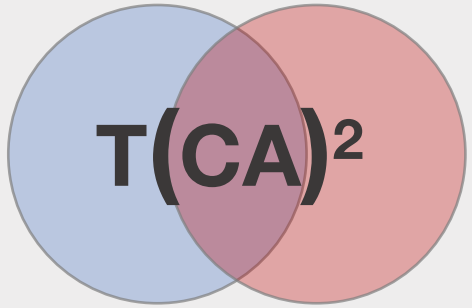
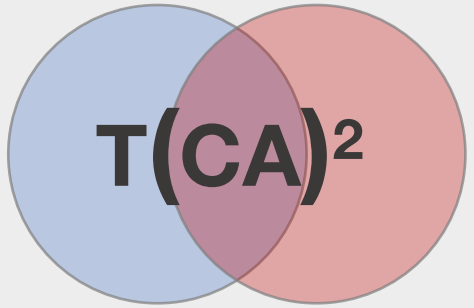
Audiovisual Data Collection Principles



“New” data collection principles

- Existing audiovisual data collection principles exist (e.g., Derry et al., 2010; Goldman et al., 2007), but these principles assume that a person is watching & analyzing the data
- Often “good enough” for a person is not sufficient for successfully using computational techniques



A bit about video formats

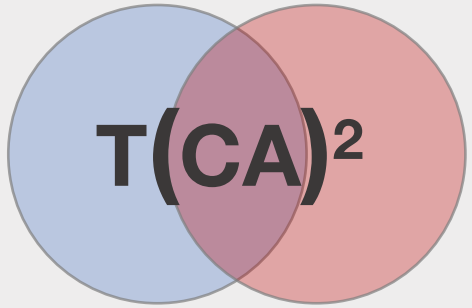
Container (e.g., AVI – Audio Video Interleave)

- Specifies metadata about the video
- How many streams are there? What format are they?
- How long is the video?

Stream:

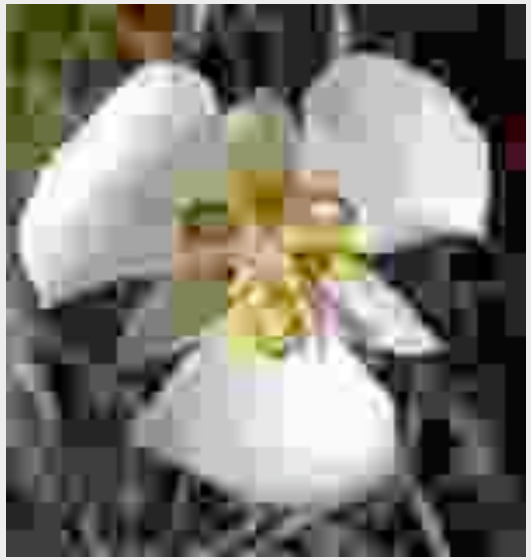
- Actual video/audio data
- Common video formats like H.264, MPEG-2, VP9, etc.

Frame rate (frames per second; FPS)

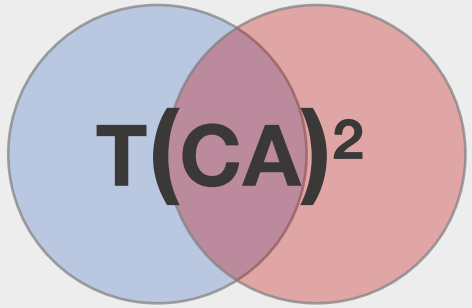


Bitrates for video streams

- *Compression:* ways that video data are abbreviated to reduce size, usually incurring some small distortion to the video
- *Bitrate:* how much data is budgeted per second to store the video



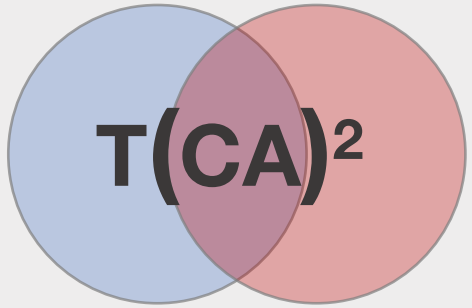
From: https://en.wikipedia.org/wiki/Compression_artifact



Size (resolution) of video streams

- Common sizes
  - 4K (4096x2160)
  - 1080p (1920x1080)
  - 720p (1280x720)
  - DVD (720x480)
  - VGA (640x480)

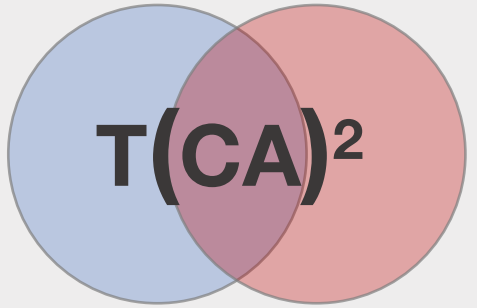
- Best not to change from your original source when possible



Frame rate (FPS)

- Common values
  - 24 (common film rate)
  - 23.976 (NTSC TV standard), = 24 * 999 / 1000
  - 25 (PAL TV standard)
  - 30
  - 29.97 (NTSC), = 30 * 999 / 1000
  - 60

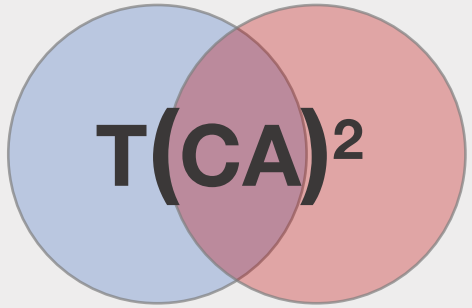
- This is also something to avoid changing when possible
- Pay attention to rounding errors!



Choosing resolution + bitrate

- Resolution
  - If you already have data, keep the resolution whatever it is (if possible)
  - Otherwise, do test recordings at the max resolution and see how well your hardware/software handles it, then back down if necessary

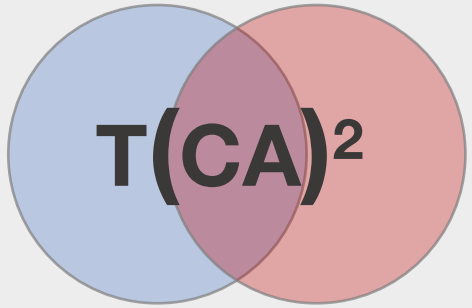
- Bitrate
  - Very difficult to select, which has led to constant rate factor (CRF) options that try to manage this automatically
    - With CRF, lower numbers are higher quality; usually 18 is considered a reasonable high-quality setting (somewhat software-dependent)



What are good formats to use?

- Whatever format your camera outputs

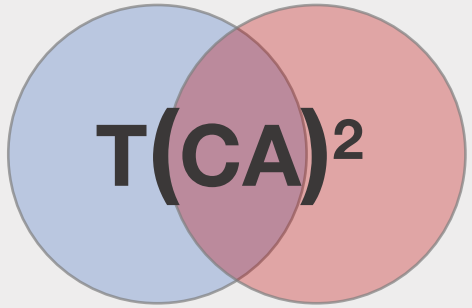
- H.264 is one of the most standard stream formats, and there is good free software to encode it
  - Most often used in an MP4 container format
  - Has many options, including for making videos that are compatible with most web browsers



Converting and editing video

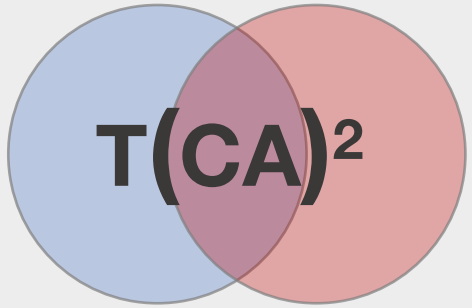
- Important keywords:
  - *Encode/re-encode:* video data are converted to images, then re-compressed (copy of a copy)
  - *Pass-through:* data are not modified

- Re-encoding video should be avoided when possible
  - Changing the container format but not the stream
  - Cutting
  - *When it's not possible:* resizing, changing the stream format
  - And keep the originals!



Hardware/physical considerations

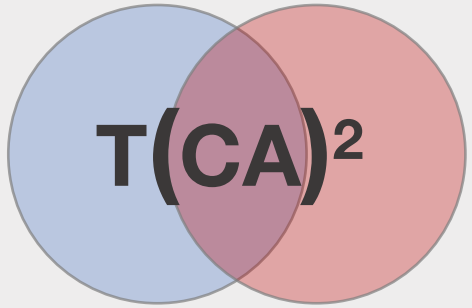
- Lighting
- Frame duplication (cheaper cameras)
- Time synchronization
  - Internet/computer-connected cameras
  - Physical options (recording a precise clock; clapping)



Institutional review board considerations

- Four common levels of consent for video:
  - None
  - Project team access only
  - Access for research or educational purposes (e.g., presentations)
  - Public access (“media release”)

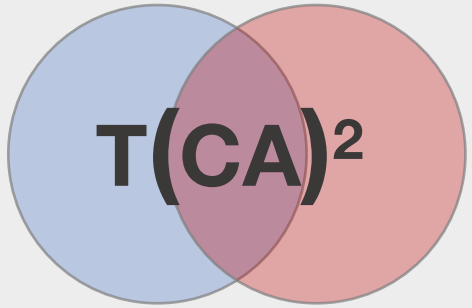
- Privacy-informed camera positioning is occasionally necessary



Audio Data Collection

- Main considerations:
  - Microphones
  - Audio settings and formats
  - Placement and setting

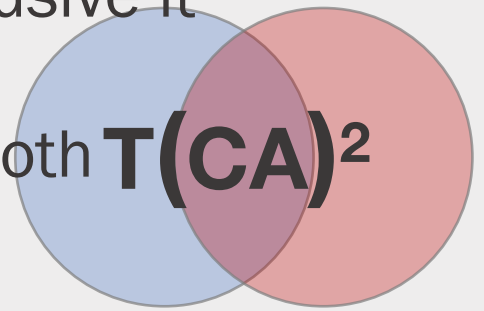
- Lots of questions to consider about what data you want to collect, how much you're collecting, and where you're collecting it.



Types of microphones

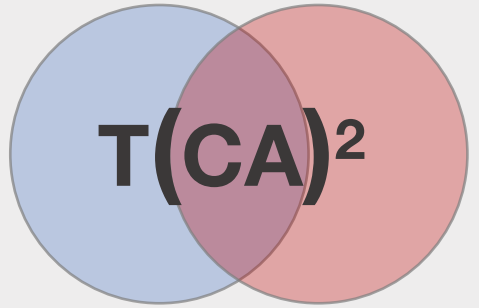
- Lapel mics
- Close-talking headworn mics
- Table mics
- Microphone arrays
- Microphones embedded in phones/tablets/computers
- Microphones that are connected to a video camera
- Special microphone devices (e.g., LENA)

- Consider the amount of space the mic takes up and how obtrusive it might be for the students
- The right microphone will dramatically impact your analysis (both what is possible to do and how well it does it)



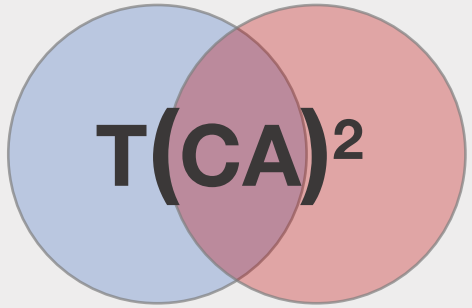
Where to put your microphone?

- What kind of microphone is it?
- What audio do you want to capture?
  - One person, small group, whole class?
- What audio do you NOT want to capture?
  - How much background noise is there?
- Is the person/people going to move their body or turn their head a lot? Will they walk around the room?
- What kind of audio recorder are you using? (and where is that going?)
- Do you need to plug it (or the recorder) in?



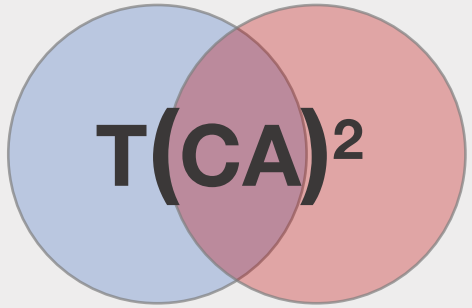
Wired or Wireless?

- Wireless seems like a great option. But… there are a lot of limitations with it as well.
- How many mics will you have recording at one time?
  - There is a limit to how much bandwidth is available and how much data you can record at once.



Batteries or power cords? Or both?

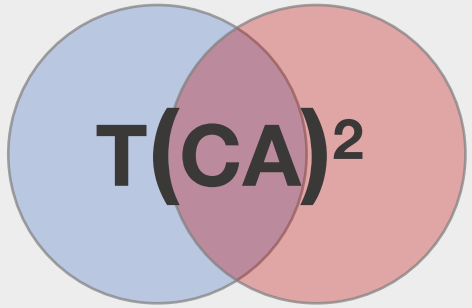
- The audio recorder needs power.
- If you're recording for a whole class period you might not want to rely on batteries.
- Battery back-up is always good to have, just in case.
- Need to think about where the power outlets are in the classroom and whether or not you need extension cords (and to consider if you're worried about children tripping on the cords)



Do you need to synchronize?

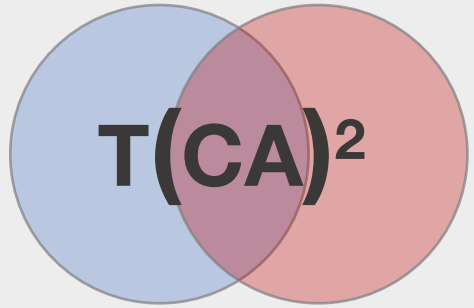
- All channels on one recorder will be on the same clock
- Across multiple recorders?
- Across different data streams?

- Clapboards, audio signal, software



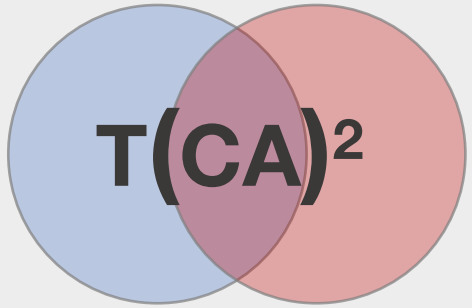
Audio formats and recording rates

- You almost certainly want to record in an uncompressed format like .wav
- Sampling rate:
  - Number of samples collected per second
  - 48 kHz is good
  - This is now a fairly standard rate for professional grade equipment
  - If you're focused on human speech, you don't need to go higher than this
- Bit depth
  - How much information is in each sample
  - 24 bits is good



IRB considerations

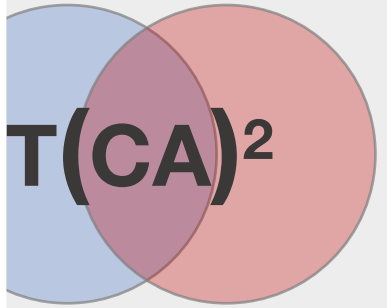
- Privacy of students
- Will students know when they are being recorded?
- Some newer systems are able to create metrics in close to real-time and store those so that the raw data is never saved.
- Analysis methods to preserve privacy
- Do you need to do speaker registration?



Hardware recommendations

- Close-talking microphone:
  - Good for capturing individual speaker and blocking out nearby speakers and background noise
  - Wired (XLR) and wireless versions
  - https://www.amazon.com/Audio-Technica-Hypercardioid-Dynamic-Headworn-Microphone/dp/B0002H068M

- Audio recorder:
  - Built-in stereo microphones, four XLR inputs
  - https://www.amazon.com/Zoom-H6-Six-Track-Portable-Recorder/dp/B00DFU9BRK/



Summary

- Don’t base decisions about audio & video quality on what you can distinguish; use highest/recommended bit rates
- Keep original files and avoid changing the format or re-encoding
- Use equipment and position it to best capture what you’re interested in
  - Preserve visual contrast & reduce occlusion
  - Preserve speech volume differences & avoid background noise
-