

Pembangunan Monolingual Word Alignment Pada Terjemahan Al-Quran Berbahasa Indonesia

Galih Rizky Prabowo¹, Moch. Arif Bijaksana²

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

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

Abstract

For centuries the Qur'an is present in the midst of civilization and human society consisting of 6236 verses. To measure the semantic similarity between the translations of Al-Quran verses that aim to understand more deeply the meaning related to the verses of the Koran requires a method one of them with *monolingual word alignment*. *Monolingual alignment* is a word alignment method that identifies the similarity between words in existing pairs of sentences. In addition to the use of monolingual alignment methods in measuring the similarity of existing words, it also needs a dataset that serves as a collection of a document whose content is the semantic relationship between sets that exist. But the dataset *monolingual word alignment* for Indonesian in format MSR is still very limited in volume. In this research, some features in the method of *monolingual alignment* are applied in the development of the dataset *monolingual word alignment* Indonesian in format MSR, which is *textit align identical words*, *align PFA* and *align word sequences* by generating an F1 value of 86.94 %. For the best results F1 is generated from some combination of *alignment* features with *align identical words* and *align PFA* with F1 result of 88.83 %.

Keywords: Al-Quran, *Monolingual Alignment*, MSR

