

Intisari

Perkembangan teknologi informasi disertai dengan pertumbuhan jumlah pengguna internet. Hal tersebut akan membuat aliran data baik data terstruktur maupun tak terstruktur akan meningkat drastis. Proses pencarian dokumen akan menjadi masalah serius apabila tidak diiringi dengan sistem pencarian dokumen yang handal. Namun, seringkali hasil pencarian dokumen ditampilkan dalam bentuk daftar panjang (*list view*) yang menyulitkan pengguna untuk mendapatkan informasi secara cepat. Visualisasi *clustering* hasil pencarian menjadi solusi untuk mengatasi masalah tersebut.

Penelitian ini mengimplementasikan sistem untuk menampilkan hasil pengelompokan dokumen dengan visualisasi berbasis hirarki. Apache Solr dan digunakan sebagai *platform* dalam pencarian dan Carrot² digunakan untuk *clustering* hasil pencarian tersebut. *Clustering* dokumen hasil pencarian menggunakan algoritme *bisecting k-means*. Kemudian, hasil pengelompokan dokumen divisualisasikan menggunakan Carrot² *visualization* yang terdiri dari visualisasi *foamtree* dan *circles*.

Hasil dari penelitian ini adalah implementasi sistem yang dapat menampilkan hasil *clustering* dokumen menggunakan visualisasi *foamtree* dan *circles*. Setiap kelompok akan memiliki label yang mewakili dari setiap anggota kelompok yang berdasarkan hasil pengelompokan menggunakan algoritme *bisecting k-means*. Evaluasi dari hasil penelitian ini mendapatkan nilai presisi hasil pencarian dan kualitas label mencapai 0.94 dan 0.8 dari skala 0 sampai 1 serta adanya efisiensi, kebaruan dan stimulan yang tinggi berdasarkan evaluasi pengalaman pengguna menggunakan metode UEQ.

Kata kunci : *Clustering, Bisecting k-means, Apache Solr, Carrot², Visualisasi.*

Abstract

The development of information technology is accompanied with the increasing number of internet users. This increasing number of users will make the data flow, either structured or unstructured data, increases significantly. The document search process will be a serious problem if it is not accompanied with a reliable document search system. However, the document search result is often displayed in the form of list view which can be inconvenient for users who want to quickly obtain informations. Search result clustering visualization is the solution to solve this problem.

This research implemented a system to visualize document clustering result with a hierarchy-based visualization. Apache Solr is used as a search platform, while Carrot² is used for the search result clustering. The document search result clustering process uses bisecting k-means algorithm. The document clustering result is then visualized using Carrot² visualization which consists of foamtree and circle visualization.

The result of this research is the system implementation which can visualize document clustering result using foamtree and circle visualization. Each cluster will have a label that represents each cluster's documents according to the clustering result using bisecting k-means algorithm. As an evaluation of this research result, the precision value of the search result and the label quality was up to 0.94 and 0.8 on a scale of 0 to 1. Based on the users experience evaluation using UEQ method, there were also efficiency, novelty, and high stimulant in this research result.

Keywords : *Clustering, Bisecting k-means, Apache Solr, Carrot², Visualization*