

Tan Malaka bapak Republik Indonesia/karya Muhammad Yamin

Yamin, Muhammad

Deskripsi Dokumen: <http://lib.ui.ac.id/opac/themes/libri2/detail.jsp?id=4421&lokasi=12>

Abstrak

Jumlah dokumen yang semakin besar seiring dengan perkembangan informasi di Indonesia dan penggunaan istilah yang berbeda untuk konsep yang sama akan mempersulit pencarian dokumen berdasarkan istilah tertentu. Tesaurus diharapkan dapat membantu mengatasi kesulitan tersebut dengan memberikan sekelompok istilah sehingga semakin banyak dokumen relevan yang dapat ditemukan kembali. Penyusunan tesaurus perlu dilakukan secara otomatis, karena hal ini akan mempercepat proses pengelompokkan istilah, terutama jika jumlah istilah tersebut besar. Proses penyusunan tesaurus dalam penelitian ini dimulai dengan mengambil sejumlah istilah yang berasal dari beberapa dokumen. Proses berikutnya adalah menyusun Matriks Istilah-Dokumen dengan menghitung bobot istilah dalam setiap dokumen masukan. Matriks tersebut digunakan untuk menghitung kemiripan setiap pasangan istilah berdasarkan perhitungan dalam model ruang vektor (Vector Space Model), seperti Cosine atau Jaccard. Nilai kemiripan tersebut disimpan dalam matriks kemiripan (Similarity matrix) dan kemudian digunakan untuk mengelompokkan istilah-istilah secara hirarki dengan menggunakan teknik Kluster hubungan lengkap (Complete Link Cluster). Penelitian ini diharapkan dapat memberikan gambaran struktur hirarki istilah dalam dokumen-dokumen masukan dan sejauh mana metode yang digunakan dapat menyusun tesaurus secara otomatis.