Connected Media Experiences: Web Based Interactive Video using Linked Data

Lyndon Nixon
STI International & Seekda GmbH
Neubaugasse 10/15
1070 Vienna, Austria
+43 1 2364002 13
lyndon.nixon@sti2.org

Matthias Bauer
STI International
Neubaugasse 10/15
1070 Vienna, Austria
+43 1 2364002 13
matthias.bauer@sti2.org

Cristian Bara
Seekda GmbH
Neubaugasse 10/15
1070 Vienna, Austria
+43 1 2364002 13
cristian.bara@seekda.com

ABSTRACT
This demo submission presents a set of tools and an extended framework with API for enabling the semantically empowered enrichment of online video with Web content. As audiovisual media is increasingly transmitted online, new services deriving added value from such material can be imagined. For example, combining it with other material elsewhere on the Web which is related to it or enhances it in a meaningful way, to the benefit of the owner of the original content, the providers of the content enhancing it and the end consumer who can access and interact with these new services. Since the services are built around providing new experiences through connecting different related media together, we consider such services to be **Connected Media Experiences (ConnectME)**. This paper presents a toolset for ConnectME – an online annotation tool for video and a HTML5-based enriched video player – as well as the ConnectME framework which enables these media experiences to be generated on the server side with semantic technology.

Categories and Subject Descriptors
H.5.4 [Information Interfaces and Presentation]: Hypertext/hypermedia - Architectures

Keywords
Hypervideo, clickable video, Web media, Linked Data, media linking, annotation, enrichment

1. INTRODUCTION
ConnectME is a project which began in June 2011 as a nationally funded project in Austria. The participating partners are STI International, Salzburg Research, PS Media and Yoovis GmbH. The goal of ConnectME is to develop a hypervideo platform based on open Web standards for the delivery of interactive video experiences and Web services which support the conceptual annotation of video, Web-based linkage between concepts and content, and on-the-fly augmentation of video with content including aspects of personalisation and contextualisation. In this submission, we present the Web based annotation tool for video, which generates storable and sharable RDF based media annotations, the Web based hypervideo player, and the ConnectME framework, which extends an existing system known as the Linked Media Framework, which handles the server side processing from the media annotations to the final content for the enriched video. To the best of our knowledge this is the first Web based video annotation tool supporting use of Linked Data as well as first Web based hypervideo player dynamically enriching videos with content based on Linked Data annotations.

WWW 2013 Companion, May 13–17, 2013, Rio de Janeiro, Brazil.
ACM 978-1-4503-2038-2/13/05.

2. BACKGROUND & RELATED WORK
While hypervideo – the idea of hyperlinking to content from within video – has been around since the 1980s[1], the combination of online video, semantic annotation and Web linking in ConnectME is to the best of the authors' knowledge unique in the field. Online video is a clear trend in media consumption, yet the automated association of videos to related Web material is still a subject of technology demos like Mozilla's Popcorn[2] which uses textual tags associated to video to link into Wikipeda articles, maps and so on. Semantics could solve the inherent ambiguities of textual tagging. Work on semantic annotation of video has focused on using the rich metadata captured in improving multimedia indexing, search and retrieval, but the role it could play in enabling an enriched playout of the video is taken up anew in ConnectME. Traditionally, multimedia presentation systems [1] have indeed relied on formal knowledge about the multimedia but not agreed on a shared model for that knowledge. Earlier work on the Cuypers presentation engine [2] did explore use of RDF based knowledge models [3]. The emergence of Linked Data has meant semantic annotations can refer to freely accessible Web based metadata which can be reused in UIs for content selection and browsing, but work has gone not much further than the limited media linked to directly from Linked Data descriptions [4]. Automated linking from semantic annotations to online content related to the annotation needs to incorporate Multimedia Information Retrieval techniques [5] and benefit from increased publication of media metadata in a structured/Linked Data form [6]. The state of the art in Web hypervideo today does not have answers to these issues being addressed by research in ConnectME, and hence focuses on manual annotation and linkage to other content in the video (see Web based offers by companies such as WireWax, Videoclix, Overlay.TV or Klickable). The LinkedTV project (www.linkedtv.eu) is addressing automation of the video analysis and annotation for hypervideo playout: STI

[1] Systems such as Hypersoap (www.media.mit.edu/hypersoap/) demonstrated the possibility of interactive product placement in a broadcast setting
[2] http://webmademovies.etherworks.ca/popcorndemo/

International is the scientific coordinator and brings its experience from ConnectME into this project.

2. CONNECTME WORKFLOW

The ConnectME workflow uses a set of executable Web services called from a server side platform which also provides for the workflow's data storage and retrieval in order to generate, from the starting point of a semantic annotation of an online video, a final set of content linked to spatial and temporal moments in the video that can be played out as a form of dynamic content enrichment in the ConnectME hypervideo player. Figure 1 provides a high level view of this workflow. The main steps in the workflow, printed on the left, are to identify objects in video, annotate them with (Linked Data[3]) concepts and make use of this annotation to link the video objects to other Web content.

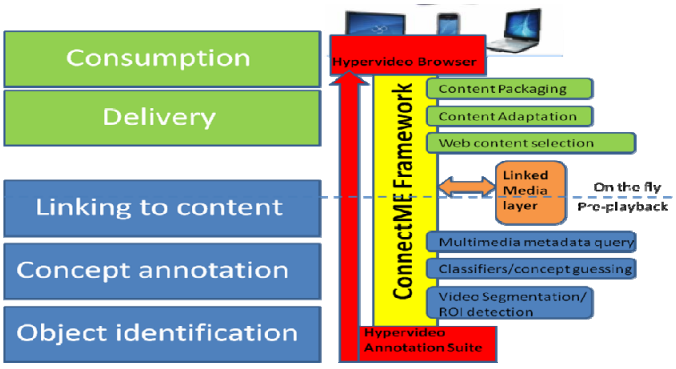

Figure 1. ConnectME workflow

We have implemented a working set of tools to demonstrate Connected Media Experiences by (Section 3.1) annotating a video with concepts, (Section 3.2) letting the server side platform use concept metadata to extract links to Web content which is relevant, and (Section 3.3) an interactive video player for enabling a viewer to access and browse the enrichment information during the program.

3. THE CONNECTME TOOLS

3.1 Annotation Tool

ConnectME has developed a Web-based hypervideo annotation tool in PHP. The user interface uses HTML5, the Video.js player and jQuery with extensions to provide for video loading and manipulation, such as selecting spatial and temporal parts of the video, and hence works across all latest versions of Web browsers. Using the HTML 5 video tag for embedding video files allows playing videos without need for any Flash-based plug-in and the Video.js library provides several useful video control methods. The hypervideo annotation tool provides an easy and intuitive timeline based user interface. The timeline allows adding annotations to the video as well as editing existing ones. Furthermore, changing annotation times, concepts or spatial regions are user-friendly because most actions are made by dragging elements or clicking on them.

Ajax powers as-you-type concept suggestion from the DBPedia concept base to support annotators in quickly finding the right concept: also when a concept is selected, the annotation tool shows some explanatory text (the DBPedia concept's abstract) to help annotators be sure they choose the correctly intended concept. Furthermore, the annotation tool supports searching for geographic locations in Geonames. The map preview displayed by using the Google Maps JavaScript API V3 helps ensure annotators choose the intended place. Locally the annotation being made is stored as JSON in a file on the server and the file itself gets referenced by using a Cookie, so that even if the browser page is reloaded or closed/opened again, the annotation task can be continued without loss of information.

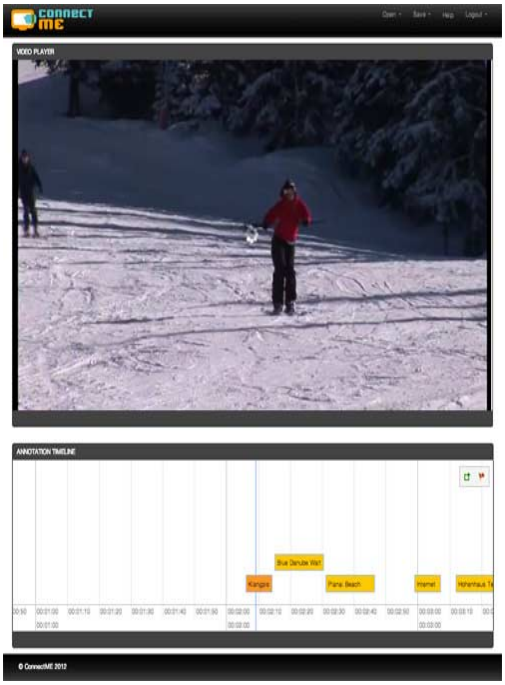

Figure 2. ConnectME annotation tool. Below the timeline which visualises which concepts are annotated where in the video.

The annotator can also save the created annotations locally in a file and select between the serialisations N3, RDF/XML and Turtle. This allows previewing the work before submitting it to a ConnectME Framework instance or storing the annotation file locally on the annotators computer.

When the annotation is saved, the tool is directly connected to an instance of the ConnectME framework which stores the annotations and makes them available to the ConnectME workflow when the video is requested from the hypervideo player. For the video annotation schema a RDF based format has been selected which is based around the W3C Media Ontology[4] with some extensions for enabling an annotation of the annotations (who made them, when, what are the rights for re-use) re-using the Open Annotation Data Model[5], and a backwards compatible ConnectME specific extension for describing how concepts are represented by the video object, which is leveraged in the framework when determining the linkage to Web content. Video fragments are referenced using the W3C Media Fragments recommendation[6].

[3] http://linkeddata.org provides a Web based concept space with URI based metadata look-up for more information about concepts

[4] http://www.w3.org/TR/mediaont-10/

[5] http://www.openannotation.org/spec/core/

[6] http://www.w3.org/TR/media-frags/

3.2 ConnectME Framework

ConnectME draws on the concept of *Linked Media* to enable a Web based connection between concepts from the Linked Data concept space and Web content which, for the purposes of this linking, have been annotated in terms of Linked Data concepts. The key principles of Linked Media are:

- Web media needs to be annotated in terms of its online parts
- Web media needs to be annotated with terms which represent a shared understanding of a domain or identification of a thing
- Web media needs to be annotated using a media ontology which supports the above two issues
- The expressed representation of different concepts by different media fragments in different ways shall be the basis to interlink media across the Web

The first three points are covered in the annotation tool (W3C Media Fragments specification, Linked Data as concept namespace, W3C Media Ontology and extensions as annotation scheme). The fourth point is part of a Linked Media implementation in our framework. The Linked Media Framework[7] is an easy-to-setup server application that bundles central Semantic Web technologies to offer advanced services. The Linked Media Framework consists of LMF Core and LMF Modules. The core component of the Linked Media Framework is a Linked Data Server that allows to expose data following the Linked Data Principles. The Linked Data Server implemented as part of the LMF goes beyond the Linked Data principles by extending them with Linked Data Updates and by integrating management of metadata and content and making both accessible in a uniform way.

In order to implement the Connected Media Framework, we chose to build upon the Linked Media Framework, which already offered out of the box much of the necessary basis functionality such as storage and retrieval of the semantic media annotations, as well as a means to access media or its metadata in a straightforward manner, following Linked Data principles. Since additional functionalities are plugged in via modules, ConnectME develops its own specific modules to turn the Linked Media Framework into a Connected Media Framework: The *concept extraction module* supports the video annotation tool by suggesting concepts to link to the video via textual analysis of available subtitling or transcript files for that video. For this, an instance of Apache Stanbol has been specifically trained to handle the particular corpus of concepts in ConnectME materials. The *Linked Media engine* is a specific component implementation which exposes multimedia object descriptions to the ConnectME workflow in a common structured metadata format. The provision of usable multimedia object descriptions on the Web as Linked Data is referred to in the project as the Linked Media layer. To find objects relevant for any concept in the video annotation, media repositories need to be queried and their responses provisioned as Linked Media. Hence the engine incorporates a semantic service middleware which brokers between ConnectME and heterogeneous media sources (Web APIs, SPARQL endpoints, etc.).

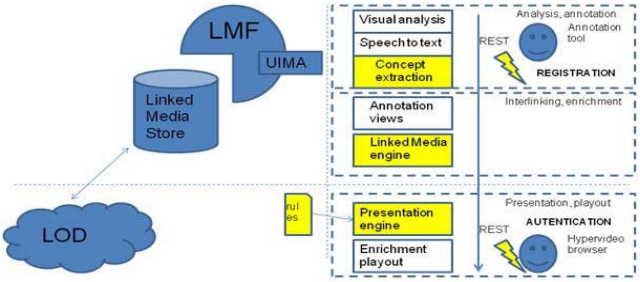

Figure 3. ConnectME extending the Linked Media Framework.

3.3 ConnectME Hypervideo Player

The current implementation of the hypervideo player is based on the open source LIME player[8] (making us of the HTML5 video tag, video.js, JQuery, CSS3, JSON2, backbone.js, underscore.js, RDFQuery and VIE library) and runs in any latest version of any of the main browsers. The (desktop) Web based version which responses to mouse events has been adapted to both run in a smartphone/tablet (with touch interaction) as well as over Google TV (reacting to remote control events). The player incorporates support for the W3C Media Fragment syntax that should allow video to be accessed not as an entire media resource but in terms of a temporal and/or spatial part thereof. As the video plays, a Javascript code checks for annotations on the next active video segment, and enables access to additional content when it is relevant to the concept annotating that segment via a plugin and widget architecture. Annotations refer to Linked Data resources and the ConnectME framework has collected links to content relevant to those resources. The hypervideo player has a core that sustains the video playback mechanism and connects to the ConnectME Framework to retrieve the annotations in an initialization phase. A set of plugins is then attached to the core, each of which is specialized in recognizing a certain type of annotation resource. Plugins will retrieve and render relevant content for given resources and display them in the form of widgets. Widgets appear and disappear (as the related concept is present, and no longer present, in the video) from the right hand side list, by default only active widgets are shown (to reduce the distraction of the viewer from the video) but for navigation the full list can be accessed at any time. To additionally support the viewer to navigate within hypervideo, markers on the timeline indicate when concepts are present and by bringing a marker into focus (e.g. mouseover) a pop-up shows the viewer which concept(s) are present in that point of the video. In Figure 4a, a geolocation plugin has rendered geolocation-related widgets for the concept of Schladming (a place in Austria) that produce weather maps, route maps and satellite imagery maps; an info plugin fetched the abstract and concept label fields from DBPedia that describes Schladming and composed an information widget. Each plugin marks it's widgets with a specific icon. Since plugins can be configured for any Linked Data source, the player architecture is very flexible regarding the content selected and displayed in a widget. The video player allows concept based viewing. For this we defined the lifecycle of the widgets as follows: 1) a widget is instantiated in a dormant state called "inactive" in the initialization phase of the video player. This widget will have a gray icon 2) a widget becomes "active" when the timeline cursor of the player is within the time interval

[7] http://www.newmedialab.at/LMF

[8] https://github.com/tkurz/lime

described by it's associated annotation. Active widgets are described by colored icons. 3) after the timeline cursor exits the time interval of the annotation, the widget becomes again inactive. Only active widgets display their extended information, e.g, Figure 4b shows a extended active information widget for the concept of Schladming.

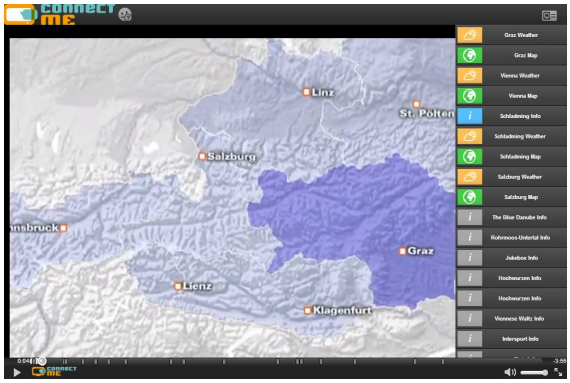

Figure 4a (above). ConnectME hypervideo player. Information about the locations mentioned/shown in the video is available.

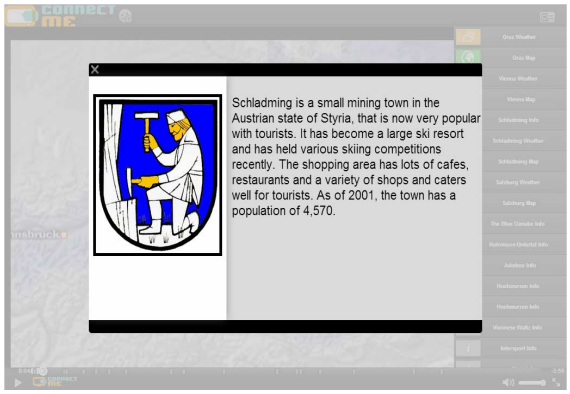

Figure 4b (below). ConnectME hypervideo player. Info widget click/tap results in a overlapping information panel

4. CONCLUSION

The ConnectME annotation tool, framework and player act as a proof of concept for semantics-based dynamic enrichment of videos based on Linked Data annotations. In the project we continue to explore how we can better automate and simplify the annotation step for the user, maximize the flexibility and relevance of the linkage to online media resources, and improve the intuitiveness of the user interaction with the hypervideo player in order to easily and effectively access and browse video enrichments, on desktop, tablet and SmartTV platforms. For the demo session we can show the annotation tool and the hypervideo video player, and if desired the administration interface of the ConnectME framework can also be shown. **A screencast of the annotation tool and hypervideo player can be seen at** http://bit.ly/1511uvs

5. ACKNOWLEDGMENTS

Our thanks to the ConnectME project partners Salzburg Research, Yoovis GmbH and PS Media who also contribute to the technical work and scenario evaluation. Linked Media Framework and LIME Player are open source projects led by Salzburg Research. This work is also partially supported by the Integrated Projects LinkedTV (www.linkedtv.eu) and ExperiMEDIA (www.experimedia.eu), funded by the European Commission through the 7th Framework Programme (FP7-287911, 287966)

REFERENCES

[1] L. Nixon, L. Aroyo and L. Miller. "NoTube: the television experience enhanced by online social and semantic data", at the 1st International Conference on Consumer Electronics (ICCE 2011), Berlin, Germany, September 2011.

[2] Bordegoni, M.; Faconti, G.; Maybury, M.T.; Rist, T.; Ruggieri, S.; Trahanias, P.; Wilson, M., "A Standard Reference Model for Intelligent Multimedia Presentation Systems", Computer Standards & Interfaces, 18:477-496, 1997.

[3] Ossenbruggen, J. R., Cornelissen, F. J., Geurts, J. P., Rutledge, L. W., and Hardman, L. 2000 *Cuypers: a Semi-Automatic Hypermedia Generation System*. Technical Report. UMI Order Number: INS-R0025., CWI (Centre for Mathematics and Computer Science).

[4] Lindley, C. A., Davis, J. R., Nack, F., and Rutledge, L. W. 2001 *The Application of Rhetorical Structure Theory to Interactive News Program Generation from Digital Archives*. Technical Report. UMI Order Number: INS-R0101., CWI (Centre for Mathematics and Computer Science).

[5] Schreiber, G., et al., Semantic annotation and search of cultural-heritage collections: The MultimediaN E-Culture, Demonstrator. Journal of Web Semantics (JWS), 6(4):243–249, 2008.

[6] A. Hanjalic, R. Lienhart, W.-Y. Ma, and J. R. Smith. "The Holy Grail of Multimedia Information Retrieval: So Close or Yet So Far Away?" Proceedings of the IEEE, 96(4):541–547, April 2008.