The informasjonsvitskap assignment.

Overview

Your assignment is to enhance the usefulness of web sites with semantics, and at the same time to learn about your degree!

The web site is of course:

http://www.uib.no/studieprogram/BASV-INFO#uib-tabs-construction

The assignment does not involve programming, and you will not be making an application as such. What you will do is provide the basic data models and semantic markup that could be used as a basis for a new web site and third party applications.

Your deliverables will be HTML snippets with semantic markup, data models (ontologies), and sample queries that will demonstrate how applications could use the data.

Approach

You will start at the page listing the course of studies - Studieløp - where you will find all the courses in the degree. You will need to define information about these subjects by

a) making an ontology and populating it with data from the web site
b) finding information from the UiB open data initiative
c) adding the information to web sites using RDFa and JSON-LD
d) adding schema.org properties
e) writing SPARQL queries that answer interesting questions from the ontology

Making an ontology

There are many different kinds of information available about the subjects.
You can look at the topics covered in each course - Programming/Languages/ Philosophy/Logic/.. Methods/Statistics/… Media/… Data/Internet/…
These topics form an ontology of topics, and each course covers some of these topics.

You should have an ontology of subjects and subject areas, and indicate which subject areas are covered in each subject. For example:
INFO132 covers Programming/Java/Simulation/Object-oriented programming/
In addition the courses have prerequisites, exam types, assignments, schedules. These should also be included in the ontology.

Remember that when making an ontology there are several design stages. You must look at the terms in your domain and pick out key terms you want to model. You can also make use of existing ontologies for ideas, or to import the entire ontology if it is useful.
Here are some interesting resources, but obviously there are more:
The list of competencies at LinkedIn
http://vocab.org/aiiso/schema
http://www.cs.vassar.edu/~weltyc/papers/subjects/subject.html
http://linkeduniversities.org/lu/index.php/vocabularies/

Finding information

The University has an open data initiative, which has made available several APIs for data retrieval. Their main site is here http://data.uib.no
The list of public APIs is linked from the page http://opendata.app.uib.no
You can use the API without any programming by simply pasting a URL into a web browser according to the API specification. (First at least one

member from each group needs to request an API key). So for example I can get information about INFO116:

https://fs.data.uib.no/MY KEY/json/basisinfo/emne/INFO116/2014H

(Note that all the JSON examples are formatted via a browser plugin
JSONView for Chrome: https://chrome.google.com/webstore/detail/jsonview/chklaanhfefbnpoihckbnefhakgolnmc
JSONView for Firefox: http://jsonview.com/
JSON Formatter for Safari in the plugins store)


```
{
  "@attributes": {
    "timeToLiveSeconds": "86400"
  },
  "emne": {
    "emnekode": "INFO116",
    "emnenavn_bokmal": "Semantic Technologies",
    "emnenavn_engelsk": "Semantic Technologies",
    "emnenavn_nynorsk": "Semantic Technologies",
    "fagkode_sortering": "INFVITENSK",
    "fagnavn_bokmal": "Informasjonsvitenskap",
    "fagnavn_engelsk": "Information science",
    "fagnavn_nynorsk": "Informasjonsvitskap",
    "faknavn_bokmal": "Det samfunnsvitenskapelige fakultet",
    "faknavn_engelsk": "Faculty of Social Sciences",
    "faknavn_nynorsk": "Det samfunnsvitskaplege fakultet",
    "faknr_reglement": "15",
    "institusjonsnr_reglement": "184",
    "instituttnavn_bokmal": "Institutt for informasjons- og medievitenskap",
    "instituttnavn_engelsk": "Department of Information Science and Media Studies",
    "instituttnavn_nynorsk": {},
    "instituttnr_reglement": "17",
    "nsdnivakode": "LN",
    "studienivakode": "100",
    "studienivanavn_bokmal": "Grunnleggende emner, niv\u00e5 I",
    "studienivanavn_engelsk": "Foundation courses, level I",
    "studienivanavn_nynorsk": {},
    "studiepoeng": "10"
  }
}
```


And you can get scheduling information for 116
https://timeplan.data.uib.no/MY KEY/json/timeplanliste/now/INFO116

```
{
  "@attributes": {
    "timeToLiveSeconds": "900"
  },
  "info": {
    "semester": "h",
    "topic": "INFO116",
    "uri": "http:\/\/www.uib.no\/emne\/INFO116",
    "year": "2014"
  },
  "row": [
    {
      "activity": "INFO116",
      "dates": {
        "row": [
          "1408623300",
          "1409228100",
          "1409832900",
          "1410437700",
          "1411042500",
          "1411647300",
          "1412252100",
          "1412856900",
          "1413461700",
          "1414066500",
          "1414674900",
          "1415279700",
          "1415884500",
          "1416489300",
          "1417094100"
        ]
      },
      "dates_iso": {
        "row": [
          "2014-08-21",
          "2014-08-28",
          "2014-09-04",
          "2014-09-11",
          "2014-09-18",
          "2014-09-25",
          "2014-10-02",
          "2014-10-09",
          "2014-10-16",
          "2014-10-23",
          "2014-10-30",
          "2014-11-06",
          "2014-11-13",
          "2014-11-20",
          "2014-11-27"
        ]
      },
      "description": "Forelesning",
      "duration": "6300",
```

(In addition, there are a series of APIs which might give some useful supplementary information about subjects, study programs, and locations. You should look at them to see if there is anything useful in them: http://hackathon.b.uib.no/data-og-idear/uibdata/)

This information could go into your ontology, but it will also need to go into the web site using mark up. You decide what you think should go into the ontology!

Adding the information to web sites using RDFa and JSON-LD

Much of the information you get from the API would be useful to third party applications if it was available on the web site. For example scheduling information could be extracted automatically. So for example you could imagine an application that helped you chose courses by automatically storing the schedules of any subject you chose, to make sure there are no clashes.

You should figure out what information you want to add, and how to add it. The main site for example could contain all the JSON-LD facts for all the subjects. The individual sites could contain the JSON-LD that is only relevant to them, and also contain some RDFa markup (see next section). You should mark up the main site and five additional subject sites.

Adding schema.org properties

Schema.org contains some relevant entities for this task. For example event has schedule information, which could be used to mark up the time table for each subject. For a selection of 5 subjects, you should add schema.org markup for scheduling and anything else you think might be interesting. You should add both RDFa version and add the information to your JSON-LD snippet in the header.

You should test the markup with the google testing tool (http://www.google.com/webmasters/tools/richsnippets#) and show the result of the

analysis. Make sure you know how to use the tool by copying and pasting some sample code from the schema.org examples into the snippet tool, and look at the structured data output.

Here are some resources to help you write and test schema.org:
http://schema.org/Event
https://support.google.com/webmasters/answer/4620133?hl=en
(http://webdesign.tutsplus.com/articles/getting-started-with-the-structured-data-markup-helper--webdesign-12953 Markup helper itself not working at the moment. But this address works:
https://www.google.com/webmasters/markup-helper/u/1/?hl=en)
http://www.google.com/webmasters/tools/richsnippets#

Writing SPARQL queries that answer interesting questions from the ontology

You will not be expected to make user interfaces to present the added data in any way. What you will do is to construct SPARQL queries over the RDF description of the site, and show its output. For example you could write queries that show what other subjects need a particular subject as a pre requisite. (So if you chose a subject, what will it enable you to take in the future ..)

You should make 10 interesting queries and show its output.

Deliverables

This is a fairly complicated assignment. A good pass can be obtained with appropriate and technically correct schema.org markup, a sensible basic ontology, adding some data to the web site with RDFa and JSON-LD, and getting a few basic SPARQL queries working. The highest marks will go to groups with creative uses of the markup, and more extensive and complete

uses cases covered, and good justifications for their particular use of markup.

The deliverables will be

1. A written assignment (explained below)
2. All of the annotated HTML pages
3. The ontology as an OWL file
4. All SPARQL queries and result sets

The assignment should include a written report as well as a sample of the html pages, SPARQL queries, etc. The report should present an overview of what has been achieved. What work did each group member contribute? Why was the ontology constructed in the way it was? What data sources did you use to gather information for the markup and ontology? What kinds of questions can be answered by the ontology (the competency questions you used)? What can the web site do with the added semantics (e.g. third party applications)? The report should also include examples of the expected rich snippet from the markup (from the rich snippet tool). The report should be no longer than 15 pages, including figures, example markup, queries, etc.