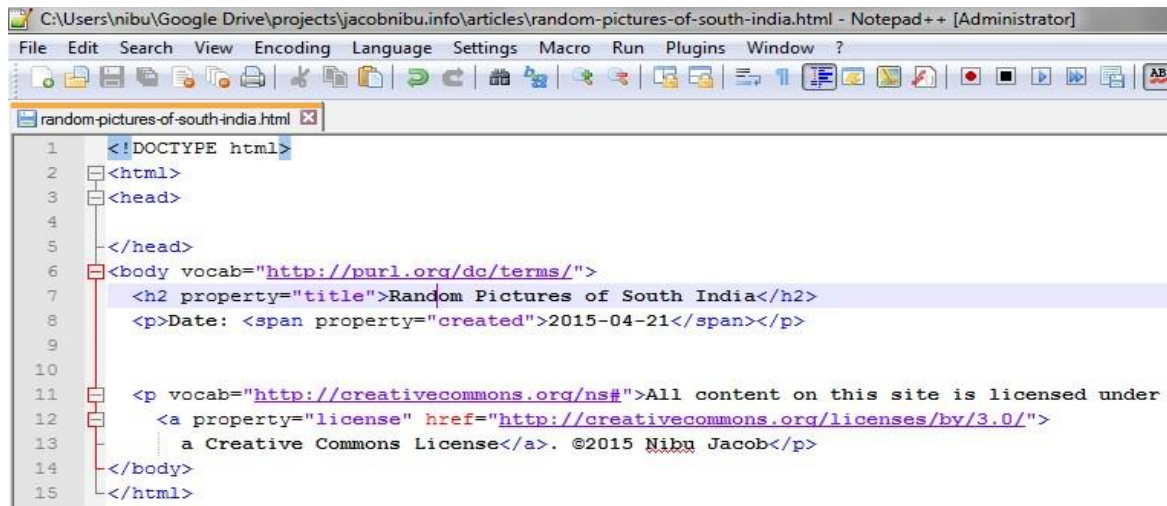# Creating and Deploying Linked Open Data

Task: Create an example of Linked Open Data: two or more datasets, adding connections to other datasets (e.g., Dbpedia, Freebase) using owl:sameAs or rdfs:seeAlso
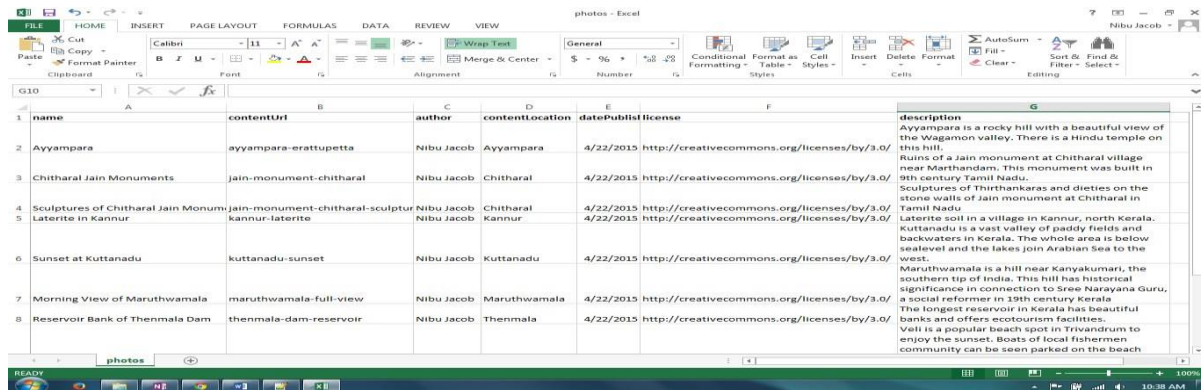
Implementation

As per W3 best-practices [1], two datasets uniquely collected and created by me were selected for this assignment. One dataset consists of 14 photographs taken by me of interesting places in South India. This data could be connected to relevant geographical and historical data (the second dataset) which in turn could be connected to resources in existing public datasets using owl:sameAs or rdfs:seeAlso links.

In order to make an open data that is both machine-readable and human-readable, RDFa was chosen as the RDF format to use for this assignment. An HTML5 page containing RDFa code was created in Notepad++ following the examples given by RDFa primer [2]. The basic HTML structure with document metadata (using DC terms) and license information was coded:

For photographs, the ImageObject schema provided by Schema.org was selected as vocabulary to be used [3]. A CSV file containing selected properties of this schema along with their values was created using Excel (this would be the second dataset):

This CSV dataset was converted into RDF data using RDF Refine extension [4] of OpenRefine (Google Refine 2.5) [5] [6]. Only 4 out of 8 contentLocation values could be reconciled to DBpedia URIs using Google Refine:

This RDF data was then incorporated into the HTML5 page using RDFa:

Terms in 'Description' column of the CSV data file were explored in DBpedia to find corresponding and related resource URIs. For the four contentLocation values that have no corresponding DBpedia URIs, connections to existing resources were explored. The search service of Sindice was useful for this [7].

To find existing sameAs links for the resources in the page, the search service at sameAs.org was used [8]. The dataset was linked to existing open datasets using owl:sameAs and rdfs:seeAlso properties.

Since multiple vocabularies (dc, schema, owl, rdfs) are used in the same document, the vocab attribute in body tag was replaced by a prefix attribute and predefined prefixes were used wherever possible [9].

The license information was updated using customized HTML code from Creative Commons site [10].

A part of the RDF data contained in the HTML5 page is displayed below using RDFa Play [11]:
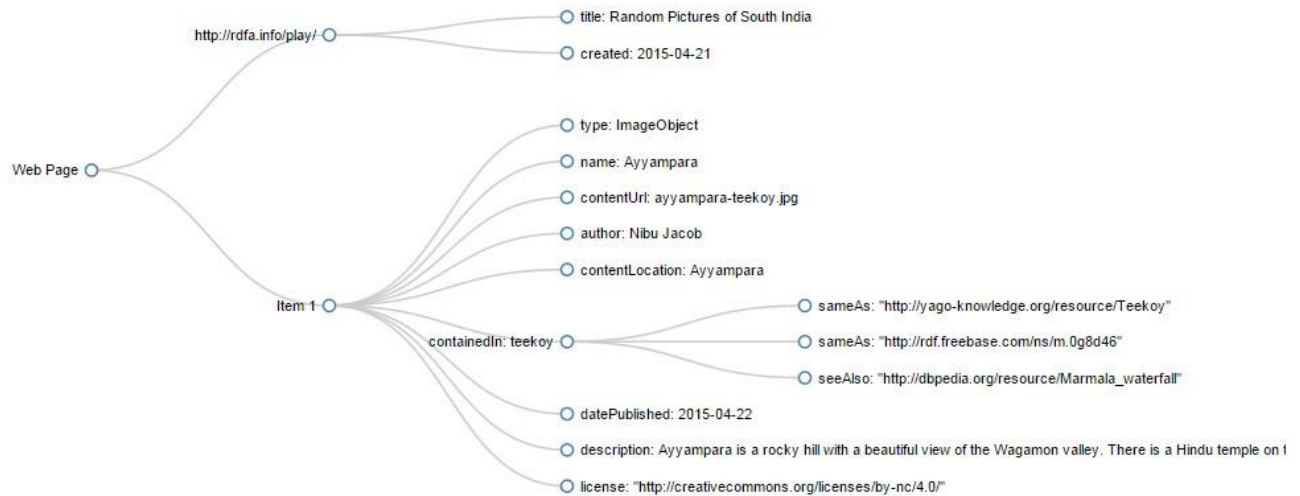
For the data to get unique URIs, all photographs were placed in a folder named "photos" in a personal server "jacobnibu.info" and the HTML page was placed in the same folder to make the URIs dereferenceable. The human-readable data URI is http://jacobnibu.info/photos/

References
1. http://www.w3.org/TR/ld-bp/

2. http://www.w3.org/TR/xhtml-rdfa-primer/

3. http://schema.org/ImageObject

4. http://refine.deri.ie/

5. http://openrefine.org/

6. https://www.youtube.com/watch?v=XdpzmGxA33U

7. http://sindice.com/

8. http://sameas.org/

9. http://www.w3.org/2011/rdfa-context/rdfa-1.1

10. http://creativecommons.org/choose/

11. http://rdfa.info/play/