Panel: Semantic Metadata as Linked Data Building Blocks

Joseph A Busch
Taxonomy Strategies
Washington, D.C.
jbusch@taxonomystrategies.com

Marcia Lei Zeng*
Kent State University
Kent, OH
mzeng@kent.edu

Gail Hodge
Information International Associates
ghodge@iiaweb.com

ABSTRACT
In the Linked Data movement, value vocabularies (such as the Linked Data version of a thesaurus or a classification scheme) and metadata element sets (which are generally made concrete through RDF Schemas or OWL ontologies) have become the building blocks of any dataset that joins the CKAN Linked Data hub of the Open Knowledge Foundation (2012). Evidence can be seen either from analyzing the datasets in the Data Hub or by browsing through the long list of the value vocabularies and metadata element sets included in one of the three deliverables of the W3C Library Linked Data Incubator Group Final Report (2011). This panel will introduce the concept of semantic metadata and the two types of vocabularies involved, discuss how value vocabularies are transformed in order for machine to process and understand, present new use cases of the conventional knowledge organization systems (KOS) in the Semantic Web environment, and report the findings regarding publishing semantic assets, describing and accessing them, and enabling their interoperability in the Semantic Web environment.

Keywords
Knowledge Organization Systems, Metadata, Value vocabularies, Linked Data, Semantic metadata, Information standards, Semantic assets

PRESENTATIONS AND PANELISTS
The presenters in this panel have been active in the areas of networked knowledge organization systems and more recently on Linked Data activities. Their short biographies and presentation descriptions are listed below, arranged in presentation order. The panel will introduce concepts, applications, standardization initiatives related to semantic metadata in the context of Linked Data movement, and open up for questions and discussions.

1. Semantic metadata: A Tale of two types of vocabularies

Joseph A Busch, Taxonomy Strategies

Presentation description: Semantic metadata is metadata that is expressed using a standard syntax that can be commonly processed by applications and tools. There is always an implied statement in any description or "classification" of an object, for example, <News Item><Topic><US Presidential Election 2012>. This is a subject-predicate-object triple, or more specifically, a class-attribute-value triple. The first two elements of the triple – class, attribute – are metadata elements with a defined semantic relationship. The third element is a value, from a controlled vocabulary. This talk will focus on: The two types of vocabularies involved with semantic metadata, the class.attribute vocabulary, and the value vocabulary. Examples of standard metadata vocabularies such as Dublin Core and FOAF, and canonical lists of named entities (people, organizations, places, events and things) especially well-branded names such as products and services will be shown.

Short biography: Mr. Busch is the Founder and Principal Consultant of Taxonomy Strategies. Taxonomy Strategies guides global companies, government agencies, international organizations and not-for-profits such as Nike, Oracle, the U.S. Environmental Protection Agency, the International Monetary Fund and Harvard Business Publishing in developing metadata frameworks and taxonomy strategies. Before founding Taxonomy Strategies, Joseph Busch held management positions at Interwoven, Metacode Technologies, the Getty Information Institute and PriceWaterhouse. He is a Past President of the ASIS&T, and past member of the Board of Directors of the Dublin Core Metadata Initiative.

2. Working with big data in concise presentations using taxonomies

Marjorie Hlava, Access Innovations

Presentation description: In creating and consuming Linked Data, one of the barriers to a dataset provider is to find out and decide which dataset to link to. Evidence
shows that majority of the datasets in the Linked Data Cloud connect to dBpedia while seldom to link to others. To find out what dataset would be valuable to link in a particular subject domain needs to have a deep understanding of its content. At this presentation Marjorie Hlava will provide insight of leveraging KOS for data mining of very large databases. In an example, she will present the methodology and results of data mining using three different vocabularies as a basis for semantic analytics instead of free-text analytics. The results indicates that, beyond the conventional use of KOS for information processing, KOS can be invaluable in discovering the overlap and directions of science, revealing trends and the edges of the fields, assisting in the decision of linking to other Linked Data datasets.

**Short biography:** Marjorie M.K. Hlava is President, Chairman, and founder of Access Innovations, Inc. She is past president of NFAIS and past president of the ASIS&T and the 1996 recipient of ASIST’s prestigious Watson Davis Award. She has published more than two hundred articles and several books on information science topics. She also holds three patents. Her research areas include furthering the productivity of content creation and the governance layer for information access through automated indexing, thesaurus development, taxonomy creation, natural language processing, machine translations, and machine aided indexing.

**3. The state of KOS in the Linked Data movement – the publishing, management, and interoperating of KOS for the Semantic Web**

Marcia Lei Zeng, Kent State University

**Presentation description:** Building on a research and three international working group experience, Marcia Zeng will report on the state of KOS in the Linked Data movement, including: (a) an analysis of the value vocabularies in the Data Hub that have been published as Linked Data using SKOS and the extensions; (b) an introduction to the efforts of describing and accessing semantic assets in registries through defining an Asset Description Metadata Schema (ADMS) as well as a parallel work of developing a DCMI Application Profile for KOS Resource (which all the panelists have been the key developers); and (c) a focused discussion of the new content in the recently published ISO 25964 Part I that includes a thesaurus data model and an alignment effort with SKOS.

**Short biography:** Marcia Zeng is professor of Kent State University. She has chaired and served on standards committees and working groups for IFLA, SLA, ASIST, NISO, and ISO. She has been a Working Group member of the ISO 25964 Thesauri and Interoperability with Other Vocabularies. She recently completed her role as an invited expert for the W3C Linked Library Data Incubator Group and for the Asset Description Metadata Schema (ADMS) of the European Union. She currently serves as a member of the Executive Board of the International Society for Knowledge Organization (ISKO) and the Dublin Core Metadata Initiative (DCMI) Advisory Board, She is also Director-at-large of ASIS&T.

**MODERATOR:**

Gail Hodge, Information International Associates

**Short biography:** Gail Hodge is a senior information scientist with Information International Associates, Inc. (IIa). She has been a consultant and contractor for government and private sector organizations, specializing in database design, metadata, taxonomy and controlled vocabulary development, process reengineering, standards and information policy. Gail was Director-at-large of the ASIS&T. She is also the co-chair of the Task Force in developing the DCMI-KOS Application Profile for the KOS resources and serves on the DCMI Advisory Board.

**REFERENCES**


* To whom all correspondence should be sent.