Background:
- Our sponsors Preethi Raghavan and Xu Zhe are building an Opinion Mining Toolkit which works on the user comments retrieved from websites such as usnews.com, to classify documents.

Project Goals
- The project aims at creating web services that allows an administrator to add patterns (set of html tags that surround the comments) into a database using which comments can be retrieved from URLs.
- User can provide feedback based on the comments retrieved and this helps the administrator modify/add/delete the patterns stored.
- Along with the comments text, the other details of comments like author name, comment title and time stamp should also be retrieved.

Challenges
- Retrieving the user comments given a URL is not trivial. Each website follows its own pattern of storing comments.
- In some cases, the source of the webpage does not show user comments.
- Parts of the comments can be images.

Technologies Used
- Java 6, Tomcat, Axis 2 and MySQL.

The Application provides authentication for both the normal and admin user and has a rich user interface.
- It allows new user registration.
- It lets the admin user view the source of any URL and to obtain the relevant patterns to extract comments.
- Provides a feedback mechanism which allows the application to correct the erroneous patterns.
- Handles multiple patterns for a single URL – providing UI feature to select the appropriate pattern for the given webpage.