Interlinking Multimedia – Principles and Requirements

Tobias Bürger[1], Michael Hausenblas[2]

[1] Semantic Technology Institute, STI Innsbruck, University of Innsbruck, 6020 Innsbruck, Austria,
tobias.buerger@sti2.at
[2] Institute of Information Systems & Information Management, JOANNEUM RESEARCH, 8010 Graz, Austria,
michael.hausenblas@joanneum.at

Abstract. The linked data principles have gained a huge momentum by providing means to interlink datasets and by that contributing to a rich user experience on the Web. Methods to interlink data however still do not cover multimedia content in a sufficient way, as *Interlinking Multimedia* requires more than just putting resources globally in relation to each other. In order to close this gap, we propose a set of principles and requirements for interlinking multimedia content on the Web. As one major source we have identified user interaction to establish static or dynamic links between (parts of) multimedia resources.

1 Introduction

In early 2007, the W3C launched the Linking Open Data (LOD) community project[3] whose goal is to bootstrap the Semantic Web by publishing datasets using RDF and to publish and interlink open data on the Semantic Web. This is either done by using already existing sets of open data or by creating new linked datasets. The LOD project currently includes over 30 different datasets: From one billion triples and 250k links in mid-2007 the LOD dataset has grown to over two billion triples and 3 million links in early 2008, representing a steadily growing, open implementation of the Linked Data principles[4].

Several approaches exist for semantically linking data: RDF links can either be set manually or generated by automated linking algorithms for large datasets [1]. Advanced approaches such as those described in [1] make use of extended literal lookups or graph matching algorithms which are used to disambiguate similar matches.

Recent development in the linked data community is well documented by the proceedings of the Linked Data on the Web workshop (LDOW2008) [2], and submissions received by the Triplification challenge[5] including a proposal for "User Contributed Interlinking" [3] (UCI) of multimedia content [11].

What however can be observed is, that recent approaches for linking data mainly focused on the automated integration of textual resources and the interlinking of resources

[3] http://esw.w3.org/topic/SweoIG/TaskForces/CommunityProjects/LinkingOpenData
[4] http://esw.w3.org/topic/LinkedData
[5] http://triplify.org/Challenge

as a whole. However, referring to Sir Tim Berners-Lee *the next generation Web should not be based on the false assumption that text is predominant [...] The Web is a multimedia environment, which makes for complex semantics* [4]. This fact has to be taken into account when we think about future directions for linked data. The envisioned situation is a *Interlinked Multimedia Web* in which objects or sequences in multimedia resources are linked to each other based on their semantic relationships.

Only recently Web 2.0 based applications emerged, in which image metadata is augmented by user generated tags. However, the possibility to set typed links between resources, or objects that are part of these resources, is still immature. YouTube launched a first facility[6] to annotate parts of videos spatio-temporally and to link to particular time points in videos[7] which is a promising start. However typed links between fragments can not be established.

The contribution of this paper is an analysis of the envisioned situation and a proposal of a set of requirements for *Interlinking Multimedia* which has only recently been discussed in the linked data community[8]. Furthermore we propose a set of principles based on the semantics of multimedia content and the interaction with multimedia content that can be used to interlink multimedia resources on a semantic level (section 2). We especially identify intended and monitored user interaction as a source for high quality links (section 3).

2 Interlinking Multimedia – Principles and Requirements

The interlinking of resources and parts of it, shares similarities with Hypermedia research: A hypermedia document such as defined in [5] refers to a collection of information units including information about synchronization between these units and about references between them. Typically temporal and a spatial dimensions are included, whereas references can be made between parts in both dimensions. *Interlinking Multimedia* is not an attempt to resurrect hypermedia but rather a light-weight, bottom-up approach to interlink multimedia content on the Web.

As only recently demonstrated by the BBC[9], interlinking of music related information, which may be publicly available on the Web or in closed archives, can significantly contribute to an enhanced end user experience. Moreover, as summarized in [6], there is a demand in several other communities for annotation tools to specify links between whole objects or segments within these objects and the typing of these links or relationships: Not only media researchers that want to relate and annotate segments between books, or screenplays and different films or film versions demand for facilities to interrelate rich content [6]. In order to realise the envisioned situation in which multimedia resources are semantically interlinked on a fine-granular level, one should take the following principles and requirements into account:

[6] http://youtube.com/watch?v=UxnopxbOdic
[7] http://www.techcrunch.com/2008/10/25/youtube-enables-deep-linking-within-videos/
[8] http://community.linkeddata.org/MediaWiki/index.php?InterlinkingMultimedia
[9] http://www.bbc.co.uk/music/beta

1. In order to be become part of the LOD cloud, ***Interlinking Multimedia*** **must follow the linked data principles** [7]:
   (a) All items should be identified using URIs;
   (b) All URIs should be dereferenceable and it should be possible to lookup the identified items using HTTP;
   (c) When looking up an URI, that is, an RDF property is interpreted as a hyperlink, it leads to more data;
   (d) Links to other URIs should be included in order to enable the discovery of more data.
2. Solutions should take into account the characteristics of multimedia whose semantics – when watched by a user – are typically derived based on the experiences and background of a human being. Thus **solutions should consider provenance information**; who says what and when is an important contextual aspect to represent the semantics of content (even if statements or references were created by machines).
3. **Metadata descriptions have to be interoperable** in order to reference and integrate parts of the described resources. This issues are discussed in [8], addressed by recent proposals like ramm.x[10] and by the W3C Media Annotations Working Group[11].
4. As discussed in [9] recently, **localizing and identifying fragments** is essential in order to link parts of resources with each other. It is essential to provide means to mark up spatial or temporal fragments, then to provide a mechanism to specify URIs for those fragments and finally to draw links between those fragments. This issue is particularly researched in the recently started W3C Fragments Working Group[12].
5. Furthermore **Interlinking Methods**, which we discuss in section 3, are essential in order to manually or (semi-) automatically interlink multimedia resources.

3 Interlinking Methods

Due to the inherent characteristics of multimedia content, the implementation of interlinking methods is far from being trivial. This is mainly due to the Semantic Gap, i.e. the large gulf between the low-level semantics which are derivable by machines and the high level semantics a user is typically interested in. This gap significantly hinders automation in the establishment of high quality links. As only little work is available at time of writing, we propose a set of interlinking methods that could close this gap:

Automatic Interlinking (AI) can be applied in situations in which quality metadata information is available that can be used to identify objects and their semantics. While AI methods[13] have demonstrated to yield fair results for global, textual resources [13], for fine-grained interlinking of multimedia content we doubt that this is the preferred path to follow.

[10] http://sw.joanneum.at/rammx/
[11] http://www.w3.org/2008/WebVideo/Annotations/
[12] http://www.w3.org/2008/WebVideo/Fragments/
[13] http://esw.w3.org/topic/TaskForces/CommunityProjects/LinkingOpenData/EquivalenceMining

Emergent Interlinking (EI) can be based on the principles of Emergent Semantics whose underlying principle is to discover semantics through observing how multimedia information is used [12]. This can be essentially accomplished by putting multimedia resources in context-rich environments being able to monitor the user and his behavior. In these environments, two different types of context are present: (i) static or structural context, which is derived from the way how the content is placed in the environment (e.g. a Web page) and (ii) dynamic context, which is derived from the interactions of the user in the environment (e.g. his browsing behavior, which links he follows, or on which object he zooms). As stated in [12], in appropriate environments, the browsing path of a user is semantically coherent and thus allows to derive links between objects which are semantically close to each other.

User Contributed Interlinking (UCI) – a term which has been coined in [3] – is an approach for manually creating high-quality interlinks. The advantage of the application of UCI for the interlinking of multimedia is that it is based on end users as sources of qualitative information. First steps have already been made for UCI-based interlinking methods, such as available in the still image concept demonstrator CaMiCatzee [11] or in Henry[14] (for interlinking temporal audio fragments).

Game Based Interlinking (GBI) can be based on the principles set forward by Louis van Ahn with his *games with a purpose*[15] [14]. By that, interlinking of resources or parts of these resources could be hidden behind games. This is related to UCI but with the main difference that the user is not aware of him contributing links, e.g. his task is hidden behind a game. GBI seems to be a promising direction for multimedia interlinking. The most interesting examples to build on are Louis van Ahn's *ESP Game* in which users are asked to describe images or *Squigl*[16] in which users are asked to trace objects in pictures. Another interesting approach is followed by *OntoGame* [15] whose general aim is to find shared conceptualizations of a domain. During the game, players are asked to describe images, audio or video files. Users are awarded if they describe content in the same way. These approaches together with appropriate browsing interfaces for multimedia objects could be a promising starting point to let users draw meaningful relations between objects and parts thereof.

The methods can be arranged in a three-dimensional matrix with the dimensions time, quality and amount of annotations as depicted in Figure 1: While UCI might reach the highest quality and needs the highest amount of time from an end user perspective, automatic interlinking might produce the greatest amount of annotation and thus links with the least amount of time and manual effort needed.

4 Conclusion and Further Challenges

In this paper we discussed a future direction for linked data and pointed out to several issues with respect to *Interlinking Multimedia*. Besides the requirements that we

[14] http://dbtune.org/henry/
[15] http://www.gwap.com/
[16] http://www.gwap.com/gwap/gamesPreview/squigl/

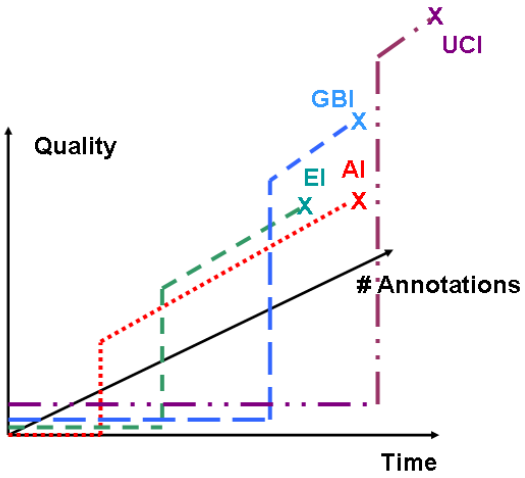

Fig. 1. Multimedia Interlinking – methods

formulated in section 2, a few other challenges have to be faced. These include generally applicable challenges for LOD like *Discovery & Usage* which has recently been addressed with voiD, the vocabulary of interlinked datasets[17], *Performance & Scalability* or *Privacy & Trust* which is addressed in another position paper for this workshop. We particularly identify user interaction is an essential ingredient to address a fourth challenge: *Quality* of links.

We believe that with the realization of the *Interlinking Multimedia* – principles a further step can be taken to a truly rich experience on the *Web of Data*.

Acknowledgements: The research leading to this paper was partially supported by the European Commission under contract IST-FP6-027122 "SALERO".

References

[1] Raimond, Y., Sutton, C., and Sandler, C.: "Automatic Interlinking of Music Datasets on the Semantic Web" In: Proceedings of Linked Data on the Web (LDOW2008), Beijing, China, 2008.

[2] Bizer, C., Heath, T., Idehen, K., and Berners-Lee, T. (eds.): "Proceedings of the Linked Data on the Web Workshop", Beijing, China, April 22, 2008, CEUR Workshop Proceedings, ISSN 1613-0073, online CEUR-WS.org/Vol-369/

[3] Halb, W., Raimond, Y., and Hausenblas, M. "Building Linked Data For Both Humans and Machines" In: Proceedings of Linked Data on the Web (LDOW2008), Beijing, China, 2008.

[4] Berners-Lee, T., Hall, W., Hendler, J. A., O'Hara, K., Shadbolt, N. and Weitzner, D. J. "A Framework forWeb Science" In: Foundations and Trends in Web Science, Vol. 1, No 1, 2006, pp. 1130

[17] http://community.linkeddata.org/MediaWiki/index.php?VoiD

[5] Hardman, L. "Modelling and Authoring Hypermedia Documents" PhD thesis, CWI, Amsterdam, 2004. http://homepages.cwi.nl/ lynda/thesis/

[6] Schroeter, R. and Hunter, J. "Annotating Relationships between Multiple Mixed-media Digital Objects by Extending Annotea" In: Proceedings of the European Semantic Web Conference (ESWC2007), 2007.

[7] Berners-Lee, T. "Linked Data", Design Issue Note, 27.07.2006, online http://www.w3.org/DesignIssues/LinkedData.html

[8] Tzouvaras, V., Troncy, R., Pan, J. Z. (eds.) "Multimedia Annotation Interoperability Framework" W3C Incubator Group Editor's Draft, 14 August 2007. http://www.w3.org/2005/Incubator/mmsem/XGR-interoperability/

[9] Troncy, R., Hardman, L., van Ossenbruggen, J., and Hausenblas, M. "Position Paper on Identifying Spatial and Temporal Media Fragments on the Web". W3C Video on the Web Workshop. Dec 2007.

[10] Hausenblas, M., Bailer, W., Bürger, T., Troncy, R. "Deploying Multimedia Metadata on the Semantic Web" In: Posterproceedings of the 2nd International Conference on Semantics And digital Media Technologies (SAMT 07), 2007.

[11] Hausenblas, M., Halb, W. "Interlinking Multimedia Data" Linking Open Data Triplification Challenge at the International Conference on Semantic Systems (I-Semantics08) at TRIPLE-I, September 2008.

[12] Grosky, W. I., Sreenath, D. V., and Fotouhi, F. "Emergent Semantics and the Multimedia Semantic Web". SIGMOD Rec. 31, 4 (Dec. 2002), 54-58.

[13] Hausenblas, M., Halb, W., Raimond, Y., and Heath, T. "What is the Size of the Semantic Web?". I-Semantics 2008: International Conference on Semantic Systems, 2008.

[14] von Ahn, L. "Games with a Purpose"'. IEEE Computer 39(6): 92–94 (2006)

[15] Siorpaes, K. and Hepp, M. "Games with a Purpose for the Semantic Web". IEEE Intelligent Systems 23(3): 50–60 (2008)