

# An Audiovisual Metadata Platform to Support Mass Description



Jon Dunn, Indiana University / @jwdunn  
Tanya Clement, UT Austin / @tanyaclement  
Chris Lacinak, AVPreserve / @avpreserve  
AMIA / December 1, 2017



THE  
ANDREW W.  
**MELLON**  
FOUNDATION



## Outline

1. Background and context - Jon
2. Workshop outcomes - Tanya
3. Next steps - Chris



## Background

- AMIA 2016:
  - Chris Lacinak and Jon Dunn. “From Mass Digitization to Mass Description: Indiana University’s Strategy To Overcome The Next Great Challenge.”
- [go.iu.edu/1Pvj](https://go.iu.edu/1Pvj)



## The Challenge

- Growing AV collections
  - Digitization
  - Explosion of born-digital
- Increased expectations for access





## The Challenge

- Many AV collections lack metadata
  - Discovery
  - Identification
  - Navigation
  - Rights
- Institutions lack resources for large cataloging/inventory/rights clearance projects



## The Opportunity

- Mass digitization approach extended to AV
  - “Digitize first”
- Emergence and continued improvement of machine learning and other automated tools
- How can we leverage the best of automated tools and human expertise?



## Existing Work in the context of AV Archives

- Application of specific machine learning tools
  - e.g. speech-to-text, named entity recognition
- “Black box” systems
  - One size fits all, brute force approach to automated metadata generation
- Customized workflows
  - e.g. MICO Platform

## Context: Indiana University

- MDPI: Media Digitization and Preservation Initiative
  - 280,000+ AV items; 25,000+ films
- 80+ different units
- 20+ different physical formats
- Partnership with Memnon
- Variety of existing (or nonexisting) metadata
- Avalon Media System access platform



avalon  
MEDIA SYSTEM





## Context: Indiana University

- ◉ Consulting engagement with AVPreserve in 2016 to identify metadata and rights workflows
- ◉ Phased approach
- ◉ Identification of MGMs: Metadata Generation Mechanisms
- ◉ Need for platform to support workflows, metadata warehouse

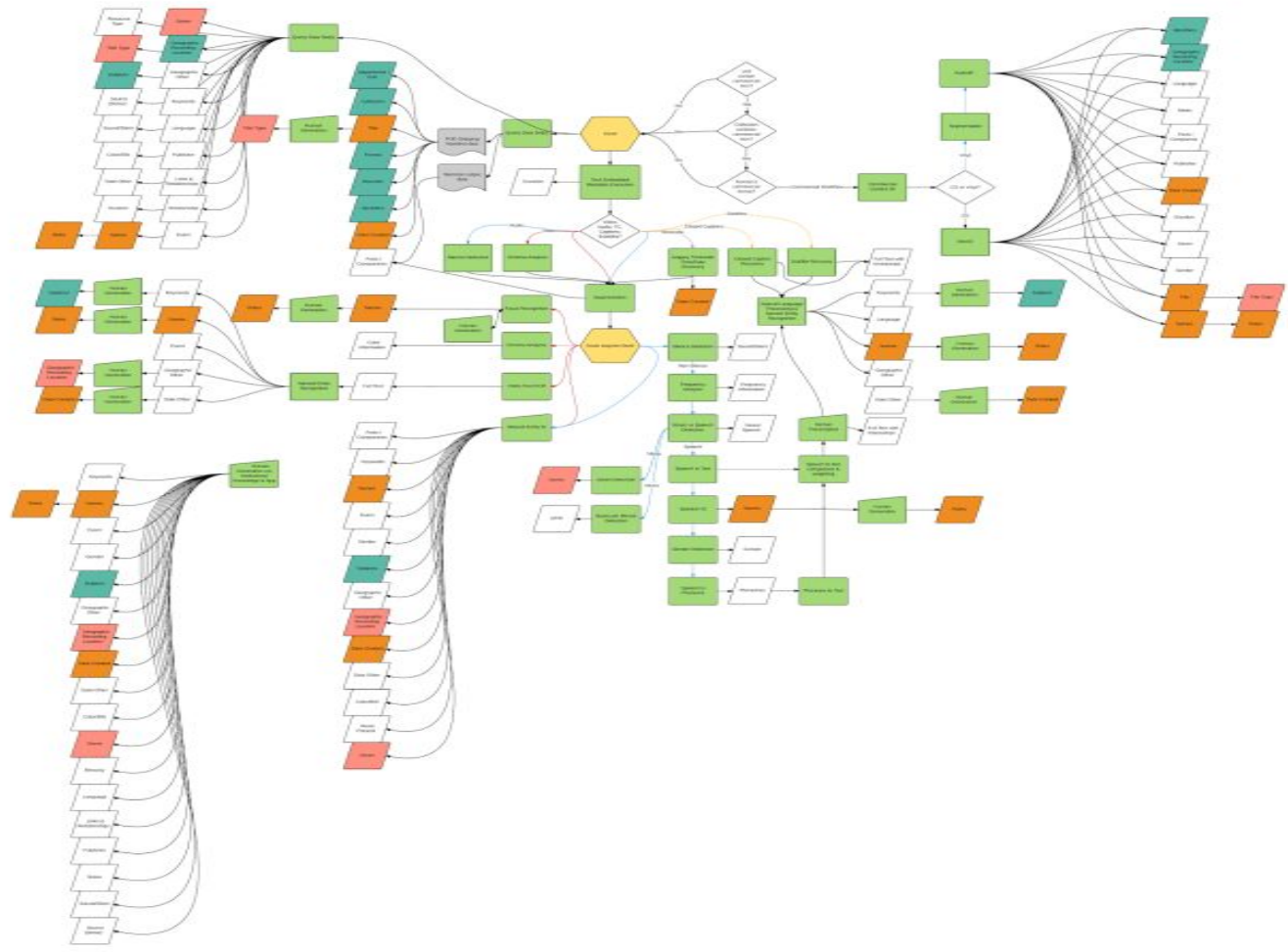
**DOCUMENT INFO**  
Raffles University Sustainability Strategy  
Project:  
Sustainability Generation Process and  
Evaluation  
Revision Draft | 2024-09-07  
A/R: Mohammad

**LINE LEGEND**

- Green: Main Process
- Orange: Sub-Process
- Red: Key Milestone
- White: Information
- Grey: Strategic Objective
- Yellow: Key Milestone

**SHAPE LEGEND**

- Hexagon: Main Process
- Rectangle: Sub-Process
- Diamond: Decision Point
- Cloud: Strategic Objective
- Triangle: Key Milestone
- Circle: Information
- Parallelogram: Information



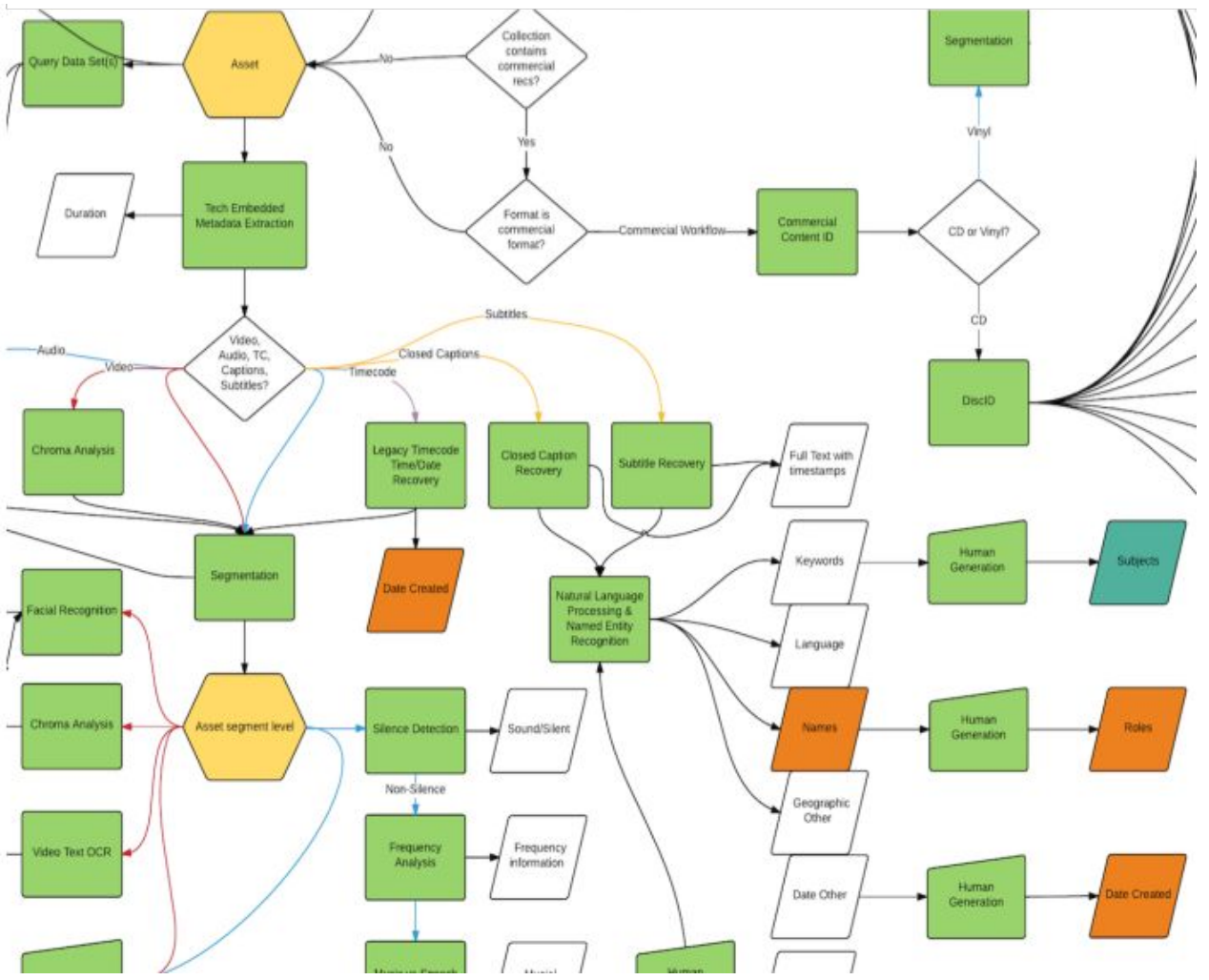
**DOCUMENT INFO**  
 Indiana University Metadata Strategy  
 Project  
 Metadata Generation Process and  
 Outputs  
 Review Draft | 2016-03-07  
 AV Metadata

**LINE LEGEND**

- Data —>
- Audio —>
- Video —>
- Captions/Subtitles —>
- Timecode —>

**SHAPE LEGEND**

- Document (grey rounded rectangle)
- Source Item (yellow hexagon)
- Decision (white diamond)
- Automated WM (green rounded rectangle)
- Required Field (orange parallelogram)
- Highly Volatile Field (red parallelogram)
- Manual WM (green rounded rectangle)
- Physical Field (teal parallelogram)
- Other Fields (white rounded rectangle)





## Context: UT/HiPSTAS

- ◉ *High Performance Sound Technologies for Access and Scholarship*
- ◉ An assessment of scholarly requirements for analyzing sound
- ◉ An assessment of technological infrastructures needed to support discovery
- ◉ Preliminary tests that demonstrate the efficacy of using such tools in humanities scholarship
- ◉ Developing a freely available, open-source, API-driven application for general use



## AMP: Audiovisual Metadata Platform

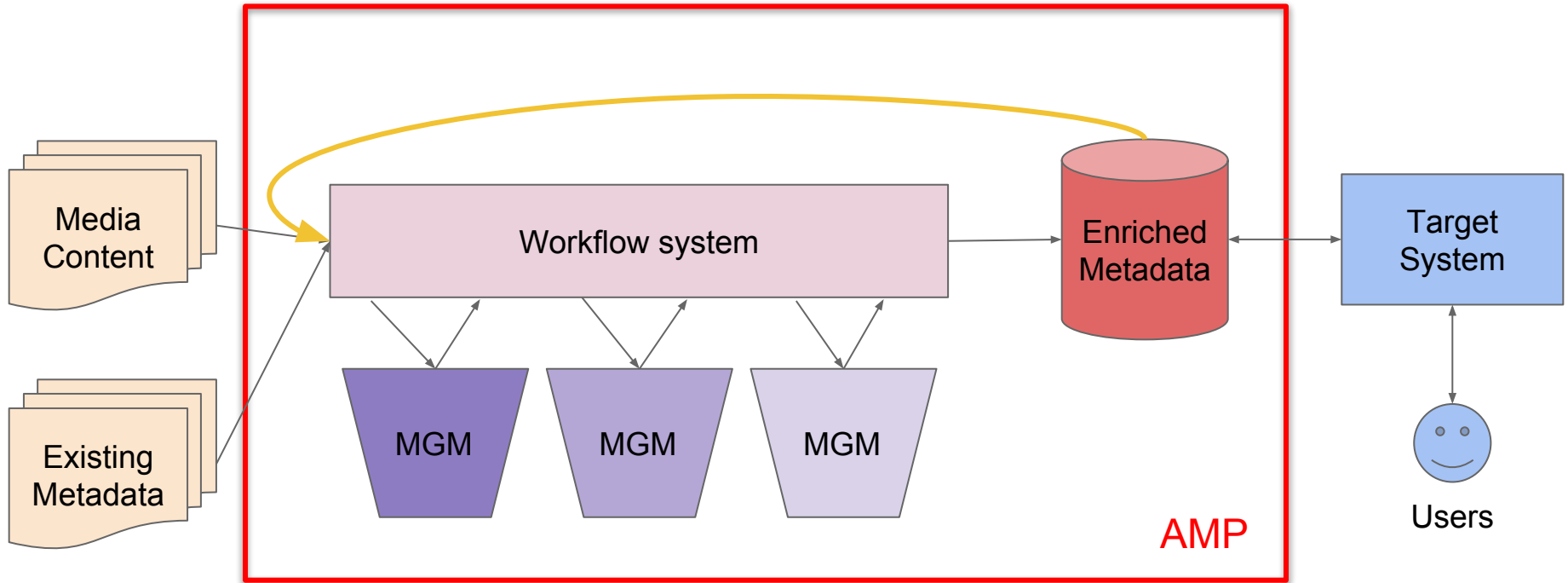
- *Audiovisual Metadata Platform*
- Planning grant from Andrew W. Mellon Foundation (July 2017 - January 2018)
  - Focus on technical architecture
- In-person workshop (September 2017)
- Planned deliverables:
  - White paper
  - Draft proposal for implementation and pilot test



## AMP: Audiovisual Metadata Platform

- Open source software platform to support metadata creation for AV collections
- Design and execute workflows combining automated and human steps
- Integrate multiple MGMs
  - Automated, manual
  - Local, HPC, cloud

# AMP Conceptual Diagram

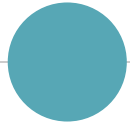




## Core AMP Team

- Indiana University Libraries
  - Jon Dunn
  - Julie Hardesty
- University of Texas at Austin School of Information
  - Tanya Clement
- AVPreserve
  - Adeel Ahmad
  - Chris Lacinak
  - Amy Rudersdorf





# Mellon-funded workshop

3 Days

16 People

1 Tech Platform Plan





# Stakeholders, User Requirements, & Personas |





# Workshop Participants

- Kristian Allen, UCLA Library
- Jon Cameron, IU Libraries
- Maria Esteva, Texas Advanced Computing Center, UT at Austin
- Mike Giarlo, Stanford University Libraries
- Brian McFee, Music & Audio Research Laboratory, NYU
- Scott Rife, Packard Campus for AV Conservation, Library of Congress
- Sadie Roosa, WGBH Media Library & Archives
- Felix Saubier, TIB/German National Library of Science & Technology
- Brian Wheeler, IU Libraries
- Maria Whitaker, IU Libraries



## Workshop logistics


---

- ⦿ **Day one:**
  - overview with framing and meeting goals
  - review of requirements, user personas (actors); current technical landscape
  - focus on non-technical criteria
- ⦿ **Day two:**
  - focus on technical component candidate system identification and ranking
- ⦿ **Day three: presented four metadata generation scenarios**

# Actors

- System administrator
- Content owner
- Target system
- Target user

**AUDIOVISUAL METADATA PLATFORM PLANNING**  
Metadata librarian persona



**NAME** Melinda Farias  
**TITLE** Metadata librarian for an access repository  
**AFFILIATION** Large academic institution

**MOTIVATIONS**  
Melinda is highly motivated by quality, but understands the need to balance this with efficiency and cost considerations. For that reason, automation is intriguing to her—she is willing to work with imperfect data and clean it up, if it means that she can provide more access points for her patrons.

Metadata that adheres to standards and taxonomies (e.g., LCSH) is important to her. PBCore, MARC, PREMIS, and rightsstatements.org are the types of standards she follows to create metadata. These are also used by the library's discovery and preservation systems, so when data conforms to them her job is easier.

Melinda sees the value in finding relationships among assets in the collection. She wants metadata to help make those connections easier and more obvious so patrons' find exactly what they need as quickly as possible.

**GOALS**  
Produce accurate metadata records for audiovisual collections  
Provide the most granular description of audiovisual content as possible  
Easily ingest assets and metadata into library systems  
Make patrons' search experience as efficient and successful as possible  
Enable clear communication of intellectual property and copyright restrictions and use  
Share metadata with data aggregators, partners, or researchers in various formats

**METADATA ACCESS METHOD**  
Structured data formats (XML, json, csv, Excel)

**TECHNICAL PROFICIENCY**  
High

*"Metadata is costly and time consuming to produce, so I'm interested in finding ways to provide users with the greatest access to audiovisual material as possible, for the lowest cost and greatest accuracy."*

Melinda Farias is a librarian who works in the metadata department in a large academic university library. She wants to be able to capture lots of details about the audiovisual collections (both physical and digital) she is responsible for describing. She is often frustrated because so little information is available about the physical materials, in particular, and it is difficult for her to view or listen to them (due to lack of playback machines and workflow ROI). Full description of digitized av is also extremely time consuming, so she tends to create basic descriptions in MARC or Dublin Core.



## **Business Requirements**

---

- ⦿ Automate analysis of AV content & metadata
- ⦿ Provide an intuitive interface
- ⦿ Leverage best-of-breed tools in a single workflow
- ⦿ Generate metadata with minimal errors
- ⦿ Offer a variety of metadata output
- ⦿ Build a community of developers
- ⦿ Offer easy-to-use APIs
- ⦿ Process multi-TB batches of content at a time
- ⦿ Support collaborative efforts
- ⦿ Control what metadata is made public



---

## **Functional requirements**

- **Asset management**
- **Copyright and security**
- **Storage**
- **Metadata standards**
- **Modularity**
- **Multi-tenancy**
- **Scalability**
- **Usability**



## Technical Requirements

- ⦿ Data Model (42)
- ⦿ APIs (42)
- ⦿ Open source (41)
- ⦿ Scalable (41)

Queueing	38
Data output	38
Access to metadata	36
Authentication	36
Versioning	36
Modularity	35
SEO	33
GUI/User Interface	32
Ease of configuration	31
Multi-tenancy	30

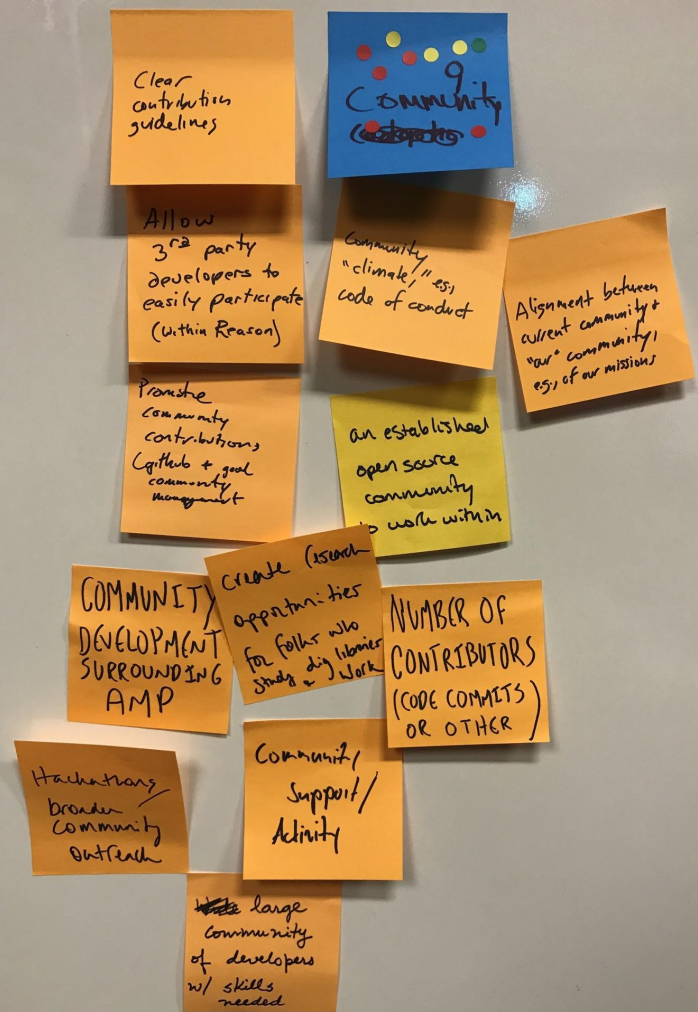


## Non-functional Requirements

- Clear service/cost model
- Wide adoption
- Robust documentation
- Active dev community

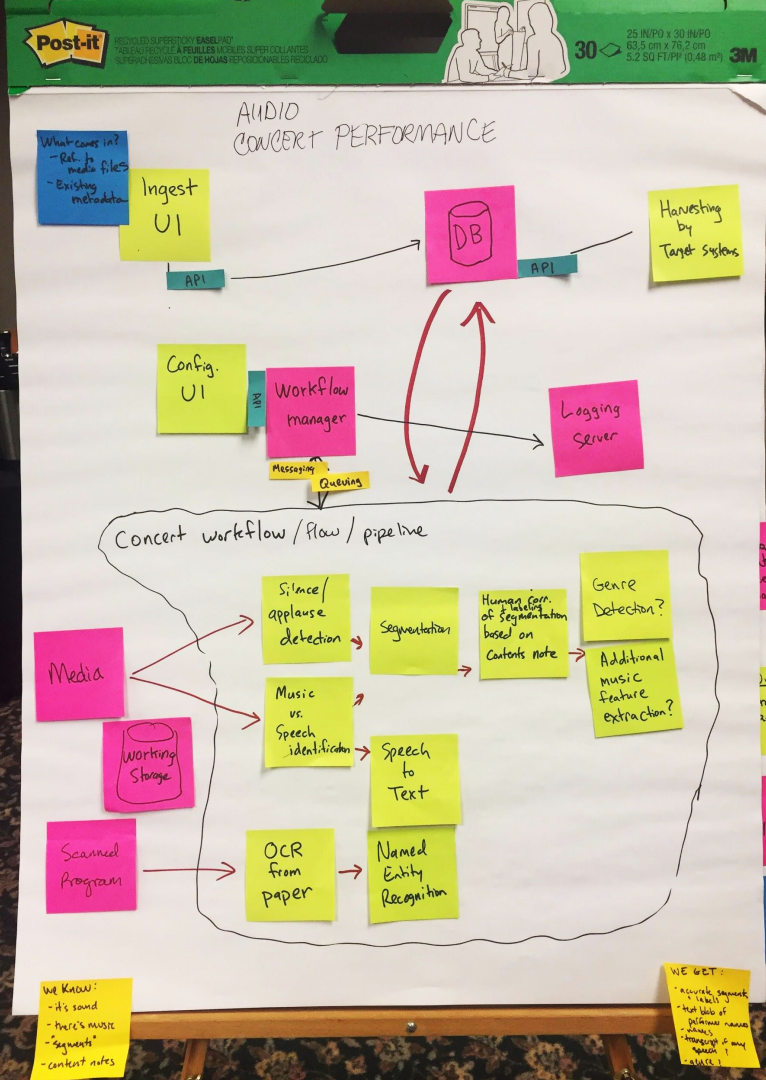
Also:

- Proven stability of the product
- Open licensing
- Established governance model
- Robust outreach & marketing program
- Training available
- Semantic versioning
- Support for internationalization



## Metadata generation scenarios

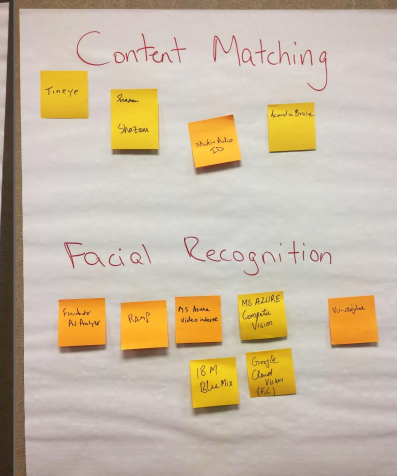
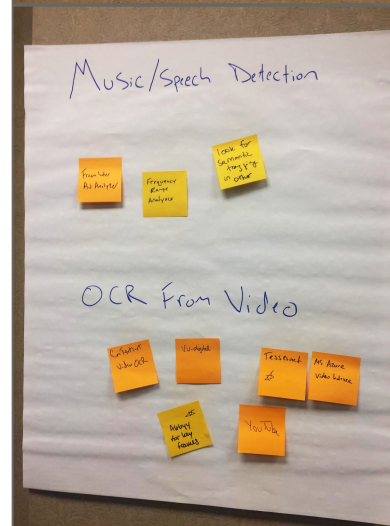
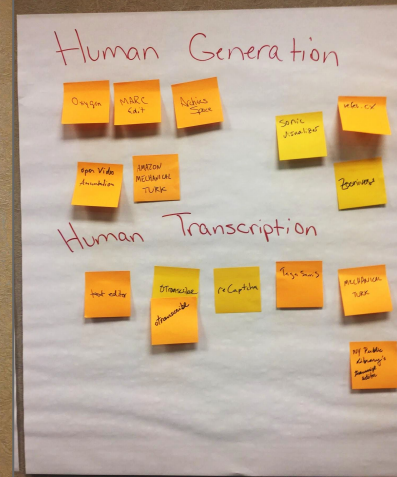
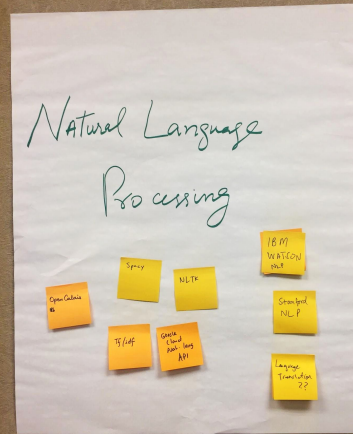
Given actual media and asked to map out how metadata would be generated from media and related documentation ingest / creation to export.





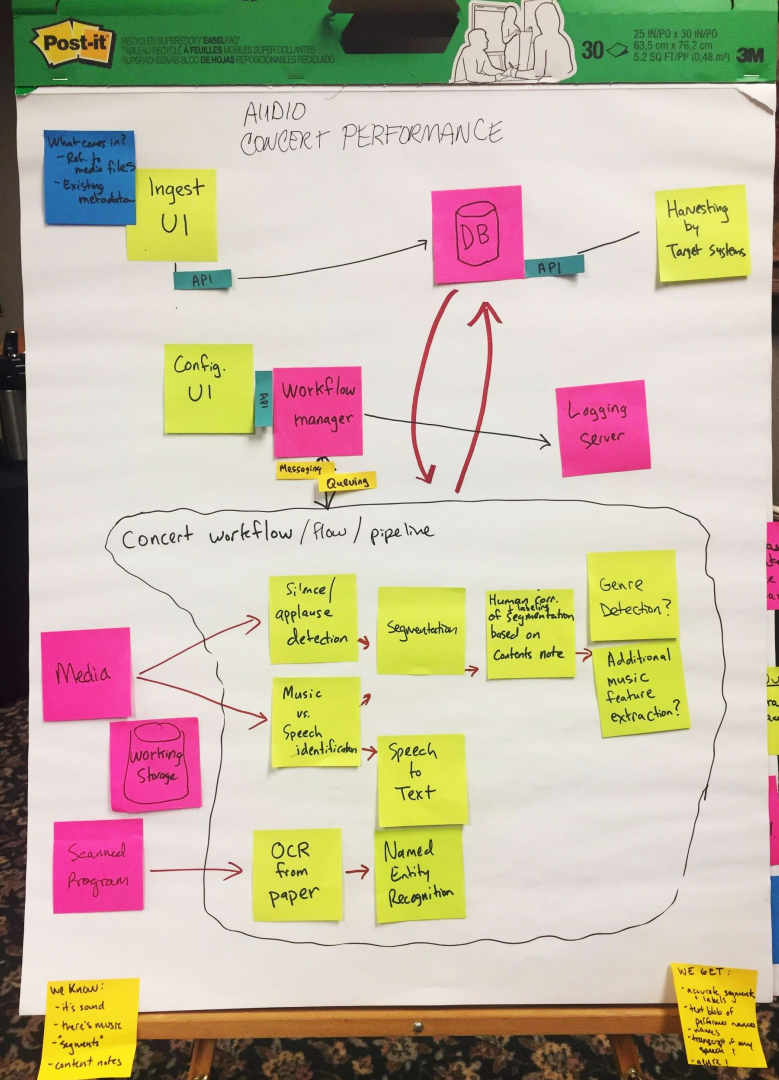
## Metadata Generation Mechanisms (MGMs)

Identified systems that could perform NLP, content matching, facial recognition, OCR from video, music/speech detection, and more



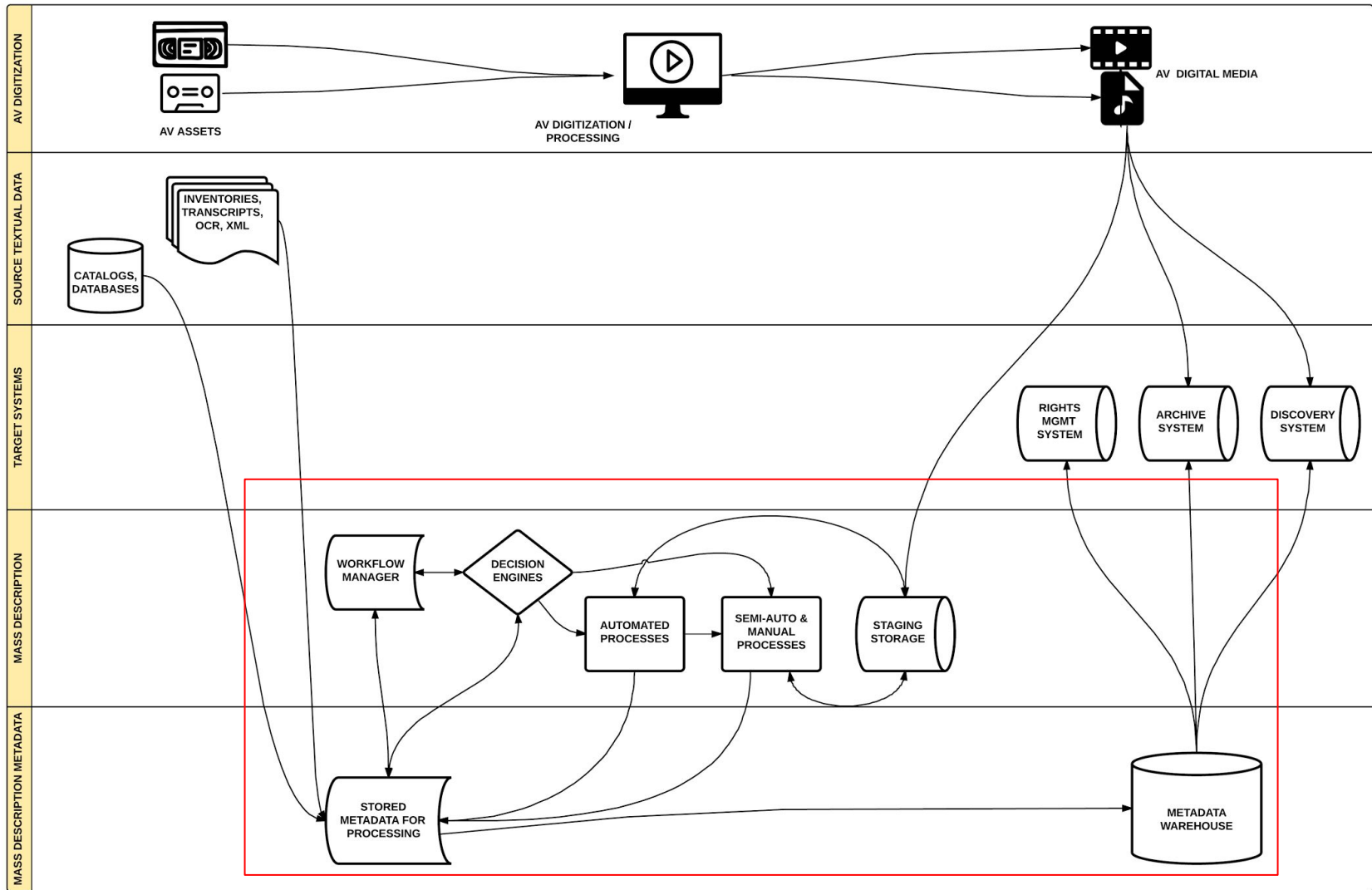
# Metadata generation scenarios

- audio music concert performance
- audio oral history - Italian and Finnish interviews and songs
- video oral history - Yiddish interviews
- video promo segments aired during halftime on college basketball TV broadcasts

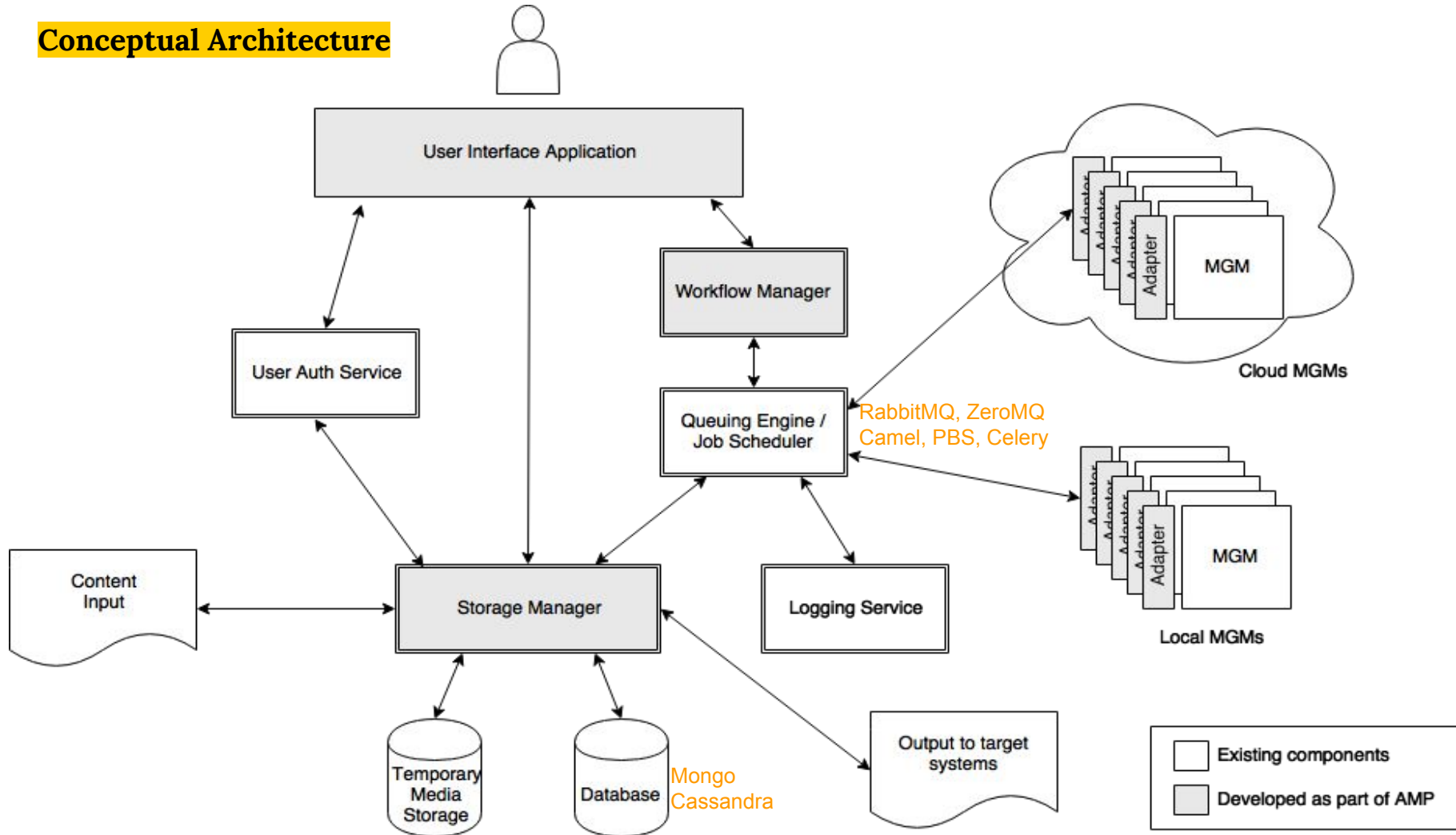


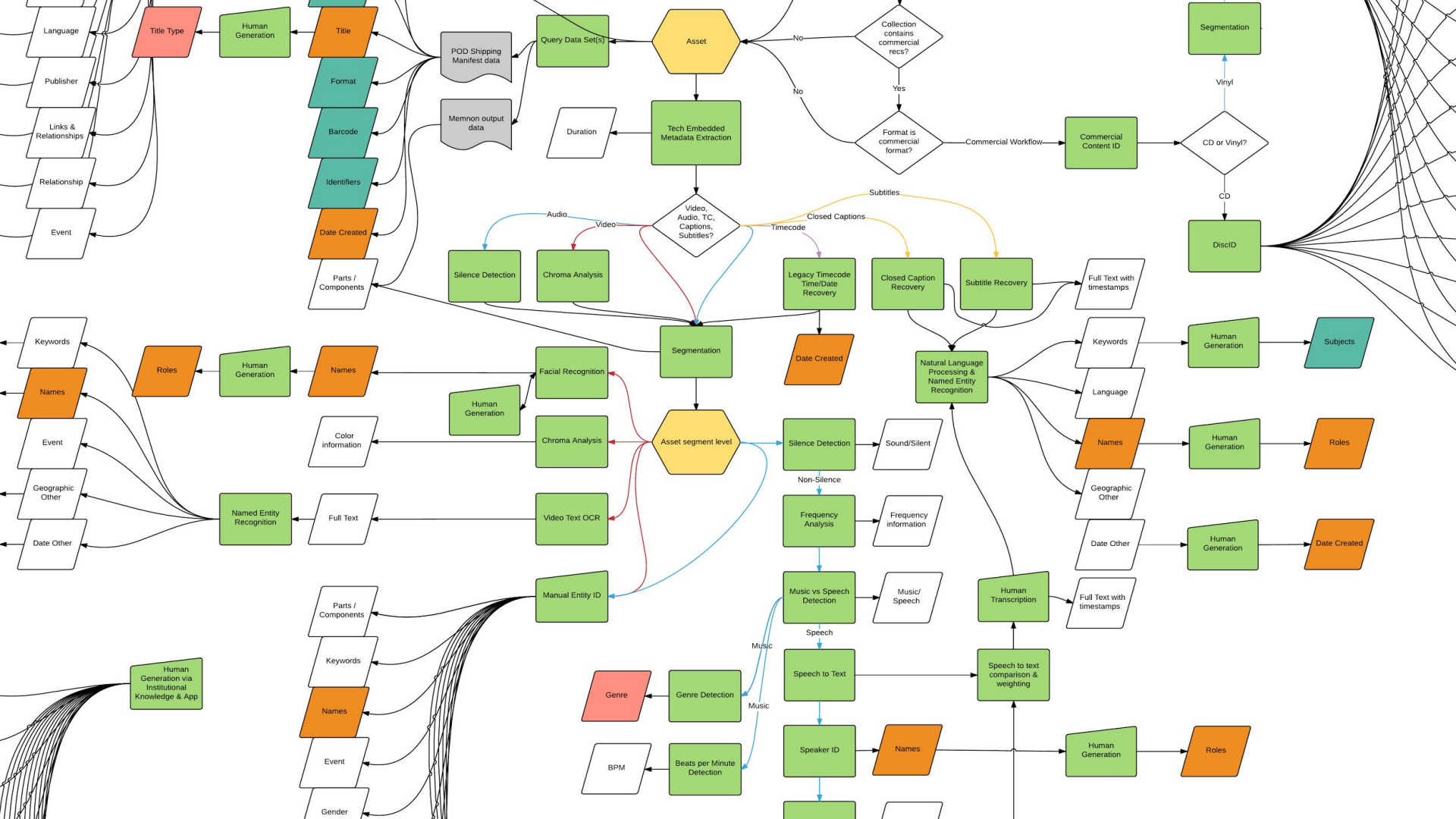


# Next Steps



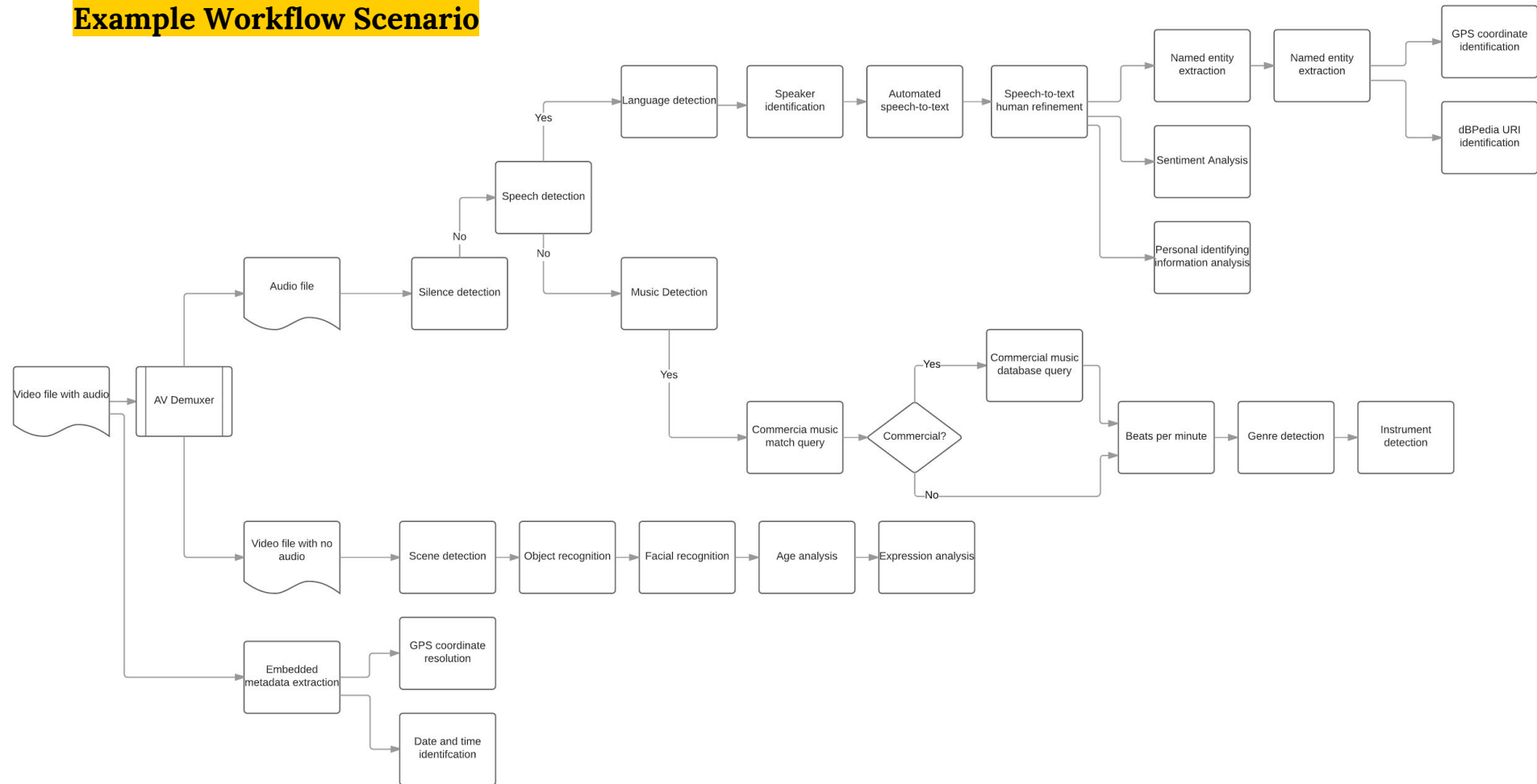
# Conceptual Architecture



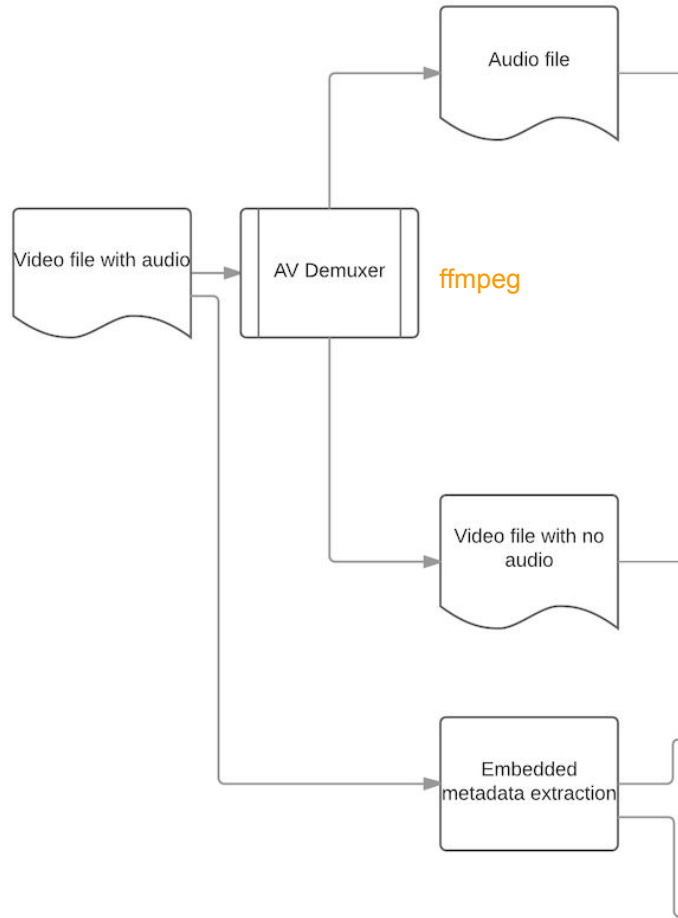




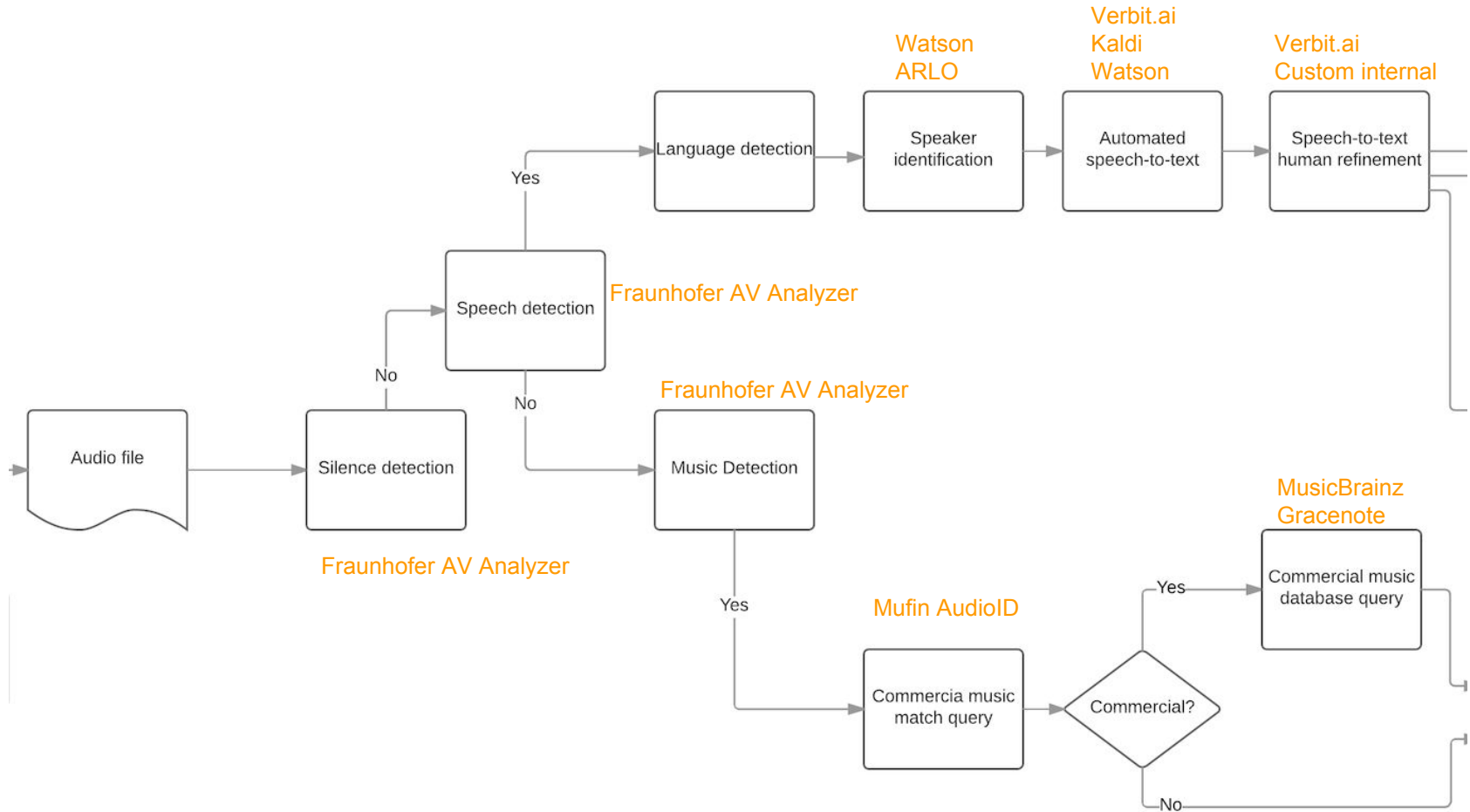
# Example Workflow Scenario



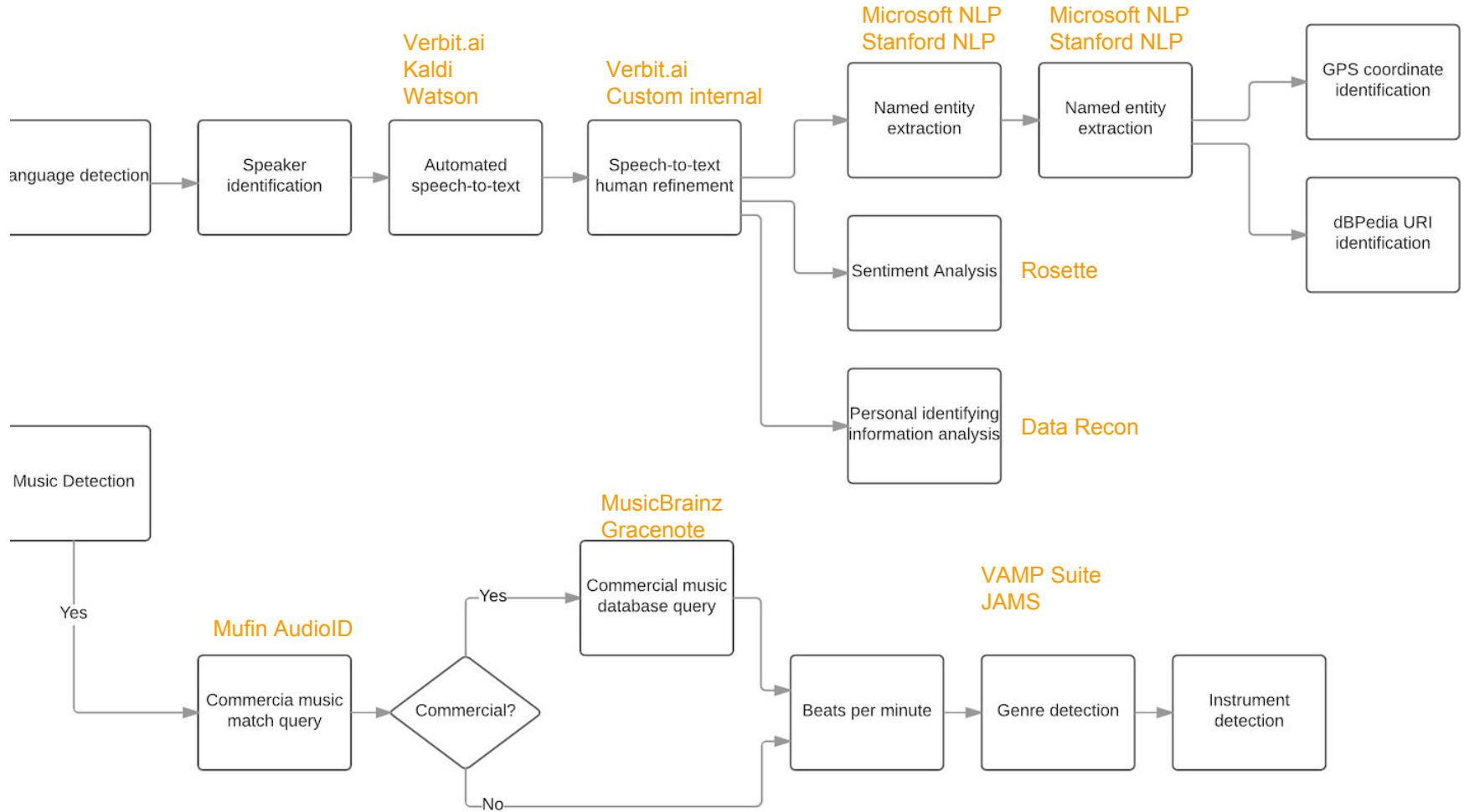
## Example Workflow Scenario



# Example Workflow Scenario

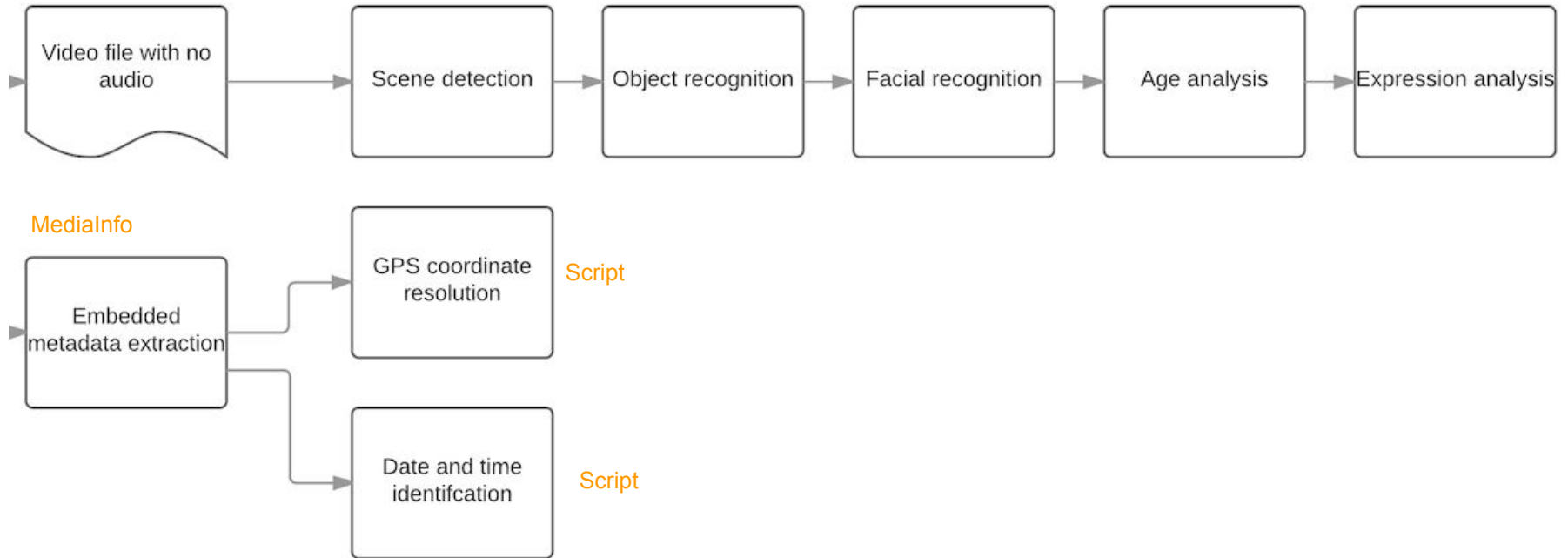


# Example Workflow Scenario



## Example Workflow Scenario

Fraunhofer AV Analyzer  
VU Digital



## Borrow, build, buy considerations



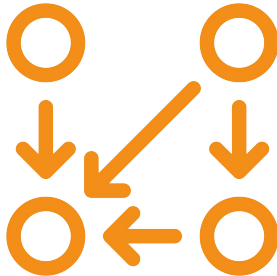
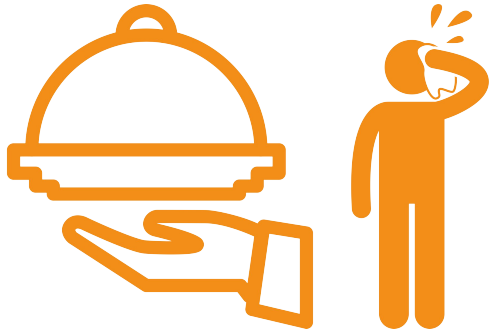
- Ownership of “the machine”
- Black boxes

- Very limited options
- Metadata cultivation concept
- Focal point



- Limited capabilities demonstrated
- Exploring building on top of MiCO

## Anticipated challenges & considerations





**Stay  
Tuned!**





**Thanks!**

Jon Dunn

jwd@iu.edu | @jwdunn

Tanya Clement

tclement@utexas.edu | @tanyaclement

Chris Lacinak

chris@avpreserve.com | @avpreserve