

Erik Susilo, 2014. *Sistem Temu Kembali Informasi Nama Tanaman Ilmiah Menggunakan Algoritma Phonix dengan Peranan N-Gram Subtitution dan Code Shift*. Skripsi ini dibawah bimbingan Badrus Zaman, S.Kom., M.Cs. dan Indah Werdiningsih, S.Si., M.Kom. Program Studi S1 Sistem Informasi. Fakultas Sains dan Teknologi, Universitas Airlangga.

---

## ABSTRAK

Nama adalah salah satu faktor penting dalam sistem temu kembali informasi, karena digunakan sebagai kata kunci pada pencarian. Dalam melakukan pencarian, nama yang digunakan sebagai kunci terkadang terjadi kesalahan pengejaan. Sehingga nama yang dicari tidak dapat ditemukan dalam pencarian. Beberapa pendekatan telah dikembangkan untuk memecahkan masalah tersebut, salah satunya adalah dengan algoritma berdasarkan fonetik. Algoritma Phonix adalah algoritma yang diciptakan berdasarkan kesamaan fonetik. Tujuan dari penelitian ini adalah untuk mengukur kinerja N-Gram Subtitution dan Code Shift dalam upaya meningkatkan Recall dan Precision pada algoritma Phonix.

Sistem temu kembali informasi nama ilmiah tanaman ini dilakukan dalam tiga tahap. Tahap pertama adalah pengumpulan data dan informasi berupa nama ilmiah tanaman anggrek. Tahap kedua adalah analisis sistem yang meliputi, proses penyeragaman menggunakan metode N-Gram Subtitution, algoritma Phonix, proses Code Shift, serta proses pencocokan menggunakan Dice Coefficient. Tahap ketiga adalah perancangan dan pembangunan sistem yang meliputi uji coba sistem serta evaluasi sistem. Dalam uji coba dan evaluasi sistem, dibandingkan antara algoritma Phonix dengan menggunakan metode N-Gram Subtitution dan Code Shift, algoritma Fuzzy Soundex, dan algoritma Phonix tanpa metode N-Gram Subtitution dan Code Shift.

Berdasarkan uji coba, algoritma Phonix dengan menggunakan metode N-Gram Subtitution dan Code Shift menghasilkan nilai Recall sebesar 0.9, Precision sebesar 0.5, dan F-Measure sebesar 0.57 pada Threshold 70%. Hasil tersebut lebih baik dari algoritma Phonix yang tanpa menggunakan metode N-Gram Subtitution dan Code Shift maupun Fuzzy Soundex pada Threshold yang sama.

**Kata kunci :** *algoritma Phonix, N-Gram Subtitution, Code Shift, Sistem Temu Kembali Informasi*

Erik Susilo, 2014. *Information Retrieval System of Scientific Names of Plants using Phonix Algorithm with Role of N-Gram Substitution Method and Code Shift*. This skripsi was under guidance Badrus Zaman, S.Kom., M.Cs. and Indah Werdiningsih, S.Si., M.Kom. Bachelor Degree Information System Study Program. Faculty of Science and Technology, Airlangga University.

---

---

## ABSTRACT

The name is one of the important factors in information retrieval systems, because it is used as a keyword to search. In conducting a search, the name that used as a key sometimes occur spelling errors. So the name is sought cannot be found in the search. Several approaches have been developed to solve these problems, one of which is the algorithm based on phonetic. Phonix algorithm is an algorithm that is created based on phonetic similarity. The purpose of this study was to measure the performance of N-Gram Substitution and Code Shift in an effort to improve the Recall and Precision on Phonix algorithm.

Information retrieval scientific name of the plant was carried out in three stages. The first stage was the collection of data and information in the form of a scientific name of orchids. The second stage was the analysis of systems which include, homogenization process using N-Gram Substitution, Phonix algorithm, Code Shift, as well as the process of matching using the Dice Coefficient. The third stage was the design and construction of systems that includes system testing and evaluation system. In testing and evaluation of the system compared to the Phonix algorithm using the N-Gram Substitution and Code Shift, Fuzzy Soundex algorithm, and the Phonix algorithm without the N-Gram Substitution and Code Shift.

Based on trials, Phonix algorithm using the N-Gram Substitution and Code Shift produces Recall 0.9, Precision 0.5, and F-Measure 0.57 at the Threshold 70%. The results are better than the Phonix algorithm without using the N-Gram Substitution and Code Shift as well as Fuzzy Soundex at the same Threshold.

**Keyword :** *algorithm, N-Gram Substitution, Code Shift, Information Retrieval*