

Structuring and Accessing Semantic Multimedia Data

Steffen Staab

Uni. Koblenz
Koblenz, Germany
staab@uni-koblenz.de

1 Bio

Steffen Staab is professor for databases and information systems at the University of Koblenz-Landau, leading the research group on Information Systems and Semantic Web (ISWeb). His interests lie in researching core technology for ontologies and semantic web as well as in applied research for exploiting these technologies for knowledge management, multimedia and software technology. He has participated in numerous national, European and intercontinental research projects on these different subjects and his research has led to more than 100 refereed contributions in journals and conferences. Dr. Staab held positions as researcher, project leader and lecturer at the University of Freiburg, the University of Stuttgart/Fraunhofer Institute IAO, and the University of Karlsruhe and he is a co-founder of Ontoprise GmbH. He is on several journal editorial boards and is incoming Editor-in-Chief of Elsevier's Journal of Web Semantics. For more information see: <http://isweb.uni-koblenz.de/> and <http://www.uni-koblenz.de/staab/>

2 Abstract

Navigating large heterogeneous media repositories is difficult because of the heterogeneity of possible data representations and because it requires frequent search for the 'right' keywords, as traditional searching and browsing do not consider the semantics of multimedia data. To resolve these issues we use ontologies and semantic data in two core areas of multimedia storage and access. First, we represent multimedia data in our Core Ontology for MultiMedia - COMM. COMM is an ontology content design pattern able to accommodate heterogeneous representations of low- and high-level multimedia data in a homogeneous manner. Second, we allow for searching and browsing by exploiting semantic background knowledge at a large scale using our Networked Graphs reasoning infrastructure. The methods are integrated into the SemaPlorer prototype allowing for the easy usage of Flickr data based on semantic sources such as DBpedia, GeoNames, WordNet and personal FOAF files. The computational demands are met by federating data access on Amazon's Elastic Computing Cloud (EC2) and Simple Storage Service. Thus, SemaPlorer remains scalable with respect to the amount of distributed components working together as well as the number of triples managed overall.