



UNIVERSITY OF BERGEN
COURSE EXAM UNDER SOCIAL SCIENCE DEGREE

June 9th 2015, 09.00 – 12.00

Answer four of the six questions below in Norwegian or English. Each question weighs 25%.

Question 1 (25%).

- (a) What is the “AAA slogan” discussed in Allemang and Hendler's textbook? What does each “A” mean?
- (b) What is the "closed world assumption" (CWA)?
- (c) Does the web of data (semantic web) at large rely on CWA, or does it rely on the alternative “open world assumption” (OWA)?
- (d) How and why are CWA/OWA relevant for the web of data?
- (e) Which constructs in RDFS are related to the CWA/OWA distinction?

Question 2 (25%).

- (a) Draw a sketch of an OWL ontology that describes the situation below. For example, you can draw the ontology as a labelled graph, where the labels are URIs and literal values.

“In a team chess tournament (“lagsjakkturnering”), each Team consists of Players. A Player is either Junior or Adult and no one is both. A Match between two Teams consists of Games. Each Game involves two Players, one playing the white pieces, the other the black pieces. A Game has a result, which is either white win, black win, or draw (uavgjort eller remis). Teams, Players, Matches and Games are all mutually disjoint.”

- (b) Define the ontology from task (a) *either* using SPARQL Update *or* using TURTLE. (No need to do both.)
- (c) A Team has at least four Players. Define this restriction, ~~for example~~ using TURTLE *or* Description Logic notation.
- (d) A JuniorTeam is a Team that has only JuniorPlayers. Define the class JuniorTeam, ~~for example~~ using TURTLE *or* Description Logic notation.
- (e) Outline Jena code (using Java or Java-like pseudocode) that defines the two classes from tasks (c) and (d).

Question 3 (25%).

Assume you have an RDF graph that uses the *dc:title* predicate for all resource titles. Some of the resources represent hotels. They have the *rdf:type schema:Hotel*.

- (a) Write a SPARQL query that lists the *dc:titles* (containing the names) of all hotels.

The RDF graph also uses the *schema:address* predicate to describe the address of each hotel. An address resource has *rdf:type schema:PostalAddress*. We will pretend each postal address has a *schema:city* predicate that gives the hotel's city as a string literal.

- (b) Write a SPARQL query that lists the titles (or names) of all hotels along with their cities.
- (c) Write a SPARQL query that lists, for each city, the titles (names) of all hotels in that city. The cities should be presented in alphabetical order.

Each hotel in the graph also has Geo coordinates for latitude and longitude.

- (d) Write a SPARQL query that lists all nearby hotels with latitude and longitude that are within 0.006 degrees latitude and 0.01 degrees longitude of the hotel named “Hotel Saint-Germain-des-Pres”.
- (e) Make the query from task (d) more efficient by only considering hotels in the city of “Paris”.

Some DBpedia resources also have Geo coordinates.

- (e) For each hotel in Paris, list the URIs of nearby DBpedia resources, using DBpedia's public SPARQL endpoint. (For example, nearby can mean within 0.006 degrees latitude and 0.01 degrees longitude of the hotel.)

Question 4 (25%).

You are brought in as an advisor for a group of small and medium-sized companies that design and produce furniture (“møbler”). They want to start using semantic technologies to make information sharing quicker and easier, but there are no dedicated vocabularies available, except for a few general ones, like schema.org, that describe products in general.

How would you advise them to go about developing a new vocabulary for the companies? For example, how would you organise the work into phases and steps, who will you involve, and what will you do in each phase/step? How will you make sure you reuse as much as possible of existing vocabularies? How will you ensure that the vocabulary you propose is any good?

Question 5 (25%).

You are involved in a local music festival in the Bergen area. You lead a project that will use semantic technologies to organise and enrich the festival's data about bands and artists, playlists, concerts and other events, venues (“spillesteder”) and stages (“scener”). One aim is to improve the festival's web pages, but the semantic data will be used for other purposes too.

(a) Which existing vocabularies would you use to organise the data? How are they connected? Give examples of terms you would like to use. You can use the following situation as an example:

“The band *Jaga Jazzist* will play at the *Bastionen* scene in *Bergenhus Festning* Saturday June 13th 2016 from 15:15 to 16:15. They will perform the songs *Oslo Skyline* and *Swedenborgske Rom* in that order. Lars and Martin Horntveth are both members of the band. Lars is the composer of Oslo Skyline. The track is ~~taken from~~ ~~published on~~ the album *What We Must* ~~published~~ by *Ninja Tune*. The following concert at Bastionen is *Kyla La Grange* from 17:30 to 18:15.”

(b) What other kinds of data would you like your semantic data set to include or link to, and how?

(c) Would you need to invent new vocabularies or vocabulary terms yourselves? If so, give examples.

Question 6 (25%).

(a) Show the general architecture of a semantic application in a drawing. Try to show many alternative ways in which the application can import, store, process, present and export data.

(If you want to, you can use a semantic application you know as an example – but don't use your programming project, because anonymity may be broken.)

(b) Explain each of the main components in the architecture from task (a) briefly.

(c) For each component in task (a), give examples of common standards, protocols and technologies.