Week 2 Lab – Exploring Web APIs & Schema.org

Task 1: Exploring DBpedia REST APIs:
http://wiki.dbpedia.org/projects/dbpedia-lookup#
DBpedia is a crowd-sourced community effort to extract structured information from Wikipedia and make this information available on the Web. DBpedia allows you to ask sophisticated queries against Wikipedia, and to link the different data sets on the Web to Wikipedia data. The "Keyword Search" API is one of the REST based endpoints for accessing the DBpedia dataset. (There are many other API endpoints using protocols such as SPARQL, but this is beyond the scope of this lab, for those interested take a look at http://wiki.dbpedia.org/OnlineAccess).

- Create queries for the following and execute them as url in a browser. In the resulting XML response, examine the capabilities and limitations of this API and the crowd-sourced data it holds (i.e., if the data always accurate/available):
  - Cardiff University
  - Informatics
  - Cardiff School of Computer Science and Informatics
  - Cardiff University School of Medicine

Task 2: Adding Semantics through Microdata:

First, explore the hierarchy of schemas at http://schema.org/docs/full.html to get an idea of the types of entities and their properties that can be described. Examine at least the following two schemas in more detail:
- http://schema.org/Person
- http://schema.org/CreativeWork

Second**,** following the example in todays' lecture, create a plain text webpage for the following piece of text from wikipedia:

> *The Hobbit, or There and Back Again is a children's fantasy novel by English author J. R. R. Tolkien. It was published on 21 September 1937 to wide critical acclaim, being nominated for the Carnegie Medal and awarded a prize from the New York Herald Tribune for best juvenile fiction. The book remains popular and is recognized as a classic in children's literature.*
>
> *The Hobbit is set in a time "between the Dawn of Færie and the Dominion of Men", and follows the quest of home-loving hobbit Bilbo Baggins to win a share of the treasure guarded by Smaug the dragon. Bilbo's journey takes him from light-hearted, rural surroundings into more sinister territory.*

Use microdata and schemas from schema.org to make as much of the information in this text as possible accessible to a computer program. Use the online tool

available at https://search.google.com/structured-data/testing-tool to ensure that your annotations can be recognised. You can paste your code and run the test under "code snippet".

Third, as in task 1, use the DBpedia API to search for information about
- The Hobbit
- J.R.R. Tolkien
- Carnegie Medal
- Smaug

Examine the results, and compare them to the structured information you created. Does DBpedia provide any information that is contained in the text above, but which you did not annotate? Does your annotation provide any information that is not in the DBpedia results?