Tutorial at [ISWC 2008](#)

Title:

A Semantic Multimedia Web: Create, Annotate, Present and Share your Media

- Tutorial home page: http://www.cwi.nl/~media/iswc08/
- Presenter contact information: **Raphaël Troncy**, <Raphael.Troncy@cwi.nl>, http://www.cwi.nl/~troncy/
- Duration: **Half-day**
- Scope: **Semantic Web technologies for multimedia content** ; **Search, query, and visualization of the Semantic Web** ; **Ontologies** ; **Interface Design** ; **Evaluation of Semantic Web technologies**
- Keywords:
  Multimedia Ontologies, Media Creation and Annotation, Canonoical Processes, Facetted Browser, Novel Multimedia Interfaces, Use Cases
- Categories and Subject Descriptors:
  - H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems,
  - H.5.3 [Group and Organization Interfaces]
- The tutorial materials will be provided to attendees and available on the web under the Attribution-Noncommercial-Share Alike Creative Commons licence. The tutorial web page will be regularly updated with material, examples, use cases and tools that will be demonstrated.

Tutorial Abstract

The success of content-centered (social) Web 2.0 services contributes to an ever growing amount of digital multimedia content available on the Web. Video advertisement is becoming more and more popular and films, music and videoclips are largely consumed from legacy commercial databases. Re-using such multimedia material is, however, still a hard problem. Why is it so difficult to find appropriate multimedia content, to reuse and repurpose content previously published and to adapt interfaces to these content according to different user needs?

This tutorial proposes to cover these questions. Based on established media workflow practices, we describe a small number of fundamental processes of media production. We explain how multimedia metadata can be represented, attached to the content it describes, and benefits from the web that contains more and more formalized knowledge (the Web of linked data). We show how web applications can benefit from semantic metadata for creating, searching and presenting multimedia content.

Learning Objectives, Scope and Target Audience

This tutorial is designed for practitioners, researchers and PhD students who work in creating, searching and presenting multimedia content for exchanging and sharing over the Web. The target audience will learn how to understand the semantics of various media, how to describe them, and how to make use of such descriptions in the whole multimedia creation process including management, distribution, delivery and reuse. The tutorial also targets multimedia content providers, such as TV broadcasters and news agencies, who want to sell and expose their content on the web, and industries who supply added value services in content enrichment and organization.

While the tutorial is focused on *Multimedia Semantics on the Web*, it should also be of interest to people working in: Multimedia Ontology Engineering, Multimedia on the Web, Multimedia User Interface Design, Content-Based Indexing and Retrieval, TREC Video Retrieval and Multimedia Information Retrieval.
The tutorial will include lectures, use cases and demonstrations. Being partially funded by the EU K-Space Network of Excellence, the tutorial will be widely advertised in mailing lists and among related EU research projects for maximizing participation.

Tutorial Full Description

Working with multimedia assets involves their capture, annotation, editing, authoring and/or transfer to other applications for publication and distribution. There is substantial support within the multimedia research community for the collection of machine-processable semantics during established media workflow practices. An essential aspect of these approaches is that a media asset gains value by the inclusion of information (i.e. metadata) about how or when it is created or used, what it represents, and how it is manipulated and organized. For example, users sharing photos on Flickr or Picasa Web would like to keep control of the tags and metadata associated to the media in order to automatically generate digital photo books for a specific event. Semantic search of news require new models and interfaces that could aggregate media from several sources and personalize the news to the user interests and location.

In this tutorial, we consider the use of Semantic Web technologies for improving the multimedia user experience on the Web. We explain how multimedia metadata can be represented, attached to the content it describes, and benefits from the web that contains more and more formalized knowledge. We show how web applications can benefit from semantic metadata for creating, searching and presenting multimedia content.

While many multimedia systems allow the association of semantic annotations with media assets, there is no agreed-upon way of sharing these among systems. As an initial step, and based on established media workflow practices, we identify a small number of fundamental processes of media production, which we term **canonical processes** (see Figure 1). The tutorial introduces these processes, defined in terms of their inputs and outputs and regardless of whether these processes can, or should, be carried out by a human or a machine. We illustrate these processes with two systems coming from both academic and industrial research communities: CeWe Photobook – an online photobook creation web application and Vox Populi – a system for automatic generation of argumentation-based video sequences.

Semantic descriptions of non-textual media can be used to facilitate retrieval and presentation of media assets and documents containing them. Existing multimedia metadata standards, such as MPEG-7, provide a means of associating semantics with particular sections of audio-visual material. While technologies for multimedia semantic descriptions already exist, there is as yet no formal description of a high quality multimedia ontology that is compatible with existing (semantic) web technologies. We therefore present four proposals for MPEG-7 based ontologies, and we provide a comparison of them. We describe COMM in detail, a Core Ontology of MultiMedia for annotation that extends the DOLCE upper ontology. We explain how semantic multimedia metadata can be represented, attached to the media itself and linked to other vocabularies defined in the Semantic Web. We demonstrate a semi-automatic ontology-based annotation tool for producing semantic annotations of image, audio and video content.

COMM has been designed for representing multimedia metadata, but with different media – such as text, image, video, audio – and with different applications – such as news or cultural heritage – come also a lot of different specific metadata standards and vocabularies, and the situation we found today is a web hosting a plethora of formats. For example, for still images, we find many different standards ranging from EXIF headers in photographs and MPEG-7 image descriptors to XMP/IPTC semantic information or simple user-defined tags from a Web 2.0 application. This makes life difficult for end users and application developers. We show with several use cases how web applications benefit from using multiple metadata formats. We explain how metadata interoperability can be achieved by using Semantic Web technologies to combine and to leverage existing multimedia metadata standards.

Multimedia metadata are therefore heterogenous in formats and types and Semantic Web technologies help in integrating them semantically. Underlying technologies are insufficient in their own right and users require interfaces to access these more complex data. Facet browsing and auto-completion have become popular as a user friendly interface to data repositories. Users should be able to select and navigate through facets of resources of any type and to make selections based on properties of other, semantically related, types. We present various facet browser interfaces developed within academic research projects but deployed more and more in commercial web applications. We show novel search and presentation techniques which make use of interoperability between the data and between the vocabularies, using two demonstrators in the Culturage Heritage and the News domains.
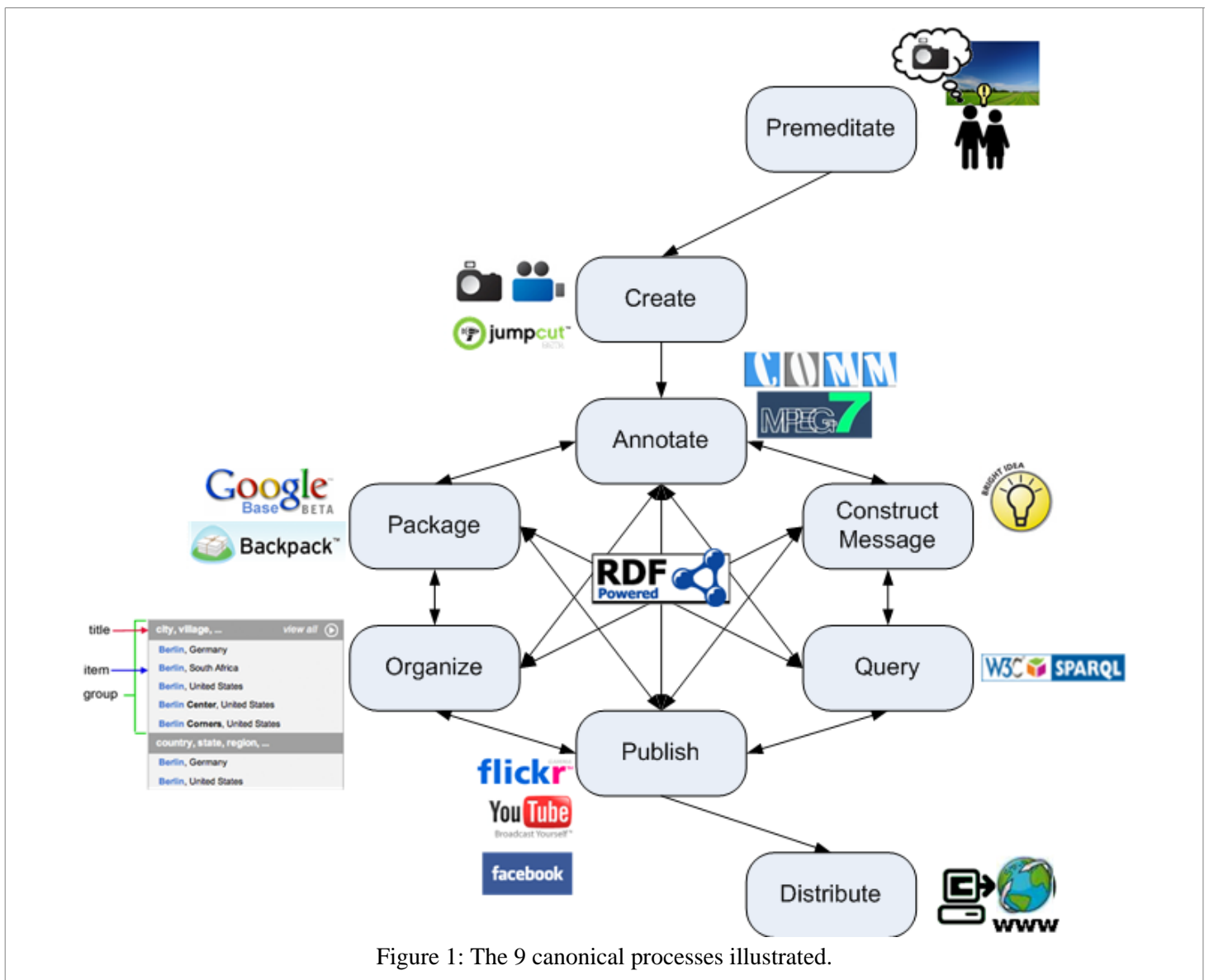
Figure 1: The 9 canonical processes illustrated.

The schedule of the tutorial is as follows:

1. **Welcome, Introduction, and Overview (5 minutes)**
   Welcome participants, find out who they are and what they want, provide overview of tutorial goals and schedule.
2. **Understanding Multimedia Applications Workflow (55 minutes)**
   - Illustrative example: an online photobook creation web application from CeWe Color
   - Illustrative example: Vox Populi, a system for automatic generation of biased video sequences using rhetorical presentations patterns together with semantic metadata
   - The Canonical Processes of Media Production 101
3. **Semantic Annotation of Multimedia Content (30 minutes)**
   - Multimedia interoperability issues: how to combine multiple metadata formats?
   - Use Case: Creating Personal or Professional Photo Books
   - Use Case: Facetting your Music Songs
   - MPEG-7 based Multimedia Ontologies
4. **Coffee Break (30 minutes)**
5. **Semantic Annotation of Multimedia Content (cont.) (30 minutes)**
   - COMM: a Core Ontology of MultiMedia for annotation
   - Annotation Tools examples and demonstrations
6. **Semantic Search and Presentation of Multimedia Content (55 minutes)**
   - Link your data!
   - Facet Browsing interfaces, auto-completion search and ranking algorithms
   - Browsing multimedia datasets: the eCulture and the News domain
7. **Wrap up, Conclusion and Q/A (5 minutes)**

History and References

Tutorial history:

This tutorial was given together with Prof. Lynda Hardman during the 17th World Wide Web Conference on April 21st, 2008 in Beijing (China). It follows the lectures Providing Flexible Interfaces to Annotated Multimedia Repositories and Multimodal Interaction
given by Prof. Lynda Hardman during the K-Space Summer School on Multimedia Semantics (SSMS) organized in Chalkidiki, Greece (2006) and Glasgow, UK (2007)
respectively. It is built on a number of real applications and use cases developed within the W3C Multimedia Semantics Incubator Group (June 2006 - August 2007). The tutorial includes theoretical work presented during the Workshop on *Multimedia for Human Communication - From Capture to Convey* at ACM Multimedia 2005 or discussed during the Panel on *The role of multimedia metadata standards in a (Semantic) Web 3.0* at WWW 2007.

Relevant references:

- Richard Arndt, Raphaël Troncy, Steffen Staab, Lynda Hardman and Miroslav Vacura. COMM: Designing a Well-Founded Multimedia Ontology for the Web. In 6th International Semantic Web Conference (ISWC'2007), Busan, Korea, November 11-15, 2007.
- Stefano Bocconi, Frank Nack and Lynda Hardman. Automatic generation of video documentaries. In Journal of Web Semantics, 2008 (*to appear*).
- Lynda Hardman. Canonical Processes of Media Production. In Proceedings of the ACM Workshop on Multimedia for Human Communication - From Capture to Convey (MHC 05), November 11, 2005.
- Michiel Hildebrand, Jacco van Ossenbruggen and Lynda Hardman. /facet: A Browser for Heterogeneous Semantic Web Repositories. In 5th International Semantic Web Conference (ISWC'2006), pages 272-285, Athens (GA), USA, November 5-9, 2006.
- Raphaël Troncy, Jacco van Ossenbruggen, Jeff Z. Pan and Giorgos Stamou. Image Annotation on the Semantic Web. W3C Multimedia Semantics Incubator Group Report (XGR), 14 August 2007.
- Vassilis Tzouvaras, Raphaël Troncy and Jeff Z. Pan. Multimedia Annotation Interoperability Framework. W3C Multimedia Semantics Incubator Group Report Editor's Draft, 14 August 2007.

Acknowledgments:

Biography of the Lecturer

Raphaël Troncy obtained his Master's thesis with honors in computer science at the University Joseph Fourier of Grenoble, France, after one year spent in the University of Montreal, Canada. He benefited from a PhD fellowship at the National Audio-Visual Institute (INA) of Paris where he received his PhD with honors in 2004. During his PhD, he taught undergraduate courses in the University René Descartes – Paris 5 (FR), and gave lectures in the INTD Bachelor of documentation on *audio-visual documentation and databases*. He has also given invited lectures at the University of Amsterdam and Glasgow University.

He was awarded ERCIM Post-Doctorate Research Associate in the National Research Council (CNR) in Pisa, Italy in 2005, and in the Centre for Mathematics and Computer Science (CWI) in Amsterdam, the Netherlands in 2006 where he is currently employed. Raphaël Troncy is co-chair of the W3C Incubator Group on Multimedia Semantics, and an active participant in the EU K-Space Network of Excellence.

His research interests include Semantic Web and Multimedia Technologies, Knowledge Representation, Ontology Modeling and Alignment. Raphaël Troncy is an expert in audio visual metadata and in combining existing metadata standards (such as MPEG-7) with current Semantic Web technologies. He also works closely with the IPTC standardization body and the relationship between the NewsML language and the Semantic Web.

- Email: Raphael.Troncy@cwi.nl
- Homepage: http://www.cwi.nl/~troncy/