

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,

Abstrak

Telah berabad-abad Al-Qur'an hadir di tengah-tengah peradaban dan pergaulan umat manusia yang terdiri dari 6236 ayat. Untuk mengukur kesamaan semantik antar terjemahan ayat Al-Quran yang bertujuan untuk memahami lebih dalam terhadap makna yang terkait pada ayat Al-Quran dibutuhkan suatu metode salah satunya dengan *monolingual word alignment*. *Monolingual alignment* adalah metode penyejajaran kata yang mengidentifikasi kesamaan antar kata dalam pasangan kalimat yang ada. Selain penggunaan metode *monolingual alignment* dalam pengukuran kesamaan kata yang ada, dibutuhkan pula sebuah *dataset* yang di *align* yang berfungsi sebagai kumpulan dari objek yang isinya adalah hubungan semantik antar sets yang ada. Namun *dataset monolingual word alignment* untuk bahasa Indonesia berformat MSR masih sangat terbatas volumenya. Dalam penelitian ini, diaplikasikan beberapa fitur dalam metode *monolingual alignment* dalam pembangunan *dataset monolingual word alignment* bahasa Indonesia berformat MSR yaitu *align identical words*, *align PFA* dan *align word sequences* dengan menghasilkan nilai F1 86.94 %. Untuk hasil F1 terbaik dihasilkan dari beberapa gabungan fitur *alignment* yaitu dengan fitur *align identical words* dan *align PFA* dengan hasil F1 sebesar 88.83 %.

Kata kunci : Al-Quran, *Monolingual Alignment*, MSR

