### 

### Service Layer: *Đầu ra của hệ thống, giúp người dùng thực hiện các tương tác với hệ thống*

**Customer Web UI:** Giao diện giúp người dùng thông thường có thể thực hiện các tương tác với trang web và tiếp nhận kết quả hệ thống trả về

### Processing Layer: *Thực hiện các xử lý phục vụ cho các tính năng người dùng của hệ thống.*

**Router Handler + Authentication:** Quản lý việc tiếp nhận yêu cầu của người dùng và chuyển yêu cầu về đúng module xử lý, sau đó trả kết quả về đúng người dùng sau khi nhận kết quả từ các module xử lý.

**Pre-recommend Processing:** Quản lý việc xử lý dữ liệu để tạo ra các thông tin có ý nghĩa phục vụ cho mục tiêu đề xuất:

* **Build Knowledge Graph:** Dựa theo những tính năng đề xuất, chọn ra những node liên quan để xây dựng đồ thị tri thức tương ứng. (Candidate Graph)
* **Apply Recommender Algorithms:** Áp dụng các thuật toán đề xuất để chọn sẵn một số khóa học/chương trình học gợi ý tổng quát và lưu vào dữ liệu tạm **Cache (Result)**.

**Recommendation Controller:** Thực hiện các yêu cầu người dùng bằng cách sử dụng các thuật toán hoặc kỹ thuật xử lý dựa trên dữ liệu người dùng từ *serving storage* và *cache* ở *Pre-recommend Processing*để lọc lại kết quả (phù hợp hơn với năng lực hiện tại của người dùng) và trả về cho người dùng (kết quả cũng được lưu lại vào ***Cache*** để dùng cho những truy vấn tương tự trong tương lai gần)

**User Management Controller:**

* **User Behavior Collecting:** Quản lý việc ghi nhận các thao tác cập nhật của người dùng trên trang web
* **User Profile Management:** Xử lý các tính năng liên quan đến việc quản lý hồ sơ người dùng

*Quá trình xử lý tại layer này có thể mô tả thông qua sơ đồ sau:*

### 

### Storage Layer: *Lưu trữ dữ liệu và xử lý các biến đổi dữ liệu nhằm phục vụ các hoạt động của hệ thống.*

**Data Lake:** Lưu trữ các thông tin hiện của từng thành phần sau khi thu thập, làm sạch và rút trích đặc trưng.

* **Career:** Lưu trữ các dữ liệu về thông tin nghề nghiệp và năng lực yêu cầu. (hiện tại là csv files)
* **Course:** Lưu trữ các dữ liệu về thông tin khóa học trực tuyến và năng lực liên quan. (hiện tại là csv files)
* **User:** Các thông tin về người dùng user profile.

**Mapping/Importing Module:** Xử lý và đổ dữ liệu vào DW theo kiến trúc hiện có.

**Data Warehouse:** Lưu trữ tất cả những dữ liệu đã thu thập từ các nguồn và thông tin lịch sử trong quá trình hoạt động. Các dữ liệu này được tổ chức để dễ dàng lưu trữ lâu dài và thực hiện các tính năng thống kê hoặc đầu vào cho các thuật toán.

**Extract Module:** Rút trích dữ liệu từ DW sang Serving Storage để phục vụ cho việc vận hành trang web.

**Serving Storage:** Dữ liệu cơ bản phục vụ cho hoạt động bình thường trên ứng dụng, với data model là kiến trúc graph bao gồm đầy đủ các node với dữ liệu mới nhất vừa được trích từ warehouse.

**Semantic Annotated Datasourcce:** Dữ liệu thô đã được chú thích về mặt ngữ nghĩa được dùng để cải thiện độ chính xác, cũng như tính phù hợp của các Deep Learning Model ở thời điểm tư vấn.

### Data Ingestion Layer: *Thu thập và rút trích dữ liệu có ý nghĩa cho hệ thống tư vấn.*

**Crawl Controller:** Điều khiển các luồng hoạt động và xử lý của các Crawler

**Raw Data Zone:** Các dữ liệu thô sau khi Crawl

* **Raw MOOCS:** chứadữ liệu sau khi Crawl của các khóa học trực tuyến (csv files).
* **Raw Jobs:** chứadữ liệu sau khi Crawl của các công việc đăng tuyển (csv files).

**Feature Extraction Module:**

* **Competency Handler:** sử dụng mô hình học sâu *Competency Recognition Model* để tiến hành nhận diện các thực thể năng lực được đề cập trong các mô tả công việc cũng như mô tả khóa học.
* **Synonym Detector:** sử dụng kho từ vựng về năng lực *Vocabulary Mapping Model* để tiến hành nhận diện các năng lực mang cùng một ngữ nghĩa để dữ liệu đầu ra được sạch hơn.
* **Connection Weight Calculator (Course level):** tính toán/đánh giá mức độ thành thạo một năng lực/kỹ năng có thể đạt được sau một khóa học.
* **Competency Statistic (Career probability):** thống kê độ phổ biến của năng lực trên một tiêu đề nghề nghiệp, ở thời điểm thu thập dữ liệu.

**User Activity Handler:** Quản lý việc cập nhật hồ sơ và hoạt động người dùng hệ thống vào kho dữ liệu. (có thể triển khai realtime)