TIB|AV-Portal: Integrating Automatically Generated Video Annotations into the Web of Data

Jörg Waitelonis[1], Margret Plank[2], and Harald Sack[3]

[1] yovisto GmbH, August-Bebel-Str. 25-53, 14468 Potsdam, Germany
joerg@yovisto.com
[2] German National Library of Science and Technology, Welfengarten 1 B, 30167 Hannover, Germany
margret.plank@tib.eu
[3] Hasso-Plattner-Intitute, Prof.-Dr.-Helmert-Str. 2-3, 14482 Potsdam, Germany
harald.sack@hpi.de

Abstract. In the poster presentation the challenges and the chosen strategy for Linked Data-based metadata export of the TIB|AV-Portal and its multimedia content are introduced. In difference to other library metadata the TIB|AV-Portal deploys automated metadata extraction and named entity linking to provide time-based semantic metadata regulary published as RDF data dump files. The underlying data model for the published dump files has been simplified to enable efficient accessibility and interlinking with further information resources and learning platforms.

1 Introduction

In order to improve the accessibility, citability and the sustainable use of scientific videos, the German National Library of Science and Technology (TIB) in cooperation with the Hasso Plattner Institute (HPI) has developed the TIB|AV-Portal[4][4]. In this paper we present the strategy and implementation for the semantic metadata export of the TIB|AV-Portal, which provides access to high grade scientific videos from the fields of technology/engineering, architecture, chemistry, information technology, mathematics and physics in English and German. In addition to reliable authoritative metadata (Dublin Core[5]) time-based metadata is generated by automated media analysis. Based on text-, speech- and image recognition text-based terms are extracted and mapped to subject specific GND[6] subject headings with named entity linking (NEL). The cross-lingual retrieval uses inter-language links based on an ontology mapping (DBpedia[7], Library of Congress Subject Headings[8], e. a.). These technologies improve the search for and the re-use of scientific videos by e.g. enabling pinpoint access to individual video segments.

[4] http://av.tib.eu/
[5] http://dublincore.org/
[6] http://www.dnb.de/EN/gnd
[7] http://dbpedia.org/
[8] http://id.loc.gov/authorities/subjects

To enable efficient metadata re-use and interlinking it was decided to regularly publish metadata data dump files encoded with RDF[9] according to the Linked Data principles. The general difference to other Linked Data based publication of library metadata, e.g. [7, 1, 6, 3, 2], lies in the fact that the video resources of the TIB|AV-Portal are subject to automated metadata extraction and semantic analysis resulting in time-based metadata that contain extracted content descriptions including provenance and – if available – confidence information. A simpler metadata model has been derived from the complex internal metadata representation to foster efficient access and readability.

2 RDF-Export for the TIB|AV-Portal

For sake of simplicity, clarity, and to avoid potential redundancies, it is recommended only to use a suitable subset of potentially available vocabularies. The following general requirements had to be considered:

1. How many metadata can be covered by which of the available vocabularies?
2. How well does a vocabulary item's semantics fit the intended meaning of the metadata item?
3. A high degree of generality should be achieved, but which items have to be better chosen to be more specific?
4. How well do datatypes match? Are there any conversions necessary or is there any loss in accuracy?
5. How popular is the vocabulary (in a certain community only or in general)?
6. When only using individual parts of a vocabulary in combination with other vocabularies, it has to be ensured that no logical contradictions occur.

Table 1 summarizes vocabularies used for the mapping of standard as well as temporal metadata of the TIB|AV-Portal. `dcterms` and `dctypes` are generic and widely used vocabularies that have been chosen to represent generic items as e. g. the subject or the language of a resource. All major search engines support the `schema`.org vocabulary for structured data. It defines a standard set of type and property names covering a wide range of topics. The TIB|AV-Portal website embeds RDFa annotations with `schema.org` to improve the accessibility for search engines. Since all TIB|AV-Portal resources are referenced by DOIs, the library specific `bibframe` vocabulary is used to model DOI as well as specific title properties, e. g. subtitle to reference subheadings. The `foaf` vocabulary is mainly used to name persons as well as organizations.

As a result of the TIB|AV-Portal's automated multimedia analyses different kinds of spatio-temporal metadata is created. To annotate a video segment the Open Annotation (OA) Data Model[10] has been chosen. The primary aim of the OA Data Model is to provide a standard description mechanism for sharing annotations between systems. In connection with W3C's Media Fragment Identifier the reference to arbitrary spatio-temporal segments of a video can easily be established.

[9] https://www.w3.org/RDF/
[10] http://www.openannotation.org/spec/core/

Table 1. Overview of namespaces and prefixes used for the RDF export of standard metadata (top), and spatio-temporal metadata (bottom)

Prefix	Namespace	Name
`bibframe`	http://bibframe.org/vocab/	Bibframe Vocabulary
`dbp`	http://dbpedia.org/resource/	DBpedia Resources
`dcterms`	http://purl.org/dc/terms/	DCMI Metadata Terms
`dctypes`	http://purl.org/dc/dcmitype/	DCMI Type Vocabulary
`foaf`	http://xmlns.com/foaf/0.1	Friend of a Friend Vocabulary
`gnd`	http://d-nb.info/gnd/	Integrated Authority File (GND)
`schema`	http://schema.org/	Schema.org Vocabulary
`tib`	http://av.tib.eu/resource/	TIB|AV-Portal Resources
`cnt`	http://www.w3.org/2011/content#	Representing Content in RDF
`itsrdf`	http://www.w3.org/2005/11/its/rdf#	Internationalization Tag Set (ITS)
`nif`	http://persistence.uni-leipzig.org/nlp2rdf/ontologies/nif-core#	NLP Interchange Format
`oa`	http://w3.org/ns/oa#	Open Annotation Data Model

An important challenge was to also include the semantic text annotations which was realized by including the NLP interchange format (NIF2) vocabulary that aims to achieve interoperability between Natural Language Processing (NLP) tools, language resources, and annotations. Fig. 1 illustrates a simplified example as the combination of OA Model and NIF2. The `tib:annotation/yxyx` is the dedicated RDF resource of the annotation referring to the body as well as the target. The annotation target refers to the URI of the video including media fragment identifier(s). The annotation body refers to a resource `tib:annotation/yxyx#char=0,24` of type `nif:Context`. The URI is extended by a fragment identifier specifying the character range of text. This resource is the reference context for the semantic annotations typed by `nif:Phrase`. Each annotation holds all the additional information created by the generating algorithms (e. g. NEL), such as the text ranges, confidence values, the knowledge-base entity URI, as well as the version of the algorithm itself. NIF2 enables instantly to connect the dataset to other NEL annotators, which allows to asses the quality of the semantic analysis objectively compared to other systems [5].

3 Further Challenges and Conclusion

Since automated extraction methods always produce potential errors, regular cycles of manual data cleansing are conducted to maximize metadata quality before preparing a new version of the data dumps. Inaccuracy and potential faults of the automated metadata extraction are subject of current and ongoing research. Multimedia analysis technology as well as semantic analysis are constantly refined to obtain higher quality results. However, the possibility of a manual override is currently evaluated. Manual correction of faulty low level results, as e.g. in the speech-to-text analysis imply a subsequent reevaluation of the thereby changed context.

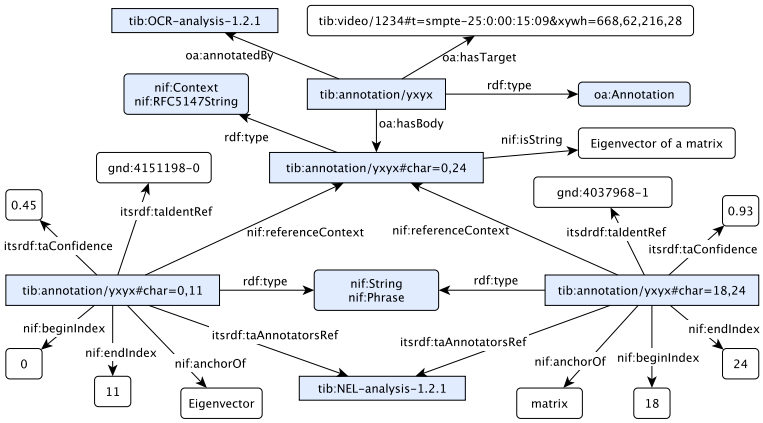

Fig. 1. Example of temporal annotation of a video frame with NIF2 and OA.

Since the context has changed, the semantic analysis might reconsider previously taken decision for the NEL. Furthermore, a manually confirmed or changed result should be considered for future context analysis and entity disambiguation. The general goal of these efforts is to allow for high quality automatically generated metadata, which in turn enable more precise and complete search results.

In this paper, we have presented the strategy and implementation for the Linked Data-based metadata export of the TIB|AV-Portal. Automated metadata extraction provides time-based metadata that are regulary published in form of RDF data dump files. The underlying data model for the published dump files has been simplified to enable easy accessibility and readability.

References

1. M. Fernández, M. d'Aquin, and E. Motta. Linking data across universities: An integrated video lectures dataset. In *Proc. of ISWC 2011*, LNCS, vol. 7032, pages 49–64. Springer, 2011.
2. L. Nixon and R. Troncy. Survey of semantic media annotation tools for the web: Towards new media applications with linked media. In *The Semantic Web: ESWC 2014 Satellite Events*, pages 100–114. Springer, 2014.
3. Y. Raimond, T. Scott, S. Oliver, P. Sinclair, and M. Smethurst. *Linking Enterprise Data*, chapter Use of Semantic Web technologies on the BBC Web Sites, pages 263–283. Springer US, Boston, MA, 2010.
4. H. Sack and M. Plank. AV-Portal - The German National Library of Science and Technology's Semantic Video Portal. *ERCIM News*, 2014(96), 2014.
5. R. Usbeck et al. GERBIL – general entity annotation benchmark framework. In *Proc. of 24th WWW conference*, 2015.
6. J. Waitelonis and H. Sack. Augmenting Video Search with Linked Open Data. In *Proc. of Int. Conf. on Semantic Systems 2009*, 2009.
7. H. Q. Yu, C. Pedrinaci, S. Dietze, and J. Domingue. Using linked data to annotate and search educational video resources for supporting distance learning. *Learning Technologies, IEEE Trans. on*, 5(2):130–142, 2012.