

Conference Report

**SWIB13- Semantic Web in Libraries Conference
November 25–27, 2013**

Hamburg, Germany

BY

**DR. G. GNANA SEKARI
Library and Information Officer,
CCRAS, JANAKPURI, NEW DELHI, INDIA**

AND

**MISS SHWETA DHINGRA
Library Consultant, CCRAS, JANAKPURI, INDIA**



What is this conference?

SWIB13 took place at at Bürgerhaus Wilhelmsburg, Hamburg, north part of the Federal Republic of Germany.

Linked Open Data (LOD) has become a widespread method for the publication and management of data on the web. Many libraries and related institutions have already run projects or launched products that comply with the LOD paradigm. New projects, services and tools are emerging continuously.

Why the 2013 Annual Conference?

The SWIB conference is a joint event of ZBW-Leibniz Information Centre for Economics and North Rhine-Westphalian Library Service Centre (hbz). It aims to provide substantial information on linked open data developments relevant to the library world and to foster the exchange of ideas and experiences among practitioners. SWIB encourages thinking outside the box by involving participants and speakers from other domains, such as scholarly communications, museums and archives, or related industries.

Linked data is a technology that is ready to roll out in libraries. There are a lot of projects that seem to be doing the same things independently of one another. In this situation there are different insights in each of these implementations. At some point someone should aggregate these and make a killer product.

In addition to its content, the conference provides ongoing opportunities to discuss emerging issues and best practices with colleagues and consultants across a broad continuum of semantic web issues. For example, in a single venue, you can discover best practices from libraries of leading institutions, investigate new ways to look at old problems, hear experts address codes and standards, and talk one-on-one with experienced vendors. Past attendees have told their colleagues that the value of the networking opportunities alone exceeds the price of admission.

Over the years, I have found that SWIB conferences deliver the best value in professional development. The professional value of the SWIB Annual Conferences that I have attended has been consistent and cost-beneficial.

What was SWIB13 about?

Presentations' themes:

- Mappings and Mashups,
- Libraries and Beyond,
- Ontology engineering,
- Contributing to Europeana,
- Base Technology: the Web, Repositories enhanced

What did I learn?

I will provide a brief set of details for each of the sessions I attended over the three days of this year's conference. In addition, I will share the conference proceedings with my colleagues, including presentations to work groups and implement in the Libraries.

Detailed Session Information

Monday, November 25—Pre-Conference Workshop

Introduction to Linked Open Data **Felix Ostrowski/Adrian Pohl**

This introductory workshop aims to introduce the fundamentals of Linked Data technologies on the one hand, and the basic legal issues of Open Data on the other. The RDF data model will be discussed, along with the concepts of dereferencable URIs and common vocabularies. The participants will continuously create and refine RDF documents to strengthen their knowledge of the topic. Linked Data tenets such as publishing RDF descriptions in a web environment and utilizing Content-Negotiation will be demonstrated and applied by the participants. Aggregating data from several sources and querying this data will showcase the advantages of publishing Linked Data, and RDF Schema will be introduced as an effective way of data integration. On a side track, Open Data principles will be introduced, discussed and applied to the content that is being created during the workshop.

General Session Presentations

Tuesday, November 26

Soylent SemWeb Is People! Bringing People to Linked Data

By: Dorothea Salo

Salo spoke about how hard it is to teach linked data to librarians because linked data is defined in terms that are outside the experience of librarians — it takes comparatively little time to teach HTML, XML, but RDF seems to be very hard to

explain. The lack of tools, reliable library data resources, etc. means that linked data take-up is slow. A chief worry is that there will be too little traction for linked data in libraries because it isn't easy — Salo used the term "negative path dependence" whereby a lesser, but easier technology that is inferior to some competitor wins out because not because it is good, but because it can be seen to work. MARC was seen as the inferior technology. It was pointed out that in order for library linked data to work, the tools to do the job need to be better.

The talk left me with a few questions: I wonder why librarians don't interest themselves in the key Internet technologies more? An understanding of the basics of HTTP is a good thing for librarians that work with URLs anyway... Why should libraries move off the already embedded technology (MARC), when it seems to largely be fit for purpose? What things motivate the use of linked data in libraries? The only "new" need Salo identified was "sharing", but perhaps MARC can be more easily shared.

[Automatic Creation of Mappings between classification Systems for Bibliographic Data](#)

By: Magnus Pfeffer

Pfeffer gave a talk about ontology matching in classification systems using matching of manifestations of the terms in bibliographic data. There are many similarities to the work of Knut Hegna ##REF## The statistical method used was comparison of intersection and union of matches. The perceived problem with this method was the lack of negative inferences (where there are definitely no links). The links are evaluated by measuring recall and precision against a "gold standard", feedback from this evaluation is fed back into the gold standard. Pfeffer identified issues in SKOS related to qualification of relations.

This talk was interesting for me as it relates directly to work carried out at University of Oslo/NTNU.

Cross-Lingual Semantic Mapping of Authority File

By: Nadine Steinmetz/ Magnus Knuth/ Harald Sack

Again, this talk was about matching in authority files, this time between a bibliographic data set and DBPedia. Here, issues involving how concepts are treated (broad strokes) in DBPedia, while classification systems have largely narrow concepts.

Mash-up for Book Purchasing

By: Philipp Zumstein

Zumstein created a browser plugin that helped subject specialists in their book-ordering workflow; rather than the traditional, multi-search process, the data from various resources (holdings, catalogues, OPAC, etc) is aggregated and presented together so that time is saved. A mix of different technology is used: RDF, Z39.50, etc.

As the system uses ISBNs, the traditional issues are noted (ISBN is not an ID, linking editions, etc.)

AgriVIVO: A Global Ontology-Driven RDF Store Based on a Distributed Architecture

By: Valeria Pesce/ Johannes Keizer/ John Ferreira

Talked about motivations and specifications that caused them to choose VIVO; how they adapted it to work for them by modifying core VIVO and using Drupal on top. Problems in the data: lack of identifiers. A lot of examples of how small OSS projects are sometimes problematic — one guy knows a lot and there's little documentation, but I suspect that using a commercial product that was "fixed" would leave a much bigger mess (cf. all experience with custom solutions and upgrades in the history of IT).

I have some issues with the approach taken in VIVO; as far as I can see, using normal linked data techniques (like vocabulary re-use) from domains outside

the core VIVO ontology makes upgrades break — unless one jumps through vocabulary importing hoops. As this is the case, I wouldn't call VIVO a linked data tool, but rather an ontology driven tool.

HEAL-LINK Activities and Plans on Annotating, Organizing and Linking Academic Content

By: Nikolas Mitrou/ Nikolaos Konstantinou/Dimitrios Kouis/ Periklis Starou

Spoke about his projects on authorities and e-books.

Linked Data for Libraries: Great Progress, but What Is the Benefit?

By: Richard Wallis

What is the benefit for libraries? What is the benefit for users? No web of library data yet. Traction is coming in the form of Google's knowledge graph — RDF-like structures (entities). Why is this relevant to libraries? BnF: 80+% of search comes from search engines; people aren't using search interfaces we provide. This is just true, so exposing data on the web is very important. Changing from cataloguing to catalinking (Eric Miller). Use the Web. Linked data. Schema.org Schema.org is a good tool for search engines. BIBFRAME is companion vocabulary for libraries.

BIBFRAME: Libraries Can Lead Linked Data

By: Julia Hauser/ Reinhold Heuvelmann

BIBFRAME is an exchange format built on linked data principles. (A direct replacement for MARC). Linked data is more agnostic than MARC, data can be anything; but it should support different models, support RDA and other cataloguing rules. Be extensible for new material types. BIBFRAME must be enriched and stabilized. What about version-controlled? GNL implements RDA and FRBR, how well does BIBFRAME transport RDA? BIBFRAME can replace MARC, but lots of changes must happen in order for this to work out.

Building a National Ontology Infrastructure

By: Matias Mikael Frosterus/ Sini Pessala

Session summary:

Large-scale ontology project. Centralized. The reverse of distributed data. Improve interoperability across the spectrum of users. Lightweight SKOS ontologies intended for annotations. They provide an upper ontology. Ontologies are easy to link. Rigid definitions. Can be explicitly related to one-another. Ontologies used to link different resources together, but the changes are difficult to manage directly because there are many users and systems. The upper ontology provides an abstract layer that is rarely updated; links between the domain ontologies are reduced. Interaction happens via the upper ontology.

Trilingual ontology. YSO top level based on DOLCE. Cognitive. Draft. Culture before language. Represent information in the way Finnish culture represents them, not like Swedish/English language.

On the Way to a Holding Ontology

By: Carsten Klee/ Jacob Vob

The speech on Holding ontology was quite impressive and of great use in implication .

Wednesday, November 27

Decentralization, Distribution, Disintegration – towards Linked Data as a First Class Citizen in Libraryland

By: Martin Malmsten

Conversion to RDF from various formats used in library data helps to give an understanding of the data — RDF-thinking helps you to get what the data is modeling, not just simple description. Next logical context: data -> linked data (more data with links). Sick of ineffectiveness of MARC in workflows because it

didn't support what they needed to do — converting data between the formats was a pain in the arse, so they started creating a linked-data-based system. Linked data system adds the functionality they needed. Use JSON-LD because it's a tool non-RDF people can use. RDF lets you solve problems as they arise, while other formats (MARC) can't extend to cover new functionality.

Open source, self-programmed. Important distinction is that there's no difference between you data and other people's data. You need to be able to handle the distributed model.

They have MARC, but it's generated from RDF. Co-existence with the legacy data is necessary. MODS is also there. The formats need to be kept away from the core system — they are exchange formats. Inside, they have linked data, outside they have many data formats. This separation keeps the core system safe from the influence of the bad ideas.

Data aggregates from publishers/vendors breaks the idea of linked data — everyone should publish their own data, if not the aggregate must be transparent. Linked data needs interfaces for users. In order to convince people, they created an interface and helped institutions create their own linked data. TBLs 5-star model isn't enough, actually using the data makes it real; to be useful it needs to be used. Linked data and a UX is useful data. Do things quickly, try to understand what people are trying to achieve. RDF isn't a format; it's a way of representing data.

The “OpenCat” Prototype: Linking Public Libraries to National Datasets

By: Agnes Simon

Have produced data. It is used a bit by non-library people. It is also used by libraries. Created an OPAC+ based on linked data. Users are happy with results like FRBR-ization, enrichment, navigation changes.

Semantic Web Technology in Europeana

By: Antoine Isaac/ Valentine Charles

Spoke about the Europeana project and its tools and workflows.

Specialising the EDM for Digitised Manuscripts

By: Steffen Hennieke/ Violeta Trkulja

Digitized manuscripts to Europeana. Tool chain for data migration, Pundit and the openglam community. Heterogenous data formats, TEI, MARC, MAB2, etc. Europeana Data Model is generic in order to tackle heterogenous data formats. I have obvious problems with this approach. Application of LOD to Enrich the Collection of Digitized Medieval Manuscripts at the University of Valencia Jose Manuel Barrueco Cruz

Spoke about the process of converting traditional metadata to linked data.

ResourceSync for Semantic Web Data Copying and Synchronization

By: Simeon Warner

A nice introduction featuring some good – and realistic — ideas around library resource management, Warner gave a coherent, smart and understandable overview of ResourceSync.

From Strings to Things: A Linked Open Data API for Library Hackers and Web Developers

By: Fabian Steeg/ Pascal Christoph. and

Enhancing an OAI-PMH service Using inked Data: A Report from the

Sheet Music Consortium

By: Stephen Davison

These were great presentations that captured my attention even at this last stage of the conference. Some of the methods of creating context seemed very well designed.

Application of LOD to Enrich the Collection of Digitized Medieval Manuscripts at the University of Valencia

By: Jose Manuel Barrueco Cruz/ Cristina Garcia Testal and lastly

Exposing Institutional Repositories as Linked Data – a Case – Study

By: Vitali Peil/Najko Jahn

The Valuable data provided in these sessions was very useful.

Conclusion

The library profession is constantly adjusting to new technical developments, standards and innovative new services. In developing countries like India, it is difficult for librarians and libraries to invest in continuing training and professional development, and hence, the so attended Swib13 conference on Semantics Web Libraries had been a good opportunity for me in terms of knowledge sharing and training.

During these 3 days I have learnt many new professional concepts such as RDF and Ontology by attending the workshop on “introduction to linked open data” and the other lectures.

Thanks to the BI International for the grant I received for attending this conference. I wish I would be able to submit papers for

presentation, in the next upcoming conferences at Germany and other European countries with the help of BI International.
