

COMBINING AND CURATING AUTOMATED METADATA FROM MULTIPLE SOURCES

Lauri Saarikoski
DataTV2020 webinar
2020-09-14



MeMAD

Methods for Managing
Audiovisual Data

MeMAD project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 780069. This presentation has been produced by the MeMAD project. The content in this presentation represents the views of the authors, and the European Commission has no liability in respect of the content.



Project objectives

**Novel tools for
digital storytelling**

Expand media audience

**Understanding of multi-modal
/-lingual
content analysis and linking**

**Standardize models and
distribution formats
for describing enriched
audiovisual data**

About: The MeMAD project (2018-2021)

- Funded by the EU Horizon 2020 programme:
Tools for smart digital content in the creative industries
- Consortium is a mix of universities and research centers, SMEs with products for the media industry and content creators



We have worked on...

**Technologies for
metadata extraction and
machine translation**

**Combining data sources
and structures**

**Multimodality by
combining analyses**

**Developing media tools and
workflows**

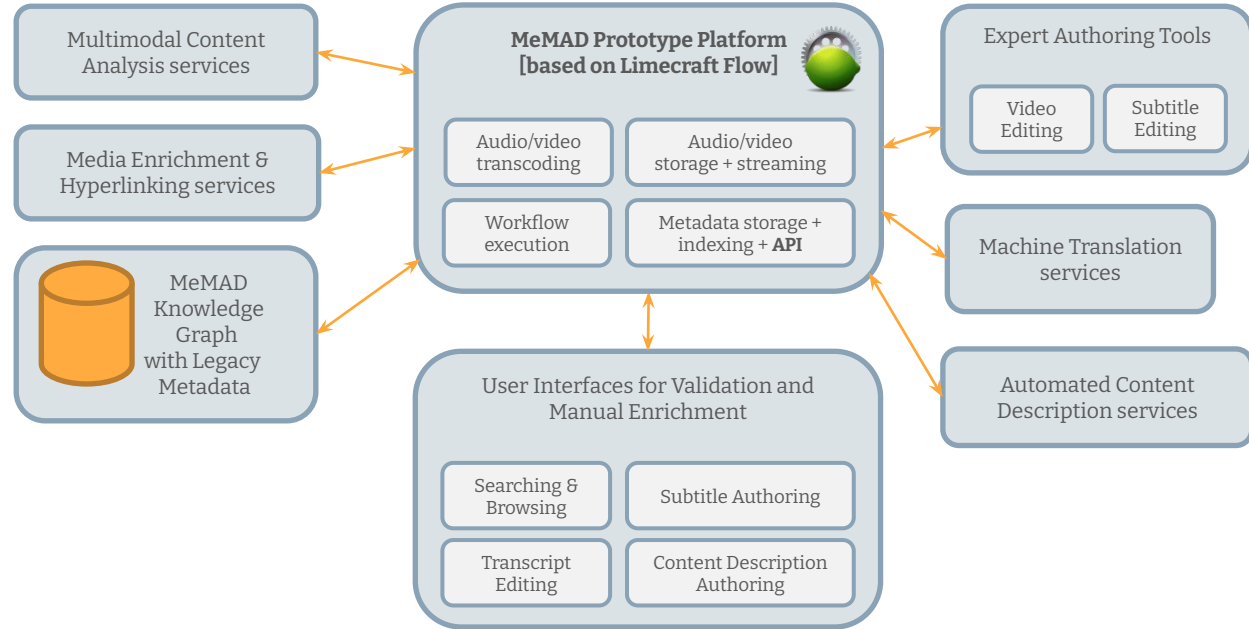
Consumer evaluations

**Comparing human
perception with machine
learning approaches**

MeMAD is building a prototype platform

Used for

- metadata storage
- GUIs
- media manipulation tools
- workflow coordination
- integration with post-production tools (e.g., Avid, Adobe, Wincaps, etc.)



MeMAD AME services and metadata

1. Audio segmentation
 2. Audio classification
 3. Automated speech recognition (ASR)
 4. Face detection and recognition
 5. Shot-cut detection
 6. Video captioning
 7. **Machine translation (MT)**
 8. Named entity recognition and disambiguation
 9. Semantic enrichment
 10. Subtitle generation
 11. Content description generation
 12. Content segmentation
 13. Relevant TV moment detection
 14. Spoken language segmentation and classification
- Combination of commercially available services (e.g., ASR/NER for common languages) and technologies developed by the consortium
 - Trials of multi-modal fusion
 - Many are available as open source tools
<https://github.com/MeMAD-project/mmca>

Processing 'legacy' archive metadata

- To assess the usefulness of AME, we want to compare it to the equivalent of what is available now: existing 'legacy' metadata already in archives
 - Can we obtain similar and usable results from AME, compared to the legacy metadata?
- Legacy metadata provided by INA and Yle, (+1600 hours of media) was processed and consolidated according to the EBUCore/MeMAD ontology and put into the MeMAD **Knowledge Graph**
 - Knowledge Graph is a Virtuoso-based triple store which can be queried using SPARQL and gathers 'finished' MeMAD metadata
 - Custom converters transformed data from YLE and INA metadata into Knowledge Graph RDF entries

Automated metadata interchange formats

<https://github.com/MeMAD-project/interchange-formats>

- Formats defined for a number of metadata types
- Easier integration with other applications

Example:

- The ASR format builds upon individual transcript paragraphs with fine-grained properties and the EBUCore data model

```
"@id": "http://memad.eu/test-trans-para",  
"@type": [  
  "ebucore:TextLine",  
  "memad:SpeechParagraph"  
],  
"ebucore:textLineContent": "Yeah, I think ...",  
"ebucore:textLineStartTime": 37035,  
"ebucore:textLineEndTime": 60000,  
"ebucore:hasTextLineRelatedCharacter": « Charles",  
"ebucore:textLineLanguage": "en",
```


Developing media tools and workflows

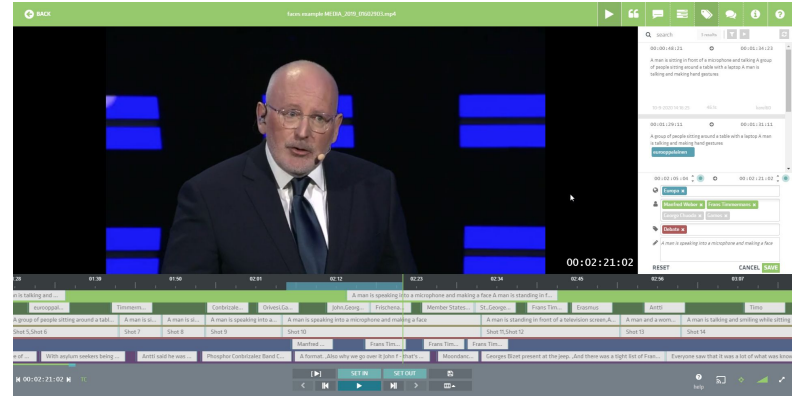
Prototype tools, concept design and end-user **evaluations**

Automated metadata for:

- **Media archive** management (search and documentation)
- **Video editing**

Machine translation, ASR and NLP for:

- **Subtitling**
- Media archive searches



Prototype GUI:
Semi-automated content description
(curating AME & MT results)

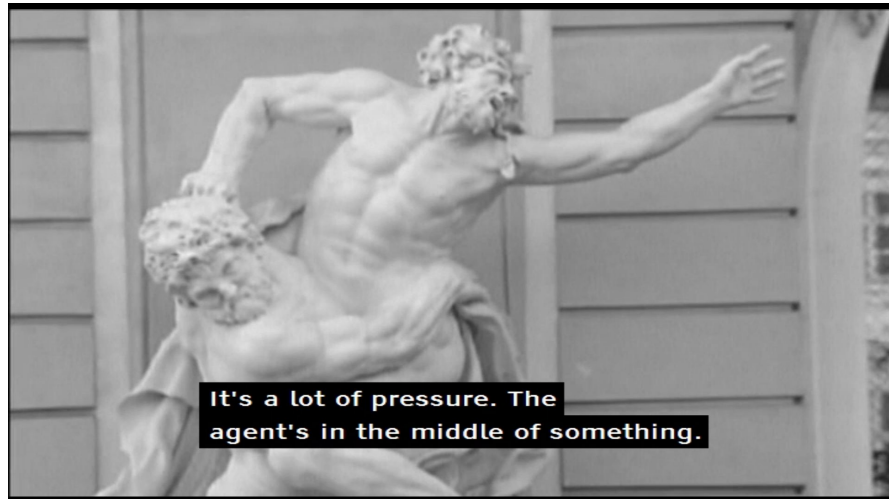
Consumer evaluations

Machine-translation and ASR in subtitles

Acceptance? Accessibility gains?

Enriched metadata for discovering and navigating in video content

How to make use of rich metadata and the knowledge graph?



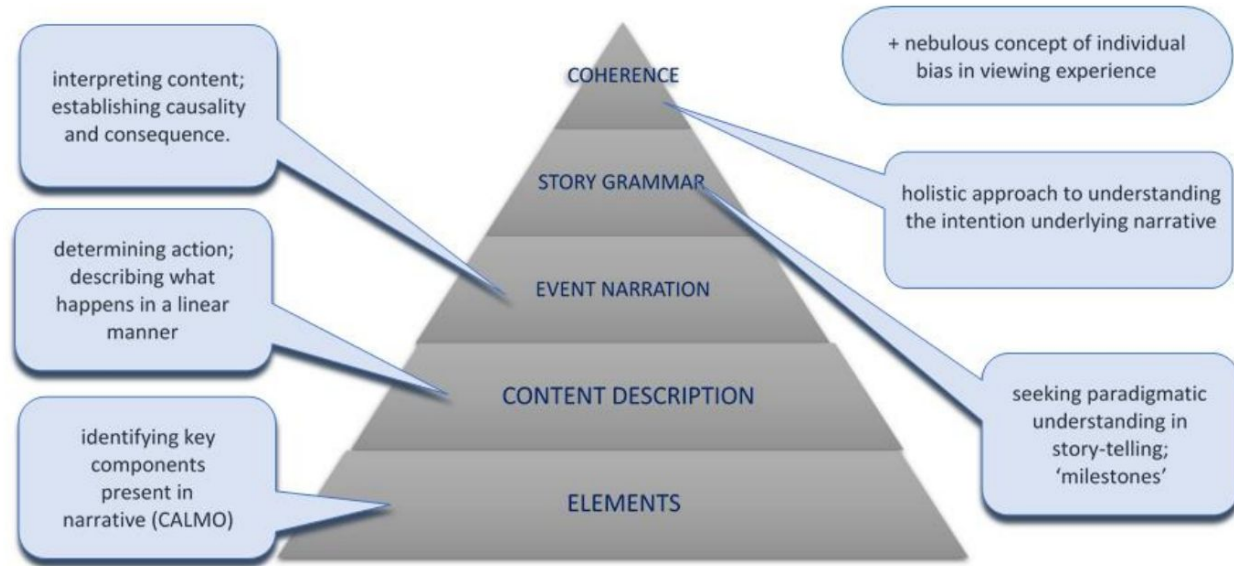
Audience panel sample:
machine-translated subtitles

Comparing human perception and machine learning approaches

Do we annotate and analyse the right things?

→ Feedback to machine learning development

→ Guidelines for producing useful [multi-purpose] annotations



MeMAD project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 780069. This presentation has been produced by the MeMAD project. The content in this presentation represents the views of the authors, and the European Commission has no liability in respect of the content.



Currently focusing on

- Multimodal story- and topic-based content segmentation
- Improving automatic video captioning with cues from face recognition, speaker diarization etc.
- Final end-user evaluations

Some findings (for a media company)

- Translation and subtitling workflows benefit from MT and NLP
- Current production tools may struggle with large data quantities
- Automated audio description based on video captioning needs further work

Thank you!

Lauri Saarikoski

Development Manager, Archives
Yle, the Finnish Broadcasting Company

lauri.saarikoski@yle.fi

Latest news and results:

[Memad.eu](https://memad.eu)

<https://github.com/MeMAD-project/>



MeMAD project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 780069. This presentation has been produced by the MeMAD project. The content in this presentation represents the views of the authors, and the European Commission has no liability in respect of the content.



MeMAD

Methods for Managing
Audiovisual Data