# Connected Media Experiences: interactive video using Linked Data on the Web

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 paper presents a set of tools and an extended framework with API for enabling the dynamic enrichment of online video with Web content via Linked Data. 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 the ConnectME approach, which is to annotate video with Linked Data concepts, dynamically bind annotations to related Web content and then provide a flexibly enriched video playout. Using Linked Data allows us to decouple video annotations from hardcoded enrichments, meaning a video can be annotated once but its enrichment is always up to date.

## 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 (research.sti2.org), Salzburg Research (www.salzburgresearch.at), PS Media and Yoovis GmbH (www.yoovis.tv). 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. The latter 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.



LDOW 2013. May 14 2013, Rio de Janiero, Brazil

## 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[2], yet the automated association of videos to related Web material is still a subject of technology demos like Mozilla's Popcorn[3] 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

[1] Systems such as Hypersoap (www.media.mit.edu/hypersoap/) demonstrated the possibility of interactive product placement in a broadcast setting

[2] e.g. http://www.reelseo.com/video-50-bandwidth-consumption/ reports from 2010 how video has reached over 50% of US Internet bandwidth consumption

[3] http://webmademovies.etherworks.ca/popcorndemo/

the Cuypers presentation engine [2] did explore use of RDF based knowledge models [3]. The emergence of Linked Data has meant media annotations can refer to freely accessible Web based metadata about the referred-to concept, which can be reused in UIs for content selection and browsing, but work has gone not much further than the limited associated media linked to directly from Linked Data descriptions [4] such as with `foaf:depiction`. 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). On the other hand, in the Semantic Web / Linked Data community, the NoTube project promoted semantic annotation of TV programming for personalised recommendations [7], and Mediaglobe also provides a RDF based hypervideo player for cultural heritage information in video archives [8], but neither go as far as to consider how semantic annotations of media could be used to automatically interlink media information in a Web system.

## 2. CONNECTME WORKFLOW

The ConnectME workflow generates, 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) concepts and make use of this annotation to link the video objects to other Web content.

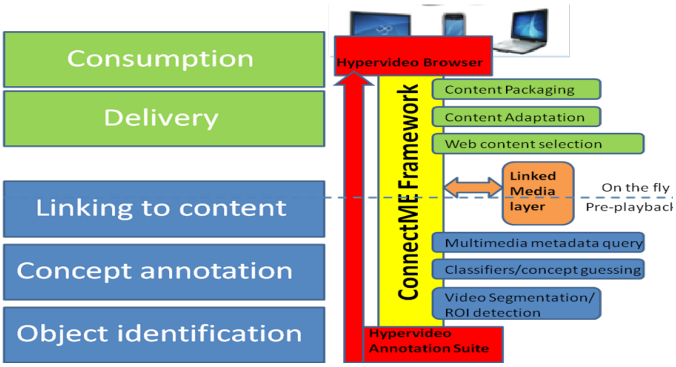



Figure 1. ConnectME workflow

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 (http://annotator.connectme.at). 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. Figure 2 shows the tool being used to annotate a video about the Austrian tourism resort of Schladming.

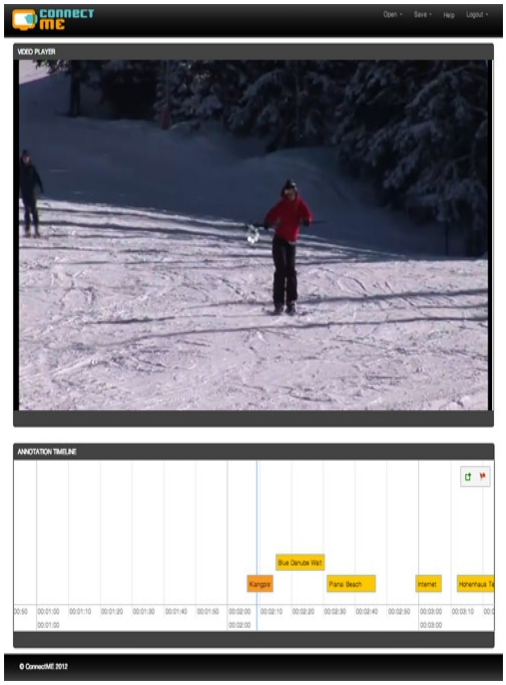



Figure 2. ConnectME annotation tool. Below the timeline which visualises which concepts are annotated where in the video.

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].

## 3.2 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 core component of the Linked Media Framework is a Linked Data Server that allows to expose data following the Linked Data Principles and integrates management of metadata and content and making both accessible in a uniform way. As extension for the LMF Core, a number of optional modules can be used to extend the functionality of the Linked Media Server:

- LMF Permissions implements and extends the WebID and WebACL specifications for standards-conforming authentication and access control in the Linked Media Framework
- LMF Media Interlinking will implement support for multimedia interlinking based on the work in the W3C Multimedia Fragments WG and the W3C Multimedia Annotations WG
- LMF Reasoner implements a rule-based reasoner that allows to process Datalog-style rules over RDF triples; it is based on the reasoning component developed in the KiWi project, the predecessor of the LMF
- LMF Versioning implements versioning of metadata updates; versioning itself is already carried out by LMF Core, but the management of versions will be carried out by this module
- LMF Enhancer offers semantic enhancement of content by analysing textual and media content; the LMF Enhancer will build upon UIMA, Apache Tika, and the semantic lifting engine of the Apache Stanbol framework[8].

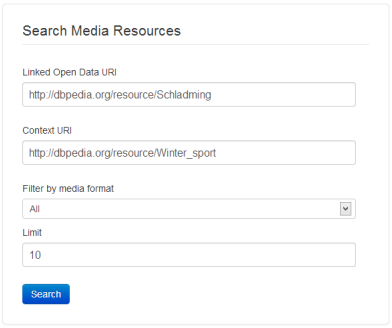
Since additional functionalities are plugged in via modules, ConnectME develops its own specific modules to turn the Linked Media Framework into a Connected Media Framework.

The media interlinking and enrichment step is handled by a specific component implementation which retrieves conceptually relevant online media and exposes their descriptions to the ConnectME workflow in a common structured metadata format. To find media objects relevant for any concept in the video annotation, different media repositories need to be queried and their responses aggregated and provisioned with a common metadata description. Hence for ConnectME we have developed a specific instance of semantic service middleware which brokers between ConnectME and heterogeneous media sources (Web APIs, SPARQL endpoints, etc.), called the *Linked Services Infrastructure (LSI)* (http://production.sti2.org/lsi).

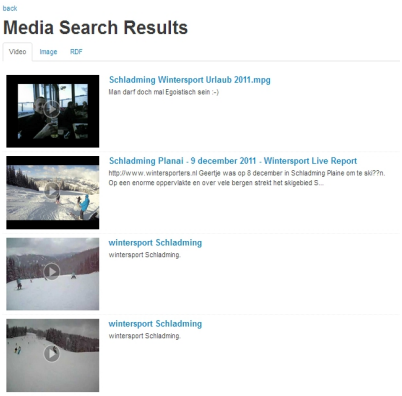
The idea behind Linked Services [9] is to make use of Semantic Web Service technology in describing, selecting and executing Web APIs. The iServe[9] [10] platform, developed in the Open University, acts a repository for the descriptions of the Web APIs. These descriptions are themselves semantic, using the specifications hRESTs[11] and MSM[10] to generate a RDF based machine processable description of the Web APIs inputs and outputs. Lifting and lowering rulesets (in XSLT extended by SPARQL based inline queries) define how a Linked Data URI may be mapped into a request to the Web API, and how the Web API response can be mapped to the LSI's RDF model (based on W3C Media Ontology) for a consistent and aggegrated response of relevant media resource descriptions to the client. The REST API allows a single service to be queried (e.g. based on its description, such as only services who have content for a particular type of concept), a concept to be mapped to media resources by a single service call, or a concept to be mapped to media resources aggegrated from all services called (based on their relevance to the concept in the query). Parameters allow that only image or video content is returned, a limit placed on the amount of media contents in the response, and whether Zemanta[11] is used for Named Entity Recognition in the media resource titles and descriptions (this can reduce the latency in waiting on a response from LSI). Named Entity Recognition is used with an optional additional "context" parameter: the idea is that aggregating media content on the basis of a single concept can be very general (e.g. "Schladming"), while the provision of an additional 'contextual' concept (e.g. "Winter sports") enables more focused media content responses from the LSI. To determine sets of concepts related to a media item, we extract named entities as Linked Data URIs from the free text metadata of those media items such as their titles, descriptions, or user-provided tags. We

[4] http://www.w3.org/TR/mediaont-10/
[5] http://www.openannotation.org/spec/core/
[6] http://www.w3.org/TR/media-frags/
[7] http://www.newmedialab.at/LMF
[8] http://incubator.apache.org/stanbol/
[9] http://iserve.kmi.open.ac.uk/
[10] http://iserve.kmi.open.ac.uk/wiki/index.php/IServe_vocabulary
[11] http://developer.zemanta.com/

work on using the Linked Data graph to map contextual concepts to related concept sets, for example "Winter Sport" as the context to a query on "Schladming" should return images and videos related to instances of winter sports at the venue, e.g. a video annotated with the concepts "Schladming + skiing" will match a LSI query for "Schladming + Winter sports", based on the semantic closeness of the concepts within the DBPedia graph (we map instances to DBpedia categories and use the path between respective categories to calculate instance closeness). Currently we have implemented the Linked Services Infrastructure with support for 6 public Web APIs[12]: Foursquare Venues, Ookaboo, YouTube, Play.fm (audio mixes), Instagram, and Flickr.




Above: LSI Web interface to the RESTful API, query on Schladming and Winter Sport




Above: video results for the query "Schladming" and "Winter Sport"

## 3.3 ConnectME Hypervideo Player

The current implementation of the hypervideo player is based on the open source LIME player[13] (making use of the HTML5 video tag, video.js, JQuery, CSS3, JSON2, backbone.js, underscore.js, RDFQuery and VIE library) and runs in the latest version of any of the main browsers (http://player.connectme.at). The (desktop) Web based version which responds 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 using the Linked Services Infrastructure. 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 Schladming (dbpedia.org/resource/Schladming) 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, see Figure 4b. Each plugin marks its 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.

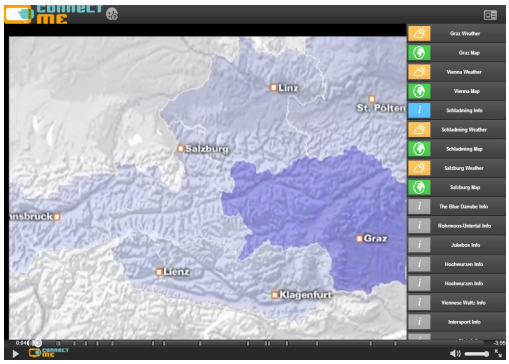



Figure 4a. ConnectME hypervideo player. Information about the locations mentioned/shown in the video is available.

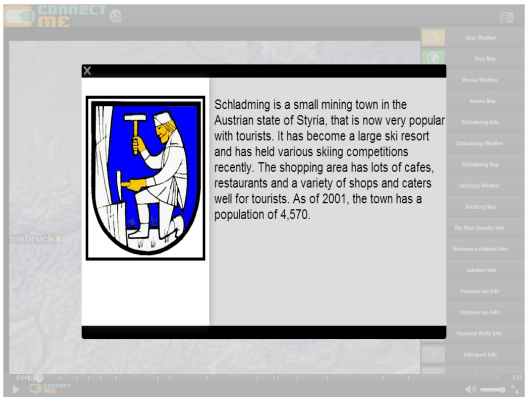



Figure 4b. ConnectME hypervideo player. Info widget click/tap results in a overlapping panel containing DBpedia information

[12] http://production.sti2.org/lsi/documentation.html
[13] https://github.com/tkurz/lime

**A screencast of the annotation tool and hypervideo player can be seen at http://bit.ly/1511uvs**

## 4. CONCLUSION & FUTURE WORK

Linked Data has proven in many use cases to be useful as a means to unambiguously identify concepts in Web systems and make processable metadata about those concepts available. However, the resulting informational display based on Linked Data annotations has been typically reduced to pretty-printing the content of triples, with limited media content connected to the Linked Data URI[14]. A novelty of our approach is the *Linked Services Infrastructure* which connects online media resources with the conceptual annotation. The ConnectME annotation tool and player demonstrate how LOD and HTML5 can make dynamic enrichment of videos based on Linked Data annotations accessible to non-technical users, both to the video annotators and the end consumer accessing the enriched video. The work will be evaluated both from the side of the consumer (consuming enriched video) and the administrator (creating enriched video), with a focus on tool usability and the end user experience, with a winter experiment at Schladming leading to some initial insights [12]. The player UI is currently being redesigned to better reflect the outcomes of this first trial. In the context of the larger EU project LinkedTV we collaborate with partners on video analysis, speech to text, face and object detection techniques. These may be able to support an automated pre-annotation of video in the ConnectME workflow, thus lowering the manual annotation effort to that of results correction. We also plan to expand the Linked Data and online Web media sources used in the workflow, in order to leverage other – perhaps more specialised – sources for metadata and content.

## 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. ConnectME is funded by the Austrian funding agency FFG (http://www.ffg.at). Linked Media Framework and LIME Player are open source projects led by Salzburg Research. Linked Services Infrastructure is programmed by Dong Liu based on work done at the Open University. The Schladming demo was created in the context of the MediaCONNECT experiment (http://www.experimedia.eu/experiments/media-connect). 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] 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.

[2] 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).

[3] 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).

[4] 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.

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

[6] Nixon L., Dasiopoulou S., Evain JP., Hyvönen E., Kompatsiaris I. and Troncy R. "Multimedia Broadcasting and eCulture", Chapter in the book "Handbook of Semantic Web Technologies", Springer, 2011. ISBN 978-3-540-92912-3

[7] Nixon L., Aroyo L. and Miller L. "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.

[8] Hentschel C., Hercher J., Knuth M., Osterhoff J., Quehl B., Sack H., Steinmetz N., Waitelonis J. and Yang H. "Open up Cultural Heritage in Video Archives with Mediaglobe". At the 12th International Conference on Innovative Internet Community Systems (IICS 2012), Trondheim, Norway, June 2012.

[9] Pedrinaci, C. and Domingue, J. (2010) Toward the Next Wave of Services: Linked Services for the Web of Data, Journal of Universal Computer Science. http://oro.open.ac.uk/23073/

[10] Pedrinaci, C., Liu, D., Maleshkova, M., Lambert, D., Kopecky, J. and Domingue, J. (2010) iServe: a Linked Services Publishing Platform, Workshop: Ontology Repositories and Editors for the Semantic Web at 7th Extended Semantic Web Conference.http://oro.open.ac.uk/23093/

[11] Kopecky, J., Gomadam, K. and Vitvar, T. (2008) hRESTS: an HTML Microformat for Describing RESTful Web Services, IEEE/WIC/ACM International Conference on Web Intelligence, Sydney, Australia. http://oro.open.ac.uk/24831/

[12] Grubert J., Reitmayr G., Nixon L. and Bara C., MediaConnect Experiment Progress Report. Published online at http://de.scribd.com/doc/132245233/D4-5-2-MediaConnect-Experiment-Progress-Report-v1-0

[14] e.g. compare http://sig.ma/search?q=schladming with Figs 3, 4b