

Software Requirements Specification

Efficient Algorithm for recommendation of
data visualization tools

Group Members:

Cristobal Leiva

Ahmad Amayri

Jorge Ortiz

Supervisor:

Fabrizio Orlandi

Project Overview:

Objective

- Develop an algorithm to recommend data visualization tools (Charts) based on selected Datasets / Properties.
- Intended to be use by systems like LinDA Project that provide ways to explore Linked Data Datasets.

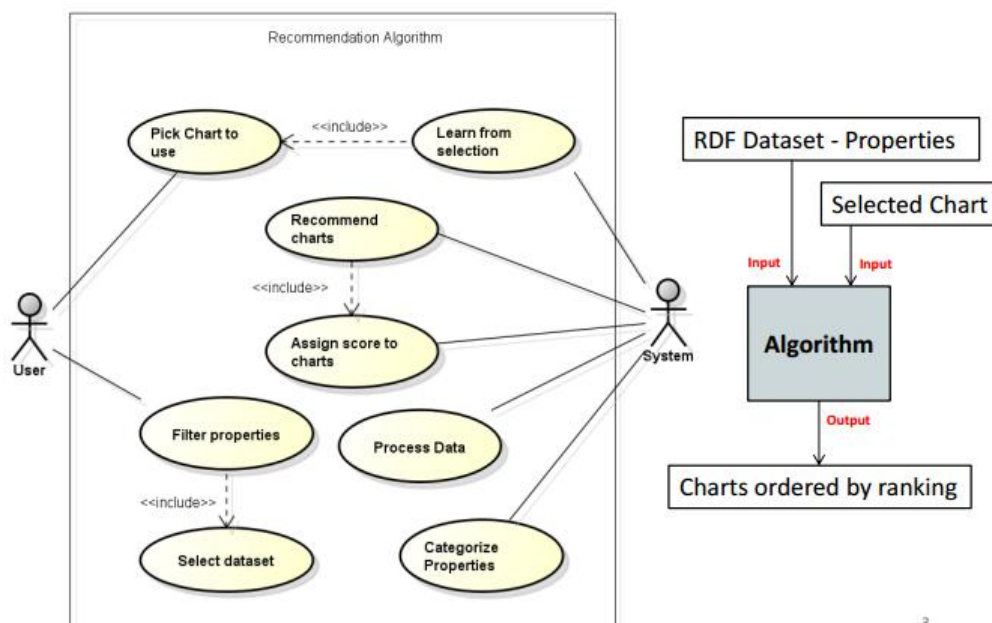
Charts recommendation

- Ranked list of recommended charts
- Google Charts such as Bar, Bubble, Line, Geo Charts...

Algorithm

- Google Charts scope / Analyze data formats
- Categorize Properties (Data type, Scale of Measurement, etc...)
- Learn from previous selections (feedback)
-

Use-Case diagram



Requirements orientation:

Since the project is focused in the implementation of the algorithm. The highest priority remains in the back-end of the application. The front-end also takes a relevant part but it should be adaptable to different charts and graphs sources. The requirements are classified as it follows:

Functional requirements:

1. Information to analyze.

Description: The information to analyze should be imported like datasets with its respective properties. The datasets should have a good format and enough information to be able to create a graph from it. In addition in properties should be able to give information about the type of graph that could be appear either temporal, statistical or special.

Priority: High.

2. Data processing:

2.1. Data Reader:

Description: The system requires a reader for the datasets.

Priority: High.

2.2. Properties Reader:

Description: The system requires a reader for the properties of the dataset. This reader should be able to read the types of properties.

Priority: High.

2.3. Algorithm logic:

Description: The algorithm is able to classify the information based in the content and the properties. It is able to distinguish between statistical, special, temporal data between others and suggest the best 4 charts for the information.

Priority: High.

3. Graphs Database (Google charts):

3.1. Library: The project requires the charts for drawing the information of the datasets.

Priority: Medium.

3.2 Properties: The properties of the charts take part in algorithm for the classification of the most accurate graphs.

Priority: Medium.

4. Comparing Algorithms:

Description: The system requires third part algorithms to be able to compare and get results about the reliability and validity of our results.

Priority: Medium.

5. Exception handling:

Description: The algorithm needs an exception system for cases where the datasets do not have a proper format or the information in the properties cannot be well read.

Priority: Medium.

Non-functional Requirements

1. Scalability: The system should support different size of datasets and should work consistently.

Priority: Medium.

2. Performance Requirement: The application should keep a constant time performance for different scales of information.

Priority: Medium.

3. **Security:** The application should have enough security to let the datasets been analyzed only for the system.

Priority: low.

4. **Interoperability:** The system should behave in a similar way for different platforms.

Priority: low.

References:

http://en.wikipedia.org/wiki/Requirements_engineering
http://en.wikipedia.org/wiki/Software_requirements