**Knowledge Discovery and Management  
Problem Set (PS-1A)**

Name: Kalyan Kilaru

Class ID: 10

1. Describe your knowledge about knowledge graph.

Knowledge graph is the systematic way of putting facts, people and places together, to create interconnected results that are more accurate and relevant.

More specifically, the [knowledge graph](http://searchengineland.com/library/google/google-knowledge-graph) is a database that collects millions of pieces of data about keywords people frequently search for on the World wide web and the intent behind those keywords, based on the already available content.

2. Why do we want to build such a knowledge graph?

We build the knowledge graph so that the users can get information about people, facts and places that are interconnected in one way or the other. We can also explore via intuitive structure and knowledge graph is like a combat information overload.

3. What steps are required? Show your own workflow for this task.

To build a knowledge graph for a text, we need to follow some steps as discussed below:

Step 1: Natural Language Processing

Natural language processing (NLP) is the ability of a computer program to understand human speech as it is spoken. NLP is a component of artificial intelligence. The NLP is subdivided into categories like sentence segmentation, tokenization, lemmatization, parts of speech tagging and parsing.

Step 2: Extracting the Data

We extract the information by chunking, entity coreference and relational extraction.

Step 3: Grouping the Information

After the chunking of data, we segment them based on the topics like Arts, people, places etc.

Step 4: Knowledge graph construction

For the graph construction, we incorporate ontological constraints and relational patterns and discover statistical relationships within knowledge graph.

4. What are the challenges?  
  
The interconnected units of knowledge powers and enhances multiple backend features such as Disambiguating and recognizing entities in context, Data expansion to enrich semantic search, connecting entities to content and data sources and inferential reasoning. ­