Adding Wings to Red Bull Media

Search and Display semantically enhanced Video Fragments

Thomas Kurz
Salzburg Research
Jakob-Haringer Str. 5/3
A-5020 Salzburg
thomas.kurz@
salzburgresearch.at

Sebastian Schaffert
Salzburg Research
Jakob-Haringer Str. 5/3
A-5020 Salzburg
sebastian.schaffert@
salzburgresearch.at

Georg Güntner
Salzburg Research
Jakob-Haringer Str. 5/3
A-5020 Salzburg
georg.guentner@
salzburgresearch.at

Manuel Fernández
Red Bull Media House
Oberst Lepperdinger Str. 11
A-5020 Salzburg
manuel.fernandez@
redbullmediahouse.com

ABSTRACT

The Linked Data movement with the aims of publishing and interconnecting machine readable data has originated in the last decade. Although the set of (open) data sources is rapidly growing, the integration of multimedia in this Web of Data is still at a very early stage. This paper describes, how arbitrary video content and metadata can be processed to identify meaningful linking partners for video fragments - and thus create a web of linked media. The video test-set for our demonstrator is part of the Red Bull Content Pool[1] and confined to the *Cliff Diving* domain. The candidate set of possible link targets is a combination of a Red Bull thesaurus, information about divers from www.redbull.com and concepts from DBPedia[2]. The demo includes both a semantic search on videos and video fragments and a player for videos with semantic enhancements.

Categories and Subject Descriptors

H.5.1 [**Information Systems**]: Multimedia Information Systems; H.5.4 [**Information Systems**]: Hypertext / Hypermedia; D.0 [**Software**]: General

Keywords

linked data, semantic web, semantic search, media player, media fragments

1. INTRODUCTION

In 2007 the Linking Open Data (LOD) community project was initiated by the W3C [1]. The main goal was to bootstrap the Semantic Web by publishing datasets using standards like the Resource Description Framework (RDF)[2]. Ideally this approach will foster the interlinking of (open) data on the semantic web by identifying and using already existing sets of open data available on the World Wide Web (WWW), and of course by creating new linked datasets [3]. Unfortunately multimedia assets like videos in the context of Linked Data (LD) played until recently a subsidiary role. In order to improve this situation, the W3C has initiated the Video in the Web activity, where the associated Working Groups have already published first results: The W3C Media Fragments Working Group[3] is working on a recommendation to address temporal and spatial media fragments in the Web using URIs. To find a common media description for many different media objects and formats on the web, the W3C Media Annotations Working Group[4] provides an ontology [4] and API [5] designed to facilitate cross-community data integration of information related to media on the Web.

Interlinking multimedia data and metadata on the Web in a standardized way is the basis task of Linked Media. For the Web of textual data there are several interlinking frameworks, which try to detect related resources in different datesets and to create links to them. In [6] several frameworks are compared to each other with respect to their functionalities. Because the common interlinking methods are used on resources dominated by text, they are usuallly not sufficient for standalone multimedia data, but there are approaches which use media surrounding [7] e.g. the title of the page, descriptions above or beneath, etc. Also there are first steps aligning multimedia and events [8] and some other special use cases like the linking of image libraries and semantic resources [9].

Considering these research results, we applied the methods and standards in a real world scenario, namely the video content in the Red Bull Content Pool with the following objectives a) improving search on video content and b) providing relevant background information for videos during consumption. In the next sections we describe our approach and give an overview on our running demo application.

[1]http://www.redbullcontentpool.com/
[2]http://dbpedia.org/

[3]http://www.w3.org/2008/WebVideo/Fragments/
[4]http://www.w3.org/2008/WebVideo/Annotations/

WWW 2012 Companion, April 16–20, 2012, Lyon, France.
ACM 978-1-4503-1230-1/12/04.

2. SEMANTIC VIDEO ENHANCEMENT

To weave common video content into the Web of Data there is a need for a semantic-lifting preprocessing step. In this section we describe, how data and metadata in the Red Bull Content Pool is represented today and how we use existing information for linking semantic concepts to time- and region-based video fragments.

2.1 Red Bull media content

The Red Bull Content Pool designed and operated by Red Bull Media House GmbH is the central repository of media content related to sports events organized by Red Bull, e.g. the Air Race, the Cliff Diving Competition, or the Red Bull Rampage mountainbike race. Media content is mostly raw or processed video material that Red Bull offers to other media providers in different formats and quality for further use. Typically, the content is also annotated with the event, the year, the location, and the involved athletes. Furthermore there are attached documents representing transcriptions of spoken text, describing shots, music-lists etc.

Even there is a lot of additional information, it is not well structured and mostly text-based. More importantly, besides rudimentary metadata like title and keywords, it is neither used for advanced search nor displayed to users to get a deeper understanding of the resource. To overcome these problems we applied a semantic analysis process and build a semantic search and player application that integrates all available information.

2.2 Fragmentation and Analysis

The semantic lifting process of Red Bull video content can be separated into two steps: a) the identification of media fragments based on structured metadata (e.g. shot descriptions) and b) the extraction of semantic concepts like persons, locations etc. from textual content. As the basic concept set should integrate diverse datasets (e.g. redbull.com, DBPedia, etc.) we use a semantic lifting engine that adds semantic information to 'non-semantic' pieces of content. A comparison of several lifting engines (Apache Stanbol Enhancer[10], DBPedia Spotlight[5], Lupedia[6]) with test data from our real-world scenario has shown that none of them can be used for full-automatic annotation naturally (the results are mostly bad). Tuning the engines with integrated datasets increases the number of true positives. In our case we combined a Red Bull specific thesaurus with Open Data sets about locations, buildings etc. We extracted the thesaurus from existing information on the Red Bull Content Management System, which is actually used to generate the B2C platform redbull.com. This integrated corpus builds the base-index for our enhancement process implemented by a customized Stanbol instance. We link the outcome to resources and their fragments which results in a graph structure like depicted in figure 1. We use the fragments representation taking hashes as recommended by W3C Media Fragments WG. The Media Ontology 1.0 (media-ont prefix in the sample) link fragments and concepts. To represent annotations plus additional information (like author etc.) we use the Open Collaboration Annotation model [11] (prefix oac).

[5]http://spotlight.dbpedia.org/demo/index.html
[6]http://lupedia.ontotext.com/

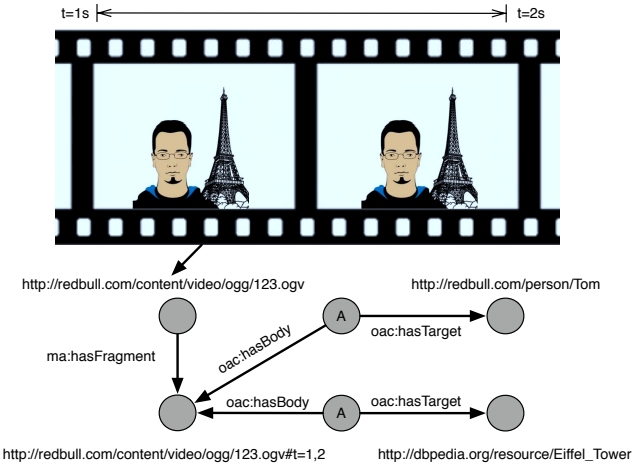

Figure 1: A sample graph for enhanced video content.

3. SEARCH AND PLAYER

The enhancement steps described in the previous section produce an RDF-graph containing the video resource itself, video fragments and extracted concepts. Because we link internal and external content this graph may expand on many external data sets and thus include lots information that can be utilized in many different ways. For our demo we implemented an improved search on videos and a player that shows both the video and the aligned semantic data, which promises a richer user experience.

3.1 Semantic video search

Considering semantic relations improves the search through big datasets. Especially for videos, which have per se a closed character and can not be overviewed as simple as e.g. images, result quality can additionally be increased by considering video fragments. E.g. if a user searches for 'helicopter', she most probably tries to find scenes where a helicopter occurs, not the whole videos where the helicopter is hidden within seconds 180 to 184.

In our demo (a sample search: http://labs.newmedialab.at/RedBull/video/redbull/search.html#*:*) we take into account semantic concepts to provide a faceted search. To narrow down the result set users can select one or more facet properties on the right side. It is also possible to reduce the results to fragments (based on its type). The result set is displayed in the center (using paging) and supplies previews as well as information on types, duration and summary. The single results link directly to the extended video player.

3.2 Extended video player

Semantic enhanced media extends / enriches the user experience when consuming multimedia content on the web. In this section we describe our proof of concept of an html5 based video player that is able to display both, video and metadata. As described in figure 2, the central part of the linked media player is the video itself. Around the video the user is able to retrieve more information about the current action. This information is based on the time- and region-based metadata that was extracted in the semi-automatic pre-processing. Regions within the video that correspond to

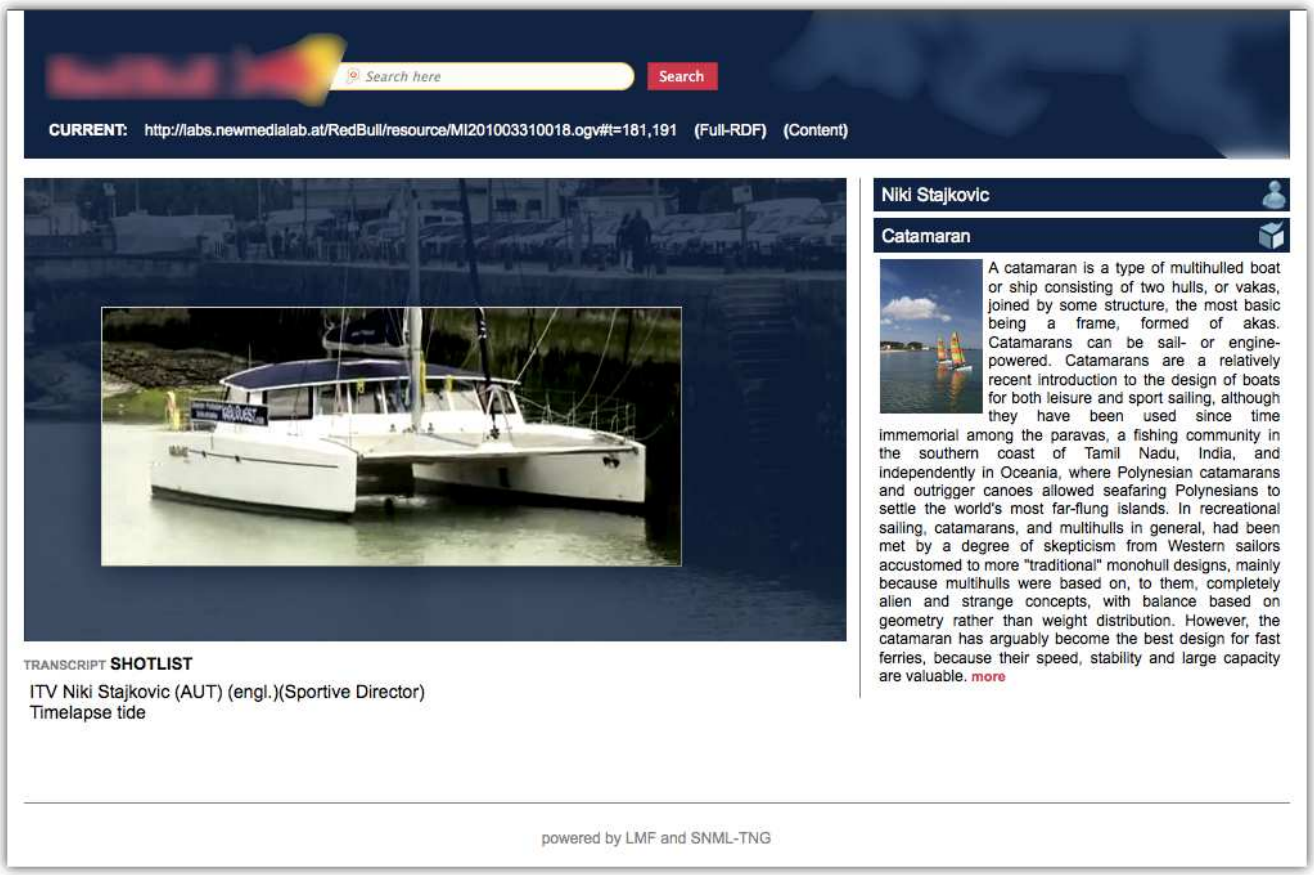

Figure 2: A video player surrounded with structured and ordinary metadata[8]

an annotation are marked with a border (like the *catamaran* in the screenshot) and (hyper-)link to a user-friendly representation of metadata on the right side (description about *catamaran* from DBpedia[7]). There are also time depending notifications to current concepts, like persons appearing or speaking currently in the video.

4. IMPLEMENTATION

As backend we selected the Linked Media Framework (LMF) [12], a modular Linked Data server that is developed in a research centre called Salzburg NewMediaLab[9]. The clients (search and player) are lightweight javascript implementations using RESTful webservices for client-server communication.

4.1 Backend: The Linked Media Framework

The Linked Media Framework is an easy-to-setup Linked Data server application. It bundles central Semantic Web technologies and offers advanced services. LMF allows to interact with resources in a RESTful way. In addition it offers several modules that can be used for annotation and retrieval processes (like a Reasoning module, a Semantic Search module etc.) and is easy extendable.
As recommended for Linked Data[13], LMF uses semistructured data representation (namely RDF) and HTTP URLs as uniform resource identifier. That allows both, an integration of ontologies and the implementation of Media Fragments innately. LMF offers a freetext-based index structure (SOLR) as well as a graph-based index and so allows data retrieval in several ways. Through an intelligent caching mechanism the framework is able to index external data that is reachable over Linked Data, SPARQL[14] Endpoints or proprietary webservices.

4.2 Frontend: ajax and html5

The search application is implemented in javascript (plus HTML and CSS). It communicates with SOLR webservice and thus benefits from the js-solr-client library ajax-solr[10]. The client application uses available plugins for result-templating, faceting and paging.

The extended video player makes use of the html5 video tag and Media Fragment implementation of latest web browsers. The player itself is a simple video event handler that passes time events to a number of registered extensions. At the time of writing the player has extensions for concept, tran-

[7]http://dbpedia.org/resource/Catamaran
[8]http://labs.newmedialab.at/RedBull/video/redbull/player.html#http://labs.newmedialab.at/RedBull/resource/MI201003310018.ogv%23t=181,191
[9]http://www.newmedialab.at/
[10]http://evolvingweb.github.com/ajax-solr/

scription, shotlist and metadata visualization. Each extension queries for required data by its own (using resource access methods like SPARQL or Linked Data).

5. DEMO OUTLINE

The demo is online available and described in the testbed of Salzburg NewMediaLab[11]. To be able to check out the fragment feature, it is recommended to use Firefox version 9 or higher, because the browser already integrates (parts of) the W3C fragments recommendation.

Within the demo we show, how semantic enhancements can improve the search for videos in content pools. We also demonstrate how the employment of video fragments (using the related standard) allows a quicker decision making, if search results fulfill users intentions. E.g. searching for the term 'helicopter' results in four short video fragments that can be browsed much quicker than 3-minutes videos. We demonstrate, how search results can be narrowed down via faceted search, which is based on semantic annotations. The search results directly link to a player where single results are displayed. We show, how linked information from different datasources can be represented together and synchronous to video playback and thus enrich the users experience. Further, to get more into details, we take a look 'behind the scenes' and browse semistructured metadata as well as explain the backend system and its supplemental features.

6. CONCLUSION

In this paper we presented a proof of concept for the consumption of semantically linked videos. We demonstrate our approach (search as well as display) with sample data from the Red Bull Content Pool. In further steps we will integrate more automatic components into the enhancement process. Furthermore we plan a user-friendly tool to control semi-automatic semantic lifting of huge video-content sets. This tool will be based on the WYSIWYG paradigm and thus bring text and video editing in the web closer together.

7. ACKNOWLEDGMENTS

The semantic video search and the player for annotated videos described in this paper were planed and developed within the Austrian research center "Salzburg NewMediaLab - The Next Generation" (SNML-TNG). The center is funded by the Austrian Federal Ministry of Economy, Family and Youth (BMWFJ), the Austrian Federal Ministry for Transport, Innovation and Technology (BMVIT) and the Province of Salzburg. The demo videos are taken from the content pool of Red Bull Media House GmbH, a partner of SNML-TNG.

8. REFERENCES

[1] Linking Open Data: W3C SWEO Community Project. http://esw.w3.org/topic/SweoIG/TaskForces/CommunityProjects/LinkingOpenData, 2010.

[2] C. Bizer et al. Linked Data - The Story So Far. International Journal on Semantic Web and Information Systems (IJSWIS), Vol. 5, Issue 3, 2009.

[3] C. Bizer, et al. Interlinking Open Data on the Web. 2007.

[4] W. Lee, et. al. Ontology for media resource 1.0. W3C Candidate Recommendation. http://www.w3.org/TR/2011/CR-mediaont-10-20110707/,2011.

[5] F. Stegmaier, et. al. Api for media resources 1.0. W3C Working Draft. http://www.w3.org/TR/2011/WD-mediaont-api-1.0-20110712/,2011.

[6] F. Scharffe and J. Euzenat. Alignments for data interlinking. http://melinda.inrialpes.fr, 2009.

[7] T. Steiner. SemWebVid - Making Video A First Class Semantic Web Citizen. Poster & Demo Session ISWC2010. Shanghai. 2010.

[8] A. Fialho, et. al. What's on this evening? Designing User Support for Event-based Annotation and Exploration of Media. 1st International Workshop "EVENTS 2010" at SETN 2010, Athens, Greece, May 2010.

[9] L. Hollink et. al. Semantic annotation of image collections. In: Proc. of the 4th Int'l. Workshop on Knowledge Markup and Semantic Annotation, 2003.

[10] Violeta Damjanovic, Thomas Kurz, et al. Semantic Enhancement: The Key to Massive and Heterogeneous Data Pools. In: 20th International Electrotechnical and Computer Science Conference (ERK 2011), September 2011.

[11] B. Haslhofer, R. Simon, R. Sanderson, and H. V. de Sompel. The Open Annotation Collaboration (OAC) model. Workshop on Multimedia on the Web (MMWeb2011) in conjunction with iSemantics2011, September 2011.

[12] Thomas Kurz, Sebastian Schaffert and Tobias Bürger: LMF - A Framework for Linked Media. In: Workshop on Multimedia on the Web (MMWeb2011), September 2011.

[13] Tim Berners-Lee. Linked Data. W3C Design Issues, 2006.

[14] E. Prud'hommeaux and A. Seaborne. Sparql query language for RDF. W3C Recommendation, 2008.

[11] http://labs.newmedialab.at/RedBull/video/redbull.html