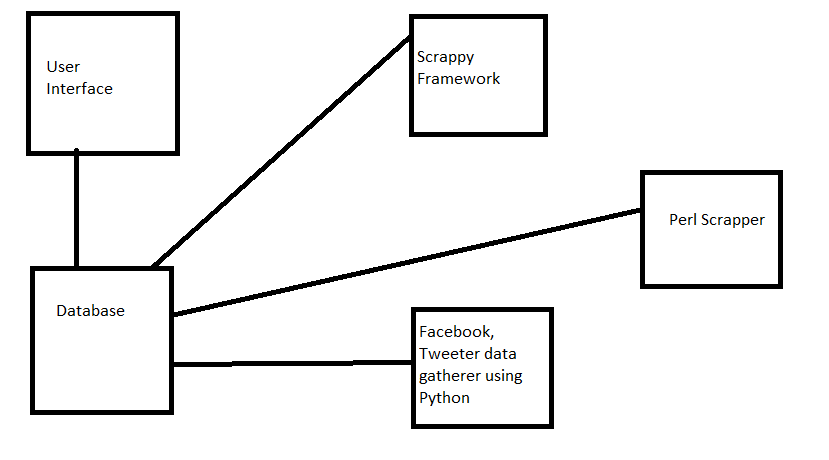
Architecture Document

Symc – Crawler is divided into 5 parts architecture wise.



1. User Interface: JSP is used for this purpose to render pages.
2. Database: MySql database is used for storage of data.
3. Scrappy Framework: It is used for spider purpose. Urls fetched are passed to perl scrapper.
4. Perl Scrapper : Blogs, article and comments are scrapped using perl Scrapper.
5. Facebook and Tweeter data gatherer using python: This is used for bringing data from apis of facebook and tweeter and feed to database.