I – Introduction

* Trong thời đại 4.0 hiện nay, với nhu cầu về việc lưu trữ tài liệu đang tăng cao thì câu hỏi được đặt ra là làm sao để có thể lưu trữ tài liệu một cách hiệu quả và lâu dài? Từ đó nhằm phục vụ các nhu cầu về lưu trữ tài liệu của chúng ta một cách rõ ràng, mạch lạc theo từng chủ đề cụ thể, đặc biệt có thể cho một lượng lớn người sử dụng một cách lây dài như trong các công ty, nhà trường, thư viện số,… Ngoài ra còn có một vấn đề đáng quan ngại đó là sự trùng lắp dữ liệu khi phải lưu trữ một lượng lớn tài liệu như vậy, từ những nhu cầu trên, nhóm chúng em xin có ý tưởng về đồ án “Xây dựng hệ thống quản lý tài liệu và kiểm tra tài liệu trùng lắp trên kho dữ liệu chung”.

1. Purpose

* Tạo ra một hệ thống trang web có thể lưu trữ một lượng lớn dữ liệu một cách lâu dài, có thể cho một lượng khá lớn người truy cập và sử dụng, giúp cho việc tìm kiếm tài liệu theo chủ đề, quản lý tài liệu và kiểm tra trùng lắp một các dễ dàng và hiệu quả.

1. Scope

* Hệ thống quản lý tài liệu có thể kiểm soát tốt các hoạt động lưu trữ tài liệu theo từng chủ đề, xử lý tốt việc upload tài liệu và kiểm tra các tài liệu trùng lắp, tương tự, tìm kiếm tài liệu nhanh chóng và chính xác.

1. Sofware Process
2. Description

* Một trang web quản lý dữ liệu cho một công ty nhỏ, để có thể sử dụng mỗi nhân viên phải có một tài khoản để đăng nhập vào trang web, tài khoản này sẽ được admin đăng kí cho khi nhân viên đó vào làm tại công ty.
* Với mỗi tài khoản mà có thể đăng nhập vào trang web thì người dùng có thể xem tài liệu, dowload tài liệu trên kho dữ liệu của trang web. Ngoài ra người dùng còn có thể upload tài liệu lên kho dữ liệu để lưu trữ, đặc biệt khi upload tài liệu lên thì hệ thống sẽ kiểm tra xem tài liệu đó đã có trong kho dữ liệu chung hay chưa để có thể tránh các trường hợp upload lên quá nhiều tài liệu trùng hay tương tự nhau.
* Riêng với tài khoản admin thì có thêm các chức năng đăng kí các tài khoản mới cho các nhân viên mới trong công ty và sửa đổi mật khẩu của các tài khoản cũ theo yêu cầu của nhân viên.

1. Functional

* Người dùng phải xác thực vào hệ thống bằng tài khoản của mình và sử dụng theo quyền hạn của tài khoản đó.
* Tất cả người dùng sau khi đã xác thực được vào hệ thống thì có thể xem tài liệu trên kho dữ liệu, có thể dowload tài liệu đó, chỉnh sửa nó và có thể upload tài liệu lên kho dữ liệu theo từng chủ đề riêng biệt, đặc biệt tài liệu upload lên hệ thống sẽ được hệ thống kiểm tra xem có trùng lặp hay tương tự như tài liệu nào trong kho dữ liệu hay không và thông báo cho người dùng biết. Ngoài ra, người dùng có thể tạo các chủ đề riêng biệt để phân loại các tài liệu một cách rõ ràng, có thể tìm kiếm tài liệu một cách nhanh chóng qua tên của tài liệu hoặc các nội dung liên quan tới tài liệu đó.
* Người quản trị viên ngoài các chức năng của một người dùng bình thường như trên thì còn có thêm chức năng đăng kí tài khoản mới cho các nhân viên khi mới trong hệ thống, có thể phân quyền cho tài khoản đó và đổi mật khẩu cho các tài khoản củ theo mong muốn.

1. Non – Functional
2. Usability

Mục tiêu:

* Người dùng của hệ thống có thể sử dụng các chức năng một cách dễ dàng.
* Hệ thống được tổ chức làm sao để lỗi người dùng là ít nhất.

Non-Functional Requirements có thể kiểm tra được:

* Việc Training có thể sử dụng thành thạo mọi mọi chức năng hệ thống sau khi được đào tạo sau 1 giờ. Sau thời gian đào tạo, số lỗi trung bình tạo ra bởi người dùng có kinh nghiệm không vượt quá 2 lỗi cho mỗi giờ sử dụng hệ thống.

1. Reliability

* Tính bảo mật: Mỗi nhân viên sẽ có một tài khoản để xác thực vào hệ thống với những phân quyền khác nhau.

VD: Người quản lý có thể tạo thêm các tài khoản mới và phân quyền cho các tài khoản đó.

* Thời gian trung bình để sửa chữa hệ thống khi có lỗi: Non-Functional Requirements cho phép hệ thống ngừng hoạt động trong vòng 12 giờ sau khi có lỗi.
* Lỗi hệ thống: Hệ thống phải được kiểm tra thường xuyên. Nếu có lỗi, thì đó phải là lỗi nhỏ, không ảnh hưởng đến hệ thống và phải được sửa chữa trong ngày

1. Performance

* Thời gian đám ứng cho một hành động tối đa là 8s
* Số lượng giao dịch trên cùng một thời điểm là đủ cho một công ty vừa phải sử dụng

1. Supportability

* Dễ sửa lỗi: Phát hiện lỗi hệ thống trung bình 30 phút.
* Dễ bảo trì: Thêm chức năng mới nhanh chóng mà không ảnh hưởng đến các chức năng đã có.

1. Design Constraints

* Hệ thống được viết bằng Express Framwork
* Cơ sở dữ liệu được thiết kế trên MongoDB

1. Purchased Components

* Frontend: React – Visual Studio Code
* Backend: NodeJS - Visual Studio Code

1. Interfaces

* Giao diện phải rạch mạch, rõ ràng, dễ nhìn dễ sử dụng để người dùng có thể hiểu rõ cách hoạt động của nó,hạn chế tối đa những lỗi không đáng có ảnh hưởng đến trải nghiệm người dùng.
* Người dùng không thể chỉnh sửa đến cấu trúc của hệ thống như cơ sở dữ liệu, code hoặc giao diện.
* Dưới đây là giao diện của phần mềm:

1. SYSTEM IMPLEMENT
2. Development Enviroment

* Hệ thống được phát triển trên NodeJS – Express

1. Tool and teachnique

Ngôn ngữ: Javascript

Phần mềm được thiết kế và phát triển trên phần mềm Visual Studio Code

Áp dụng React và Express ( Express Framework )

Cơ sở dữ liệu được xây dựng trên MongoDB

Kĩ thuật kết nối với cơ sỡ dữ liệu là mongoose

1. Open Source/ Component

* [React](https://reactjs.org/) là thư viện [JavaScript](https://www.hostinger.vn/huong-dan/javascript-la-gi/) phổ biến nhất để xây dựng giao diện người dùng (UI). Nó cho tốc độ phản hồi tuyệt vời khi user nhập liệu bằng cách sử dụng phương pháp mới để render trang web. Components của công cụ này được phát triển bởi [Facebook](http://facebook.com/). Nó được ra mắt như một công cụ JavaScript mã nguồn mở vào năm 2013. Hiện tại, nó đã đi trước các đối thủ chính như [Angular](https://angular.io/docs) và [Bootstrap](https://getbootstrap.com/), hai thư viện JavaScript bán chạy nhất thời bấy giờ.
* Express là một framework giành cho nodejs. Nó cung cấp cho chúng ta rất nhiều tính năng mạnh mẽ trên nền tảng web cũng như trên các ứng dụng di động. Express hỗ rợ các phương thức HTTP và midleware tạo ra môt API vô cùng mạnh mẽ và dễ sử dụng. Có thể tổng hợp một số chức năng chính của express như sau:
* Thiết lập các lớp trung gian để trả về các HTTP request
* Định nghĩa router cho phép sử dụng với các hành động khác nhau dựa trên phương thức HTTP và URL
* Cho phép trả về các trang HTML dựa vào các tham số.
* MongoDB là một dạng phần mềm cơ sở dữ liệu sử dụng mã nguồn mở NoSQL. Nó có thể hỗ trợ trên nhiều nền tảng khác nhau và được thiết kế với mục đích hướng đến đối tượng. MongoDB hoạt động dựa vào các khái niệm Collection và Document. Đồng thời, nó có hiệu suất cao cùng với tính khả dụng tốt và dễ dàng mở rộng.
* Các Collection trong MongoDB có cấu trúc cực kỳ linh hoạt. Điều này cho phép dữ liệu không cần thiết phải tuân theo bất kỳ một dạng cấu trúc nào. Vì thế, MongoDB có thể lưu trữ những dữ liệu có cấu trúc đa dạng và phức tạp. Dữ liệu trong MongoDB được lưu bằng định dạng kiểu JSON.
* Mongoose là một thư viện mô hình hóa đối tượng (Object Data Model - ODM) cho MongoDB và Node.js. Nó quản lý mối quan hệ giữa dữ liệu, cung cấp sự xác nhận giản đồ và được sử dụng để dịch giữa các đối tượng trong mã và biểu diễn các đối tượng trong MongoDB.