

Pembangunan Data Latih Menggunakan Monolingual Word Alignment Pada Persamaan Kisah Nabi Dalam Al-Quran dan Alkitab

Nadia Aruma Dita Suwandi¹, Moch. Arif Bijaksana²

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

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

Abstrak

Telah berabad-abad Al-Qur'an dan Alkitab hadir di tengah-tengah peradaban umat manusia sebagai panduan kehidupan penganutnya dan ternyata terdapat ayat-ayat yang memiliki makna yang sama / memiliki kesamaan kisah dan saling berkaitan di antara keduanya. Untuk melakukan penelitian pengukuran kesamaan semantik antar terjemahan Al-Quran dan Alkitab tersebut, dibutuhkan suatu data latih tentang persamaan kisah dalam Al-Quran dan Alkitab itu sendiri. Dalam melakukan pembangunan data latih tersebut dibutuhkan metode yang tepat untuk digunakan, salah satunya adalah dengan menggunakan metode *monolingual alignment*. *Monolingual alignment* adalah metode penyejajaran kata yang mengidentifikasi kesamaan antar kata dalam pasangan kalimat yang ada dalam satu bahasa yang sama. Dalam penelitian ini, digunakan persamaan kisah nabi yang terkandung dalam Al-Quran dan Alkitab sebagai dataset untuk pembangunan data latih (korpus) yang *di-align* dan berfungsi sebagai kumpulan dari objek . Selain itu diaplikasikan pula beberapa fitur dalam pembangunan data latih menggunakan *monolingual word alignment* pada persamaan kisah nabi AL-Quran dan Alkitab berbahasa Indonesia dengan format MSR ini, yaitu *align identical words*, *align PFA* dan *align word sequences* dan dilakukan evaluasi intrinsik sebagai tolak ukur performa dari data latih (korpus) tersebut dan menghasilkan nilai F1 64.89 %. Dan untuk hasil F1 terbaik dihasilkan dari beberapa gabungan fitur *alignment*, yaitu dengan gabungan fitur *align identical words* dan *align PFA* dengan hasil F1 sebesar 67.62 %.

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