

ABSTRACT

The development of digital technology has encouraged the utilization of chatbots as a practical solution to accelerate the delivery of information, especially in the context of academic services. This research focuses on developing a chatbot using the RASA framework equipped with the CRFEntityExtractor method to handle frequently asked questions (FAQs) from Telkom University students. The main focus of this research is evaluating the effect of entity annotation quality on chatbot performance, both in the information extraction process and in generating relevant responses. The model was trained using two types of datasets, namely complete and consistent annotations (clean annotation), and incomplete and inconsistent annotations (noisy annotation). Evaluation was conducted using precision, recall, F1-score, BLEU, and ROUGE metrics. Experimental results show that annotation quality has a significant impact on the performance of model. On the clean annotation dataset, the precision reaches 99.5%, the recall is 99.97%, and the F1-score is 99.72%. In addition, the BLEU value reached 100 and the ROUGE F1-score was 0.9, indicating a high relevance of the answer to the question context. Datasets with good annotations are able to improve entity extraction accuracy and produce more contextually appropriate answers. This research confirms the importance of consistency and thoroughness in the annotation process as the main foundation in the development of an NLP-based chatbot system for campus information services.

Keywords: *chatbot, RASA, CRF-NER, customer service, FAQ, Telkom University*