

ABSTRACT

Keselamatan penerbangan sangat bergantung pada penyebaran informasi operasional yang tepat waktu dan akurat kepada *flight crews*, namun pembuatan *Notice to Airmen* (NOTAM) yang sesuai aturan dari *unstructured text* masih menjadi tantangan karena format yang ketat, kode khusus, dan kebutuhan konteks terbaru. Studi ini mengusulkan sebuah *framework* yang menggabungkan *large language models* (LLMs) dengan *Retrieval Augmented Generation* (RAG) untuk secara otomatis mengubah masukan bahasa alami Indonesia menjadi NOTAM yang terstruktur dan sesuai regulasi. Metode ini mengambil data kontekstual yang relevan seperti Q-codes, detail bandara, dan kondisi operasional yang tersimpan dalam *vector database*, kemudian mengintegrasikannya ke dalam *prompt* yang memandu LLM menghasilkan *output* yang akurat. *Framework* ini dievaluasi pada 179 pasangan *free-text* dan NOTAM menggunakan beberapa varian GPT, dengan metrik *accuracy*, NDCG, MRR, dan *recall*. Diharapkan bahwa pemberian konteks *domain-specific* yang presisi akan secara signifikan meningkatkan akurasi generasi dibandingkan LLM *baseline* tanpa *retrieval*. Hasil eksperimen mengonfirmasi harapan tersebut, menunjukkan peningkatan akurasi yang substansial pada semua model, seperti peningkatan akurasi Q-code dari 4,81% menjadi 19,27% dan akurasi deskripsi NOTAM dari 16,26% menjadi 60,24%, dengan model *Arctic-Embed-2-L* memberikan performa *retrieval* tertinggi (NDCG@5 sebesar 0,865, MRR@5 sebesar 0,851, Recall@5 sebesar 0,909). Temuan ini menunjukkan bahwa integrasi konteks yang terarah dapat melampaui sekadar peningkatan skala model, menawarkan solusi yang *scalable* dan andal untuk otomatisasi pembuatan NOTAM di lingkungan operasional penerbangan.

Keywords: Large Language Model, Natural Language Processing, NOTAM, Retrieval Augmented Generation.