

Introduction to Linked Open Data Workshop, SWIB 2016

When: Monday, December 4th, 13:00 - 19:00

Where: Friedrich Ebert-Stiftung Bonn, Bonner Haus, Room TBA

Google Drive with Workshop Materials: bit.ly/SWIB16LODintro

Session Details

Workshop Abstract

This introductory workshop aims to introduce the fundamentals of linked data technologies on the one hand, and the basic issues of open data on the other. The RDF data model will be discussed, along with the concepts of dereferenceable URIs and common vocabularies. The participants will continuously create and refine RDF documents about themselves including links to other participants to strengthen their knowledge of the topic. Based on the data created, the advantages of modeling in RDF and publishing linked data will be shown. On a side track, Open Data principles will be introduced, discussed and applied to the content that is being created during the workshop. Attendees are not expected to have any technical, RDF or Linked Open Data experience. We do ask that attendees bring a laptop with a modern web browser for participation.

Facilitators

Jana Hentschke: German National Library, J.Hentschke@dnb.de, @junicatalo

Uldis Bojars: National Library of Latvia, captsolo@gmail.com, @CaptSolo

Christina Harlow: Cornell University, cmh329@cornell.edu, @cm_harlow

This workshop is built off of the work of [Felix Ostrowski](#) and [Adrian Pohl](#). See their previous workshop details [here](#).

Workshop Resources

Links will be added / opened up as the resources become available for public sharing.

- [Link to this Workshop's Google Drive Folder \(containing all related resources\)](http://bit.ly/SWIB16LODintro) (bit.ly/SWIB16LODintro)
- [Link to the Session's Shared Notes](http://bit.ly/SWIB16LODintroNotes) (bit.ly/SWIB16LODintroNotes)
- [Link to Slides](http://bit.ly/SWIB16LODintroSlides) (bit.ly/SWIB16LODintroSlides)
- [Link to Etherpad for workshop data generation](http://etherpad.lobid.org/p/swib16-ws) (<http://etherpad.lobid.org/p/swib16-ws>)
 - Group 1's Etherpad: <http://etherpad.lobid.org/p/swib16-ws-a>

- Group 2's Etherpad: <http://etherpad.lobid.org/p/swib16-ws-b>
- Group 3's Etherpad: <http://etherpad.lobid.org/p/swib16-ws-c>
- Group 4's Etherpad: <http://etherpad.lobid.org/p/swib16-ws-d>
- Group 5's Etherpad: <http://etherpad.lobid.org/p/swib16-ws-e>
- Group 6's Etherpad: <http://etherpad.lobid.org/p/swib16-ws-f>
- Group 7's Etherpad: <http://etherpad.lobid.org/p/swib16-ws-g>
- Facilitators/Demo Etherpad: <http://etherpad.lobid.org/p/swib16-ws-z> (live demo etherpad)
- Backup/Complete Etherpad: <http://etherpad.lobid.org/p/swib16-ws-h>
- [LODLive](#) Visualization / [LODLive Backup](#)
 - http://graphthinking.com/~fo/LodLive/app_en.html?http://etherpad.lobid.org/p/swib16-ws-b#ag
- [Fuseki \(lobid\)](#) / [Fuseki Backup \(cmh2166\)](#) (username: fuseki / password: SW1B2016workshop)

Participant Requirements (Optional but Recommended)

Participants are strongly recommended to bring laptops with internet connection and web browser (not Internet Explorer). We are working to have a space that supports group work (so be prepared to take part and move around).

Workshop Schedule

13:00 - 13:15	General Workshop Introduction (Christina)
13:15 - 14:00	RDF Introduction: Create Your Graph (Uldis)
14:00 - 14:30	Linked Data Introduction: Link Your Graph to Your Group's (Uldis & Jana)
14:30 - 15:15	Linked Data & Semantic Web: Expanding Your Graph (Jana)
15:15 - 15:45	30 minute break
15:45 - 16:15	Linked Open Data & Data Licensing: License Your Graph (Christina)
16:15 - 17:15	Linked Data Experimentation: SPARQL & Visualization (Christina)
17:15 - 17:30	15 minute break
17:30 - 18:00	Linked Data Examples, Resources & Datasets (Christina)
18:00 - 19:00*	Wrap-up & Help Moving Forward (all)

Workshop Outline / Topics

General Workshop Introduction - 15 minutes, Christina

- Brief facilitator introductions per title slide.
- Links to workshop resources (make sure everyone is set up):
 - Google Drive
 - This Document
 - Shared Notes
 - Etherpad links
- Our expectations of you
 - Follow our workshop etiquette adapted from the [Recurse Center Rules](#), please.
 - This is an informal session, so ask questions whenever, and also reach out to facilitators.
- Schedule for the session (quick overview, apt to change)
- Recap the abstract information / workshop goals:
 - Introduce & Leverage the RDF data model
 - Create & Refine RDF Documents
 - Explain the Fundamentals of Linked Data Technologies
 - Introduce & Discuss Basic Issues of (Linked) Open Data
 - Apply Open Data principles to RDF data
 - Understand the Advantages of Modeling in RDF & Publishing Linked Data
 - Share existing LODLAM Tools, Projects, Resources, Examples, Communication channels...
- Audience's goals, domains, backgrounds:
 - Write on whiteboard or in the shared notes (whichever is available) to revisit at end to make sure we touched on them (or will follow up and contact)

RDF Introduction: Create Your Graph - 45 minutes, Uldis

- Personal Introductions & RDF: (15 minutes)
 - Have facilitators introduce themselves
 - Walk through their example of a personal graph starter cards.
 - Use initials as IDs. If same initials appear within one group note the importance of unique IDs.
 - How they enter their example on the facilitator etherpad document.
 - **Mention this introduction is creating RDF. What is RDF?**
 - **Triples**
 - **Abstract Model**
 - **Serialization is Turtle**
 - **Prefixes & Namespaces**
 - **Statements or Triples**
 - **Punctuation in Turtle (might go back to previous slide to show turtle**

structure/syntax)

- Show a diagram of the model/data being created.
- Creating your Graph: RDF Introduction: (30 minutes)
 - Make sure everyone can get access to etherpad, are in groups, have logistics down for further group work.
 - Outline folks' groups now, link each group to their separate etherpad documents.
 - Get participants to create their own name tags, first writing them down.
 - Then have each person add their personal information to their group's etherpad.
 - Review the data created in their shared etherpad.
 - If time / for those who are quick: Brief mention of classes, `rdf:type` and let participants mark themselves as `rdf:type foaf:Person`
 - If time, invite participants to try diagramming their graphs as well through pen/paper or Google Drawings (or other)

Linked Data Introduction: Link Your Graph to Your Group's Graphs - 30 minutes, Uldis, Jana

- Relating the Graphs: Linked Data introduction (10 minutes)
 - Have Jana, Uldis link themselves through `schema:knows`
 - Walk through their example of a personal graph being linked.
 - Look up `schema:knows` through its URI to show the definition.
 - Show how they enter their example on the facilitator etherpad document.
 - **Importance of Links in Linked Data**
 - **Uniform Resource Identifiers / URIs**
 - **Networks emerging / Social Networks example**
 - Show the diagram of the model/data with resources (Uldis, Jana) linked.
- First activity: Links in a single group (5 minutes)
 - For each group, first, have everyone introduce themselves to their group (if they haven't already).
 - Then have each member in the group create `schema:knows` links within their own group.
 - Add the new data to their group etherpad.
- Relating the Graphs Further: The Power of the Link (5 minutes)
 - Have Jana & Uldis link themselves with Christina through `schema:knows`
 - Walk through their example of a personal graph being extending with the new link/resource.
 - Show how they enter their example on the facilitator etherpad document.
 - **Power of Links to expand the graph**
 - Show the diagram of the model/data being expanding through linking existing and new resources.
- Linking Game (link everyone's personal graph among all the different groups): (10 minutes)
 - Have people who know each other from different groups raise their hand.

- For groups with nobody who knows another from another group, ask 1 person to switch with another, introduce themselves,.
- Have the new to that group people create further `schema: knows` links in that group's etherpad, expanding the graph. If time, bring in the new person's previous group's data too and add those links.
- Ask all who moved to switch back to their original groups.
- Make sure everyone captures newly expanded graph, and talk about traversing the graph (aka how far do you go).

Linked Data & Semantic Web: Expanding Your Graph - 45 minutes, Jana

- **Expanding the Graph: Places & URIs (5 minutes)**
 - We want to place names (literal string) with DBpedia (Wikidata, Geonames, other) reference. Facilitators demo with their data.
 - Wikidata vs DBpedia vs other sources discussion (brief)
 - How to look up these references (demo) - show exactly where / what the URI is
 - How + where to add them to the existing data
 - Resource versus literal
 - Recall introduction to HTTP URIs and use of prefixes discussed earlier.
 - Explain the data you can now pull in from those Place URIs
 - Recall work just previously on linking and graph extension
 - **Brief mention of classes, `rdf:type`, and entailment if appropriate for the group.**
 - Show how those URIs are added to the facilitator's group etherpad document.
 - Show a diagram of the model/data being created.
- **Creating a Global Graph (5 minutes)**
 - **HTTP URIs and links scaling to data globally**
 - **"Giant Global Graph"**
 - **Reference and relationship of the Global Graph to the World Wide Web**
 - **Anyone can say Anything about Anything**
- **Linked Data Principles from Tim Berners Lee. (5 minutes)**
- **Add Place URIs to your graph (10 minutes)**
 - Have participants look up and add URIs from DBpedia or another Linked Open Data vocabulary for their data.
 - Review the data created in their shared etherpad.
 - If time, invite participants to extend their diagram of their graphs through pen/paper or Google Drawings (or other)
- **Expanding the Graph further: Adding Interest Areas & LOD Vocabularies (10 minutes)**
 - Vocabularies for finding Interest, Topic or Concept URIs.
 - Facilitators demo looking this up and grabbing a URI for an interest area, applying back to their data.
 - Demo how you add `foaf:interest` plus the interest URI to your data.

- Show how those URIs are added to the facilitator's group etherpad document.
 - Show a diagram of the model/data being created.
- Add Interest URIs to your graph (15 minutes)
 - Have participants look up and add URIs from DBpedia or another Linked Open Data vocabulary for their data.
 - Review the data created in their group's etherpad.
 - If time, invite participants to extend their diagram of their graphs through pen/paper or Google Drawings (or other).
- **Vocabulary + Ontology options (and how you might decide between them).**
 - **Point out specific ones for Interest or Concept URIs.**
 - **Ontologies & Models (meta-modeling vocabularies) for GLAM (Galleries, Libraries, Archives, Museums)**
 - **Vocabularies available in RDF / as Linked Data for GLAM**
 - **Instance datasets**

Break 30 minutes

- Check for questions, state of affairs
- Have cake
- **Load triples if appropriate / ready for later querying and experimentation (facilitator)**

Linked Open Data & Data Licensing - 30 minutes, Christina

- Making your Graph Open: Applying Data License & Metadata to their dataset (5 minutes)
 - Christina walks through adding a graph creator and creation date (provenance) for the dataset using `dcterms`
 - Christina adds a license statement (CC0) to the dataset.
 - Points to make: applying to whole dataset (using the '<>' notation here, explain).
 - CC0 URI lookup and review.
 - VOiD as an alternative for asserting this provenance information.
 - W3C Recommendation to capture data license using `dcterms:rights` and a statement like CC0
 - Discuss some common data licenses, then how you might choose a license and why.
- **What is Open? What is Open Data? (5 minutes)**
 - **Access, License, Non-proprietary formats**
- **Examples of licensing data: STW (STW Thesaurus for Economics, SKOS dataset)**
 - **Show the `dcterms:rights` redirect, how to apply an `opendata` license, other dataset metadata that is captured, how this is all exposed on their HTML rendering (the public website).**
- More Examples of Licensing and Open Data from the Library Domain: (5 minutes)
 - Rightsstatement.org was mentioned: used by Europeana, DPLA, others for digital objects
 - Europeana + Getty licensing as exemplars

- DWBP best practices - here's the one on licensing:
<https://www.w3.org/TR/dwbp/#licenses>
- Open Data - there's a lot of pointers on Share-PSI localised guides page:
https://www.w3.org/2013/share-psi/wiki/Localised_Guides
- **5 Star Linked Data (5 minutes)**
 - **We are building to this through our work so far.**
- Participants now apply Data License & Metadata (10 minutes)
 - Participants add a graph creator and creation date for the dataset using `dcterms`
 - Participants add a license statement (CC0) or other to their dataset.
 - Review the data created in their group's etherpad.
 - If time, invite participants to extend their diagram of their graphs through pen/paper or Google Drawings (or other).

Linked Data Experimentation: SPARQL & Visualization - 1 hour, Christina

- Demo of data created in first part (50 minutes)
 - LODLive Visualization (10 minutes)
 - **What is LODLive?**
 - **Quick idea of the technology used, with data created**
 - Christina walks through captures of LODLive browsing of the graph created by the facilitators
 - Jumps over to live LODLive and shows same steps, but now with whole room's data
 - Show how the relationships were built, the data it pulled in was originally created by the participants
 - SPARQL & Fuseki Querying (10 minutes)
 - **What is SPARQL**
 - **Give basic overview :**
 - **Query language similar to SQL**
 - **Built of Turtle, a different RDF Serialization**
 - **General Form of a SPARQL Query**
 - **Note there is also SPARQL Updates (requires different underlying technology to accept those)**
 - Christina demos some example queries using the Fuseki SPARQL Query endpoint on the participants' data loaded.
 - Set of sample queries to run:
 - Names of workshop participants
 - Acquaintances of workshop participants
 - Acquaintances of workshop participants by name
 - Localities and Countries of workshop participants
 - Shared Interests of workshop participants

- Home cities of workshop participants
- Have groups explore their data via SPARQL Queries + LODLive (30 minutes)
 - Have the sample queries ready for them to reuse in this etherpad and in the slides: <http://etherpad.lobid.org/p/swib16-ws-sparql>
 - Have back-up visualization and data manipulation tools ready/listed in case of wifi or network issues
- Overview of some Linked Data Technologies used for understanding our data: (10 minutes)
 - Triplestores & SPARQL
 - Linked Data Fragments
 - Clients & Visualization
 - Elements of Linked Open Data stack
 - Highlight areas where we could talk more in breakouts, other spaces

Linked Data Examples, Resources & Datasets - 30 minutes, Christina

- Walk through some projects / examples (20 minutes)
 - **Linked Jazz**
 - **NYPL Browse**
 - **LD4L Editor**
 - **Oslo Public Library Catalog (Editor Videos)**
 - **Getty Vocab**s
 - **Europeana SPARQL Endpoint**
 - **Geographic Project mentioned last year...**
 - **Opaque Namespace**
 - **DPLA Color Browse (More an Open Data example)**
- Have participants play with examples or add other information to the shared notes (10 minutes)

Wrap-up & Help Moving Forward - 1 hour, probably not the full time, Christina, Jana, Uldis

- Discussion among participants: own (potential usage) of LD (30 minutes)
 - What projects could you see for using this?
 - What support would you need to make it happen?
 - Any questions or ideas to share?
- Return to the full group and have each group report back
 - Possible discussion point in this: Maintain momentum (ask participants back as a group)
 - What communication channels to use for this ongoing work
 - Document our workshop but other spaces to watch
 - How to continue to support data modeling development in our communities
- Return to Goals from the General Workshop Introduction, see how we did, how people feel

