

IMPLEMENTASI *RETRIEVAL-AUGMENTED GENERATION* UNTUK SISTEM TANYA JAWAB AL-QUR'AN BERBASIS *SIMILARITY SEARCH*

Muhammad Rezka Al Maghribi

ABSTRAK

Mahasiswa Muslim dan/atau alumni mahasiswa Muslim yang tergolong ke dalam kelompok *digital natives* kerap mencari penjelasan Al-Qur'an secara *digital*, namun masih menghadapi kesulitan memverifikasi otoritas sumber dan akurasinya, survei awal menunjukkan 50% responden mengalami masalah tersebut. Penelitian ini bertujuan merancang, membangun, dan mengevaluasi sistem tanya jawab Al-Qur'an berbasis *mobile* yang mampu menyediakan jawaban akurat, mudah dipahami, dan dapat ditelusuri ke sumber resmi. Penelitian menggunakan pendekatan rekayasa perangkat lunak dengan model *Waterfall* melalui tahap analisis kebutuhan, desain, implementasi, dan pengujian. Sistem dikembangkan dengan arsitektur *Retrieval-Augmented Generation (RAG)* yang mengintegrasikan Flutter, FastAPI, Qdrant, dan korpus resmi Kementerian Agama RI, serta menerapkan *semantic retrieval* berbasis *embedding* E5, *Strict Grounding*, *adaptive Similarity threshold*, *Negative rejection*, dan *fallback* LLM berbasis *cloud*. Evaluasi dilakukan melalui *Unit test backend* dan *frontend*, pengujian *black-box*, PSSUQ, dan validasi pakar. Hasil menunjukkan 12 *Unit test backend* dan 12 *Unit test frontend* lulus, pengujian *black-box* mencapai 100%, skor PSSUQ sebesar 1,53 termasuk *highly acceptable*, dan validasi pakar memperoleh rata-rata 3,8 dari 4. Sistem dinilai layak sebagai media bantu pembelajaran Al-Qur'an, meski masih terbatas pada korpus Kemenag RI dan belum mendukung *fallback* LLM lokal.

Kata Kunci: RAG, Sistem Tanya-Jawab, Al-Qur'an, pencarian semantik, halusinasi.

IMPLEMENTASI *RETRIEVAL-AUGMENTED GENERATION* UNTUK SISTEM TANYA JAWAB AL-QUR'AN BERBASIS *SIMILARITY SEARCH*

Muhammad Rezka Al Maghribi

ABSTRACT

Muslim university students and/or Muslim alumni who belong to the digital native group often seek digital explanations of the Qur'an, yet they still face difficulties in verifying source authority and accuracy, an initial survey showed that 50% of respondents experienced this problem. This study aims to design, develop, and evaluate a mobile-based Qur'anic question-answering system capable of providing accurate, understandable, and source-traceable answers. The research applies a software engineering approach using the Waterfall model through the stages of requirements analysis, design, implementation, and testing. The system is developed using a Retrieval-Augmented Generation (RAG) Architecture that integrates Flutter, FastAPI, Qdrant, and the official corpus of the Indonesian Ministry of Religious Affairs, while implementing E5-based semantic retrieval, Strict Grounding, an adaptive Similarity threshold, Negative rejection, and a cloud-based LLM fallback. Evaluation was conducted through backend and frontend Unit tests, Black-box Testing, PSSUQ, and expert validation. The results show that all 12 backend Unit tests and 12 frontend Unit tests passed, Black-box Testing reached 100%, the PSSUQ score was 1.53 (highly acceptable), and expert validation achieved an average of 3.8 out of 4. The system is considered feasible as a supporting medium for Qur'anic learning, although it remains limited to the Kemenag RI corpus and does not yet support a local LLM fallback.

Keywords: *RAG, Question Answering Systems, Al-Qur'an, Semantic Search, hallucination.*