

## ABSTRAK

Pencarian dokumen di dalam sebuah komputer sering menjadi hal yang sulit dilakukan. Pencari hanya dihadapkan dengan metode pencarian yang hanya mencocokkan *query* pencarian pengguna dengan nama dari file-file yang dicari. Hal ini menyebabkan pencarian menjadi sulit dilakukan jika pengguna tidak mengetahui nama file yang dicari. Sebaliknya pengguna biasanya mengetahui isi dari dokumen apa yang ingin dicari. Tujuan dari pengembangan aplikasi ini adalah membantu memudahkan pencarian dokumen dengan memperhitungkan isi dari suatu dokumen dan juga dapat menyajikan hasil pencarian secara cepat. Aplikasi ini menerapkan metode temu-balik informasi untuk membangun sebuah mesin pencarian di mana untuk setiap dokumen akan dilakukan proses *indexing* terlebih dahulu dan hasil *indexing* akan disimpan di dalam suatu basis data. Hasil penelitian menunjukkan bahwa penggunaan metode temu-balik informasi dapat membantu dalam pencarian dokumen dengan mudah dan cepat.

Kata Kunci: Temu-balik informasi, *index*, *inverted index*, mesin pencarian, *desktop search*, Linux, *daemon*

## ABSTRACT

*Usually, searching for a document in a personal computer is a difficult task to do. User is faced with a searching method where application only match use query with file name. Because of this, searching become difficult to do if user don't know the name of the file that they were looking for. In the other hand, user usually know the content or topic of the file that they were looking for. The purpose of developing this application is to help user to search for a document by taking into account the document's content and produce result quickly. This application implements information retrieval method to build a search engine where each document will be indexed and the index will be stored in a database. Research result showed that using information retrieval can help searching for a document easily and quickly.*

*Keywords: Information retrieval, index, inverted index, search engine, desktop search, Linux, daemon*

## DAFTAR ISI

|   |     |
|---|-----|
| LEMBAR PENGESAHAN.....                                | i   |
| PERNYATAAN ORISINALITAS LAPORAN PENELITIAN.....       | ii  |
| PERNYATAAN PUBLIKASI LAPORAN PENELITIAN.....          | iii |
| PRAKATA.....  | iv  |
| ABSTRAK.....  | v   |
| ABSTRACT.....   | vi  |
| DAFTAR ISI.....                                       | vii |
| DAFTAR GAMBAR.....                                    | x   |
| BAB I PENDAHULUAN.....                                | 1   |
| 1.1 Latar Belakang.....                               | 1   |
| 1.2 Rumusan Masalah.....                              | 1   |
| 1.3 Tujuan Pembahasan.....                            | 2   |
| 1.4 Ruang Lingkup Kajian.....                         | 2   |
| 1.5 Sistematika Pembahasan.....                       | 2   |
| BAB II KAJIAN TEORI.....                              | 4   |
| 2.1 <i>Software Testing</i> .....                     | 4   |
| 2.1.1 <i>Black-Box Testing</i> .....                  | 4   |
| 2.1.2 <i>White-Box Testing dan Unit Testing</i> ..... | 4   |
| 2.2 <i>Unified Modeling Language (UML)</i> .....      | 5   |
| 2.2.1 <i>Use Case Diagram</i> .....                   | 5   |
| 2.2.2 <i>Class Diagram</i> .....                      | 6   |
| 2.2.3 <i>Activity Diagram</i> .....                   | 8   |
| 2.3 <i>Information Retrieval</i> .....                | 9   |
| 2.4 <i>Desktop Search</i> .....                       | 10  |
| 2.4.1 <i>Text Transformation</i> .....                | 12  |
| 2.4.1.1 <i>Tokenizing</i> .....                       | 13  |
| 2.4.1.2 <i>Stopping</i> .....                         | 13  |
| 2.4.1.3 <i>Stemming</i> .....                         | 13  |
| 2.4.2 <i>Ranking</i> .....                            | 14  |

|   |    |
|---|----|
| 2.4.2.1 <i>Vector Space Retrieval Model</i> .....       | 15 |
| 2.5 <i>Library</i> Pendukung.....                       | 17 |
| 2.5.1 MapDB.....  | 17 |
| 2.5.2 PDFBox.....                                       | 17 |
| 2.5.3 Apache POI.....                                   | 17 |
| 2.5.4 Apache ODF Toolkit.....                           | 18 |
| 2.6 Linux dan Daemon.....                               | 18 |
| BAB III ANALISIS DAN DESAIN.....                        | 19 |
| 3.1 Analisis.....                                       | 19 |
| 3.1.1 <i>Library</i> Pendukung.....                     | 20 |
| 3.1.1.1 MapDB.....                                      | 20 |
| 3.1.1.2 PDFBox.....                                     | 20 |
| 3.1.1.3 Apache POI.....                                 | 20 |
| 3.1.1.4 Apache ODF Toolkit.....                         | 21 |
| 3.1.2 Proses <i>Tokenizing</i> .....                    | 21 |
| 3.1.3 Proses Perhitungan Manual.....                    | 23 |
| 3.2 Gambaran Keseluruhan.....                           | 27 |
| 3.2.1 Persyaratan Antarmuka Eksternal.....              | 27 |
| 3.2.2 Antarmuka Dengan Pengguna.....                    | 27 |
| 3.2.3 Antarmuka Perangkat Keras.....                    | 27 |
| 3.2.4 Antarmuka Perangkat Lunak.....                    | 28 |
| 3.2.5 Fitur-Fitur Perangkat Lunak.....                  | 28 |
| 3.2.5.1 Fitur <i>Crawling</i> Dan <i>Indexing</i> ..... | 28 |
| 3.2.5.2 Fitur Pencarian.....                            | 29 |
| 3.2.5.3 Fitur <i>Include</i> Direktori.....             | 30 |
| 3.2.5.4 Fitur <i>Exclude</i> Direktori.....             | 30 |
| 3.3 Desain Perangkat Lunak.....                         | 31 |
| 3.3.1 Pemodelan Perangkat Lunak.....                    | 31 |
| 3.3.1.1 <i>Use Case Diagram</i> .....                   | 32 |
| 3.3.1.2 <i>Activity Diagram</i> .....                   | 35 |
| 3.3.1.3 <i>Class Diagram</i> .....                      | 39 |

|   |    |
|---|----|
| 3.3.2 Rancangan Antarmuka.....                                    | 40 |
| 3.3.2.1 Form Pencarian.....                                       | 40 |
| 3.3.2.2 Form Konfigurasi.....                                     | 41 |
| BAB IV PENGEMBANGAN PERANGKAT LUNAK.....                          | 43 |
| 4.1 Form Pencarian.....   | 43 |
| 4.2 Form Konfigurasi.....   | 44 |
| 4.3 Potongan Kode Penting.....                                    | 45 |
| 4.3.1 Proses <i>Tokenization</i> .....                            | 45 |
| 4.3.2 Proses Perhitungan Kecocokkan <i>Query</i> dan Dokumen..... | 50 |
| BAB V TESTING DAN EVALUASI SISTEM.....                            | 54 |
| 5.1 Rencana Pengujian.....  | 54 |
| 5.2 Pelaksanaan Pengujian.....                                    | 54 |
| 5.2.1 Fitur Pencarian.....  | 54 |
| 5.2.2 Fitur <i>Indexing</i> .....                                 | 54 |
| 5.2.3 Fitur <i>Include</i> Directory.....                         | 55 |
| 5.2.4 Fitur <i>Exclude</i> Directory.....                         | 56 |
| 5.3 Perhitungan Kecepatan Pencarian.....                          | 56 |
| BAB VI KESIMPULAN DAN SARAN.....                                  | 59 |
| 6.1 Kesimpulan.....   | 59 |
| 6.2 Saran.....  | 59 |
| DAFTAR PUSTAKA.....   | 61 |

## DAFTAR GAMBAR

|  |    |
|--|----|
| Gambar 2.1 Contoh <i>use case diagram</i> .....                      | 6  |
| Gambar 2.2 Contoh <i>class diagram</i> .....                         | 7  |
| Gambar 2.3 Contoh <i>activity diagram</i> .....                      | 9  |
| Gambar 2.4 Proses <i>indexing</i> .....                              | 11 |
| Gambar 2.5 Proses <i>query</i> .....                                 | 12 |
| Gambar 2.6 Contoh dokumen beserta bentuk vektornya.....              | 15 |
| Gambar 3.1 <i>Use case diagram</i> .....                             | 30 |
| Gambar 3.2 <i>Activity diagram</i> Pencarian.....                    | 33 |
| Gambar 3.3 <i>Activity diagram Include</i> Direktori.....            | 34 |
| Gambar 3.4 <i>Activity diagram Exclude</i> Direktori.....            | 35 |
| Gambar 3.5 <i>Activity diagram Start/Stop Indexing</i> .....         | 36 |
| Gambar 3.6 <i>Class diagram tokenizer</i> .....                      | 37 |
| Gambar 3.7 <i>Class diagram</i> dokumen.....                         | 37 |
| Gambar 3.8 <i>Class diagram indexer</i> .....                        | 38 |
| Gambar 3.9 Desain Antarmuka <i>Form</i> Pencarian.....               | 39 |
| Gambar 3.10 Desain Antarmuka <i>Form</i> Konfigurasi.....            | 40 |
| Gambar 4.1 Tampilan pencarian yang masih kosong.....                 | 41 |
| Gambar 4.2 Tampilan pencarian saat pengguna melakukan pencarian..... | 42 |
| Gambar 4.3 Tampilan konfigurasi aplikasi.....                        | 43 |