

# **Knowledge Discovery and Management**

**Project Phase - III Report**

**By**

**Team 1:**

**Sai Venkatesh Gatiganti (Class ID: 08)**

**Karthik Reddy Vundela (Class ID: 43)**

**Chaitanya Sai Manne (Class ID: 20)**

**Sri Chaitanya Patluri (Class ID: 32)**

## Objective:

To design a Semantic Search Engine that provides search results on books based on reviews obtained from Amazon & Wikipedia.

## Expected Outcome:

To obtain search results based on context besides keywords for better accuracy. For example, if a user searches for a plot in the search engine, the books and movies with similar plot as entered by the user are shown as the search results.

## Project Domain: Movies (Plot Based Semantic Search Engine)

## Datasets:

Amazon Review Data collected by Julian McAuley, UCSD -

<http://jmcauley.ucsd.edu/data/amazon/links.html>

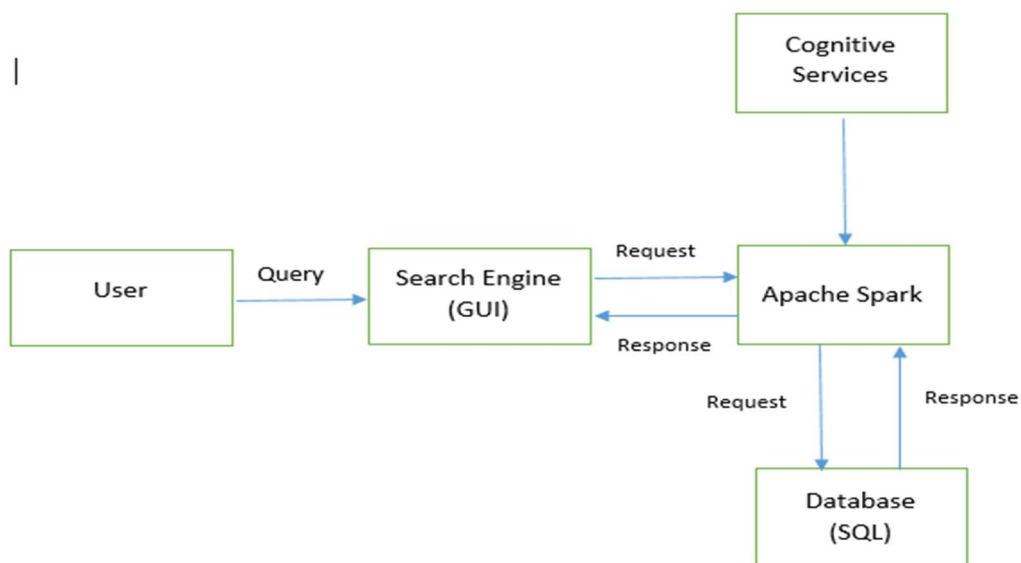
*Image-based recommendations on styles and substitutes* J. McAuley, C. Targett, J. Shi, A. van den Hengel SIGIR, 2015

*Inferring networks of substitutable and complementary products* J. McAuley, R. Pandey, J. Leskovec, Knowledge Discovery and Data Mining, 2015

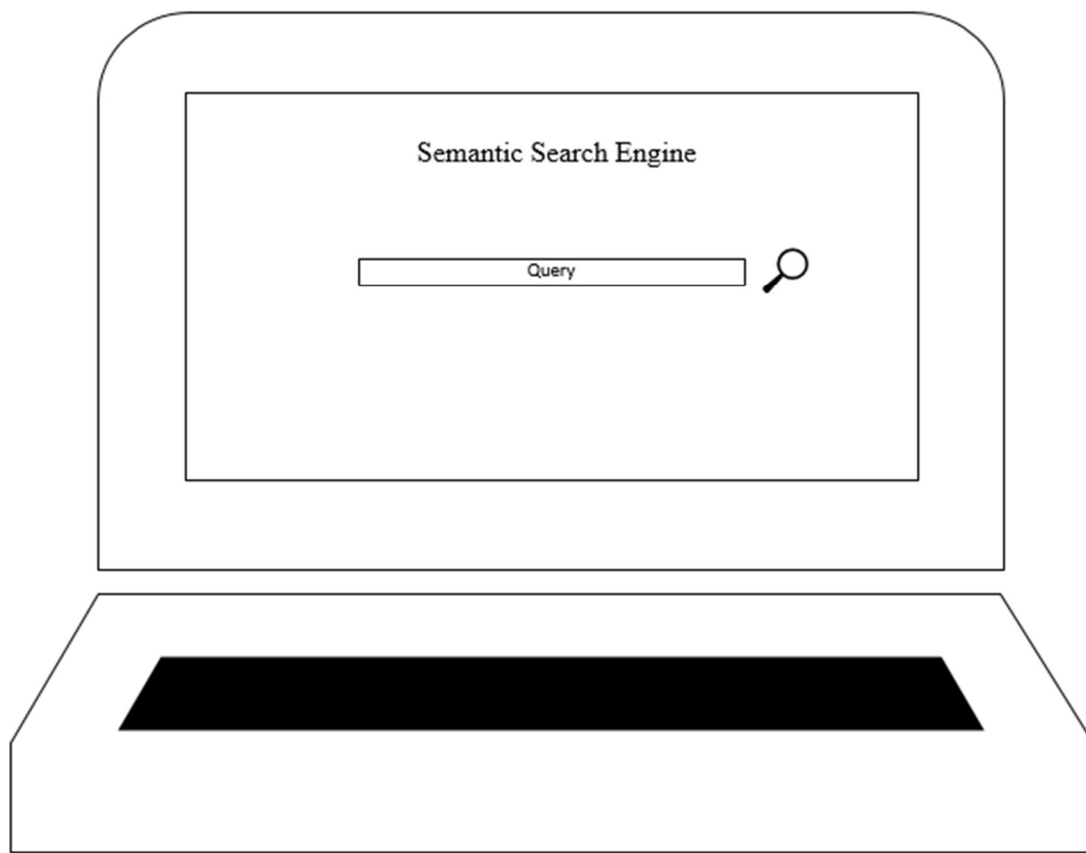
Wikipedia - <https://dumps.wikimedia.org/>

DBpedia - <http://wiki.dbpedia.org/>

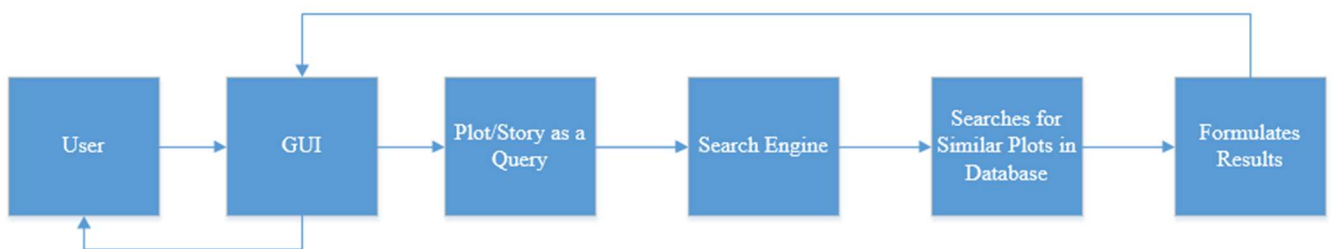
## System Architecture:



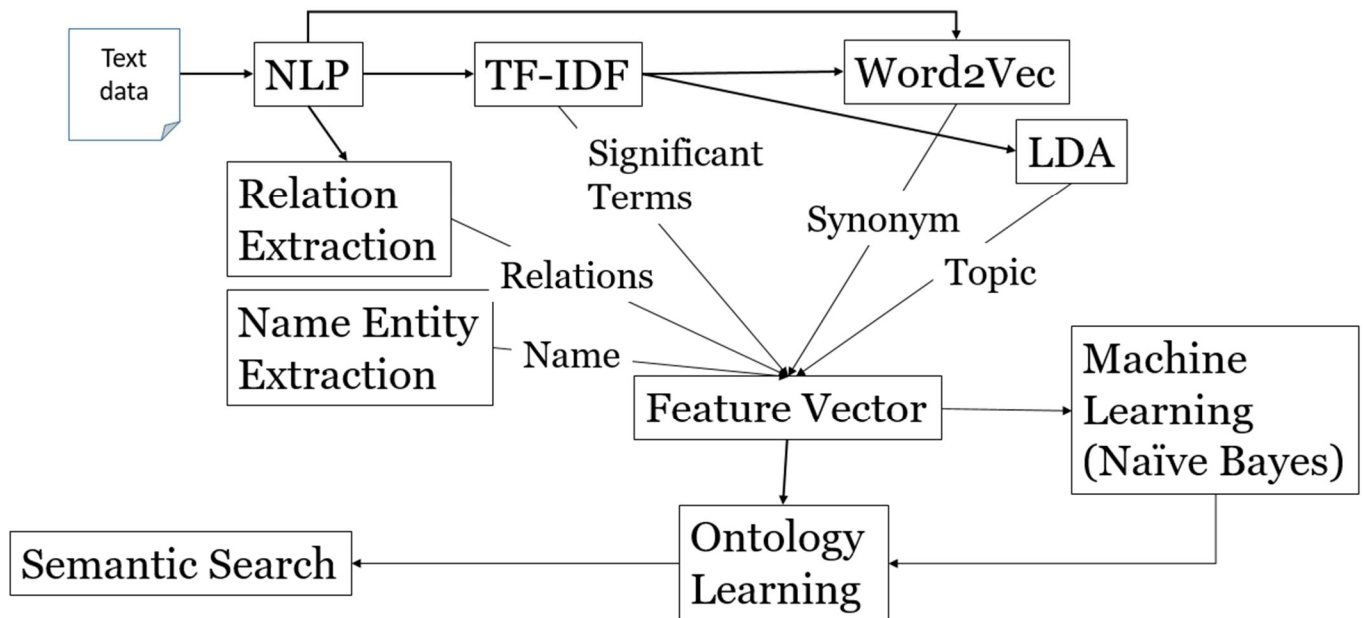
## Wireframe of Search Engine:



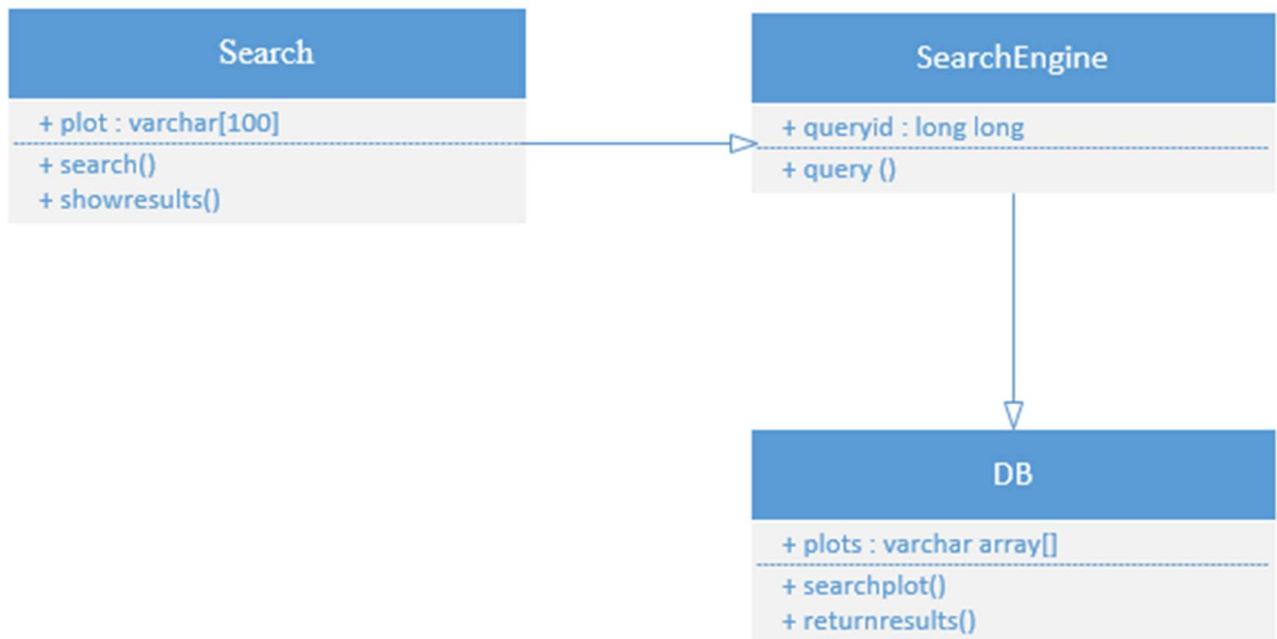
## Workflow Diagram:



## General Work-Flow:



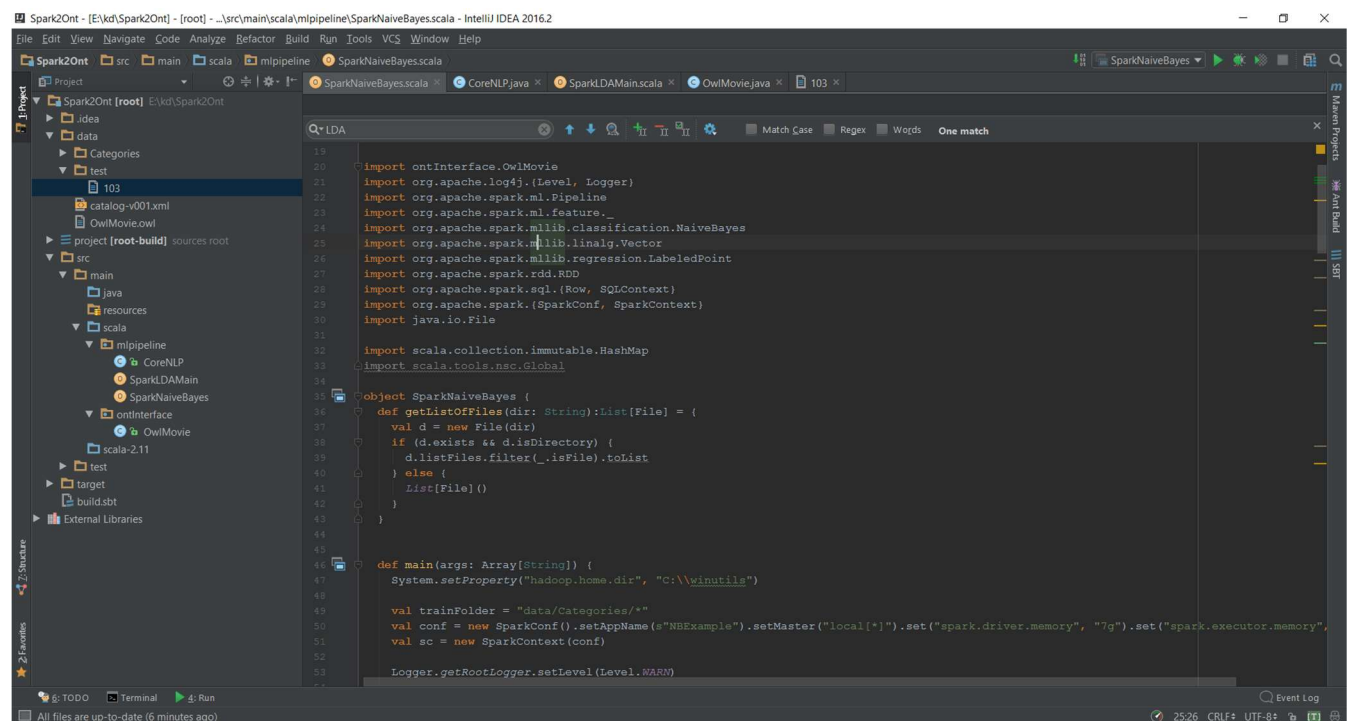
## UML Class Diagram:



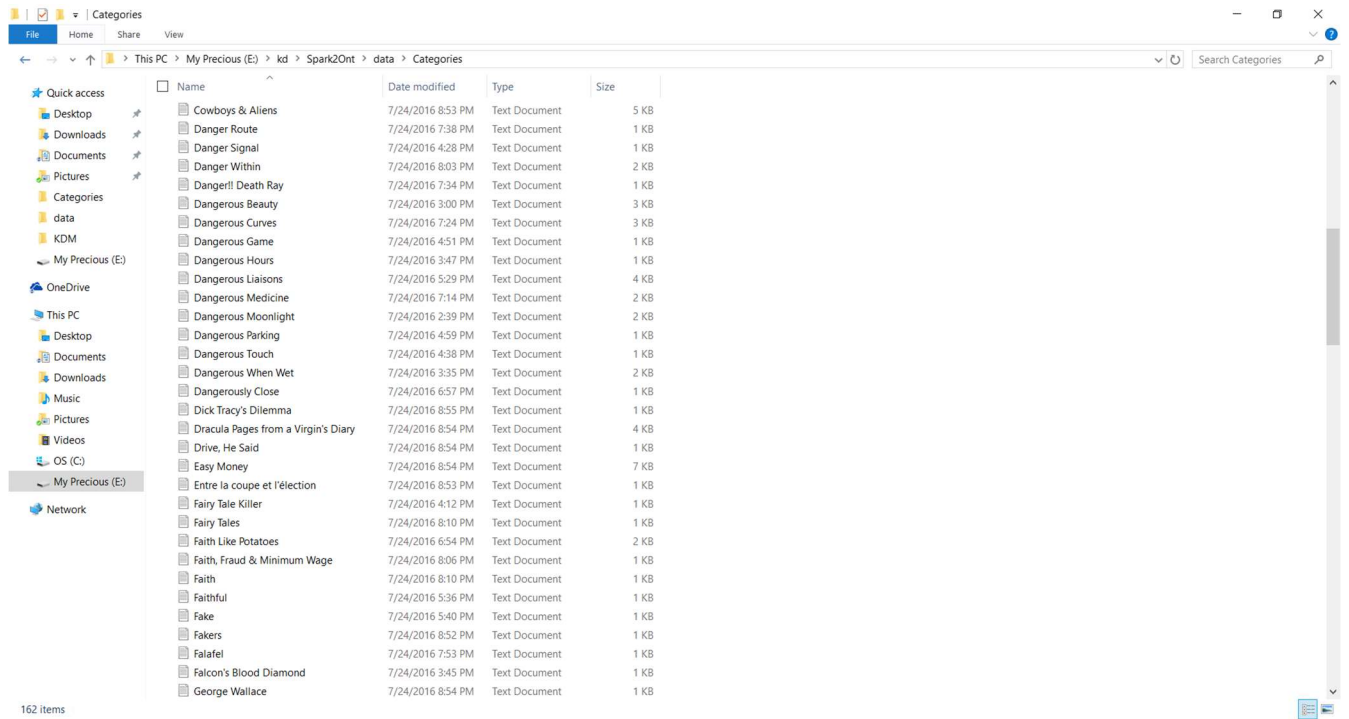
## Front-End of Search Engine:



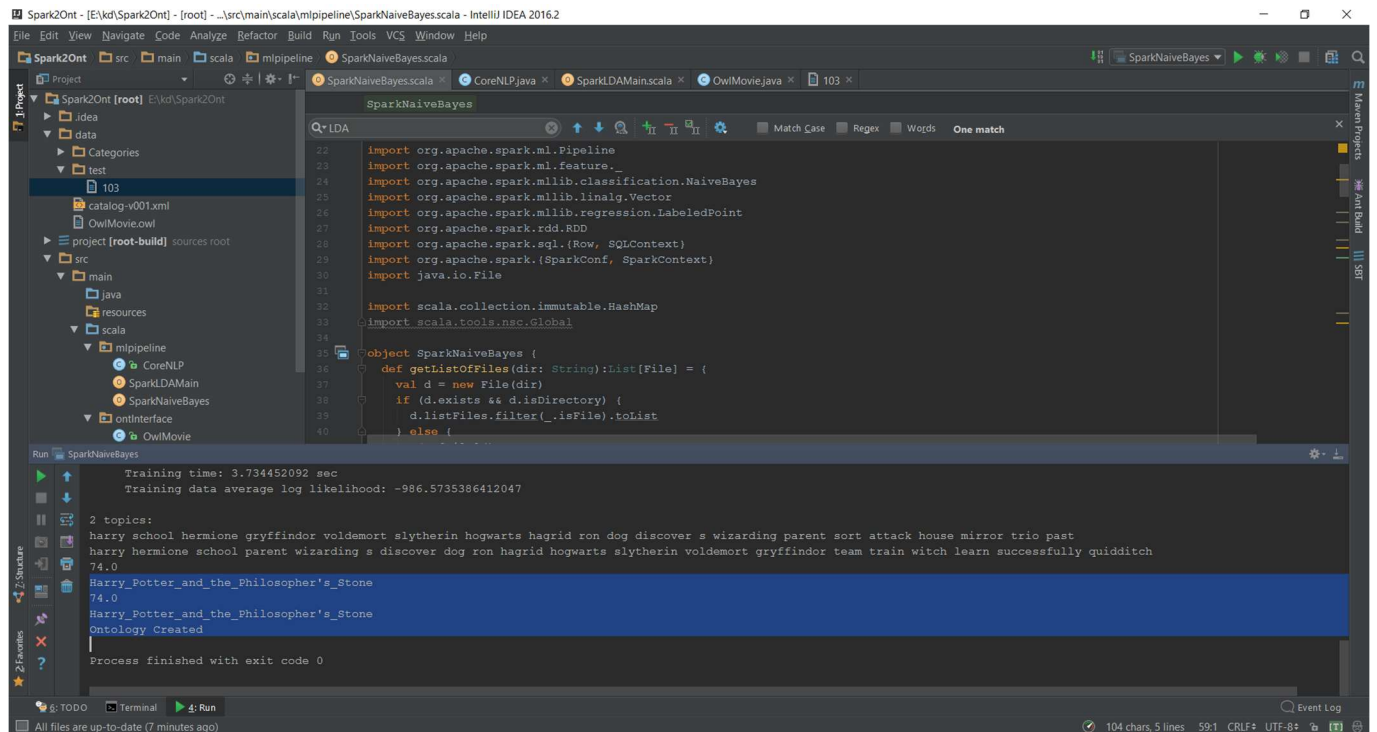
**Code:**



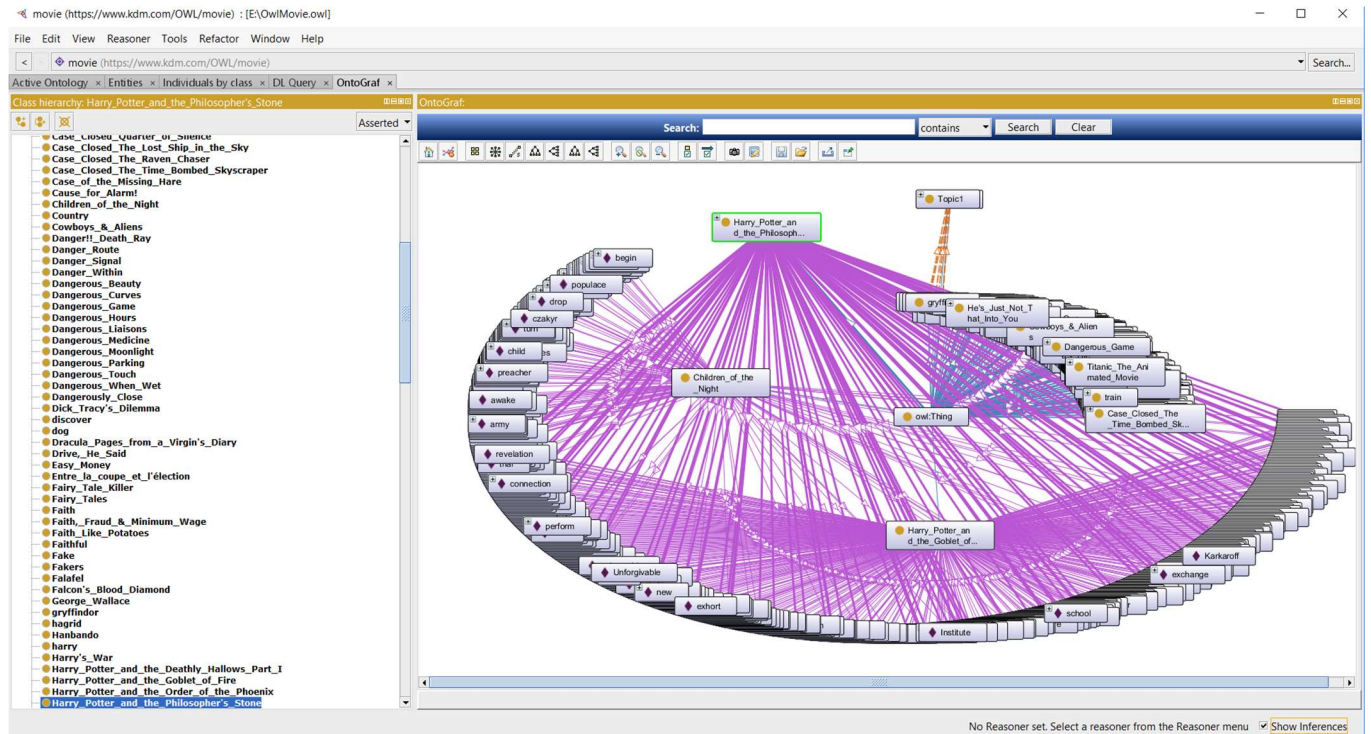
# Input Data:



# Result:

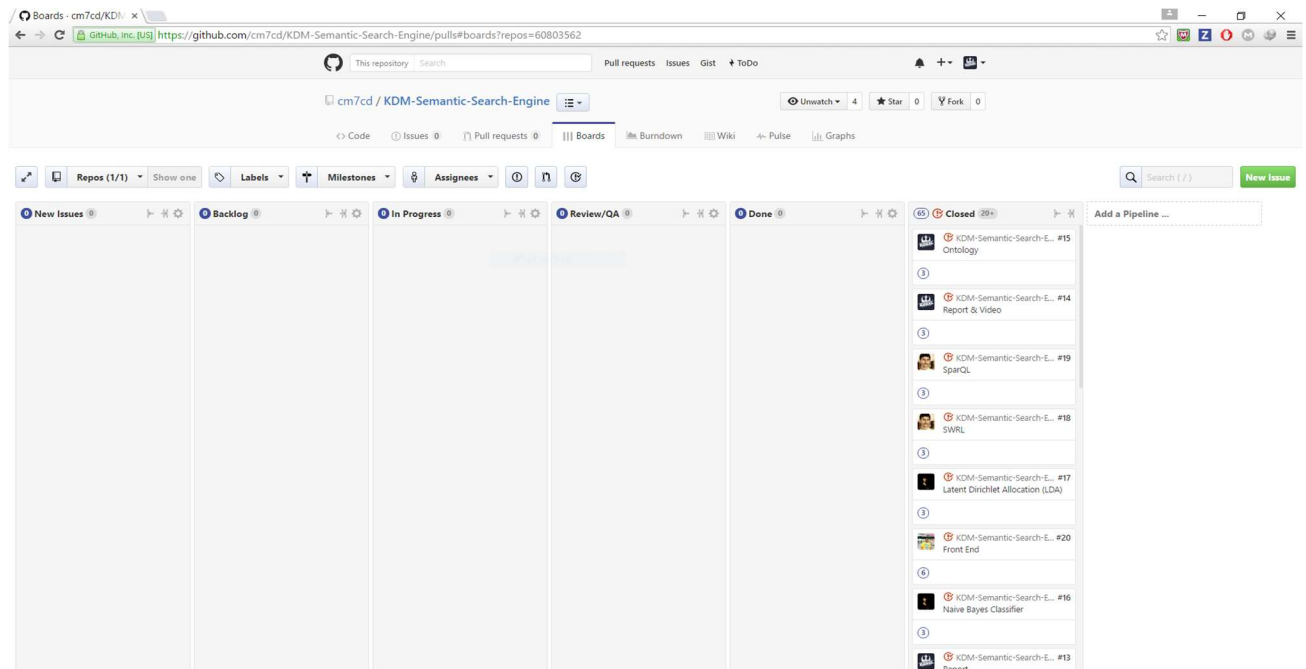


## Created Ontology:



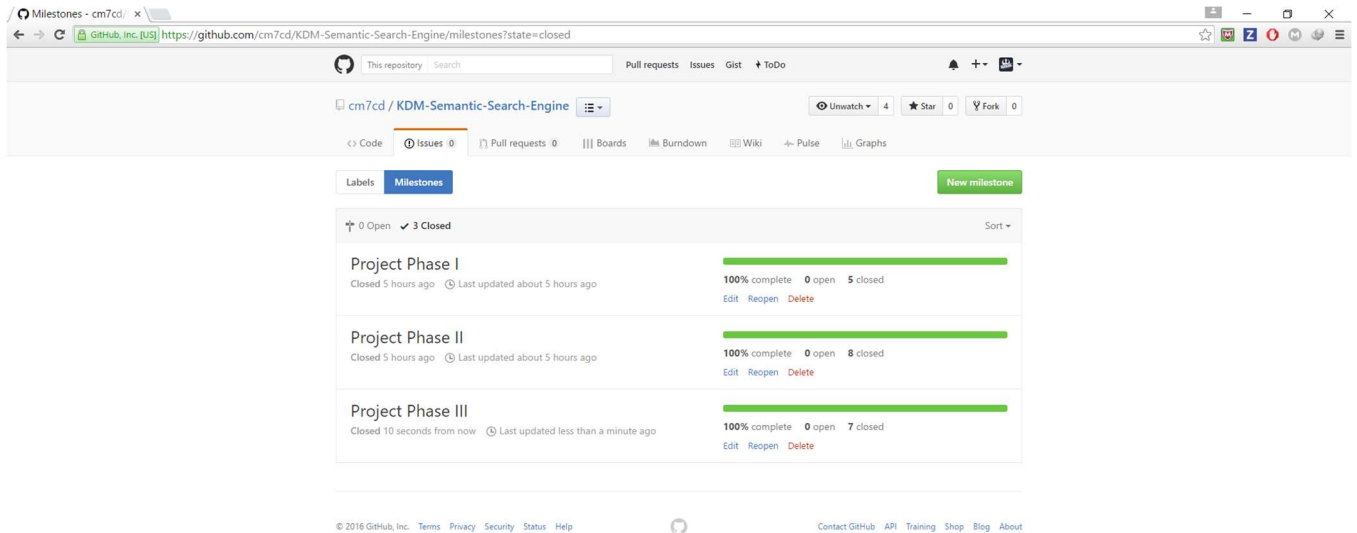
## Zen-hub Project Management:

## Issues Board:

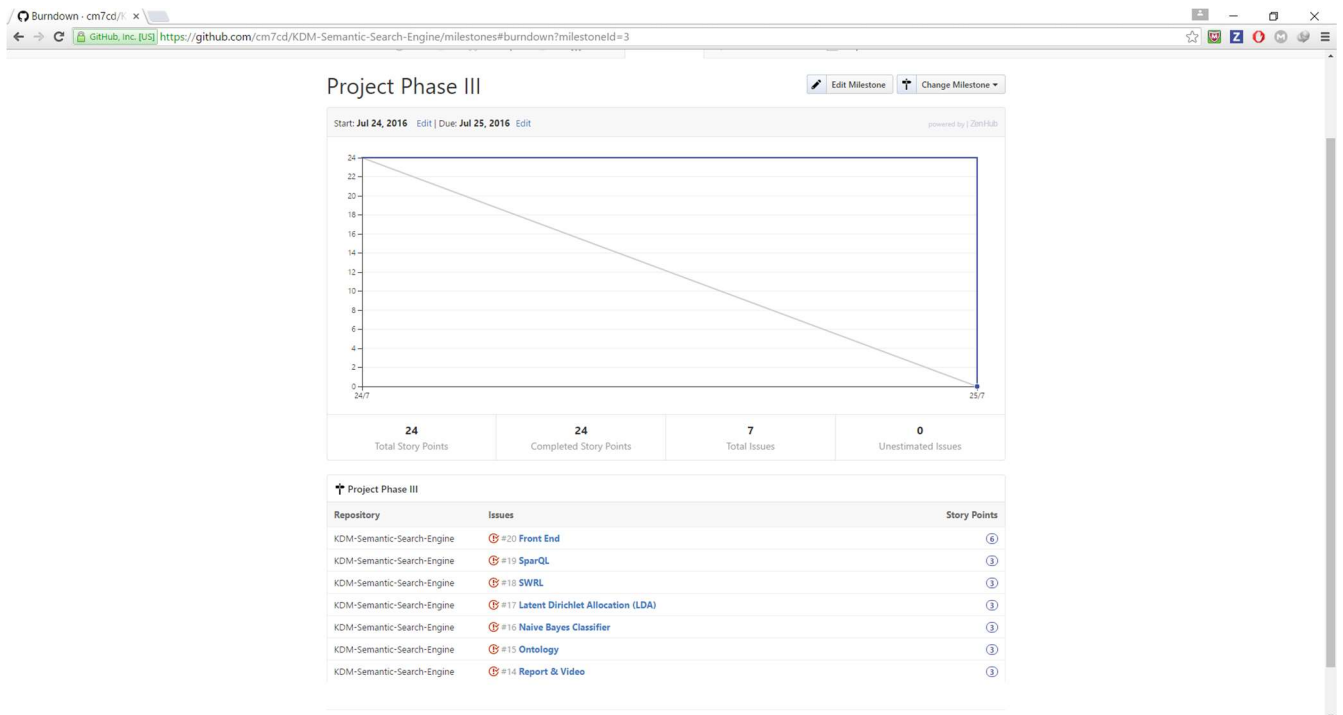




## Milestones:



## Issues & Burndown:





## **Contributions:**

Sai Venakatesh Gatiganti – Naïve Bayes Classifier, Latent Dirichlet Allocation (LDA)

Karthik Reddy Vundela – SparQL & SWRL

Chaitanya Sai Manne – Front-End

Sri Chaitanya Patluri – Report & Video.

## **Future Work:**

Our Future Work for the project includes including various other domains for the search engine and adding a dynamic web-crawler which crawls the web for related plots and displaying the results dynamically in real time.

**Youtube URL:** <https://www.youtube.com/watch?v=H3vsaDTTA0M>

**GitHub URL:** <https://github.com/cm7cd/KDM-Semantic-Search-Engine/>

## **Bibliography:**

- <http://jmcauley.ucsd.edu/data/amazon/links.html>
- <http://nlp.stanford.edu/nlp/>
- <https://en.wikipedia.org/>
- <http://wiki.dbpedia.org/>