COMP38120

Workshop 10 – Data on the Web and RDF Worksheet

Task 1: Exploring the coverage and content of the linked data space (5 min)

Explore and discuss the coverage of linked data. Is there much missing? Is the data reliable? Is it noisy?

(a) Browse DBpedia dataset using http://dbpedia.org/fct (text search). Explore the information about: (i) your home town, and (ii) your favourite band or movie/TV character.

(b) Search https://www.wikidata.org ("a free, linked database that can be read and edited by both humans and machines") for
- people
- places
- organisations
- books

Task 2: XML basics (5 min)

Make sure you understand what this XML document is about (syntactically):

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF
        xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
        xmlns:dc="http://purl.org/dc/elements/1.1/"
        xmlns:region="http://www.country-regions.fake/">

        <rdf:Description rdf:about="http://en.wikipedia.org/wiki/Oxford">
                <dc:title>Oxford</dc:title>
                <dc:coverage>Oxfordshire</dc:coverage>
                <dc:publisher>Wikipedia</dc:publisher>
                <region:population>10000</region:population>
                    <region:principaltown
                            rdf:resource="http://www.country-regions.fake/oxford"/>
        </rdf:Description>
</rdf:RDF>
```

Remind yourself about the basic XML stuff: XML elements (tags) and attributes; well-formed XML, valid XML; XML Schema Definition (XSD), XSD elements and attributes; namespaces, qualified names.

Task 3: Understanding RDF statements (5 minutes)

(a) How many RDF statements do we have below?
(b) Draw the equivalent RDF graph.

```
<?xml version="1.0"?>

 <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
          xmlns:dc="http://purl.org/dc/elements/1.1/"
          xmlns:exterms="http://www.example.org/terms/">

   <rdf:Description rdf:about="http://www.example.org/index.html">
       <exterms:creation-date>August 16, 1999</exterms:creation-date>
   </rdf:Description>

   <rdf:Description rdf:about="http://www.example.org/index.html">
       <dc:language>en</dc:language>
   </rdf:Description>

   <rdf:Description rdf:about="http://www.example.org/index.html">
       <dc:creator rdf:resource="http://www.example.org/staffid/85740"/>
   </rdf:Description>

 </rdf:RDF>
```

```
<?xml version="1.0"?>

<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
         xmlns:dc="http://purl.org/dc/elements/1.1/"
         xmlns:exterms="http://www.example.org/terms/">

  <rdf:Description rdf:about="http://www.example.org/index.html">
         <exterms:creation-date>August 16, 1999</exterms:creation-date>
         <dc:language>en</dc:language>
         <dc:creator rdf:resource="http://www.example.org/staffid/85740"/>
  </rdf:Description>

</rdf:RDF>
```

Task 4: Building some linked data (5 minutes)

Friend of a Friend (FOAF) is a descriptive vocabulary for describing people, their Identities, interests, affiliations, social networks, etc. You may wish to read further documentation about FOAF at http://www.foaf-project.org/docs and http://xmlns.com/foaf/spec/.

Visit FOAF-a-Matic (http://www.ldodds.com/foaf/foaf-a-matic) to make a FOAF file for yourself (or your sibling). Note that you will generate a FOAF file, but this site does not provide hosting. What are the predicates included in the FOAF file you created? Make sure you understand the XML syntax.

Homework

1) What are the Linked Data principles? How do they relate to the existing Web of Documents infrastructure?

2) What are the advantages of using HTTP URIs? Within the context of Linked Data, are URI aliases advantageous or not? Why?

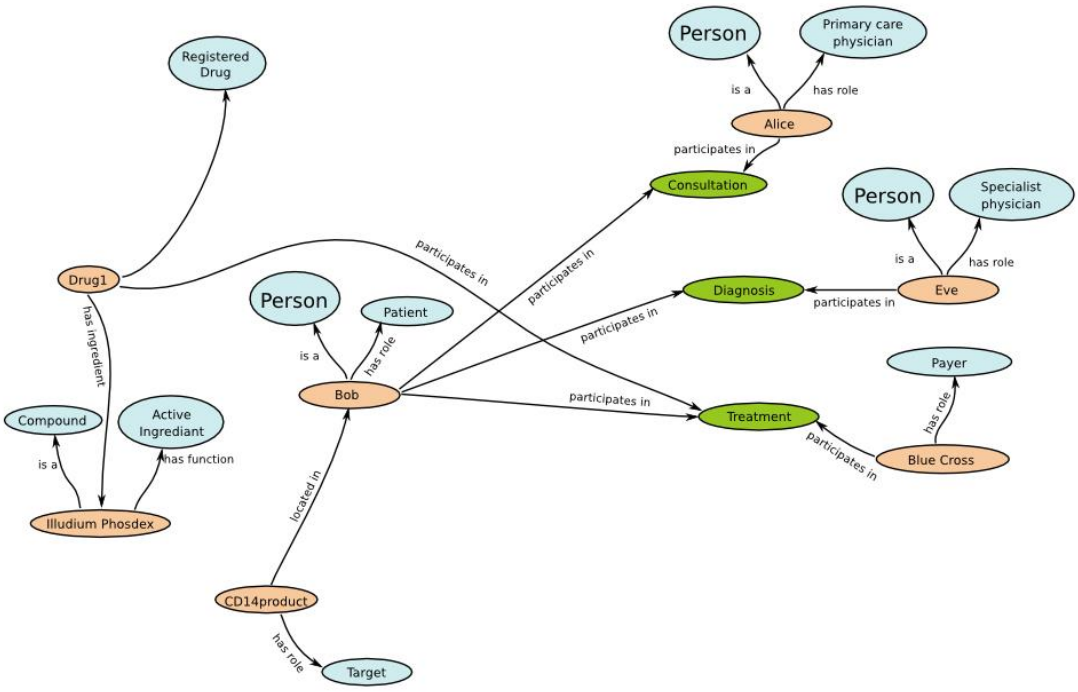
3) What are the advantages of using the RDF data model in the context of Linked Data?

4) **Internationalized Resource Identifiers** (IRIs) are a generalization of URIs that permits a wider range of Unicode characters. Every absolute URI and URL is an IRI, but not every IRI is a URI. Find out more about IRIs. Visit, for example the RDF 1.1. W3C Recommendation (http://www.w3.org/TR/rdf11-concepts/), or "What's New in RDF 1.1" (http://www.w3.org/TR/rdf11-new/).

5) Using RelFinder (http://www.visualdataweb.org/relfinder/relfinder.php), find and describe the relationships between:

Ernest Rutherford and Manchester
Alan Turing and United States

6) Explore the RDF graph below. List the resources, predicates, and all the triples.



7) Draw the triples represented by this RDF/XML description:

```
<?xml version="1.0" encoding="UTF-8"?>

<rdf:RDF
      xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
      xmlns:dc="http://purl.org/dc/elements/1.1/"
      xmlns:region="http://www.country-regions.fake/">

      <rdf:Description rdf:about="http://en.wikipedia.org/wiki/Oxford">
            <dc:title>Oxford</dc:title>
            <dc:coverage>Oxfordshire</dc:coverage>
            <dc:publisher>Wikipedia</dc:publisher>
            <region:population>10000</region:population>
            <region:principaltown rdf:resource="http://www.country-
                  regions.fake/oxford"/>
      </rdf:Description>
</rdf:RDF>
```

Generate an equivalent Turtle serialisation.