**JOBRECSYS ARCHITECTURE**

MỤC LỤC

[**Phần 1.** **KIẾN TRÚC TỔNG QUAN** 1](#_Toc455275372)

[1.1. Mô hình kiến trúc 1](#_Toc455275373)

[1.2. Mô tả kiến trúc 2](#_Toc455275374)

[**Phần 2.** **KIẾN TRÚC CHI TIẾT CÁC SUB-PROJECT** 4](#_Toc455275375)

[2.1. Project Comparer 5](#_Toc455275376)

[2.2. Project JobRecSys 7](#_Toc455275377)

[2.3. Project CreateDataSet 10](#_Toc455275378)

# **KIẾN TRÚC TỔNG QUAN**

## Mô hình kiến trúc

**SYSTEM DB**

**User DataSet**

**Spring MVC**

**Apache Mahout**

**Apache Hadoop**

**Recommender web app**

**Recommender webservice**

**Algorithms**

**Evaluation metrics**

**Visualization**

**Experimental conduction**

**Data preparation**

**System Interfaces**

**System core modules**

**System DB**

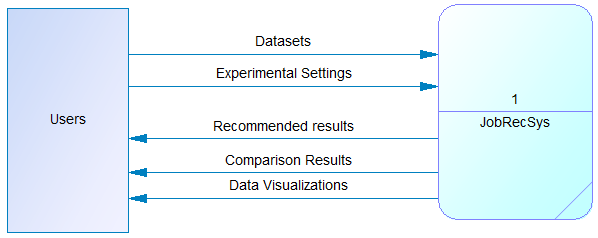
**Libraries**

## Mô tả kiến trúc

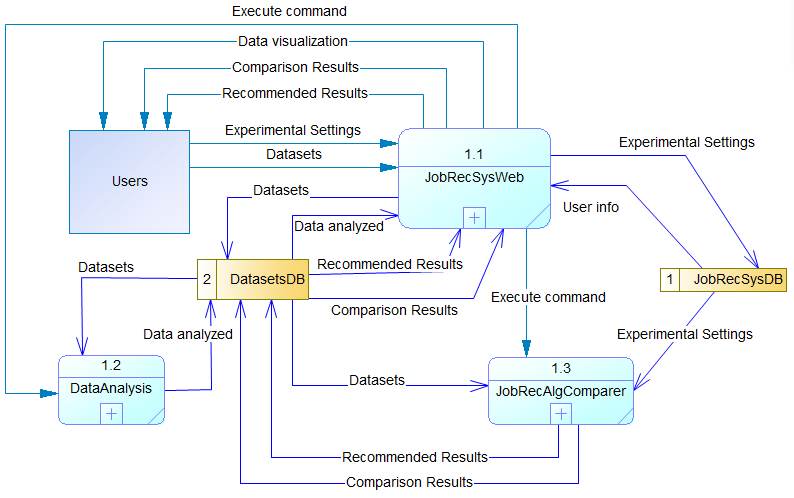
* + System interfaces
* Recommender web app: Cung cấp giao diện để người dùng có thể upload dữ liệu của họ. Sau đó có thể tiến hành thống kê hay chạy các thuật toán khuyến nghị.
* Recommender web service: Cung cấp một số API để người dùng có thể nhúng các kết quả khuyến nghị vào hệ thống riêng của họ.
  + System core modules
* Visualization: Đây là module cho phép xuất các kết quả thống kê, kết quả thực hiện của các thuật toán khuyến nghị,…
* Evaluation metrics: thực hiện các độ đo giữa các thuật toán khuyến nghị
* Experimental conduction: chạy thực nghiệm các thuật toán khuyến nghị.
* Algorithms: Chứa các thuật toán khuyến nghị như CF, CB, Hybrid,…
* Data preparation: Dùng để chuẩn bị dữ liệu đầu vào cho các thuật toán khuyến nghị để thực hiện thống kê dữ liệu trên các đặc trưng.
  + Database
* System DB: database dùng để tổ chức lưu trữ thông tin người dùng, các tác vụ thực hiện của người dùng và các thông tin thiết lập chạy thuật toán.
* User dataset: hệ thống lưu trữ files dataset do người dùng upload cùng với đó là các file kết quả khuyến nghị, kết quả so sánh và kết quả thống kê.
  + Libraries
* Spring MVC: dùng hỗ trợ xây dựng website tương tác người dùng.
* Mahout: sử dụng các thư viện recommend có sẵn trong Mahout để thực hiện khuyến nghị.
* Hadoop: sử dụng để lưu trữ dữ liệu lớn, hỗ trợ cho việc phát triển trong tương lai.

# **KIẾN TRÚC CHI TIẾT CÁC SUB-PROJECT**

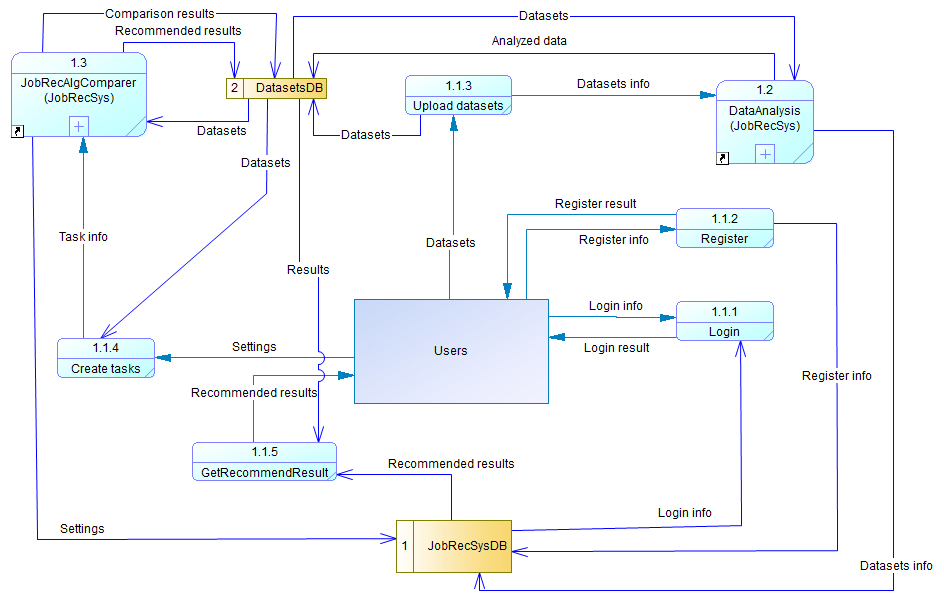
Mô hình DFD tổng quát cấp 0 của hệ thống



Mô hình DFD cấp 1

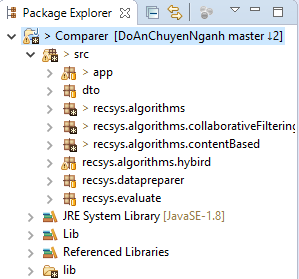


DFD cấp 2 (JobRecSysWeb)



## Project Comparer

Cấu trúc thư mục và package trong project

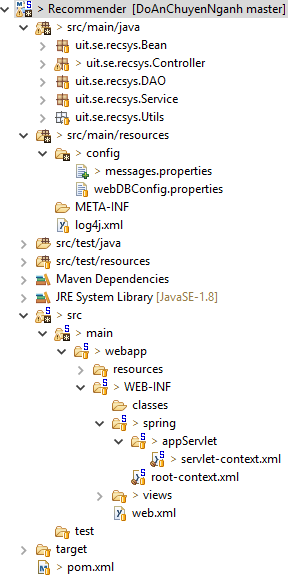


Mô tả các package và folder:

* Package src: chứa các package con sau:
  + app: chứa main class thực thi chương trình
  + dto: chứa các class dto
  + algorithms: chứa 3 package con về thuật toán gồm:
    - abstract class RecommendationAlgorithm: class cha chứa các phương thức thiết lập đầu vào và ra của dataset, class này dùng cho các class cài đặt thuật toán kế thừa.
    - collaborativeFiltering: chứa class cài đặt thuật toán collaborative filtering
    - contentBased: chứa class cài đặt thuật toán content-based
    - hybrid: chứa class cài đặt thuật toán hybrid
  + datapreparer: chứa các class tiền xử lý dữ liệu trước khi đưa vào cho các thuật toán chạy
  + evaluate: chứa các class cài đặt việc so sánh thuật toán.
* Thư mục lib: chứa các thư viện của project

## Project JobRecSys

Cấu trúc thư mục và package trong project

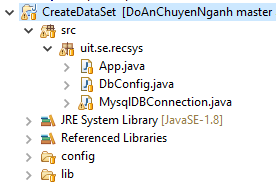


Mô tả các package và folder:

* **Package src/main/java**: là package chính chứa toàn bộ các package trong project gồm:
  + controller: chứa các file controller
  + bean: chứa các file java bean
  + dao: chứa các file model tương tác với database
  + service: chứa các service xử lý
  + utils: chứa các file tiện ích
* **Package src/main/resources**: là package chứa các file tài nguyên của chương trình. Package này chứa folder quan trọng sau:
  + Config: chứa các file config cho hệ thống gồm:
    - File message.properties: là file chứa các thông báo lỗi dùng xuất ra màn hình cho người dùng
    - File webDBConfig.properties: là file chứa các thông tin cấu hình kết nối đến mysql server.
* **Package src/main/webapp**: gồm các file và folder liên quan đến phần cấu hình project và các views trong project:
  + Folder resources: chứa tài nguyên của project sử dụng để hổ trợ lập trình giao diện website:
    - Thư mục css: chứa các file .css
    - Thư mục js: chứa các file .js
    - Thư mục images: chứa các hình ảnh
    - Thư mục libs: chứa các thư viện của các bên thứ 3 hỗ trợ làm giao diện web.
  + Folder WEB-INF: chứa các file cấu hình:
    - Spring/appServlet/servlet-context.xml: cấu hình các bean cho ứng dụng web (web bean)
    - Spring/root-context.xml: cấu hình các bean cho toàn ứng dụng (non-web bean)
    - Views: chứa các file views của web (file .jsp)
    - Web.xml: cấu hình web, là file đầu tiên được nạp vào server khi ứng dụng chạy để cấu hình và nạp các tập tin cấu hình khác.
* **File pom.xml**: là file cấu hình các dependences (thư viện: spring, JDPC, …) của ứng dụng web, các dependences này sẽ được maven tự động tải về thêm vào project khi build project lần đầu.

## Project CreateDataSet

Cấu trúc thư mục và package trong project



Mô tả các package và folder:

* Package uit.se.recsys: chứa các class sau:
  + App.java: class main thực thi chương trình
  + DbConfig.java: class cấu hình kết nối đến database
  + MysqlDBConnection.java: class tạo connection và thực thi các câu lệnh truy vấn sql.
* Thư mục config: chứa file config.txt chứa các thông số cấu hình database cần kết nối
* Thư mục lib: chứa thư viện của project