

Pengembangan Aplikasi Web Lexin untuk Model Retrieval Augmented Generation Terkait Sumber Hukum di Indonesia = Lexin Web Application Development for Retrieval Augmented Generation Model on Legal Sources in Indonesia

Robertus Aditya Sukoco, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920574367&lokasi=lokal>

Abstrak

Pencarian dokumen hukum yang relevan dan efisien merupakan kebutuhan penting bagi praktisi hukum, akademisi, dan masyarakat umum di Indonesia. Penelitian ini memperkenalkan Lexin, sebuah aplikasi web berbasis model Retrieval Augmented Generation (RAG) untuk mengolah sumber hukum di Indonesia. Dengan memanfaatkan teknologi seperti Elasticsearch untuk pencarian teks penuh dengan pengurutan dan penyaringan beserta GPT-4o Mini dari OpenAI sebagai model bahasa besar, aplikasi ini memberikan solusi inovatif untuk meningkatkan akurasi, relevansi, dan efisiensi pencarian dokumen hukum. Pengembangan Lexin mencakup integrasi antarmuka berbasis Next.js untuk frontend, FastAPI untuk backend, dan Elasticsearch sebagai penyimpanan dokumen hukum. Evaluasi dilakukan melalui User Acceptance Testing (UAT), System Usability Scale (SUS), User Experience Questionnaire (UEQ), dan pengujian beban menggunakan Locust. Hasil evaluasi menunjukkan bahwa frontend Lexin berhasil menawarkan antarmuka pengguna yang intuitif dan responsif. Evaluasi UAT menunjukkan tingkat kepuasan pengguna mencapai 93,4%, dengan umpan balik positif terkait kemudahan navigasi dan kecepatan akses fitur. Pengukuran SUS menghasilkan skor 81, yang berada dalam kategori "Sangat Baik", mengindikasikan aplikasi mudah dipahami dan digunakan oleh berbagai jenis pengguna. Evaluasi UEQ menunjukkan nilai tinggi pada dimensi efisiensi (1,8) dan atraktivitas (1,7), memperlihatkan desain antarmuka yang ramah pengguna dan fungsional. Di sisi backend, aplikasi menunjukkan performa yang baik dengan kemampuan menangani hingga 304 permintaan dalam 60 detik menggunakan 4 instance backend API, dengan rata-rata waktu respons sebesar 2.418 ms, turun 55% dibandingkan tanpa autoscaling. Kemampuan Elasticsearch dalam pencarian dokumen menunjukkan waktu respons rata-rata sebesar 94 ms pada beban rendah dan 894 ms pada beban tinggi. Secara keseluruhan, aplikasi ini berhasil memberikan solusi yang efisien untuk pencarian dan pengelolaan dokumen hukum, memenuhi kebutuhan pengguna dari berbagai latar belakang. Lexin diharapkan dapat menjadi alat yang bermanfaat dalam meningkatkan aksesibilitas informasi hukum di Indonesia.

.....Efficient and relevant legal document retrieval is a critical need for legal practitioners, academics, and the general public in Indonesia. This research introduces Lexin, a web application based on the Retrieval Augmented Generation (RAG) model to process Indonesian legal resources. By leveraging technologies such as Elasticsearch for full-text search with ranking and filtering capabilities, and OpenAI's GPT-4o Mini as a large language model, the application provides an innovative solution to enhance the accuracy, relevance, and efficiency of legal document search. The development of Lexin integrates a Next.js-based frontend, FastAPI for the backend, and Elasticsearch as the repository for legal documents. Evaluation methods include User Acceptance Testing (UAT), the System Usability Scale (SUS), the User Experience Questionnaire (UEQ), and load testing using Locust. The evaluation results show that the Lexin frontend successfully offers an intuitive and responsive user interface. UAT evaluation results show a user

satisfaction rate of 93.4%, with positive feedback on navigation ease and feature access speed. The SUS measurement yields a score of 81, categorized as "Excellent," indicating that the application is easy to understand and use by various types of users. UEQ evaluation demonstrates high scores in the dimensions of efficiency (1.8) and attractiveness (1.7), reflecting a user-friendly and functional interface design. On the backend side, the application shows robust performance, handling up to 304 requests within 60 seconds using 4 backend API instances, with an average response time of 2,418 msa 55% reduction compared to scenarios without autoscaling. Elasticsearch demonstrates an average response time of 94 ms under low load and 894 ms under high load conditions. Overall, this application successfully provides an efficient solution for legal document retrieval and management, addressing the needs of users from various backgrounds. Lexin is expected to serve as a valuable tool for improving the accessibility of legal information in Indonesia.