A System for Task-Oriented Content Analysis and Search in Media Production

Georg Thallinger (JOANNEUM RESEARCH)
NEM Summit
Istanbul, Oct. 17, 2012

The TOSCA-MP project receives funding from the European Union’s Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 287532.
Outline

- Context
- Project overview
- Scenarios and tasks
- System architecture
- Conclusion
Context

- Professional media industry
- Media life cycle
  - Production
  - Distribution
  - Archiving and reuse
- Many business processes involve searching for media items
  - News production
  - Archiving
  - Producing documentaries
  - ...
Media Production

- Searching not only in repositories of own organisation, but
  - Content providers such as other broadcasters, news agencies, ...
  - Networks such as Eurovision
  - User generated content
- Efficient annotation
  - For own reuse
  - Monetising own content by enabling it is found by others
- Strengthened collaboration paradigm, increased content syndication
  - Large-scale repositories distributed across a network
  - Multilingual
  - Heterogeneous in terms of structure and data model
  - Journalists in the field do work previously done by dedicated staff in the studio
TOSCA-MP

- Task-oriented search and content annotation for media production
- Consortium
Motivation & Objectives

- **Motivation**
  - Media production and archiving workflows are changing rapidly
  - Workflows need to be increasingly networked and distributed
  - Media search in distributed large-scale repositories in the network (multilingual and heterogeneous)
  - Off the shelf tools, able to adapt to users‘ tasks are needed

- **Objectives**
  - Enable content holders to leverage scalable advanced distributed processing
  - Enable editorial staff in networked media production to seamlessly use content from distributed heterogeneous repositories
  - Make benchmarking of indexing and search methods an integrated part of the workflow
Research areas

- Media Analysis
  - Multilingual speech metadata extraction/translation
  - Content-adaptive visual metadata extraction and enrichment
  - Aligning and linking metadata
- Task-adaptive search & retrieval and user feedback
- User interfaces for annotation and search
- Task models and tools benchmarking
- Scalable distributed repository framework (indexing and search) for digital media production workflows
- Results are integrated in a SOA based framework
Usage Scenarios

- 10 Business goals
  - High level goals to be achieved by broadcasters, service providers…

- 15 related scenarios
  - Concrete examples, how a business goal is achieved with tasks and actors involved

- Four categories
  - Content access & retrieval
  - Assisted production
  - News service distribution
  - Infrastructure (repositories)

- Details on website (D6.1)
Example: “Assisted production of news stories using distributed multilingual sources”
Task models

- Collecting a set of high-level task descriptions from users
  - e.g. specific types of annotation, specific types of search, content summarisation
  - [http://tosca-mp.eu/tasksurvey](http://tosca-mp.eu/tasksurvey)
- Create formalised task models (ConcurTaskTrees, CTT)

- Automated generation of business process descriptions for service orchestration
- Automate the validation of tools in different media production workflows (scenarios)
Logical system design

- Distributed Repository Framework (DRF)
  - Data models
  - Internal Essence Repository
  - External Repository

- Services
  - Generation & aggregation of metadata
  - Benchmarking methods
  - Semantic enrichment and linking
  - Validation of models and benchmarking methods
  - User feedback analysis
  - Networked media search engine

- Metadata Production Management Framework (MPMF)
  - GUI
    - Search and Browsing
    - Control / Orchestration
    - Review & Annotation UI
  - Orchestration
    - ESB
    - Process Engine
    - Task modelling
Technical system design

- Service oriented architecture
- Compatible with EBU/AMWA FIMS
Conclusion

- Task-adaptive content analysis, annotation and search tools
- Search in distributed heterogeneous repositories
- Performed analysis of business goals and scenarios
- Collection and formalisation of task descriptions
- Service-oriented architecture
Questions?

http://www.tosca-mp.eu

The research leading to these results has received funding from the European Union’s Seventh Framework Programme under the grant agreement no. 287532, “TOSCA-MP” (http://tosca-mp.eu).