ABSTRAK

APLIKASI SISTEM TEMU KEMBALI INFORMASI MENGGUNAKAN MODEL RUANG VEKTOR BERBASIS WEB

Oleh

RHESA RULLIF ELIAN
10103170

Berbagai perkembangan dalam sistem temu kembali informasi, mesin pencari (*search engine*) telah dipacu oleh hasil-hasil penelitian dan pengembangan teknologi yang diarahkan untuk bahasa Inggris. Dalam tulisan ini saya melaporkan masalah-masalah yang saya temukan dalam proses perancangan *mesin pencari* untuk Bahasa Indonesia, serta hasil-hasil yang telah saya capai, termasuk dalam kontribusi orisinal saya adalah aturan gramatika (*grammar rule*) untuk pemenggalan imbuhan-imbuhan (*stemming*) kata-kata Bahasa Indonesia.

Pada tulisan ini akan dijelaskan teknik sistem temu kembali informasi dan rancangan integrasi sistem ke basis hiperteks. Salah satu metode atau teknik yang digunakan dalam sistem temu kembali informasi adalah teknik *vector space model.* Yaitu salah satu tehnik untuk mengukur suatu kemiripan *query* dengan suatu dokumen. Melalui tehnik ini diharapkan agar *recall* dari sistem temu kembali informasi tersebut dapat ditingkatkan.

Kinerja berbagai teknik temu kembali informasi di ukur dengan menampilkan dokumen yang terambil berikut bobot peringkatnya. Pada perangkat lunak yang dibuat *term-term* yang ada pada suatu vektor yang diukur dengan sinonimnya. Data sinonim ini diambil dari *database* pengindeksan.

Kata Kunci : sistem temu kembali informasi, *stemming* bahasa Indonesia, mesin pencari, model ruang vektor, pengindeksan

ABSTRACT

INFORMATION RETRIEVAL APPLICATION USING VECTOR SPACE MODEL WEB BASED

by

RHESA RULLIF ELIAN
10103170

Information retrieval tools and search engines have mainly been leveraging research results and technologies developed for the English language. In this paper we report the issues and obstacles we met in the process of designing and developing a search engine for the Indonesian language, as well as our progress and results. The results include original contributions such as a grammar for stemming Indonesian words.

This article will be explained various retrieval information system technique and device integrate the system to bases hiperteks. One of the most used method or technique for in an information retrieval system is the vector space model. It's one technique for measure resemblance query with document.Vector query is an improvement on vector space model that is aimed at improving the recall value of the search results.

Performance of various technique retrieval information system and various indexing measure by presenting document taken by following weight similarity. The terms of each vectors are measurement using their synonyms. The synonyms pairing is obtained from the indexing database.

Keywords *: information retrieval, stemming Indonesian language, search engine, vector space model, indexing*