INFO 116 - Semantic Technologies

-Exploring the semantic web

By 213903, 218599, 229658, 222447, 231167

Department of Information Science and Media Studies

University of Bergen

November 2014

(This page left blank intentionally)

Introduction

In this written report, you are going to get an understanding of what has been accomplished, get a little peek into our different solutions, and how we came to the conclusions that were made. The report will go through what each member of the group has contributed, as well as a walkthrough of our methods creating the ontology. There will also be an explanation on which data sources are used to gather the information acquired for both the ontology and markup. We will show some example queries to what our ontology could be used for. By using our added semantics, anyone who want can use our added information and put it into their own semantic database.

What we did

In the beginning when we got the term paper there were a lot of talk on how we were going to take on the problem. We had to read the assignment a couple of times to figure out what had to be done, who should do what, and where to start. After the planning was finished, we could finally start on the main part of the term paper. One of the members started working on the ontology and had it finished before the next group meeting. This made it so that we could start with the markup, which involved RDFa and JSON+LD. We picked five courses that we found had enough variety between them, and were interesting enough to write about and would gild for some good markup. Two of the members took care of the RDFa, and were working together to find the best solution for our ontology. In the meantime, another member wrote the JSON+LD files, which we later added to the .html files. This is mainly what took time during the assignment. We had to redo our ontology a couple of times because of some lack of content, and to correct some wrongs to make it work correctly. Logs were written after each labmeeting, to see who came and what had been done. After all this had been done, we found some interesting queries we could get answers to using our ontology. These were written in SPARQL and the whole team contributed in finding the most interesting queries.

We ran into a couple of problems when we tried to implement the SPARQL queries. The problems which occurred happened since the whole ontology missed some crucial details. All of the individuals and data properties were missing, and we were only a few days before the deadline.

Ontology

As with any ontology, this is an ontology of things. We have a hierarchy with all the subclasses we found necessary. They are divided after what we found the best, and placed in each category without any duplicates. Each course is linked with the different subclasses that fits and is needed by that class.

BASV-INFO is the superclass for the main part of the ontology. All its children are what is needed to get that degree. This includes all the courses you can take, all the different subject areas you are going to have and learn, and all the different course info connected to the subjects. Although, not all the information about the different courses are implemented. Some of the courseinfo would just be busywork, which is not necessarily needed. Of course it would be of some use to have it all, but it would just take up to much of our time.

The courses are divided after introduction, exphil, choice, and specialization. Introduction has the different 100-courses that are needed for taking the bachelor, while the exphil class has the two different exphil-courses that you can choose for this bachelor. In the specialization there are all the specialization courses. Choice has all the 2xx-courses, which you can choose to have in the fifth and sixth semester, or before if you choose so or are missing some studycredits. These are all connected to their different subject areas, what kind of examination types there are, their schedule and what kind of assignments you can expect.

Parentclass to BASV-INFO is the institute where the tutoring is going to take place. We've included some different institutes, so if there should be another degree added to the ontology, it won't be necessary to do a whole lot of work to put it in.

We disjointed the subjects from each other in the choice and specialization. Everytime we added another siblingclass to a course, it also got all the information from it's siblings added to it. If we removed the information from one of the classes, they all lost it, but it all worked after the classes were disjointed.

An individual is an instance of the class course, and it's connected together with the instance of the other class. This gave us a lot of trouble when writing the queries, since in the beginning we only used classes, and could not understand why they were not doing as told. When we added this information, it all ran smoothly.

All individuals in the members list, is just for the normal course in taking the bachelor. Even though, the course can still be taken in both the autumn and spring semesters.

Data Sources

We used SPARQL as our protocol and query language. This is specifically design to be used with semantic for extracting data sources. All of our data sources are derived from the uib.no webpage. We used specifically the webpage for the information science study to gather our data. The JSON+LD files added some of the attributes from schema.org.

What can the website do with the added semantics?

In our assignment, we have added semantics to the webpage via the use of RDFa's and JSON-LD.

RDFa (Resource Descrition Framework in attributes) is mainly a recommendation by W3C that lets you provide a set of attributes that carries metadata in XML documents. The attributes within RDFa are as follows: **about** (the resource of the metadata specified by an URI or CURIE), relationship and reverse-relationships (**rel and rev**), partner resource specified by **src, href, and resource**. **Property** (lets you specify the property for the elements content or the partner resource). [1]

There are also optional attributes that lets you do additional specifications within other attributes. These are the following optional attributes:|

Content - Overrides the elements content when the property attribute is used. **Datatype** - An attribute that lets you specify the data type of the text in the property attribute. **Typeof** - An attribute that specifies the type of RDF in the partner resource or subject.

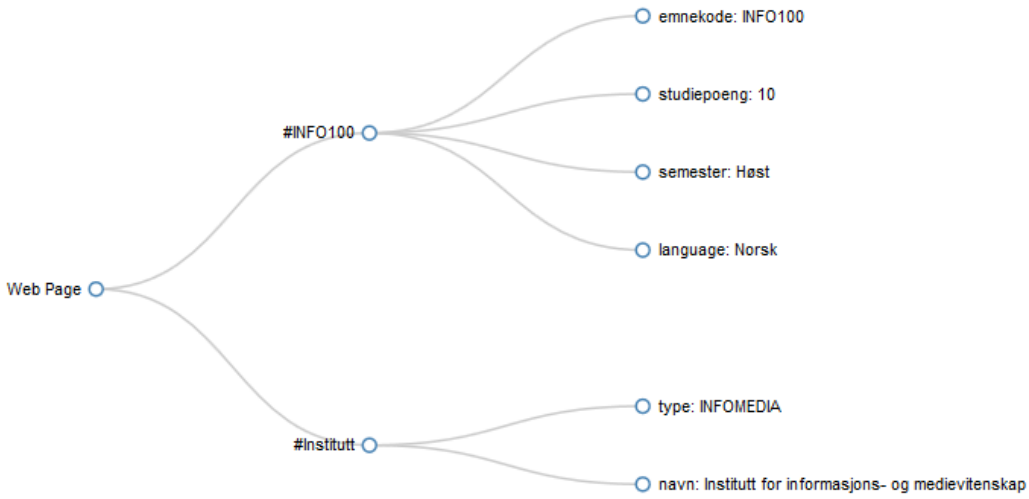
JSON-LD (Javascript Object Notation for Linked Data) is basically a format for transporting linked data to create a "tidier" and organized web. The essential priority of JSON is to convey data between a server and a web application. Linked data is easiest explained by saying you share interlinked, structured data as to connect data with other relevant data. All in all it is when data from different sources gets connected both data-linked wise and semantically. This is done in a way that both the computer and person reading the data can understand it and interpret the meaning behind it. [2]

What can the website do with the added semantics?

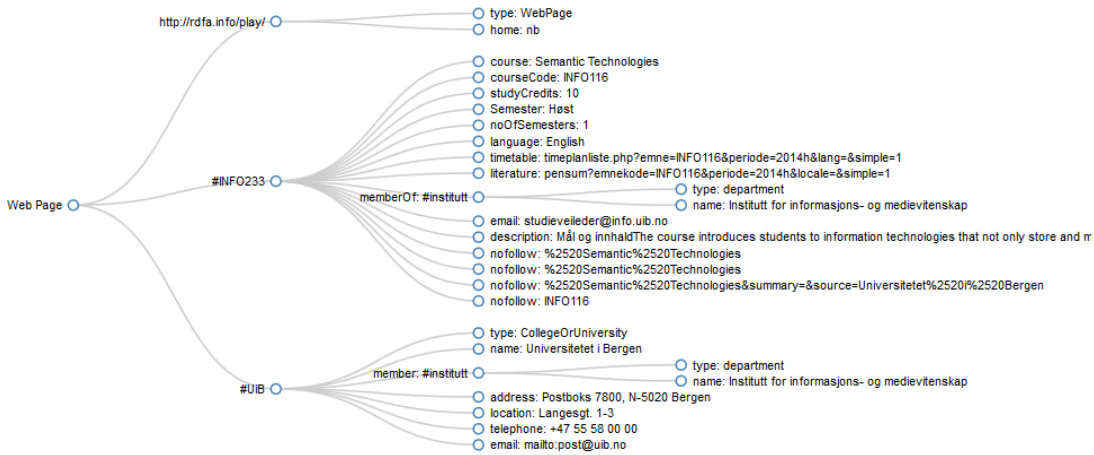
With our added semantics, the website can basically answer basic queries about the various subjects. How many of the subjects that are during the autumn, how many subjects that are during the spring and has 10 studypoints, which ones that contains programming, all courses that are in english, if there subjects with different group projects or mandatory attendance etc.

Some examples of markup

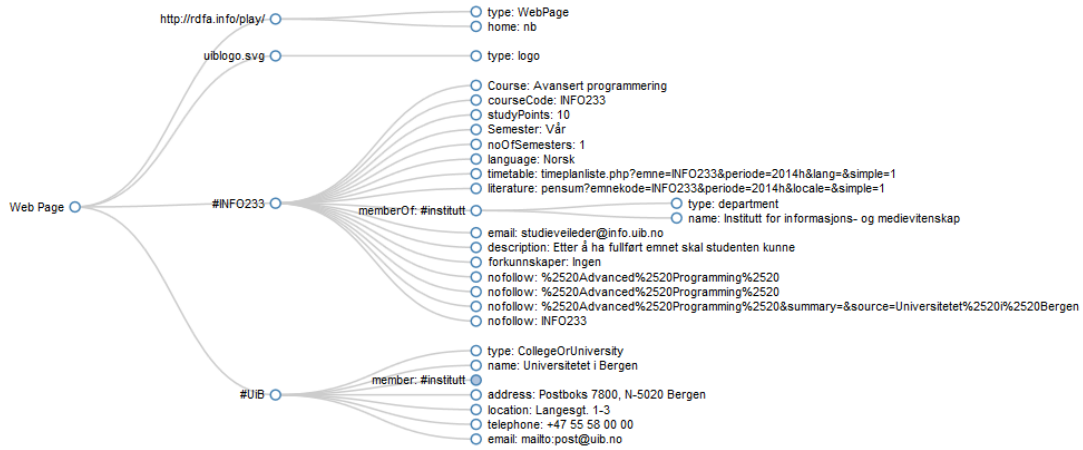
INFO100 – RDFa markup



INFO116 – RDFA markup



INFO233 – RDFa markup



SPARQL Queries

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

PREFIX owl: <http://www.w3.org/2002/07/owl#>

PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

PREFIX : <http://www.semanticweb.org/simen/ontologies/2014/9/infovitFINAL#>

1#Displays all possible subjects in BASV-INFO-program

```
SELECT  ?subject
        WHERE {
        {?subject rdfs:subClassOf :Choice. }
        UNION
        {?subject rdfs:subClassOf :Exphil.
}
        UNION
        {?subject rdfs:subClassOf :Introduction.
  }
        UNION
        {?subject rdfs:subClassOf :Specialization.
  }
```

subject
INFO216
INFO233
INFO262
INFO232
INF207
INFO212
INFO282
INFO207
SVSEM
SVEKS
INFO100
SV100
INFO125
INFO102
INFO132
INFO103
INFO110
INFO115
INFO116

2#Which ones has programming

```
SELECT ?subjName
        WHERE {
        ?subjName :covers :Programming.
}
```

subjName
INFO233V15
INFO132H14
INFO102V15
INFO232H14

3#Covers databases but NOT querylanguage

```
SELECT ?subjName
        WHERE {
        {?subjName :covers :Database} MINUS {?subjName :covers :Theory}
}
```

subjName
INFO125H14
INFO102V15

4#How many groupprojects or mandatoryAssignements

```
SELECT   ?courseName ?groups ?assignments
        WHERE {
  {?groups a :GroupProject.
  ?courseName :has ?groups}
  UNION
  {?assignments a :MandatoryAssignments.
  ?courseName :mustComplete ?assignments
}
}
```

courseName	groups	assignments
INFO116H14	INFO116H14GroupProject	
INFO212H14	INFO212H14GroupProject	
INFO262V15	INFO262V15GroupProject	
INFO115H14	INFO115H14GroupProject	
INFO216V15	INFO216V15GroupProject	
INFO216V15		INFO216V15MandatoryAssignment
SVEKSH14		SVEKSH14MandatoryAssignment
INFO282H14		INFO282H14MandatoryAssignment
INFO100H14		INFO100H14MandatoryAssignment
SV100H14		SV100H14MandatoryAssignment
INFO103V15		INFO103V15MandatoryAssignment
INFO207V15		INFO207V15MandatoryAssignment
INF207H14		INF207H14MandatoryAssignment
SVSEMH14		SVSEMH14MandatoryAssignment
INFO232H14		INFO232H14MandatoryAssignment
SVEKSV15		SVEKSV15MandatoryAssignment
INFO110V15		INFO110V15MandatoryAssignment
INFO125H14		INFO125H14MandatoryAssignmentNR2
INFO102V15		INFO102V15MandatoryAssignment
INFO233V15		INFO233V15MandatoryAssignment
INFO125H14		INFO125H14MandatoryAssignmentNR1
INFO125H14		INFO125H14MandatoryAssignmentNR3
INFO132V15		INFO132V15MandatoryAssignment

5#Which courses are at autumn

```
SELECT ?courseName
        WHERE {
                ?courseName :isAt :Autumn .
        }
ORDER BY ?courseName
```

courseName
INF207H14
INFO100H14
INFO115H14
INFO116H14
INFO125H14
INFO207V15
INFO212H14
INFO232H14
INFO282H14
SV100H14
SVEKSH14
SVSEMH14

6#Covers any of them both data-architecture and algorithms

```
SELECT ?subjName
        WHERE {
        ?subjName      :covers :Algorithms .
        ?subjNm :covers :Data-Architecture .
FILTER (?subjName = ?subjNm)
}
```

subjName
INFO233V15
INF207H14
INFO207V15

7#All courses at spring with 10 studypoints

```
SELECT ?subject
        WHERE {
        ?subject :isAt :Spring;
   :has :STC10.
```

}

subject
INFO132V15
INFO262V15
SVEKSV15
INFO216V15
INFO233V15
INFO110V15

8#Which subject areas do you learn from the course

```
SELECT ?subjArea
        WHERE {
        ?subjArea a :Subject-Areas .
  ?subject :covers ?subjArea.
FILTER(?subject = :INFO110V15)
        }
```

subjArea
Theory
Data-Architecture

9#All courses that are in English

```
SELECT ?subjName
        WHERE {
        ?subjName :is :English .
        }
```

subjName
INFO216V15
INFO282H14
INFO125H14
INFO115H14
INFO110V15
INFO262V15
INFO116H14

10#Which ones has programming at Spring

```
SELECT DISTINCT ?subjName
        WHERE {
        {?subjName :covers :Programming} UNION {?subjName :isAt :Spring.}
}
ORDER BY ?subjName
```

subjName
INFO102V15
INFO103V15
INFO110V15
INFO132V15
INFO216V15
INFO232H14
INFO233V15
INFO262V15
SVEKSV15

Sources:

[1]Authors: Ben Adida and Mark Birbeck, RDFa (Last edited The 12th of April, 2014), [Internet], Available on: http://en.wikipedia.org/wiki/RDFa

[2]Authors: Manu Sporny, Dave Longley, Gregg Kellogg, Markus Lanthaler, Niklas Lindstrom, JSON-LD (Last edited on 22 May 2014), [Internet], Available on: http://en.wikipedia.org/wiki/JSON-LD