

ABSTRAKSI

SISTEM PEMROLEHAN INFORMASI PUISI DENGAN MENGGUNAKAN *QUERY EXPANSION* BERDASARKAN *THESAURUS*

Yovita Metty Nurcahyani

NIM: 105314112

Universitas Sanata Dharma

2015

Proses pencarian dokumen, termasuk untuk pencarian dokumen puisi seperti di internet pada umumnya dirasa belum cukup membantu menemukan dokumen puisi seperti yang diharapkan. Seringkali dalam melakukan pencarian puisi, puisi-puisi yang ditampilkan hanya sesuai dengan kata kunci yang dimasukkan, namun sebenarnya kita dapat memperoleh puisi-puisi yang mengandung thesaurus.

Penulis tertarik untuk meneliti tingkat relevansi data yang diperoleh dari hasil pencarian dokumen puisi, menggunakan metode query expansion berdasarkan pada thesaurus yaitu dalam hal ini thesaurus umum dan thesaurus konteks puisi.

Beberapa tahap metode dilakukan dalam pengembangan sistem ini, tahap-tahap tersebut adalah melakukan pendalaman pustaka dari dokumen-dokumen puisi, mengumpulkan 101 dokumen puisi berekstensi .txt sebagai koleksi dokumen. Implementasi penerapan sistem pemerolehan informasi menggunakan struktur data inverted index dan pembobotan tf-idf menurut Savoy dan menggunakan Oracle sebagai media penyimpanan data. Pengujian dan evaluasi menggunakan recall and precision yang dibantu oleh lima orang responden.

Hasil pengujian menunjukkan bahwa sistem pencarian yang menggunakan Query Expansion mengalami perbaikan, baik untuk sistem yang menggunakan thesaurus umum

PLAGIAT MERUPAKAN TINDAKAN TIDAK TERPUJI

maupun thesaurus konteks puisi. Ditemukan juga bahwa sistem yang menggunakan *thesaurus umum* dapat memunculkan lebih banyak dokumen daripada sistem pencarian biasa, namun juga memunculkan banyak dokumen yang tidak relevan.



ABSTRACTION

**POETRY INFORMATION RETRIEVAL SYSTEM BY USING QUERY EXPANSION
BASED THESAURUS**

Yovita Metty Nurcahyani

NIM: 105314112

Universitas Sanata Dharma

2015

Search documents process, including poetry documents such in the internet in general are still not enough to help find poetry documents as expected. Often in searching poems, poems that appear only in accordance with the keywords entered, but in fact we can obtain the poems that contain thesaurus.

Authors interested in examining the level of relevance of the data obtained from the search results poetry document, using the method of query expansion based on a thesaurus that in this case the general thesaurus and thesaurus poetry context.

Several stages in the development of methods made this system, these stages are deepening library of documents poetry, collected 101 documents poetry as a .txt extension document collection. Implementation of the application of information retrieval system using inverted index data structure and tf-idf weighting according to the Savoy and use Oracle as data storage media. Testing and evaluation using recall and precision are assisted by five respondents.

The results show that the retrieval system using Query Expansion improved, both for systems that use general thesaurus and thesaurus poetry context. It was also found that the system uses a general thesaurus can bring more documents than the regular search system, but it also raises a lot of irrelevant documents.