

## **Abstrak**

**Bahasa Al-Quran berbeda dengan bahasa Arab sehari-hari, dalam kosa kata bahasa Al-Quran satu kata dapat diterjemahkan menjadi kata majemuk. Parafrase ialah pengungkapan kembali suatu kata atau kalimat dalam bentuk berbeda tanpa mengurangi makna, tafsir Al-Quran yang berbeda satu sama lain merupakan parafrase secara alami. Masukan berupa dataset dua terjemahan untuk disandingkan kemiripan parafrase yang ada. Pembangunan basis data parafrase untuk berbagai Bahasa sudah cukup banyak, namun untuk terjemahan Al-Quran penggunaan parafrase memiliki keunikan dan belum banyaknya basis data parafrase yang dibangun. Dengan memasukkan dataset yang ada maka system akan melakukan tokenisasi dengan menggunakan Regular Expression Filter dan menyatukan kedua tanzil yang akan dibandingkan dalam dataset menjadi sebuah basis data. Cara kerja dari Regular Expression Filter ialah mencoba memberi Pos-Tag pada tiap kata lalu disusun ulang hingga membentuk sebuah parafrase. Dengan demikian pendekatan Multi Word Expression menggunakan Regular Expression Filter cukup efisien dalam penentuan parafrase dalam terjemahan Al-Quran.**

**Kata kunci: Al-Quran, Basis data parafrase, tanzil, Qarai, Ahmed Ali**