

Report on SCA/SCB Course

Title of the course

Linked Data for Science and Innovation Studies: Using the SMS platform for enriching your research data

Venue

Vrije Universiteit Amsterdam, Amsterdam, Netherlands

Date

11-12 June 2018

Organizers

Peter van den Besselaar; Frank van Harmelen; Ali Khalili; Al Idrissou

Objectives

The aim of the course was to train researchers to get familiar with the Semantic Web and Linked Data technologies by using the SMS (Semantically Mapping Science) platform for converting, integrating, enriching and browsing data in the STI domain.

Main results

In a mixture of lectures and hands-on, participants were trained to produce and use linked data and the various possibilities provided by those data in the SMS platform. To summarize, the participants:

- i) got familiar with the SMS platform and its datasets and functions;**
- ii) learnt how to convert their own data so these can be linked to other data in the platform, as this requires a specific format**
- iii) enriched the data through entity recognition and geolocation services – the first was related to textual data where for example organization names can be recognized in a CV, or technical terms can be recognized as referring to a specific disease;**
- iv) linked one's own data to the data in the SMS data store**
- v) used the SMS faceted browser to select those data from the SMS data store that the user needs for his/her research project;**
- vi) understood how this data browsing results in a query that can be used to retrieve the data from the data store in a usable format.**
- vii) Finally, these enriched linked data then can be used for further analysis.**

List of participants: **see ANNEX 1**

Programme: **see ANNEX 2**

Assessment: **see ANNEX 3**

Materials:

1. A Quick Tour of Linked Data http://risis.eu/wp-content/uploads/2018/06/1_quickTourOfLinkedData.pdf
2. A Tour of SMS Platform http://risis.eu/wp-content/uploads/2018/06/2_smsPlatformIntro.pdf
3. Data Extraction http://risis.eu/wp-content/uploads/2018/06/3_dataExtraction.pdf
4. Data Storage and Querying http://risis.eu/wp-content/uploads/2018/06/4_dataStorageQuerying.pdf
5. Data Enrichment http://risis.eu/wp-content/uploads/2018/06/5_dataEnrichment.pdf
6. Neo4J experience <http://risis.eu/wp-content/uploads/2018/06/neo4j.pdf>
7. Data Linking <https://github.com/alkoudouss/RISIS-2018-course> and <https://github.com/alkoudouss/alignments>

ANNEX 1

LIST OF PARTICIPANTS

Name	Affiliation	Country	Email
Cezary Biele	National Information Processing Institute in Warsaw	Poland	cezary.Biele@opi.org.pl
Luis Sanz-Menendez	CSIC	Spain	luis.sanz@csic.es
Laura Cruz	CSIC	Spain	Laura.Cruz@csic.es
José Luis Ortega	CSIC	Spain	jortega@orgc.csic.es
J.C.Teelken (Christine)	VU Amsterdam	Netherlands	j.c.teelken@vu.nl
Patricia Martinkova	Czech Academy of Sciences	Czech Republic	martinkova@cs.cas.cz
Pedro Parraguez Ruiz	Technical University of Denmark	Denmark	ppru@dtu.dk
Dimitris Kosmidis	DAEM	Greece	d.kosmidis@daem.gr
Eric J. Iversen	NIFU	Norway	eric.iversen@nifu.no
Arash Hajikhani	LAPPEENRANTA UNIVERSITY OF TECHNOLOGY	Finland	arash.hajikhani@lut.fi
Yasushi HARA	CEAFJP/ L'École des hautes études en sciences sociales, Pari	France	yasushi.hara@ehess.fr
Chiara Latronico	Data Research Analyst for the project Golden Agents, University of Amsterdam	Netherlands	C.Latronico@uva.nl

ANNEX 2 PROGRAMME

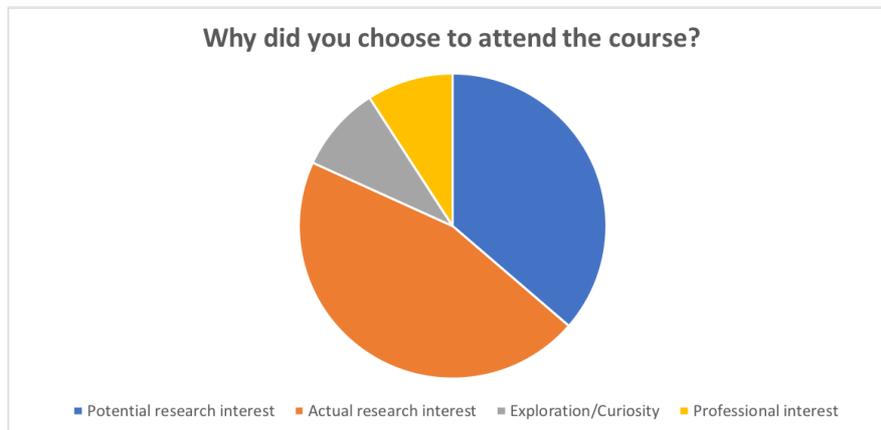
Day 1 (Monday 11th June)

9:00 - 9:30	Coffee, get to know each other
9:30 - 10:15	An introduction to the Semantic Web and Linked (Open) Data
10:20 - 11:15	Get to know the SMS platform: datasets, functionality, users
11:15 - 11:30	Short Break
11:30 - 12:30	Installing the required tools
12:30 - 13:30	Lunch
13:30 - 14:45	Converting data to RDF - CSV to RDF
14:45 - 15:00	Short Coffee Break
15:00 - 15:30	Storing data in a graph-based store (triple store)
15:30 - 16:30	Enriching data - Named entity recognition - Geo-enrichment - Google Spreadsheet - SMS geo-enricher
16:30 - 17:00	Summary and discussion
Social Event	

Day 2 (Tuesday 12th June)

9:00 - 10:15	Linking data, part 1
10:15 - 10:30	Short Break
10:30 - 12:00	Linking data, part 2
12:00 - 13:00	Lunch
13:00 - 14:00	- Query, browsing and visualization of linked data - SPARQL - Basic browsing of data - Link navigation
14:00 - 14:15	Short Coffee Break
14:15 - 15:30	Discussion of use cases
15:30 - 16:30	Summary and discussion
Course Evaluation	

ANNEX 3 ASSESSMENT



RATINGS

<i>Were the course objectives clearly defined?</i>	
<i>Were the contents of the course consistent with the course description?</i>	
<i>Has this course stimulated your interest?</i>	
<i>Was the course well organized?</i>	
<i>Was the course well structured?</i>	
<i>Have teaching materials facilitated learning?</i>	
OVERALL SATISFACTION (scale 1-10)	<div style="background-color: #FFD700; border-radius: 50%; width: 40px; height: 40px; display: flex; align-items: center; justify-content: center; margin: 0 auto;"> 8.6 </div>

BEST FEATURES

<p style="color: green; font-weight: bold;">Teachers and instructors</p> <p style="color: green;">the examples of the use of the tools, The demo of the SMS platform</p> <p style="color: green; font-weight: bold;">Geo-enrichment</p>

ASPECTS TO IMPROVE

<p style="color: red; font-weight: bold;">simplifying the installation of the tools</p> <p style="color: red;">course might have preceeded by a skype session in order to help everyone install the necessary software. Technical problems of the users broke the flow of the workshop using JP/CN data to analyze multilingual datasets. I hope this workshop could be held at Japan :)</p>
--

very innovative use of technology to create usable applications to work with STI data, hope that all the engineering efforts to build the app move to some public server so everyone can use with less effort on installation

using LD-R stardog to analyze DBpedia. Finally understand lenticular lense explorer the subject, the blend of technologies, the potential to apply in various fields, nice team! combination of the overview of components with a hadns-on approach to the individual components

there should be more dissemination activities. it is a hot aspect, should reach various segments with tools + theory

have video guide of installation and a couple of use cases in video for in-depth learning and experimentation

spend more time working on the tools and less time installation of the tools and troubleshooting