

Semantic Technologies for Multimedia Content: Foundations and Applications

Ansgar Scherp
Data and Web Science
University of Mannheim, Germany
ansgar@informatik.uni-mannheim.de

ABSTRACT

Higher-level semantics for multimedia content is essential to answer questions like “Give me all presentations of German Physicists of the 20th century”. The tutorial provides an introduction and overview to such semantics and the developments in multimedia metadata. It introduces current advancements for describing media on the web using Linked Open Data and other more expressive semantic technologies. The application of such technologies will be shown at concrete examples.

Categories and Subject Descriptors

H.2.4 [Systems]: Multimedia databases; H.5.1 [Multimedia Information Systems]

General Terms

Languages, Standardization

Keywords

Multimedia Metadata, Linked Open Data, Ontologies

1. INTRODUCTION

Multimedia content can be seen as collection of channels of information [1]. Some of these information channels are low-level and can be derived from the content itself (e. g., the color, shape, and texture). However, other channels encode higher-level semantics that cannot be derived from the multimedia content without external knowledge. Examples are the context in which some multimedia content like a photo has been taken or personal annotations that have been attached to the content by the creator.

In the context of this work, we will consider the higher-level semantics in multimedia as essential metadata information about the multimedia content that enables, e. g., to search along concepts like “Give me all presentations of German Physicists of the 20th century”. In order to describe such higher-level semantics, the idea of Linked Open

Data (LOD) has been suggested by Tim-Berners Lee some years ago.¹ LOD aims at publishing and interlinking semantic data from different origin and quality on the web using the Resource Description Framework (RDF)². The LOD movement has gained tremendous popularity in recent years. It has been adopted by large industries like the search engine providers Google, Yahoo!, Bing, and Yandex in 2011. Around the same time, Facebook has started to support the movement with its social graph protocol. The New York Times publishes metadata about its articles as LOD. The BBC is an early adopter of the LOD technology and other media organizations like the Library of Congress, the Swedish National Library, and the German National Library are also publishing LOD.

Today, all formats for encoding multimedia content like the W3C standards HTML5, SVG, and SMIL as well as the proprietary Flash format by Adobe foresee the use of RDF in their specifications to describe the metadata of the multimedia content. Thus, overall LOD can be considered as next big movement in the modeling of multimedia metadata and description of higher-level semantics of multimedia content. The tutorial is inspired and motivated by the vivid and controversial panel discussion “*Disputatio*” on the use of *Ontologies in Multimedia* [2] at ACM Multimedia in Firenze, Italy in 2010.

2. BENEFITS AND IMPACTS

The tutorial gives both beginners and intermediate audiences an introduction and overview of the developments in multimedia metadata. It provides insights into the current advancements of using Linked Open Data and other more expressive semantic technologies. Overall, the benefits of the multimedia community from offering this tutorial on the foundations and applications of semantic technologies for multimedia content are:

- Step-by-step introduction into the LOD technology.
- Understanding the relevance of LOD for multimedia semantics and multimedia metadata.
- Obtain basic knowledge in features of RDF as well as get basic knowledge in more advanced features of formal ontologies.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage, and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s). Copyright is held by the author/owner(s).

MM'13, October 21–25, 2013, Barcelona, Spain.

ACM 978-1-4503-2404-5/13/10.

<http://dx.doi.org/10.1145/2502081.2502165>.

¹<http://www.w3.org/DesignIssues/LinkedData>, last accessed: April 2, 2013

²<http://www.w3.org/TR/REC-rdf-syntax/>, last accessed: April 2, 2013

- Systematic overview of different existing works using LOD for multimedia metadata. The overview considers both industrial models, e.g., by Google and the other search engine providers and the New York Times, as well as current research works.
- Overview of today's applications and tools for modeling multimedia semantics and multimedia metadata using semantic technologies.

3. OUTLINE OF THE TUTORIAL

In the first part, the fundamentals of LOD will be introduced at the example of a fictitious media company in the UK³. Where relevant, current research questions and challenges in the context of the LOD movement are highlighted. In addition, different approaches for visualizing LOD and interacting with LOD will be provided. The introduction will be complemented by a brief look at more advanced features of RDF such as modeling class hierarchies, domain and range of properties, and lightweight inferencing techniques [3].

The second part of the lecture focuses on the use of LOD for describing the semantics of media content and multimedia content. To this end, different existing metadata models and metadata standards will be introduced and contrasted with the LOD movement. The use of LOD in the media sector will be demonstrated at examples such as the German National Library, the BBC, and the New York Times. Particular focus will be given on schema.org⁴, a metadata vocabulary for web pages including media content developed by the three largest search engine providers in the world, namely Google, Microsoft, and Yahoo!. Since its release in June 2011, schema.org has been widely adopted by content providers on the web. In November 2011, the Russian search engine provider Yandex has also started supporting the schema.org vocabulary. In addition, we will see more advanced approaches for describing the semantics of multimedia content by means of formal ontologies. These have been adopted from well-known and standardized metadata models such as MPEG-7 [4] or derived as an abstraction from the existing solutions for multimedia metadata [5].

Finally, in the third block an overview of existing tools and approaches using LOD in the context of multimedia content will be given. One example is the semantic image retrieval system by Setchi et al. [6] in the car designer domain. It makes use of ontological concepts derived from the text that is found around the images. Another example is the semantic indexing and search in educational videos conducted in the MEMORAe project [7]. Example of a semantic cross-media search and retrieval system is the semantic search engine Squiggle for images and audio [8]. Regarding the context-driven creation and semantic enrichment of personalized multimedia content, we find authoring tools like in the domain of digital photo books [9]. Further applications make use of contextual information and semantic information for exploring social media from various distributed information sources [10] and on mobile phones [11].

4. SUMMARY

The use of high-level semantics in form of Linked Open Data and ontologies is an essential ingredient of current and

future multimedia applications. Semantic technologies have made tremendous progress in the past years. Today, we witness the widespread adoption of such technologies by, e.g., search engine giants and large media companies. This tutorial aims at introducing into this next big movement of modeling multimedia metadata and higher-level semantics of multimedia content.

5. REFERENCES

- [1] K. Seluk Candan and Maria Luisa Sapino. *Data Management for Multimedia Retrieval*. Cambridge University Press, New York, NY, USA, 2010.
- [2] Simone Santini and Amarnath Gupta. "Disputatio" on the use of ontologies in multimedia. In *Multimedia*, pages 1723–1728, New York, NY, USA, 2010. ACM.
- [3] Dean Allemang and James A. Hendler. *Semantic Web for the Working Ontologist - Effective Modeling in RDFS and OWL*. Morgan Kaufmann, 2011.
- [4] Steffen Staab, Ansgar Scherp, Richard Arndt, Raphaël Troncy, Marcin Grzegorzec, Carsten Saathoff, Simon Schenk, and Lynda Hardman. Semantic multimedia. In *Reasoning Web, 4th International Summer School*, volume 5224 of *Lecture Notes in Computer Science*, pages 125–170. Springer, 2008.
- [5] Carsten Saathoff and Ansgar Scherp. Unlocking the semantics of multimedia presentations in the web with the multimedia metadata ontology. In *World wide web*, pages 831–840, New York, NY, USA, 2010. ACM.
- [6] Hiromitsu Hama, Thi Thi Zin, and Pyke Tin. Optimal crawling strategies for multimedia search engines. In *Intelligent Information Hiding and Multimedia Signal Processing*, pages 182–185. IEEE Computer Society, 2009.
- [7] Ghali Merzougui, Mahieddine Djoudi, and Amel Behaz. Conception and use of ontologies for indexing and searching by semantic contents of video courses. *CoRR*, abs/1201.5102, 2012.
- [8] Irene Celino, Emanuele Della Valle, Dario Cerizza, and Andrea Turati. Squiggle: a semantic search engine for indexing and retrieval of multimedia content. In *Workshop on Semantic-Enhanced Multimedia Presentation Systems*, volume 228 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2006.
- [9] Susanne Boll, Philipp Sandhaus, Ansgar Scherp, and Utz Westermann. Semantics, content, and structure of many for the creation of personal photo albums. In *Multimedia*, pages 641–650, New York, NY, USA, 2007. ACM.
- [10] Simon Schenk, Carsten Saathoff, Steffen Staab, and Ansgar Scherp. Semaplorer - interactive semantic exploration of data and media based on a federated cloud infrastructure. *J. Web Sem.*, 7(4):298–304, 2009.
- [11] Alexander Kleinen, Ansgar Scherp, and Steffen Staab. Interactive faceted search and exploration of open social media data on a touchscreen mobile phone. *Multimedia Tools and Applications*, pages 1–22, 2013.

³linkeddatabook.com/, last accessed: June 20, 2013

⁴<http://schema.org>, last accessed: April 2, 2013