Aladdin Event Detection Problem

• Find 15 target events in open source video:
  – Examples – Birthday, Wedding, Grooming an animal etc.
• Our Focus is on Audio
• Can Directly train classifiers for target event
• But ...
• Much of the audio is irrelevant to the event
• Not sufficient to just detect event, need to “recount” it
• Thus, need to detect sub-events (or concepts) that make up the event
Audio Concepts and Event Recognition

- Challenges
  - Event Level
    - Large number of classes and high diversity within Class
    - Weak correlation between audio and event, e.g. music and speech often irrelevant
  - Concept Level
    - Key audio concepts may not be known priori and need to be discovered
    - Identification of proper training set and generating annotation might be hard

- Proposed approach
  - Explore alternatives for each of above boxes e.g. Learn Dictionary with Random Forest + BOW, Other Agglomerative Clustering, etc.
  - Develop classifiers for Discovered and manually chosen concepts
  - Develop Event Classifiers starting with current audio content
    - Current: GMM’s, Random Forest + BOW
    - Going Forward - Refine training set to remove “irrelevant” audio
Audio Concept Detection

- Detectors of 20 audio concepts with enough annotated examples have been developed.