

# Summary of Data Collection, Processing, and Analysis Principles

## Analyzing learning with speech analytics and computer vision methods: Technologies, principles, and ethics

### Recommendations Related to Audio for Speech Analytics

#### *Data Collection*

- Aim to collect speech data at a resolution beyond what a human is able to easily notice, such as WAV at 24 bits
- Be cognizant of background noise – noise you may not notice can be more prominent in recordings
- Do not use features like auto leveling or equalizing so that differences in speech volume are preserved
- Use the recording equipment that captures the particular type of audio you are interested in (e.g., close-talking mics for single speakers, table mics for small groups, microphone arrays for larger groups)

#### *Data Processing and Analysis*

- Keep original/raw files after processing whenever feasible: this allows for the greatest flexibility with respect to analysis
- Maintain separate audio sources for individuals or groups when feasible
- Be careful using built-in audio enhancements like removing background noise or equalizing
- Existing techniques: Praat, HTK, OpenEar, Covarep, ELAN, ICSI Diarizer, LIUM Diarizer, CMU Sphinx, Google ASR, AT&T Watson ASR, Bing Speech API, Audacity, Emovoice, CMUSphinx, WebRTC

### Recommendations Related to Video for Computer Vision

#### *Data Collection*

- Use the maximum size (or resolution) of video and bitrate (or lowest CRF) when feasible
- Record in H.264 if available
- Avoid inexpensive cameras that use frame duplication
- Position cameras to avoid backlighting, when feasible, to preserve contrast
- Position cameras to reduce occlusion
- Use wide fields of view to reduce the need for re-identification (e.g. with wide angle lenses)
- Keep cameras stationary and do not pan or zoom

#### *Data Processing and Analysis*

- Re-encoding video should be avoided when feasible; the best output format to use is whatever format your camera outputs
- If re-encoding, keep the original size, resolution, and frame rate if feasible
- Keep original/raw files after processing whenever feasible
- Existing techniques: [OpenPose](#), [EduSense](#), [OpenFace](#), and [FaceReader](#)