

Penerapan *Rule-Based Noun Phrase Chunking* terhadap Al-Qur'an

Andhika Buwananda Nugraha¹, Moch. Arif Bijaksana², Arif Fatchul Huda³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹andhikanug@students.telkomuniversity.ac.id, ²arifbijaksana@telkomuniversity.ac.id,

³afh@uinsgd.ac.id

Abstrak

Pemeluk agama Islam memiliki kitab ber Bahasa Arab yang diwajibkan untuk dibaca setiap harinya, yaitu Al-Qur'an. Bahasa Arab menimbulkan banyak tantangan untuk *Natural Language Processing (NLP)*, karena secara morfologis Bahasa Arab merupakan bahasa yang rumit. Bahasa Arab kurang memiliki sumber daya dan alat beranotasi. Pada penelitian kali ini akan dilakukan penerapan *rule-based noun phrase chunking* pada Al-Qur'an. Penerapan tersebut menghasilkan akurasi sebesar 45.5 persen, nilai *precision* sebesar 20.08 persen, *recall* sebesar 86.2 persen dan *f-measure* sebesar 32.5 persen.

Kata kunci : chunking, tokenizer, pos tagging, Al-Qur'an, natural language processing, rule-based

Abstract

Muslims have Arabic books which are required to be read every day, namely the Qur'an. Arabic poses many challenges for Natural Language Processing (NLP), because morphologically Arabic is a complex language. Arabic lacks annotated resources and tools. In this study, the application of rule-based noun phrase chunking on Al-Qur'an juz 30. The application produces an accuracy of 45.5 percent, a precision value of 20.08 percent, a recall of 86.2 percent and an f-measure of 32.5 percent.

Keywords: chunking, tokenizer, pos tagging, Al-Qur'an, natural language processing, rule0based
