

## ABSTRACT

Ensuring timely and accurate dissemination of operational information to flight crews is critical for aviation safety, yet generating rule-compliant Notices to Airmen (NOTAM) from unstructured text remains a challenge due to strict formatting, specialized codes, and the need for up-to-date context. This study proposes a framework that combines large language models (LLMs) with Retrieval Augmented Generation (RAG) to automatically convert free-form Indonesian natural language input into structured and regulation-compliant NOTAM. The method retrieves relevant contextual data such as Q-codes, airport details, and operational conditions stored in a vector database, and integrates them into prompts that guide the LLM in producing accurate outputs. The framework was evaluated on 179 free-text and NOTAM pairs using multiple GPT variants, with accuracy, NDCG, MRR, and recall as metrics. It was expected that providing precise domain-specific context would significantly improve generation accuracy compared to baseline LLMs without retrieval. Experimental results confirmed this expectation, showing substantial accuracy gains across all models, such as an improvement in Q-code accuracy from 4.81% to 19.27% and in NOTAM description accuracy from 16.26% to 60.24%, with the Arctic-Embed-2-L model delivering the highest retrieval performance (NDCG@5 of 0.865, MRR@5 of 0.851, Recall@5 of 0.909). These findings demonstrate that targeted context integration can outperform model scaling alone, offering a scalable and reliable solution for automating NOTAM generation in operational aviation settings.

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