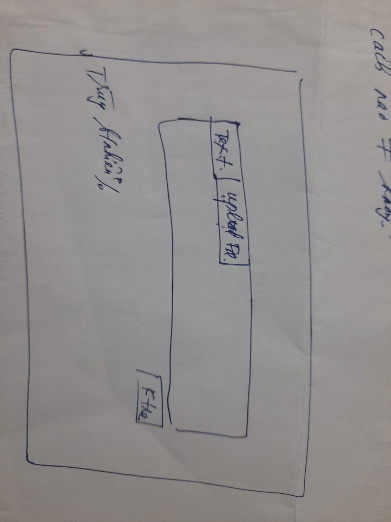
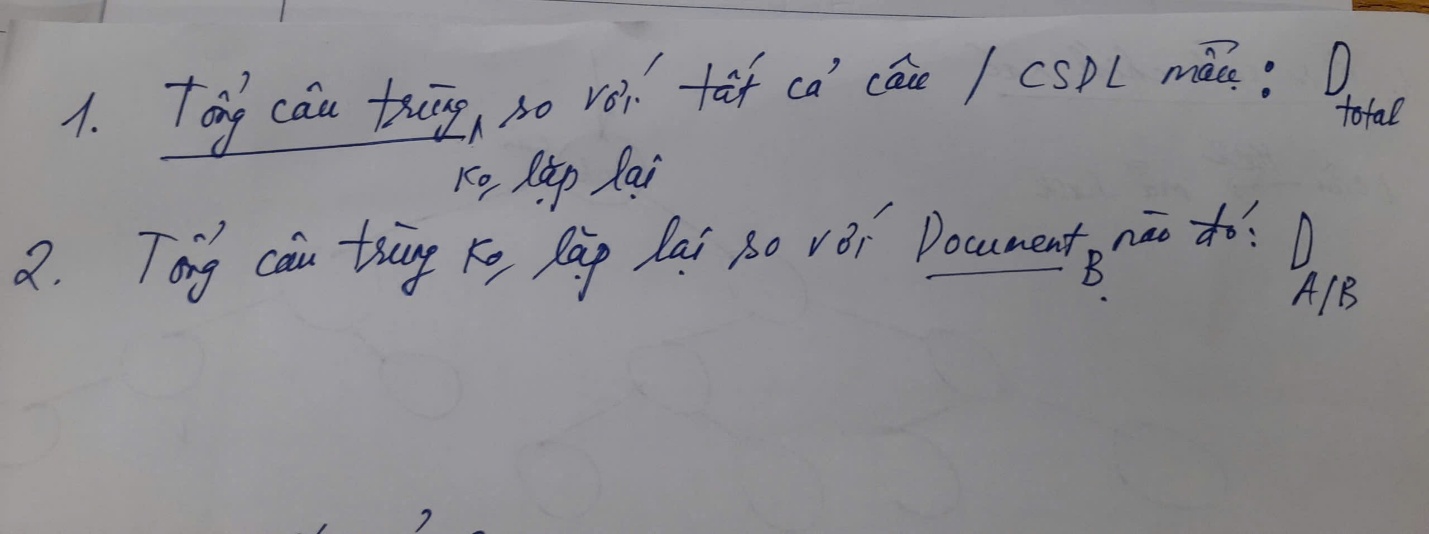
Note:

Chức năng thì lúc đầu vào sẽ là trang chủ sẽ hiển thị là bao nhiêu tài liệu, số người dùng, số lượt kiểm tra...có đăng nhập(bằng google...tài khoản ), đăng ký. Điều hướng tùy quyền người dùng.

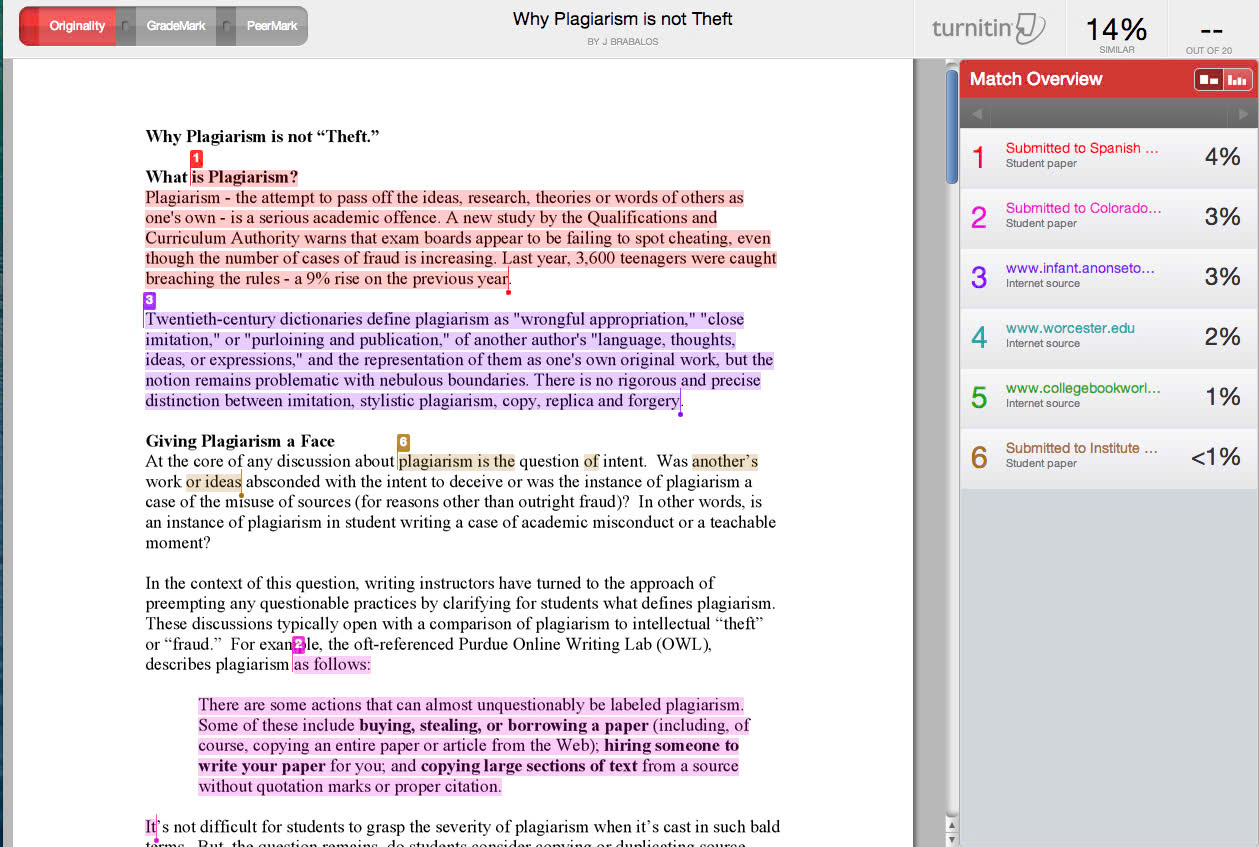
Trong trang người dùng thì. Có chỗ 2 lựa chọn là text để copy vào và tải file để so sánh.



Khi nhấn kiểm tra thì chỗ trùng bao nhiêu % sẽ hiển thị kết quả theo 2 ý



Gần giống với phần mềm Turnitin. Bấm xem chi tiết mới hiển thị (đây là ý 1, ý 2 làm tương tự)



* So sánh câu giữa 1 tài liệu với **toàn bộ cơ sở dữ liệu**.
* So sánh câu giữa 1 tài liệu với **một tài liệu khác**.
* Trả về **tỷ lệ phần trăm câu trùng**.
* Phân loại kết quả dựa trên ngưỡng phần trăm.

**+ Backend (NodeJS)**

Xây dựng API upload document: **Loại tài liệu được thêm vào dùng so sánh?**

CRUD Document

NLP:

* Tách câu: Bộ văn bản tiếng việt
* So sánh câu: Tree AVL cây cân bằng
* Dùng [**vncorenlp**](https://github.com/vncorenlp/VnCoreNLP) hoặc [**underthesea**](https://pypi.org/project/underthesea/)

**+ Frontend (ReactJS)**

Chưa có hình vẽ minh họa?

Giao diện upload document:

Hình ảnh

Hiển thị danh sách document:

Hình ảnh

Trang so sánh:

Hình ảnh

**+ Database (MongoDB)**

Lưu trữ bằng document để dễ sử dụng

**Có sử dụng mã hóa gì đặc biệt cho từ không?**

Cấu trúc

{

"\_id": "doc123",

"name": "Document A",

"uploadedBy": "user1",

"sentences": [

{ "sentence\_id": 1, "content": "Đây là câu thứ nhất.", "hash": "abc123" },

{ "sentence\_id": 2, "content": "Đây là câu thứ hai.", "hash": "def456" }

]

}

**+ Tính năng**

User:

* Uplodad tài liệu
* So sánh 2 tài liệu
* So sánh câu với tài liệu
* So sánh tài liệu với cơ sở lưu trữ

Admin:

* Quản lý tài liệu
* Thêm tài liệu
* Xóa tài liệu khởi cơ sở dữ liệu