Visualizing Mapping of Metadata Properties

Martin Höffernig, Wolfgang Weiss, Werner Bailer

JOANNEUM RESEARCH, DIGITAL –
Institute of Information and Communication Technologies, Graz, Austria
{firstName.lastName}@joanneum.at

1 Introduction

Millions of hours of audiovisual content are held by collections of dedicated broadcast, film and sound archives, institutional or corporate archives, libraries and museums. There is a large heterogeneity between the different audiovisual archives resulting from their history and tradition but also from cultural differences of the countries where those archives reside. Consequently metadata models covering the workflows and necessities in the archives differ as well. This fact and the need for metadata for various different use cases in the archives lead to a number of metadata models and standards. Thus mapping between different metadata models is inevitable in practical applications.

We are currently developing a system for automating metadata mapping by formalizing semantics of properties in the different formats and their relations [4], based on an intermediate ontology, namely the *meon* ontology [3]. In order to enable users to validate the automatically determined mappings visualization functionalities are required in the system. This paper describes the integration of the ontology visualization developed in [6] into our mapping system prototype.

Creating comprehensive, clear and intuitive visualizations of ontologies and RDF graphs is an ongoing challenge. Different approaches can be found in applications for Semantic Web engineers. An example is Protege[1], which is an open, platform independent environment for creating and editing ontologies and knowledge bases. The application is extensible by its plug-in architecture and thus provides several visualizations. IsaViz[2] is a visual tool for browsing and authoring of RDF models. Resource nodes are represented by ellipses, literals as rectangles and properties are displayed as lines with arrows. OntoSphere 3D [1] uses a collection of three-dimensional visualization techniques displaying ontologies. gFacet [2] combines the graph visualization with facet search in the graph.

These applications use different kinds of visualization techniques to present the user a possibly easy to understand and complete overview of the whole RDF graph. Using graph visualizations of RDF data especially for end users has a number of drawbacks. For example, these visualizations are flat and every node is treated as a primary node. Also, displaying a graph with hundreds of nodes and edges results in a cluttered visualization (cf. [5]). Nonetheless graph visualizations have their place, especially for Semantic Web engineers [6].

[1] `http://protege.stanford.edu/`
[2] `http://www.w3.org/2001/11/IsaViz/`

2 Implementation

The prototype[3] helps users finding, validating, and understanding metadata mappings by automatic metadata matching and appropriately visualizing mapping relations. Figure 1 shows the textual part of the user interface where the user creates a query. The first step of the user is to select an input- and output metadata format. Via the *Load button*, the application lists all available concepts from the selected formats. The next step is to select one or more concepts for which the mapping relations should be found. After confirming the selected concepts by clicking the *Ask button*, dependencies between the input and output concepts according to the defined rules are calculated and displayed. For each selected output format the information whether a mapping is feasible or not is displayed: *True*, if the output concept can be mapped from one or more of the selected input concepts, *False* otherwise. In case that there are output concepts without corresponding selected input concepts, the *Find requirements* option can be used. After selecting this option additional necessary input concepts are computed in order to establish mapping relations to the selected output concepts.

In addition to the boolean information about the feasibility of the mapping explained above, possible mapping relations are visualized in a graph visualization. The RDF-like graphs include the selected input concepts as yellow nodes, the selected output concepts as red nodes, the related *meon* concepts as green nodes and potentially missing input concepts as white nodes. It focuses on the current task of the user by displaying only necessary nodes and edges. This kind of visualization supports the user in understanding and validating the found mapping relations between input and output concepts of the metadata formats. An example of the graph visualization of the selected concepts shown in Figure 1 is depicted in Figure 2. The graph representation reveals that the input concept `mpeg7:Height` together with `mpeg7:Width` can be mapped to the selected output concept `ma:FrameSize` via `meon:Resolution`, which is part of the intermediate ontology (*meon* ontology). Beside this positive mapping relation, no appropriate mapping relation can be established to the remaining output concept `ma:Creator` from any of the selected input concepts. However, `mpeg7:UnqualifiedCreator` is a possible input concept to map to `ma:Creator`.

The mapping prototype is a Web application using standard Web technologies such as HTML and JavaScript for the user interface as well as Scalable Vector Graphics (SVG) for the graph visualization. To generate the graph we use the Java Universal Network / Graph Framework (JUNG)[4], which provides a number of layout algorithms and mechanisms to manipulate graphs. An internal evaluation has shown that the "self-organizing map layout for graphs" produces the best results for our requirements. However, this layout algorithm generates a different layout at every single run. Therefore, it is necessary to animate the graph visualization for the user. The animation helps the user to follow how

[3] `http://prestoprime.joanneum.at`
[4] `http://jung.sourceforge.net/`

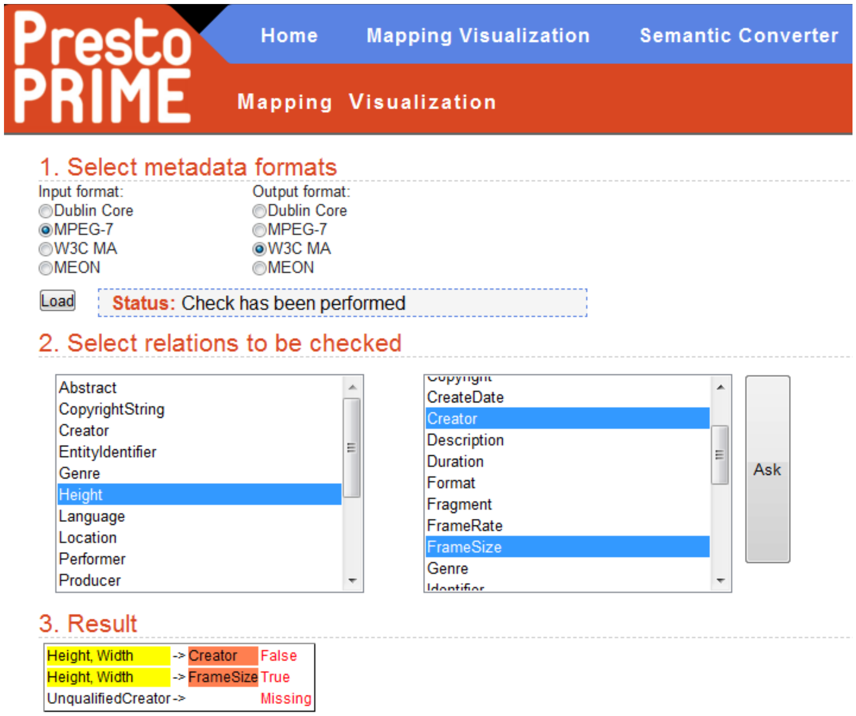

Fig. 1. Visualization interface.

the layout changed since the last run. For processing the RDF data the Jena Semantic Web framework[5] has been used.

3 Conclusion and Future Work

In this paper we have presented the visualization functionality of our metadata mapping prototype[6] which helps users finding, understanding and validating metadata mappings by automatic metadata matching and appropriately visualizing mapping relations. The visualization shows mapping relations between input and output metadata formats which are determined by the system via an intermediate ontology. It uses coloured nodes and focuses on the current user task by displaying only nodes which are necessary for the current task. The system is able to find direct metadata mappings as well as to suggest further input concepts to satisfy the desired metadata mappings. In the future the system shall support the definition of mapping rules by the user in order to improve the results in cases where incomplete or ambiguous mappings between pairs of formats exist.

[5] `http://jena.sourceforge.net/`
[6] `http://prestoprime.joanneum.at`

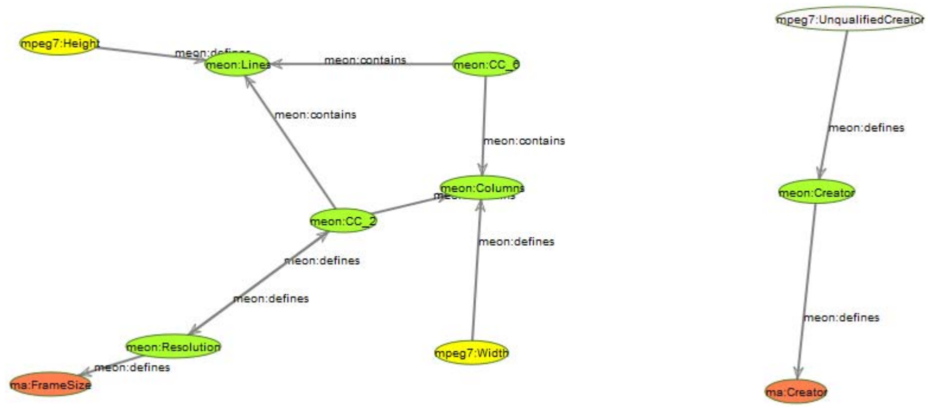

Fig. 2. Example of mapping visualization.

Acknowledgments

The research leading to these results has received funding from the European Union's Seventh Framework Programme under grant agreement nr. FP7 231161, "PrestoPRIME" (`http://www.prestoprime.eu`).

References

1. Alessio Bosca, Dario Bonino, and Paolo Pellegrino. Ontosphere: more than a 3d ontology visualization tool. In *Proceedings of the 2nd Italian Semantic Web Workshop, 14-15-16 December 2005*, number 166, University of Trento, Trento, Italy, December 2005.
2. Philipp Heim, Thomas Ertl, and Jürgen Ziegler. Facet graphs: Complex semantic querying made easy. In *7th Extended Semantic Web Conference (ESWC2010)*, June 2010.
3. Martin Höffernig and Werner Bailer. Formal metadata semantics for interoperability in the audiovisual media production process. In *Workshop on Semantic Multimedia Database Technologies (SeMuDaTe)*, Graz, Dec. 2009.
4. Martin Höffernig, Werner Bailer, Günter Nagler, and Helmut Mülner. Mapping audiovisual metadata formats using formal semantics. In *5th International Conference on Semantic and Digital Media Technologies*, Saarbrücken, DE, Dec. 2010. To appear.
5. David Karger and M.C. Schraefel. The pathetic fallacy of rdf. `http://swui.semanticweb.org//swui06/papers/Karger/Pathetic_Fallacy.html`, 2006. Retrieved 06 February 2007.
6. Wolfgang Weiss, Michael Hausenblas, and Gerhard Sprung. Visual exploration, query, and debugging of rdf graphs. In *Proceedings of the Fifth International Workshop on Semantic Web User Interaction*, April 2008.