**System Diagram**

Indexer\_time Stamp

NewsEngine

Restful Service

Client UI

Repository

NewsParser

News Engine retrieves news articles from provided URLs. These articles are filtered based on the contents using an XPath query which specifies where the required information is located on the web page and stored in a shared repository.

We use MongoDB for our shared repository. Mongo is a Document-based database which enhanced load and store overhead in comparison to traditional relational databases.

The Indexer\_time Stamp engine works asynchronously. It is invoked at certain times when new articles are fetched and compares each article to all available docs in database based on the timestamp. The index is connected to a web service API which provides RESTful service to clients.

**Component Description:**

**News Engine**

News engine is implemented in JAVA.

The data is provided by the MongoDB, and the method to start indexing is invoked using the web service .The query method is also called by web service.

**Web Service**

We used the powerful Spring Framework for the web API. The standard we used is RESTful, so we designed standardized URLs and secured our API with Oauth1.0 this allows us to protect our web service from unauthorized accesses. The controller of spring web framework responds to API calls. The framework uses dependency injection design pattern to provide programmers with simple integration of dependencies.

**Web service dependencies**

* Spring Security
* Spring Boot starter
* Apache commons
* Spring Web

**Web service API**

## /news/id/<id> Retrieve a specific article by id.

##/news/category/<cat> retrieve a specific category of an article.