Homework 9: Karma
INF 558 BUILDING KNOWLEDGE GRAPH

DUE DATE: Tuesday, 04/17/2018 @ 11:59pm on Blackboard

Ground Rules
This homework must be done individually. You can ask others for help with the tools, but the submitted homework needs to be your own work.

Summary
In this homework, you will learn using Karma to model and publish your sources under RDF related formats.

Task 1: Install and configure Karma
Download and install Karma from:
https://github.com/usc-isi-i2/Web-Karma/wiki/Installation

Task 2: Load your ontologies
You can only use the schema.org ontology or an extension of this ontology to model your data. The schema.org ontology can be downloaded from:
http://schema.org/docs/developers.html

If this ontology does not include all the classes and properties that you need to model your data, you can extend the ontology by adding more classes and properties. For example, you can add the class Student as a subclass of the class https://schema.org/Person. To extend the ontology, you don not need to edit the original ontology. You can create a new file and add your classes and properties to this file, and then load both ontologies in Karma. To load the ontologies, you can run Karma and then use the import menu to load your files. Alternatively, you can put all the ontologies in the preloaded-ontologies subdirectory of the Karma home directory and then run Karma.

Task 3: Model the data
Modeling data sources using the loaded ontologies.
https://github.com/usc-isi-i2/Web-Karma/wiki/Modeling-Data

You can use python transformations if you want to transform some columns before modeling. For example, if you want to create URI for instances.
https://github.com/usc-isi-i2/Web-Karma/wiki/Transforming-Data

Below are requirements you for your Karma model:

- You have to model two data sources. You can use the data that you collected for your project.
- You need to model at least 4 columns from each data source.

- Every class in your model need have a URI. For example, if you have the class https://schema.org/Person in your model, you should have a column in the source containing the URIs for the instances of Person. You can create URIs using python transformation. To indicate that data values in a column are the URIs for the instances of a class, you need to select "uri of class" when assigning a semantic type to the column.
- There should be at least one common class in the two models you create. You can think of this class as a means to join the data from the data sources. Therefore, you should think about how to create URIs for your classes so the same entities across two sources will have the same URI.
- You need to submit screenshots of the models. To create the screenshot, click on the "Print Model" in the menu and save the .png file. If your source has a lot of columns and you have only modeled some of them, it is better to first hide the unused columns (using the "Organize Columns" in the menu) and then print the model.

Task 4: Publish

- When you are done with modeling the data, click on the menu next to the name of the source and then click on "Publish -> Model". Karma generates a .ttl file and shows a link to it next to the name of the source. This file includes the history of your modeling task and also the R2RML mappings describing how to map the data to the ontology. Save this file. To check whether your model file is correct, you can load your data again and apply the saved model to the source (using the "Apply R2RML Model" option) to see if the loaded model is the same as the one you created before.
- In addtition to the R2RML model, you also need to export a RDF file. Click on the menu next to the name of the source and then click on "Publish -> RDF". Karma generates a .ttl file for RDF storage and shows a link to it next to the name of the source. If your URIs are not correct, there will be errors when publishing RDF file.
- Finally, you need to export a JSON-LD file. Click on the root node of your model and then select "Export JSON" (use @context from model) to publish your data as JSON-LD. Save both the JSON file and its context.

Submission Instructions

You must submit the following files and folders in a single .zip archive named Firstname_Lastname_hw9.zip and submit it via Blackboard:

- "source1.*" and "source2.*": these two file are your data sources (.json, .csv, or .xml).
- "model1.png" and "model2.png": screenshots of the models (Task 3).
- "preloaded-ontologies": this folder contains all the ontologies (Task 2).
- "r2rml1.ttl" and "r2rml2.ttl": the R2RML models you created for the data sources (Task 4).
- "rdf1.ttl" and "rdf2.ttl": the RDF file you create for the data sources (Task 4)
- "source1.jsonld" and "source2.jsonld": the JSON-LD files (Task 4).
- "context1.jsonld" and "context2.jsonld": the JSON-LD file you generated (Task 4).