

Abstract

Alquran merupakan sumber hukum utama dalam ajaran agama islam bagi umat muslim di seluruh dunia. Alquran memiliki jumlah 30 juz, terbagi atas 144 surat dan tersusun dari 6236 ayat dan memiliki bahasan topik dalam alquran yang berbeda-beda dan memiliki banyak entitas nama yang banyak sehingga seringkali seseorang kesulitan untuk memahami al-quran.

Untuk memudahkan memahami al-quran kita bisa mengidentifikasi entitas-entitas penting dalam alquran seperti nama kaum dalam al-quran. Salah satunya dengan cara mengekstraksi informasi entitas penting dalam alquran dengan NER. NER secara otomatis mengenali entitas penting seperti Nama kaum pada al-quran.

Pada pembangunan sistem pengenalan entitas nama kaum dalam alquran dengan dataset terjemahan alquran bahasa indonesia ini menunjukkan, pengujian ini menghasilkan rata- rata akurasi menggunakan f-measure yang dihasilkan adalah sebesar 0.76 untuk penggunaan data latih sebanyak 36814 data keseluruhan atau f-measure 0.51 untuk percobaan menggunakan Chunking dengan data yang sama.

Data dan Source code kami tersedia di : <https://github.com/farhandzakaryarvianto/peopleERinQuran>

Keyword: *Quran, Named Entity Recognition, Conditional Random Field*