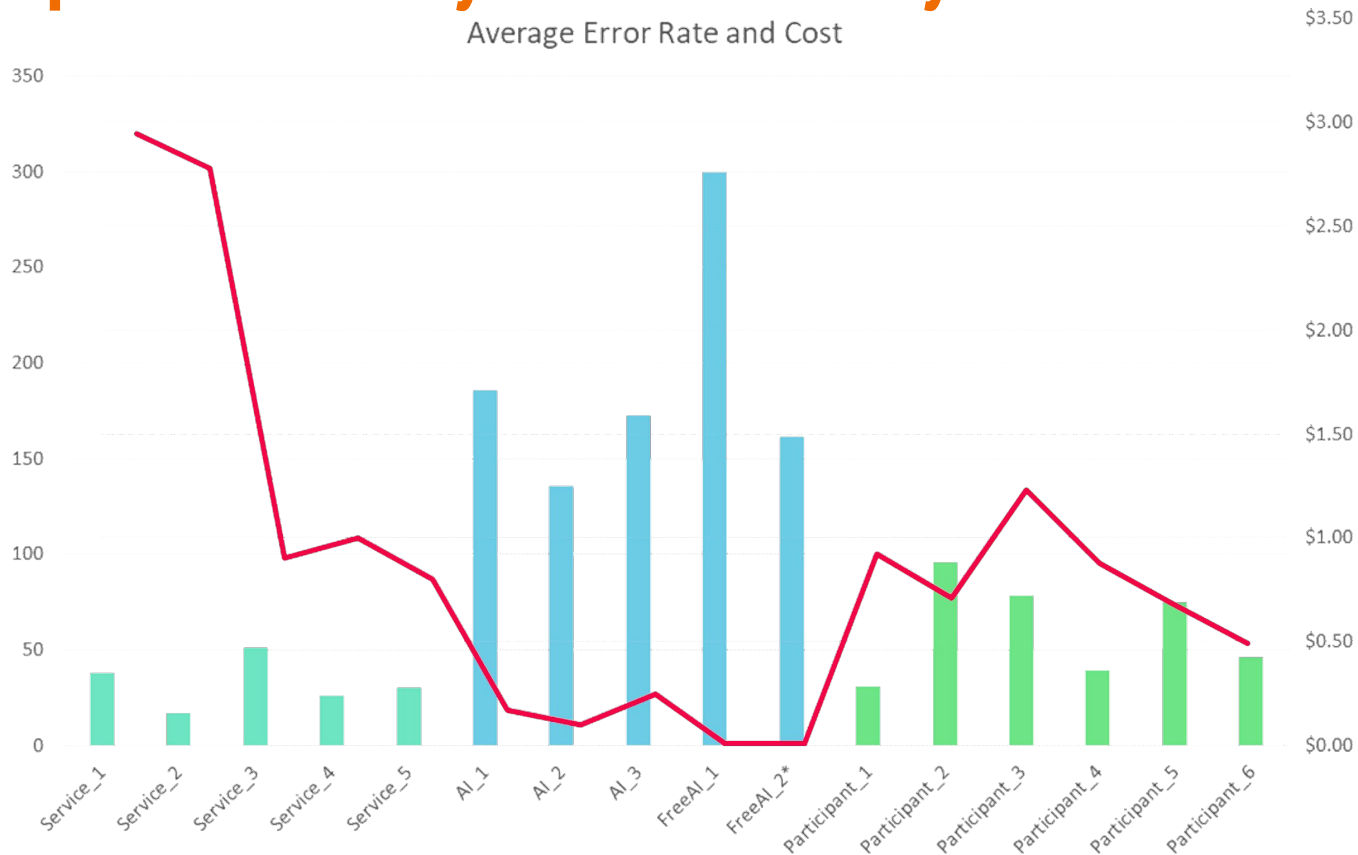
Basics of Closed Captions

Timestamp in hh:mm:ss:ttt format

Line length and line breaks will be reflected in caption formatting


```
12
00:00:52.440 --> 00:00:58.440
sit tight and not move. They
decided if one person moved, it
would be a big panic.

14
00:00:58.440 --> 00:01:02.690
So everyone said "Okay, we
won't sell our houses. We'll
just stay right here,"

16
00:01:02.690 --> 00:01:06.680
and that was kind of what they
did.
Interviewer: Okay, cut.

18
00:01:06.680 --> 00:01:09.960
Man 1: Marker.
Man 2: Two mark.
Interviewer: How did your uh,
```

Project Timeline

March 16:
UH Classes Go Online; Most Library Departments Transition to Remote Work

March 24:
Harris County Stay Home Order; All Library Depts Transition to Remote Work

March 18:
Libraries Admin Approves Project

March 26 & 27:
Zoom Training with Students

April 10:
Zoom Training with QC Students

Transcript Correction Work Underway

March 19:
First Project Managers Mtg

March 23:
Second Project Managers Mtg

Quality Control Work Underway

WEEK 1   **WEEK 2**   **WEEK 3**   **WEEK 4**

Communication

Supervisor

Project Managers

Supervisor

Students

Students
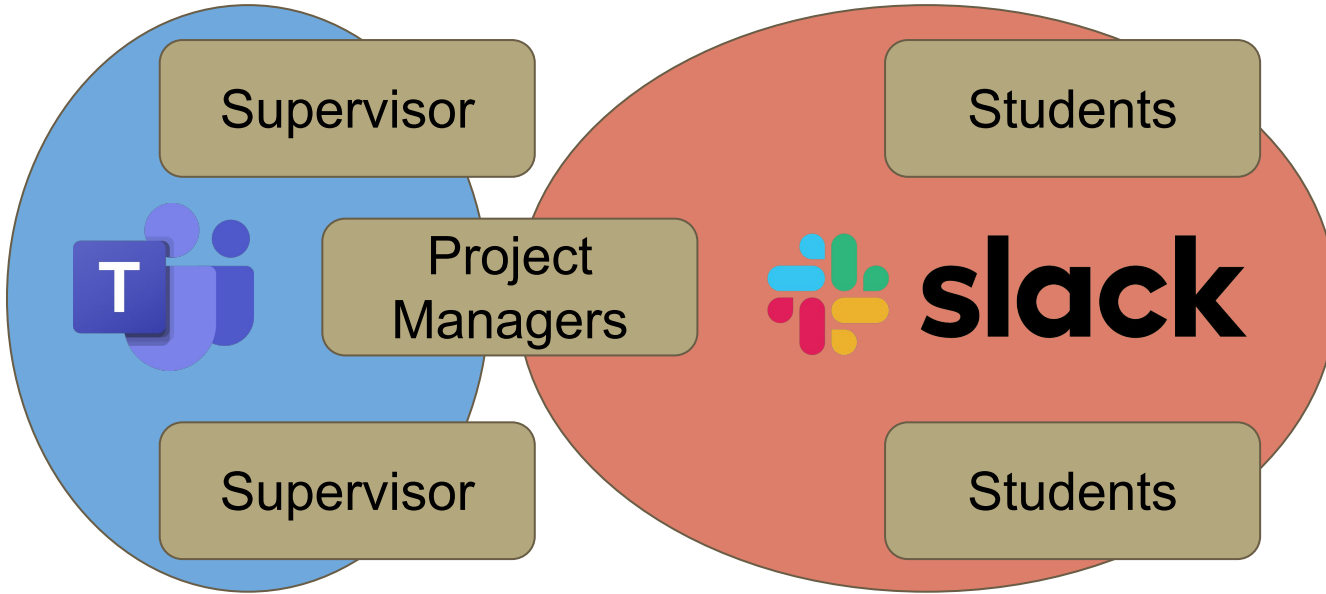
Zoom

Email

Google Calendar

Google Drive

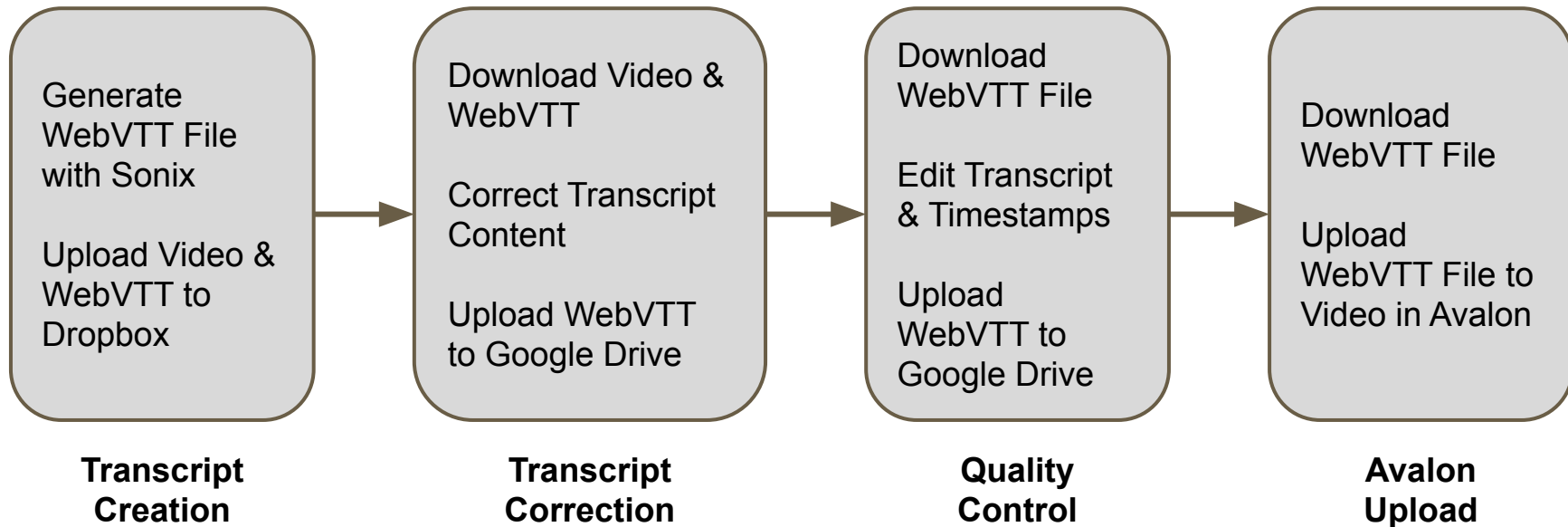# Workflow Tools

Sonix: Transcript Generation

Dropbox: Video & Transcript Distribution

Google Drive: Transcript Versioning & Progress Tracking

VLC Media Player: Video Playback

Atom: Transcript Editing

Workflow Overview

Generate WebVTT File with Sonix

Upload Video & WebVTT to Dropbox

Transcript Creation

Download Video & WebVTT

Correct Transcript Content

Upload WebVTT to Google Drive

Transcript Correction

Download WebVTT File

Edit Transcript & Timestamps

Upload WebVTT to Google Drive

Quality Control

Download WebVTT File

Upload WebVTT File to Video in Avalon

Avalon Upload

Project Successes

- Phase 1 caption correction complete for over 30% of all online videos

- Raised awareness of accessibility issues within the UH Libraries and among student employees

- Kept ~30 students employed remotely

- Developed effective workflow to continue creating accessible video content