

## Abstrak

Untuk memahami Quran dengan lebih baik, interpretasi yang benar membutuhkan lebih banyak perhatian pada ayat-ayat dalam Quran. Dari 114 surah Al-Quran, menurut al-Khatib al-Iskafi (mati 420 H./1026 M.), hanya 28 atau sekitar 25 tidak mengandung ayat yang sama atau berulang. Saat ini, untuk mendapatkan ayat berulang paling lama dalam Al-Quran membutuhkan upaya manual, dengan mencari ayat Al-Quran dari awal hingga akhir yang membutuhkan waktu lama. Ini adalah masalah Longest Common Substring (LCS) dalam Quran. Kami mengusulkan sistem menggunakan algoritma Ukkonen untuk membangun database set teks yang berulang. Algoritma Ukkonen dipilih karena metode penyelesaian masalah Longest Common Substring (LCS) terpanjang ini memiliki efisiensi yang lebih baik daripada metode sebelumnya. Kami menggunakan dua terjemahan bahasa Indonesia dan Arab. Kami menggunakan 789 ayat atau 5 surah dalam Quran. Kami membatasi kumpulan teks berulang terpanjang yang mengandung setidaknya tiga kata. Untuk Bahasa Indonesia kami menghasilkan 2341 set teks berulang, untuk bahasa Arab kami menghasilkan 2567 set teks berulang, yang kemudian disimpan dalam database. Setelah mengevaluasi sistem ini, untuk Bahasa Indonesia kami memperoleh precision 45,9 . Sedangkan untuk bahasa Arab, kami memperoleh precision 55,5 . Penelitian ini menghasilkan database baru yang sebelumnya tidak ada dalam makalah ilmiah. Basis data yang dihasilkan dapat digunakan untuk berbagai studi tentang Quran dan juga dapat digunakan untuk membandingkan dengan studi lain tentang Quran. Kode dan basis data yang dihasilkan tersedia di: <https://github.com/deviokta27/Al-Quran.git>.

**Kata Kunci:** Database, Longest Common Substring (LCS), Quran, Algoritma Ukkonen.