188.399 Introduction to the Semantic Web (VU 2,0)
2015W

Assignment 1 (10 pt.)

Instructions

Deadline

Make sure to upload **all** your results

before December 20, 2015!

What you should hand in

Upload a *single PDF document* (only one group member needs to upload) in the appropriate solution upload folder on TUWEL. Please follow the naming convention

```
SW_2015_Exercise_1_<group no>.pdf
```

Make sure to upload only a single solution per group. Furthermore, make sure to include a cover page and list your group members in the header on each page of your solution document.

Tools

- You will need to set up a Fuseki server for exercise III. The server may also be useful for testing purposes in exercises I and II.
- You will use DBPedia's SPARQL endpoint (http://dbpedia.org/sparql) in exercise IV. You may use the public web-based query interface provided by dbpedia.

Questions

Please post general questions in the TUWEL discussion forum of the course. You can also discuss problems and issues you are facing there. We appreciate if you help other students out with general problems or questions regarding the tools used (and may take that into account in case you are short a few points for a better grade). For obvious reasons, however, please do not post any solutions there.

You can also contact me directly (with specific questions) at:

```
elmar.kiesling@tuwien.ac.at
```

Please use subject line `SW_2015_<your subject>` to minimize the probability that your Email gets lost in my inbox.

188.399 Introduction to the Semantic Web (VU 2,0)
2015W

Assignment 1 (10 pt.)

I. ✐ **RDF and RDF-S syntax** [2 points]

Consider the following RDF representation

```
<rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:ex="http://example.org/">

<rdf:Description rdf:about="http://example.org/leo">
        <rdf:type rdf:resource="http://example.org/Lion"/>
</rdf:Description>

<rdf:Description rdf:about="http://example.org/bill">
        <rdf:type rdf:resource="http://example.org/Buffalo"/>
</rdf:Description>

<rdf:Description rdf:about="http://example.org/Species">
        <rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
</rdf:Description>

<rdf:Description rdf:about="http://example.org/prey_of">
        <rdf:type rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#Property"/>
        <rdfs:domain rdf:resource="http://example.org/Species"/>
        <rdfs:range rdf:resource="http://example.org/Species"/>
</rdf:Description>

<rdf:Description rdf:about="http://example.org/Lion">
        <rdf:type rdf:resource=" http://example.org/Species"/>
        <rdfs:label xml:lang="en">lion</rdfs:label>
        <rdfs:label xml:lang="de">Löwe</rdfs:label>
</rdf:Description>

<rdf:Description rdf:about="http://example.org/Buffalo">
        <rdf:type rdf:resource=" http://example.org/Species"/>
        <rdfs:label xml:lang="en">buffalo</rdfs:label>
        <rdfs:label xml:lang="de">Büffel</rdfs:label>
        <ex:prey_of rdf:resource="http://example.org/Lion"/>
</rdf:Description>
```

a) What information is encoded explicitly in this RDF/RDFS graph? What facts can be derived from the graph according to RDF-S semantics? Describe in your own words.

b) Translate the given RDF/XML document into Turtle syntax.

II. ✐ **RDF-S modeling**[1] [2 points]

Write an RDF-S model that formalizes knowledge about the domain of people, in particular the classes person, man, woman, and the properties `parent`, `mother`, `father`. Use Turtle notation and express the complete knowledge you have about such classes and properties (e.g., every man is a person, every woman is a person, every mother is a woman etc.). Include your complete RDF-S model in your solution document.

[1] Adapted from http://www.dis.uniroma1.it/~rosati/semanticweb/slides3.pdf

Assignment 1 (10 pt.)

III. **Local triple store and SPARQL querying** [2 points]

Consider the following RDF document in Turtle notation.

```
@prefix : <http://www.example.org/family> .

:gary   :isMarriedTo   :glenda .
:betty  :hasParent     :gary .
:hugh   :hasParent     :gary ;
        :hasSibling    :betty .

:adam   :isMarriedTo   :agatha .
:ben    :hasParent     :adam .
:helen  :hasParent     :adam ;
         :hasSibling   :ben .

:orene  :hasParent     :hugh , :helen .
:david  :hasParent     :ben , :betty .
```

a) *Which triples will be included in the output of the following SPARQL statement?*

```
PREFIX : < http://www.example.org/family>

CONSTRUCT { ?p :doubleCousin ?q . }
WHERE {
       ?p  :hasParent ?parent1 .
        ?p  :hasParent ?parent2 .
        ?parent1 != ?parent2.
        ?parent1 :hasSibling ?sibling1.
        ?parent2 :hasSibling ?Sibling2.
        ?q  :hasParent ?sibling1.
        ?q  :hasParent ?sibling2
}
```

b) *To verify your results, set up a local Fuseki server and start it with an empty dataset using an in-memory database. Consult the Fuseki documentation on how to serve data using an in-memory database*[2]*. Document your experiments in your solution document and provide the query results returned by your server in your preferred format.*

[2] http://jena.apache.org/documentation/serving_data/#fuseki-server-starting-with-an-empty-dataset

Assignment 1 (10 pt.)

IV. **DBPedia SPARQL queries** [4 points]

Formulate SPARQL queries that when executed against the DBPedia SPARQL endpoint (http://dbpedia.org/sparql) answer the following questions:

a) What is Marie Curie known for?
b) How many physicists that have won a Nobel price in their field and are more than 75 years old are alive today?
c) Is the Danube river longer than the Volga river?
d) What are the names of the 5 longest films directed by Stanley Kubrick?
e) List the English abstracts of all film noir movies directed by Alfred Hitchcock.
f) Which 5 films shot in Vienna had the highest gross margin (calculated as [gross – budget]/gross)?
g) Who are (notable) alumni of the Vienna University of Technology (i.e., have Vienna University of Technology as their `dbr:almaMater`*)?*
h) Suppose that all people that have been affiliated with TUV have an interest in TUV. Construct a graph that expresses this assumption using `foaf:interest`*. Consider those people affiliated that are either alumni (i.e., TUV is their* `dbp:almaMater`*) or (former) emploeyees (i.e., TUV it is among their* `dbp:workplaces` *or their* `dbp:workInstitutions`*).*

Your solution document should include:
1.) SPARQL queries in textual form
2.) Your queries as working links that return the result in HTML format
3.) The results of your queries in Turtle format