

ĐỀ CƯƠNG CHI TIẾT
KHÓA LUẬN TỐT NGHIỆP NGÀNH CÔNG NGHỆ THÔNG TIN

Họ tên sinh viên: **Võ Trọng Nghĩa**

MSSV: 110121273

Lớp: DA21TTC

Khóa: 2021

Tên đề tài: Tìm hiểu công nghệ Elasticsearch và xây dựng hệ thống tìm kiếm tài liệu thông minh

1. Mục tiêu của đề tài:

Nghiên cứu các khía cạnh lý thuyết và thực tiễn của công nghệ Elasticsearch trong việc xử lý và tìm kiếm dữ liệu. Xây dựng một website thân thiện với người dùng, cho phép tìm kiếm tài liệu hiệu quả dựa trên Elasticsearch. Hiểu rõ cơ chế hoạt động, ưu điểm và hạn chế của Elasticsearch trong xử lý tìm kiếm dữ liệu lớn. Phát triển một giao diện web trực quan, dễ sử dụng để hỗ trợ người dùng tìm kiếm và truy xuất tài liệu.

2. Nội dung thực hiện:

Tìm hiểu về công nghệ Elasticsearch:

Tìm hiểu khái niệm, kiến trúc và cơ chế hoạt động của Elasticsearch, bao gồm cách lập chỉ mục (indexing), tìm kiếm full-text và xử lý truy vấn.

Nghiên cứu các thành phần cốt lõi của Elasticsearch như index, document, mapping và query DSL (Domain Specific Language).

Khám phá các ưu điểm của Elasticsearch so với các công nghệ tìm kiếm khác về tốc độ, khả năng mở rộng và tính linh hoạt.

Tìm hiểu cách tích hợp Elasticsearch với các ứng dụng web thông qua API.

Nghiên cứu các thư viện hỗ trợ như elasticsearch.js hoặc client chính thức của Elastic để kết nối với Node.js/Express.js.

Thử nghiệm các kịch bản tìm kiếm cơ bản như tìm kiếm từ khóa, lọc theo trường dữ liệu, phân trang kết quả trên tập dữ liệu mẫu.

Phân tích và thiết kế hệ thống:

Xác định các yêu cầu chức năng: Tìm kiếm tài liệu theo từ khóa, lọc kết quả theo danh mục (loại tài liệu, ngày tạo, tác giả), hiển thị thông tin tài liệu.

Xác định các yêu cầu phi chức năng: Ứng dụng có tốc độ phản hồi nhanh, giao diện đơn giản, trực quan để sử dụng thân thiện với người dùng.

Xây dựng website và các chức năng chính:

Trang chủ: Giới thiệu hệ thống và ô tìm kiếm chính.

Trang tìm kiếm: Cho phép nhập từ khóa tìm kiếm, chọn bộ lọc (loại tài liệu, khoảng thời gian).

Trang kết quả: Hiển thị danh sách tài liệu với thông tin tóm tắt của các tài liệu đó.

Trang chi tiết tài liệu: Hiển thị toàn bộ thông tin tài liệu.

Tích hợp các tính năng tương tác như phân trang, sắp xếp kết quả, và gợi ý từ khóa.

Hiển thị kết quả: Hiển thị thông tin tài liệu dưới dạng danh sách với tiêu đề, tóm tắt nội dung, và liên kết tải. Hỗ trợ phân trang và tải thêm kết quả.

Quản lý tài liệu: Cho phép quản trị viên thêm tài liệu mới thông qua giao diện hoặc tải lên hàng loạt. Hỗ trợ cập nhật hoặc xóa tài liệu khỏi hệ thống.

Tính năng nâng cao: Gợi ý từ khóa tự động dựa trên truy vấn người dùng. Hỗ trợ tìm kiếm đa ngôn ngữ hoặc tìm kiếm theo ngữ nghĩa.

Công nghệ sử dụng:

HTML, CSS, JavaScript, để xây dựng giao diện tĩnh và tương tác ở mức cơ bản.

React để xây dựng giao diện động, giúp tối ưu trải nghiệm người dùng

Express.js để xây dựng API để xử lý yêu cầu từ giao diện người dùng và tương tác với Elasticsearch.

Node.js là môi trường để chạy backend, hỗ trợ cho Express.js.

MongoDB để lưu trữ và quản lý dữ liệu.

Elasticsearch là công cụ chính để lập chỉ mục và tìm kiếm tài liệu hiệu quả

Các công cụ hỗ trợ khác: Git để quản lý mã nguồn, Postman để kiểm tra API.

3. Phương pháp thực hiện:

Phương pháp nghiên cứu lý thuyết:

Thu thập và phân tích tài liệu từ các nguồn đáng tin cậy như sách, bài báo, tài liệu kỹ thuật của Elasticsearch, React, Express.js.

Tìm hiểu thông tin về các dự án tương tự để xác định ưu, nhược điểm và cách tiếp cận phù hợp.

Phương pháp thực nghiệm:

Sử dụng các công cụ như React, Node.js, Express.js, MongoDB, Elasticsearch để tiến hành xây dựng ứng dụng. Tích hợp Elasticsearch với các truy vấn mẫu để kiểm tra hiệu suất tìm kiếm. Cuối cùng là tiến hành kiểm thử, đánh giá hiệu suất của sản phẩm để đảm bảo đáp ứng được các yêu cầu của hệ thống tìm kiếm tài liệu thông minh.

4. Bố cục đề tài:

Tóm tắt khóa luận tốt nghiệp:

Vấn đề nghiên cứu

Các hướng tiếp cận

Cách giải quyết vấn đề và một số kết quả đạt được

Kết quả

Mở đầu:

Lý do chọn đề tài

Mục tiêu

Nội dung

Đối tượng và phạm vi nghiên cứu

Phương pháp nghiên cứu

Chương 1: Tổng quan

Giới thiệu tổng quan về đề tài

Chương 2: Nghiên cứu lý thuyết

Tổng quan về Elasticsearch

Tổng quan về ngôn ngữ lập trình

Tổng quan về công nghệ sử dụng

Giới thiệu về công cụ hỗ trợ và cơ sở dữ liệu

Chương 3: Thực hiện hóa nghiên cứu

Mô tả hệ thống

Xác định các yêu cầu chức năng của hệ thống

Thiết kế hệ thống

Chương 4: Kết quả nghiên cứu

Giao diện website hiển thị kết quả tìm kiếm

Chương 5: Kết luận và hướng phát triển

Kết quả đạt được

Hướng phát triển

5. Tài liệu tham khảo:

- [1] M. Casciaro and L. Mammino, Node.js Design Patterns, 3rd ed. Birmingham, UK: Packt Publishing, 2020.

- [2] C. Gormley and Z. Tong, Elasticsearch: The Definitive Guide. Sebastopol, CA: O'Reilly Media, 2015.

- React Official Documentation: <https://react.dev/learn>

6. Kế hoạch thực hiện đề tài:

Tuần	Từ ngày - đến ngày	Công việc thực hiện	Ghi chú
1	Từ ngày 08/04/2025 đến ngày 15/04/2025	- Viết đề cương chi tiết - Tìm hiểu về công nghệ Elasticsearch - Tìm hiểu về các công cụ, ngôn ngữ lập trình sẽ sử dụng để xây dựng hệ thống	- Hoàn thành đề cương chi tiết và có được kiến thức cơ bản về công nghệ Elasticsearch

2	Từ ngày 20/04/2025 đến ngày 27/04/2025	<ul style="list-style-type: none"> - Cài đặt các công cụ như Visual Studio Code, Elasticsearch, Node.js,... - Cài đặt môi trường phát triển với Node.js - Thiết kế cơ sở dữ liệu 	<ul style="list-style-type: none"> - Cài đặt môi trường phát triển thành công
3	Từ ngày 02/05/2025 đến ngày 9/05/2025	<ul style="list-style-type: none"> - Xác định các trang cần thiết trong website - Thiết kế giao diện cho website 	<ul style="list-style-type: none"> - Thiết kế xong giao diện cho website
4	Từ ngày 14/05/2025 đến ngày 23/05/2025	<ul style="list-style-type: none"> - Xây dựng các chức năng của hệ thống như tìm kiếm và quản lý tài liệu - Tích hợp công nghệ Elasticsearch vào hệ thống tìm kiếm tài liệu 	<ul style="list-style-type: none"> - Hoàn thiện giao diện và các chức năng chính
5	Từ ngày 30/05/2025 đến ngày 16/06/2025	<ul style="list-style-type: none"> - Viết báo cáo - Xem lại, chỉnh sửa và in quyền báo cáo - Nộp báo cáo 	<ul style="list-style-type: none"> - Hoàn thiện dự án, nộp quyền báo cáo.

GIẢNG VIÊN HƯỚNG DẪN

Trà Vinh, ngày tháng 4 năm 2025
SINH VIÊN THỰC HIỆN