

## ABSTRAK

**Nama** : Mochamad Gilang Syarief

**NIM** : 1157010042

**Judul Skripsi** : Perbaikan Algoritma Stemmer Bahasa Arab yang Diterapkan  
pada Kosakata al-Qur'an

Struktur morfologi yang kompleks dari bahasa Arab mengakibatkan beberapa peneliti tertarik untuk meningkatkan pemrosesan kata-kata bahasa Arab. Salah satunya yaitu untuk mendapatkan hasil *stemmer* yang efisien. *Stemmer* adalah langkah yang digunakan untuk memproses data teks. Tugas utama dalam *stemmer* adalah mengembalikan kata bentukan ke kata dasar (*root* atau *stem*). Algoritma *stemmer* untuk bahasa Arab terbagi menjadi dua kategori yaitu *light stemmer* dan *heavy stemmer*. Dalam penelitian ini *light stemmer* yang digunakan yaitu: *ARLStem stemmer* dan *Snowball stemmer*, serta *heavy stemmer* yang digunakan yaitu: *Khoja stemmer*, *ISRI stemmer* dan *Al-Kabi stemmer*. Evaluasi hasil percobaan dilakukan pada masing-masing algoritma *stemmer* bahasa Arab dengan menggunakan data uji kata-kata dalam Al-Qur'an surat *Al-Fil* sampai dengan *An-Nas*. Masing-masing *stemmer* memiliki kekurangan dan kelebihan dalam perlakuannya pada masing-masing pola kata dalam bahasa Arab. Hasil uji coba menunjukkan *Khoja stemmer* lebih unggul dari keempat *stemmer* lain. Dari hasil perbaikan yang dilakukan pada *stemmer* bahasa Arab: *heavy stemmer*, menunjukkan peningkatan jumlah kata benar yang dikembalikan ke kata dasar.

**Kata Kunci:** *stemmer, light stemmer, heavy stemmer, ARLStem stemmer, Snowball stemmer, ISRI stemmer, Khoja Stemmer, AL-Kabi stemmer*

SUNAN GUNUNG DJATI  
BANDUNG

## ABSTRACT

**Name** : Mochamad Gilang Syarief  
**NIM** : 1157010042  
**Title** : *The Application of Fixed Arabic Stemmer Algorithm in Al-Quran Vocabulary*

*Arabic's complex morphological structure caused the researcher's interest to increase Arabic vocabulary. One of the goals is to get the efficient result of the stemmer. Stemmer is a stage that used to process data and text. Stemmer's main task is to restore the formed words into basic words (root or stem). Arabic's stemmer algorithm consists of two categories, light stemmer dan heavy stemmer. Light stemmers that used in this study are ARLStem stemmer and Snowball stemmer. And also the heavy stemmers that used in this study are Khoja stemmer, ISRI stemmer, and Al-Kabi stemmer. Evaluation of the results of experiments carried out on each Arabic stemmer algorithm using word test data in the Qur'an from Al-Fil to An-Nas. Each stemmer has its weakness and strength in treating each Arabic's word patterns. The experimental result showed that Khoja Stemmer is superior to the other four stemmers. From the fixed result that carried on Arabic Stemmer, Heavy Stemmer showed an enhancement in the number of correct words returned to the basic words.*

**Keywords:** *stemmer, light stemmer, heavy stemmer, ARLStem stemmer, Snowball stemmer, ISRI stemmer, Khoja Stemmer, AL-Kabi stemmer*

UNIVERSITAS ISLAM NEGERI  
SUNAN GUNUNG DJATI  
BANDUNG