TRANSCRIBING TV NEWS

GOAL:
To eliminate the CBS, NBC and CNN backlog of news programs and improve the searchability of our collection by transcribing broadcast news shows

BETWEEN APRIL AND AUGUST OF 2020, 847 NEW TV NEWS EPISODES WERE TRANSCRIPTED AND CAPTIONED.

This summer project utilized Trint, a speech-to-text platform that uses A.I. to transcribe audio and video.

The summer team uploaded videos of news broadcasts to Trint and then separated the transcripts into separate segments, including the preview, news stories, commercials and the good-night segment.

The transcripts were quality checked and then exported as .srt and .xml files. The files are being uploaded to the VTNA database.
OUTCOMES

The summer project team overcame several obstacles including multiple personnel changes and the switch to remote work due to COVID-19. The number of team members varied from as many as 20 to as few as 2. Still, over 847 new TV news episodes were added to the collection. VTNA was able to provide meaningful work to staff members who would not otherwise have been able to work remotely.

NEXT STEPS

The transcription program has been ongoing since January 2019 and will continue as VTNA’s primary method of adding new content to our database. Several of the summer team members have committed to continuing their work with the program. Additionally, 5 students were added to the Trint team. A new uploading task was added so staff members could continue working.