

Creating HAVIC: Heterogeneous Audio Visual Internet Collection

Stephanie Strassel¹, Amanda Morris¹, Jonathan Fiscus², Christopher Caruso¹, Haejoong Lee¹, Paul Over², James Fiumara¹, Barbara Shaw², Brian Antonishek², Martial Michel³

HAVIC corpus overview

- LDC, with NIST, is building a large multimodal corpus of videos
- Corpus unconstrained in terms of
 - Concurrent speech, multiple languages
 - Video resolution
 - Camera motion
 - Subject topicality
- HAVIC corpus supports TRECVID Multimedia Event Detection (MED) evaluation
- Selected 75 MED events, like "Making a cake" and "Batting in a run"
- HAVIC corpus will contain thousands of hours of annotated videos:
 - Event instances - videos illustrating a MED event
 - Background clips - videos that don't illustrate a MED event
- To date 50,647 clips released to MED participants
- Total duration 1,620 hours
- Corpus slated for publication in LDC's general catalog

Example Event Kit

Event name: Getting a vehicle unstuck

Definition: One or more people work to free a motorized or unmotorized vehicle that is stuck.

Explication: A stuck vehicle is one that either cannot move, or can only move in a limited area (e.g. a parking space). Usually movement is restricted either because the vehicle is located on/in a substance that prevents the wheels of the vehicle from having enough friction to propel the vehicle forward (e.g. mud, mud covered in water, snow, etc.), or because the angle of the vehicle and/or surrounding obstacles prevent it from moving normally. Getting a vehicle unstuck can be done by maneuvering the vehicle itself or by using another device (another vehicle, an object, etc.) with the intention of getting the vehicle to a location or angle where it can again move normally.

Evidential description

scene: typically outdoors, day or night, any weather

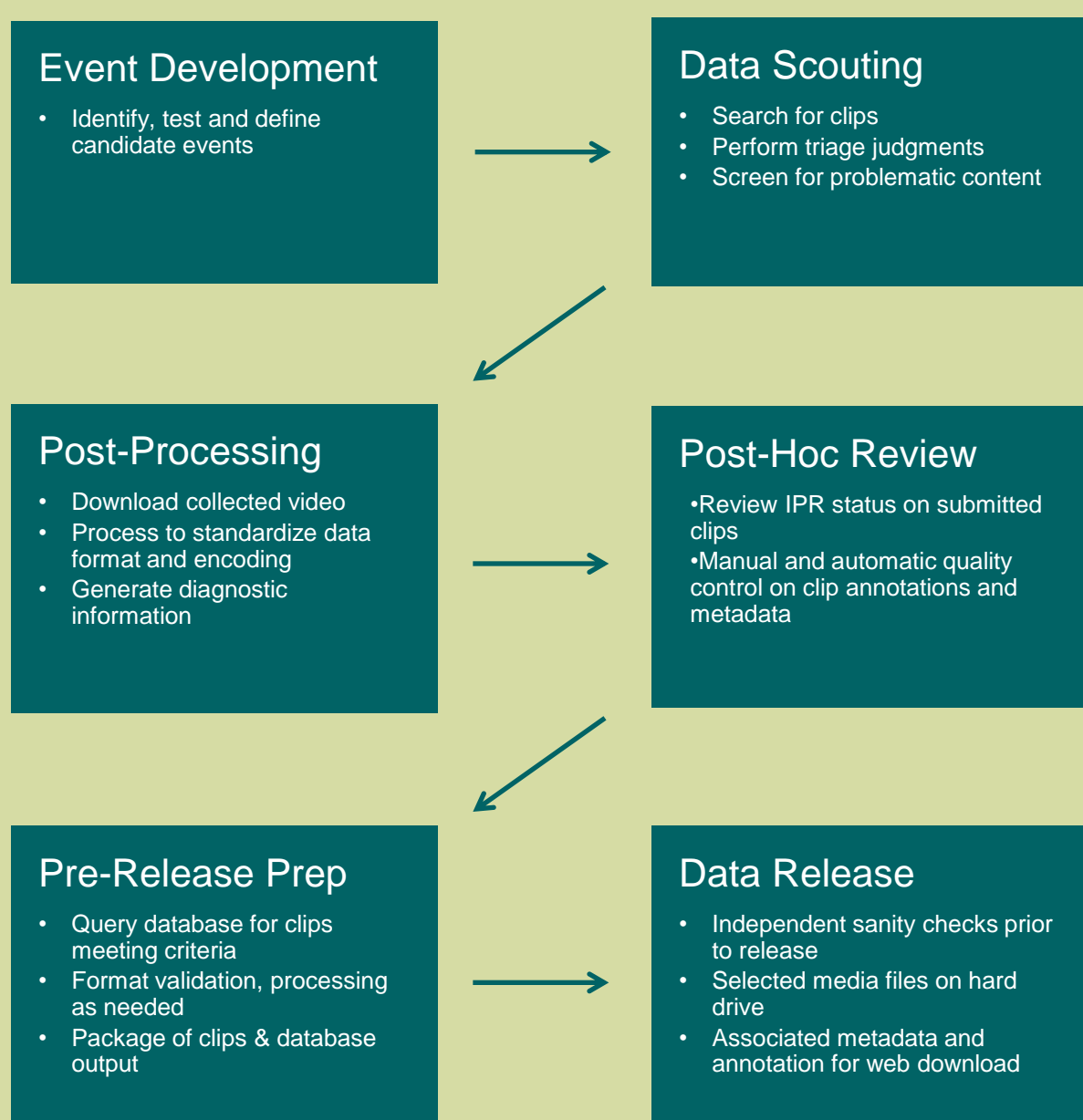
objects/people: vehicle (car, boat, bicycle, construction vehicle), person operating the vehicle, possibly other people assisting in or outside of the vehicle, possibly additional vehicle(s) assisting

activities: steering, hitching, chaining, driving, sliding, skidding, pulling

audio: vehicle tires spinning against surface; engine noise from vehicle; narration or commentary on process from participants



Data Collection Pipeline



Annotation

AScout – LDC-designed GUI

Duplicate URL checker

- Annotators use AScout to guide video searching and perform triage
- Each video is screened for objectionable content
- Required annotations :
 - Clip URL
 - Genre
 - Synopsis
 - Topic
 - License

Additional annotations required for event instances



Event Kits

We write an event kit for each MED event consisting of:

- Event definition
- Explication detailing key concepts
- Evidential description listing some audio and visual elements likely to be seen in an event instance

Corpus Challenges

- Remote annotation team
- Quality Control
 - Not all clips can be second-passed (2P)
 - Event instance videos high priority for 2P
 - 2P verifies/corrects annotation accuracy
 - Annotation database reviewed manually for potential inconsistencies
- Corpus variety
 - Variety important but difficult to measure in a large corpus
 - Data scouts search out variety along many parameters
 - Scene, number of people, objects, activities, language
 - LDC utilizes WordNet and annotation games to introduce additional variety