Authoring and Publishing Linked Open Film-Analytical Data

Henning Agt-Rickauer[1], Olivier Aubert[2], Christian Hentschel[1], and Harald Sack[3]

[1] Hasso Plattner Institute for IT Systems Engineering, University of Potsdam, Potsdam, Germany
{henning.agt-rickauer,christian.hentschel}@hpi.de
[2] LS2N UMR CNRS 6004, University of Nantes
olivier.aubert@univ-nantes.fr
[3] FIZ Karlsruhe - Leibniz Institute for Information Infrastructure, Karlsruhe Institute of Technology, Karlsruhe, Germany
harald.sack@fiz-karlsruhe.de

Abstract. Exploiting Linked Open Data for the annotation of audio-visual patterns in film-analytical studies provides significant advantages, such as the non-ambiguous use of language as well as the possibility to publish and to reuse valuable data. On the other hand, film scholars typically lack the know-how to cope with semantic annotations and, moreover, most software for annotating audio-visual material does not provide means to enter semantic annotations directly. The project presented in this paper aims to provide an ontology for film-analytical studies complemented by a video annotation software adapted for authoring and publishing Linked Open Data by non-experts.

1 Introduction

The study of audio-visual rhetorics of affect scientifically analyses the impact of auditory and visual staging patterns on the perception of media productions as well as the conveyed emotions [4]. The *AdA*-project[4] aims to explore the hypothesis of TV reports drawing on audio-visual patterns in cinematographic productions to emotionally affect viewers, by analyzing TV reports, documentaries and genre-films of the topos "financial crisis". In a large-scale corpus analysis film scientists identify and annotate low- to high-level audio-visual patterns, such as shot duration, dominant colors, major-minor tonality and depicted visual concepts. Comparison of different annotations from different scenes and genres allows film scientists to analyze this opinion-forming level of reporting. In order to avoid ambiguities within and enable reuse of and reasoning based on the generated pattern annotations, we pursue two main objectives: 1) creating a standardized annotation vocabulary to be applied for *semantic annotations* and 2) *enabling non-expert users* to adopt and benefit from these annotations.

[4] AdA-project — http://www.ada.cinepoetics.fu-berlin.de/

In the next section of this paper, we describe the vocabulary used for annotation of audio-visual material, which is based on Linked Open Data principles. The annotation process is performed by film scientists using Advene [3] — a software toolkit to annotate audio-visual documents. To accommodate specific needs of the project, Advene was extended to improve the manual annotation process, with integration of automatic and semi-automatic helpers, and with support of RDF interoperability through the import of OWL ontologies and export of RDF data. The third section of this paper will therefore present these extensions in more detail and show how Advene helps film scientists to apply semantic annotations without having to deal with the technical challenges. The last part of this paper will give a small outlook on the presented demo. We provide some screenshots and screencasts of the Advene video annotation software at `https://ProjectAdA.github.io/ekaw2018/`.

2 Open Film-Analytical Data

Film scientists carry out an in-depth corpus analysis by precisely describing feature films, documentaries and TV news according to a film-analytical annotation method called eMAEX[5]. The description involves a lot of manual effort, creating hundreds of annotations per scene as ground truth data. One goal of the project [1] is to publish this valuable data as Linked Open Data to make these annotations available to other film scientists as well as researchers from other domains. Therefore, we developed an ontology for fine-grained semantic video annotation, constructed semantic metadata for our video corpus, and implemented linked data extensions for the Advene annotation software.

The AdA ontology offers a vocabulary with a number of categories under which a movie is analyzed (e.g., camera, image composition, acoustics). Each category includes the respective concepts with which the segments of a movie are annotated (e.g., camera movement speed, field size). About 75% of the concepts have associated predefined values (e.g., long shot, medium shot, closeup and others for field size). Others are free-text annotations, such as dialog transcriptions. Currently, the AdA ontology includes 9 categories (annotation level), 78 concepts (annotation types), and 435 predefined annotation values. We provide an online version[6] and a download at the GitHub page of the project[7].

The ontology contains a data model that uses the latest Web Annotation Vocabulary[8] to express annotations and Media Fragments URIs [9] for timecode-based referencing of video material. This allows the publication of semantic audio and video annotations as Linked Open Data. Semantic metadata of the video corpus is described using the classes and properties of DBpedia, Schema.org,

[5] eMAEX - Electronically-based Media Analysis of EXpressive movements — `https://empirische-medienaesthetik.fu-berlin.de/en/emaex-system/`
[6] `http://ada.filmontology.org/`
[7] `https://github.com/ProjectAdA/public`
[8] `https://www.w3.org/TR/annotation-vocab/`
[9] `https://www.w3.org/TR/media-frags/`

and Linked Movie Database, and movies are linked to DBpedia and Wikidata if present in the respective knowledge base.

3 Bridging the Gap with Advene

Advene is a free (GPL) video annotation software aiming at a great flexibility to accommodate the various needs of different users. Some adaptations have been implemented for the project in order to facilitate the annotation task, allowing faster annotation, better collaboration and exporting data into RDF, continuing previous efforts [2] in this domain. One of the goals of the project is to allow non-technical users to manage semantic information, more specifically semantic metadata linked to audiovisual fragments. In addition to the generic constraints of video annotation, a common hurdle is the tediousness of specifying precise URIs for semantic content. The Advene platform defines a notion of package, which contains the annotations themselves, but also the definition of the annotation structure and their visualizations. In the context of a specific Advene package, URIs for all ontology elements are specified as metadata, which gives the semantic context for the non-semantic information contained in the annotations. Users are facing mostly basic types and keywords, linked to their expertise domain. When exporting annotation data to RDF, keywords are translated to proper URIs using information present as annotation type metadata.

OWL Import The Advene data model consists of user-defined annotation types and relation types. These types can be grouped into schemas that materialize a specific analysis frame. One of the goals of the project was to provide a way to map the AdA ontology into an Advene set of types and schemas, accompanied with appropriate metadata. The mapping of the AdA ontology structure translated the notion of Annotation Level into Advene schema, and Annotation Types as Advene annotation types. The annotation content can be part of a fixed vocabulary, defined by the ontology. Each predefined value from the ontology has been mapped to a simple keyword (shortname), with metadata allowing to remap it to its original URI during the RDF export phase. An OWL import plugin was developed for Advene. As AdA ontology development underwent many iterations, Advene's merge functionality has been improved to better support data migration.

RDF Export One of the characteristics of Advene is the possibility for users to define their own visualisations, through a template language initially aimed at producing XML data, but which can also produce any kind of data. This approach, which is used for the majority of Advene export filters, has initially been used to implement an RDF export of the annotation data, using the AdA ontology model. However, some project-specific complex structures like evolving or contrasting values could not easily be implemented through this simple syntax-based approach. A new `RDFLib`-based export filter has thus been developed, offering better robustness, expressiveness and performance. As with the ontology import filter, it is rather specific to the AdA ontology, but its code can be used as a reference implementation for other RDF export filters.

Interface Adaptation The annotation interface has been streamlined and extended, to allow faster input of predefined vocabularies. In order to increase the speed of manual annotation by means of automatic content classification, the interface allows for calling external webservices. As a first proof-of-concept, a convolutional neural network-based approach for visual concept detection in keyframes has been implemented. Advene tries to be agnostic about the data that can be stored in annotations. In order to ensure that newly added annotations are part of the AdA ontology, the Advene constraint checker, that gives a constant feedback, has been extended with ontology-specific checkers.

4 Demo and Conclusion

In the demo it will be demonstrated how to create semantic annotations with Advene. This includes importing the project's ontology, using Advene templates, creating annotations in the timeline view, using the constraint checker and exporting RDF. Also insights will be provided into the linked data that is already published under `http://ada.filmontology.org/`: the AdA Ontology, video corpus metadata as well as RDF annotations of the use-case movie. Finally, it will be demonstrated how this data can be queried using the SPARQL endpoint.

In this paper, we presented a small overview about how semantic annotations are applied in film-analytical studies. An ontology has been developped specifically for the purpose of annotating audio-visual patterns using Linked Open Data and it has been presented how Advene can be applied to hide the complexity of semantic annotations for non-expert users. Future work will focus on the integration of semi-automatic analysis in order to speedup the expensive annotation process.

Acknowledgments. This work is partially supported by the Federal Ministry of Education and Research under grant number 01UG1632B.

References

1. Agt-Rickauer, H., Hentschel, C., Sack, H.: Semantic annotation and automated extraction of audio-visual staging patterns in large-scale empirical film studies. In: Proceedings of the Posters and Demos Track of the 14th International Conference on Semantic Systems (SEMANTiCS 2018), Vienna, Austria (2018)
2. Aubert, O., Champin, P.A., Prié, Y.: Integration of semantic web technology in an annotation-based hypervideo system. In: SWAMM 2006, First International Workshop on Semantic Web Annotations for Multimedia, 15th WWW Conference (2006)
3. Aubert, O., Prié, Y.: Advene: active reading through hypervideo. In: ACM Hypertext'05. pp. 235–244. Salzburg, Austria (Sep 2005)
4. Scherer, T., Greifenstein, S., Kappelhoff, H.: Expressive movements in audiovisual media: Modulating affective experience. In: Müller, C., Cienki, A., Fricke, E., et al. (eds.) Body – Language – Communication. An international handbook on multimodality in human interaction, pp. 2081–2092. De Gruyter Mouton, Berlin, New York (2014)