**Semantic Web Team Project Spring 2016**

**LOGD Project Checkpoint**

**Sunish-Mayank Team**

**Project Group 3:**

* **Sunish Shah**
* **Mayank Agrawal**

**Topic: Profit Loss for various Agencies**

**Progress so far on the following topics:**

**LOGD tutorials: what was completed:**

* **Building LOGD Visualizations**: We went through the Google Visualization API and understood as too how it would be useful for our project. We decided to use the SPARQL query result obtained and then using the Google visualization API to get our visualization.
* **Understanding LOGD Data**: Completed
* **Understanding LOGD Metadata**: We understood the semantic meaning of the dataset. We used this knowledge to develop a prototype on paper on how our visualization should look and how it could be useful to majority of the audience.
* **Mashing up LOGD data with SPARQL**: We mashed up both the dataset using **“Agency Name”** as the common column between both the datasets.
* **How to find dataset using the LOGD SPARQL endpoints**: We used the following query to check the sparql endpoints:

**DataSet - 401**

PREFIX rdf: <[http://www.w3.org/1999/02/22-rdf-syntax-ns#](http://www.google.com/url?q=http%3A%2F%2Fwww.w3.org%2F1999%2F02%2F22-rdf-syntax-ns%23&sa=D&sntz=1&usg=AFQjCNHlHXopcxjLZnliePrsUehDZFFerA)>   
SELECT ?s ?p ?o  
WHERE   
{GRAPH <[http://data-gov.tw.rpi.edu/vocab/Dataset\_401](http://www.google.com/url?q=http%3A%2F%2Fdata-gov.tw.rpi.edu%2Fvocab%2FDataset_401&sa=D&sntz=1&usg=AFQjCNH_a70XvvPU6cj8R--ragTPRLI1LQ)>  
{  
?s ?p ?o  
}  
}

**Endpoints:** 207513

**DataSet - 403**

PREFIX rdf: <[http://www.w3.org/1999/02/22-rdf-syntax-ns#](http://www.google.com/url?q=http%3A%2F%2Fwww.w3.org%2F1999%2F02%2F22-rdf-syntax-ns%23&sa=D&sntz=1&usg=AFQjCNHlHXopcxjLZnliePrsUehDZFFerA)>   
SELECT ?s ?p ?o  
WHERE   
{GRAPH <http://data-gov.tw.rpi.edu/vocab/Dataset\_403>  
{  
?s ?p ?o  
}  
}

**EndPoints:** 21379

**SPARQL query progress:**

1. We were able to find the common Agency Names between the two datasets we are using. We were able to remove all the unmatched rows between these two datasets.

Example: There were some agency names which are present in dataset 401 but not in 403. We were able to group the common agencies.

1. We are facing issue in formulating the SPARQL query for adding all rows data and columns data for the corresponding Agency.

**Problems Encountered:**

We had difficulties in accessing the CSV files as it is very hard to navigate through the LOGD website and recently we encountered issues of site forbidden from access.

**Changes in the Project Plan:**

1. No changes in the data sources.
2. No change in the expected results.