Building and Querying an Enterprise Knowledge Graph

Abstract:

Knowledge workers, such as scientists, lawyers, traders or accountants, have to deal with a greater than ever amount of data with an increased level of variety. Their information needs are often focused on entities and their relations, rather than on documents. A knowledge graph is a general concept of representing entities and their relationships and there have been various efforts underway to create knowledge graphs that connect entities with each other. For instance, the Google Knowledge Graph consists of around 570 million entities as of 2014.

Existing Work:

Data modeling and storage is another important part of our knowledge graph pipeline. A data modeling mechanism should be flexible enough to allow scalable data storage, easy data update and schema flexibility. Never-Ending Language Learning and Open Information Extraction are two efforts in extracting knowledge facts from a broad range of domains for building knowledge graphs. With the extracted knowledge facts, Proposed an approach for noise removal and knowledge inference.

Disadvantages:

Proposed Work:

We present our effort in building and querying Thomson Reuters’ knowledge graph. Data in heterogeneous formats is first acquired from various sources. We then develop named entity recognition, relation extraction and entity linking techniques for mining information from the data and integrating the mined data across different sources. We model and store our data in
RDF triples, and present TR Discover that enables users to search for information with natural language questions. We evaluate and demonstrate the practicability of our knowledge graph.

Advantages:

➢ Improve The Performance Of Our NER Engine.

Modules:

➢ Service Framework Overview
➢ Data Acquisition, Transformation And Interlinking
➢ Data Modeling and Physical Storage.
➢ Querying the Knowledge Graph with Natural Language.

SYSTEM REQUIREMENTS

H/W System Configuration:-

- Processor : Pentium –III
- RAM : 256 MB (min)
- Hard Disk : 20 GB
- Key Board : Standard Windows Keyboard
- Mouse : Two or Three Button Mouse
- Monitor : SVGA

S/W System Configuration:-

- Application Server : Tomcat5.0/6.X
- Front End : HTML, Jsp
- Scripts : JavaScript.
<table>
<thead>
<tr>
<th>Server side Script</th>
<th>Java Server Pages.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Database</td>
<td>MySQL 5.0</td>
</tr>
<tr>
<td>Database Connectivity</td>
<td>JDBC</td>
</tr>
</tbody>
</table>