

ABSTRAK

Tujuan dari penelitian ini adalah mengatasi keterbatasan dalam ketersediaan situs web yang secara komprehensif menyajikan daftar rinci mengenai Isim Ma'rifat dalam Al-Quran dan Sunan Ibnu Majah. Kekurangan sumber daya yang lengkap telah menghambat kemampuan untuk mempelajari dan membandingkan Isim Ma'rifat di antara teks-teks Islam yang penting ini. Untuk mengatasi kendala tersebut, penelitian ini mengembangkan pendekatan berbasis pemrosesan bahasa alami dengan menggunakan program tokenizer Java yang terintegrasi dengan basis data MySQL yang memuat Sunan Ibnu Majah dan teks-teks Al-Quran. Program ini mengidentifikasi kehadiran awalan "alif lam" dan menghapus tanda diakritik untuk mempermudah perbandingan yang akurat antara ayat-ayat dalam kedua teks tersebut. Fokus utama penelitian ini adalah mengidentifikasi Isim Ma'rifat yang dimulai dengan "alif lam" yang hanya terdapat dalam Al-Quran (sebanyak 1183 kata), yang hanya terdapat dalam Sunan Ibnu Majah (sebanyak 2.894 kata), serta kesamaan antara keduanya (sebanyak 415 kata). Analisis ini menghasilkan pemahaman yang komprehensif mengenai perbedaan dan persamaan Isim Ma'rifat yang dimulai dengan "alif lam" antara Al-Quran dan Sunan Ibnu Majah. Hasil temuan ini memberikan masukan berharga bagi proyek Quranpedia, yang berkontribusi pada pengembangan sumber daya yang komprehensif dan mudah diakses mengenai Al-Quran dan Kutubus Sittah untuk studi Islam. Diharapkan bahwa penelitian ini akan meningkatkan pemahaman tentang Isim Ma'rifat dalam konteks agama dan bahasa, serta memberikan kontribusi penting dalam pengembangan pemrosesan bahasa alami, khususnya dalam bahasa Arab.