

ABSTRACT

ANALYSIS OF DOCUMENT CLUSTERING FOR INFORMATION TECHNOLOGY AUDIT EVIDENCE

Zata Atika

14/368545/PA/16293

Nowadays, the use of technology continuous to grow rapidly as Science and Information Technology develop in a sophisticated way. Practical use of technology is already very broad, as a result surveillance on the technology usage is required. Therefore, it needs a supervision during the process. The use of technical for enhancing services in company to help human is real. That's why there is an Information Technology Audit (IT Audit) and Control that plays an important role in one company. Audit process is required to assess whether an existing system is able to help the company to achieve the goals or not in an effective and efficient way. Currently, the documents that generated from an audit process amount to a very large number. This research aims to implement an Information Retrieval and document clustering that is able to grouping the documents based on the search results. The clustering algorithm that used is Label Induction Group (LINGO), Suffix Tree Clustering (STC) and Bisecting K-Means. It takes advantages of Apache Solr to do the indexing process of all document. The result of the research showed that the cluster label quality for search result clustering evaluation is 83.5% by using LINGO, 74.6% by using STC and 85% by using Bisecting K-Means algorithm. Based on this result, the system is able to grouping the document and resulting in relevant results from the query that entered by user and the cluster label quality is useful.

Keywords: information retrieval, text clustering, lingo algorithm, stc, bisecting k-means, it audit evidence



INTISARI

ANALISIS PENGELOMPOKAN DOKUMEN BUKTI-BUKTI IT AUDIT

Zata Atika

14/368545/PA/16293

Dewasa ini penggunaan teknologi terus berkembang pesat seiring berjalannya Ilmu Pengetahuan dan Teknologi informasi yang terjadi di dunia. Penggunaan teknologi juga sudah sangat luas pada praktiknya, sehingga dibutuhkan sebuah pengawasan dalam penggunaan teknologi tersebut. Pemanfaatan IT dalam meningkatkan layanan dan membantu proses kerja sebuah perusahaan sangatlah nyata. IT Audit memegang peranan penting dalam sebuah perusahaan. Audit dibutuhkan untuk menilai apakah suatu sistem yang ada mampu membantu perusahaan untuk mencapai tujuannya secara efektif dan efisien. Hal ini melibatkan banyak dokumen yang dibutuhkan dan dihasilkan selama proses audit berlangsung. Penelitian ini mengimplementasikan metode Information Retrieval dan Document Clustering yang dapat melakukan pengelompokan dokumen hasil pencarian. Algoritma yang digunakan adalah Label Induction Group (LINGO), STC dan Bisecting K-Means dengan memanfaatkan Apache Solr untuk melakukan proses indexing dokumen. Hasil penelitian menunjukkan bahwa nilai kualitas uji label klaster yang dihasilkan oleh LINGO adalah sebesar 83.5%, STC sebesar 74.6% dan Bisecting K-Means 85% untuk evaluasi hasil pencarian dan uji kualitas dari label yang dihasilkan. Berdasarkan hasil penelitian tersebut, sistem mampu memberikan hasil yang relevan untuk setiap query yang dimasukan oleh pengguna dan mampu menghasilkan kualitas label yang bermanfaat.

Kata Kunci: information retrieval, text clustering, lingo algorithm, stc, bisecting k-means, it audit evidence