DAFTAR ISI

LEMBAR JUDUL

LEMBAR PENGESAHAN

ABSTRAK i

ABSTRACT ii

KATA PENGANTAR iii

DAFTAR ISI vi

DAFTAR TABEL xiii

DAFTAR GAMBAR xvi

DAFTAR SIMBOL xix

DAFTAR LAMPIRAN xx

BAB I PENDAHULUAN 1

1.1 Latar Belakang Masalah 1

1.2 Identifikasi Masalah 2

1.3 Maksud dan Tujuan 2

1.4 Batasan Masalah 3

1.5 Metodologi Penelitian 4

1.6 Sistematika Penulisan 5

BAB II LANDASAN TEORI 7

Konsep Dasar Sistem 7

2.1.1 Karakteristik Sistem 7

2.1.2 Klasifikasi Sistem 9

2.2 Konsep Dasar Informasi 9

2.2.1 Siklus Informasi 10

2.2.2 Kualitas Informasi 11

2.3 Konsep Dasar Sistem Informasi 13

2.4 Sistem Temu-kembali Informasi 15

2.5 Sistem Temu-kembali Informasi Berbasis Hiperteks 15

2.6 Tokenizer 20

2.6.1 Tokenization 20

2.6.2 Parsing 21

2.6.3 Parser 21

2.6.3.1 Top-down Parsing 22

2.6.3.2 Bottom-Up Parsing 23

2.6.3.3 Gabungan Top-Down dan Bottom-Up Parsing 25

2.6.4 Grammer 26

2.6.4.1 Grammer Regular 27

2.6.4.2 Grammer Bebas Konteks 28

2.6.4.3 Lexical Analyzer 28

2.7 Stoplist 29

2.8 Stemming 29

2.9 Pembentukan Kata-kata Bahasa Indonesia 31

2.9.1 Definisi Istilah 32

2.9.2 Imbuhan (Afiks) dalam Bahasa Indonesia 33

2.9.2.1 Awalan (Prefiks) 33

2.9.2.2 Sisipan (Infiks) 34

2.9.2.3 Akhiran (Sufiks) 34

2.9.2.4 Imbuhan Terbelah (Konfiks) 35

2.9.2.5 Imbuhan Gabungan (Simulfiks) 35

2.10 *Term Weighting* (Pembobotan Istilah) 35

2.10.1 *Vektor Space Model* 35

2.11 Pengindeksan 39

2.12 Sejarah *Hypertext* 39

2.13 Pengertian *Hypertext* 43

2.14 *Search Engine* ……. 46
2.14.1 Klasifikasi Web Search Service ……... 47
2.14.1.1 *Search Engine* ……..... 48
2.14.1.2 *Directory* ……..... 48
2.14.1.3 *Library* ……..... 49
2.14.2 Mengevaluasi Aplikasi Search Engine ……... 50
2.14.2.1 Fitur-fitur Umum Search Engine Internet …….. 50
2.14.3 Anatomi Search Engine …….. 51
2.14.3.1 Crawler ……... 51
2.14.3.2 Spider ……..... 52
2.14.3.3 Indexer ……..... 52
2.14.3.4 Results Engine …… 53
2.15 Analisis dan Perancangan Sistem ……..... 53
2.16 Sejarah *Internet* …… 54
2.16.1 Fasilitas *Internet* ……. 55
2.16.2 Syarat Tergabung ke *Internet* …..…….. 57
2.17 Web Server ……... 60
2.18 Alat Bantu Pemodelan ……..... 62
2.18.1 Flowchart …… 62
2.18.2 Diagram Konteks ……..... 62
2.18.3 Data Flow Diagram (DFD) ……..... 62
2.18.4 Kamus Data …… 62
2.18.4.1 Basis Data ……... 63
2.19 HTML ……..…….. 64
2.20 CSS …….. 66
2.21 XAMPP …… 67
2.22 PHP ……..…. 69
2.23 MySQL 71

2.24 Dreamweaver ……..... 72

BAB III ANALISIS DAN PERANCANGAN …….....… 74

3.1 Analisis Masalah …… 74

3.1 Analisis Sistem *Search Engine* …….. 75

3.2 Teknik Temu Kembali Informasi …….. 76

3.4 Tahap Memasukan File Hyperteks yang akan Diproses

untuk Masukan Pengindeksan …… ... 77

3.4.1 Tokenizer ……. 77

3.4.1.1 Tokenization ……. 77

3.4.2. Parsing …… 78

3.4.2.1 Parser ……. 78

3.4.3 Stoplist Bahasa Indonesia …….....… 80

3.5 Stemming Bahasa Indonesia ……. 80

3.5.1 Algoritma Stemming Bahasa Indonesia

Mirna Adriani dan Bobby Nazief 80

3.5.2 Algoritma ini Dijabarkan untuk setiap

kata yang akan di Stemming 82

3.5.3 Peluruhan Awalan (prefiks) 85

3.6 Algoritma Model Ruang Vektor (*Vector Space Model*) …… .....… 88

3.6.1 Pembobotan Istilah (*Term Weighting*) dan Pengindeksan ..…..... 88

3.6.1.1 Pengindeksan ……. 89

3.6.2 Proses Pembobotan Kueri (*Query Term Weighting*)

dan Ukuran Kesamaan (*Similarity Measurement*) …… ..…… 92

3.6.2.1 Pembobotan Kueri (*Query Term Weighting*) ……. 92

3.6.2.1.1 Pembalikan *File* (*Inverted File*) …….. 92

3.6.2.2 Ukuran Kesamaan (*Similarity Measurement*) ……. 92

3.7 Analisis Kebutuhan Non Fungsional ……..... 96

3.7.1 Analisis Kebutuhan Perangkat Keras ……..… 96

3.7.2 Analisis Kebutuhan Perangkat Lunak …… 96
3.7.3 Analisis Kebutuhan User …… 97
3.8 Perancangan Fungsional ……..... 97
3.9 *Flowchart* Mesin Pencari (*Search Engine*) …… 98
3.10 Basis Data ……..... 98
3.11 Struktur Tabel ……..... 102
3.12 Perancangan Antar Muka …… 103
3.12.1 Perancangan Form …… .....… 103
BAB IV IMPLEMENTASI DAN PENGUJIAN SISTEM ……..... 109
4.1 Implementasi …… 109
4.1.1 Implementasi Perangkat Keras …… 109
4.1.2 Implementasi Perangkat Lunak 110
4.1.3 Implementasi Antar Muka Dan Petunjuk Penggunannya 111
4.2 Pengujian Sistem …….. 118
4.2.1 Rencana Pengujian Alpha ……..... 118
4.2.1.1 Pengujian baca *file* hiperteks dengan metode *Black Box* … .....…… 119
4.2.1.2 Pengujian proses *Tokenizer* dengan metode *Black Box* 119
4.2.1.3 Pengujian pehilangan kata sama dan
frekuensi kata dengan metode *Black Box* 120
4.2.1.4 Pengujian proses *Stoplist* dengan metode *Black Box* 120
4.2.1.5 Pengujian proses *Stemming* dengan metode *Black Box* 121
4.2.1.6 Pengujian proses pengindeksan dengan metode *Black Box* 121
4.2.1.7 Pengujian *Input Keyword* Dengan Metode *Black Box* 121
4.2.1.8 Pengujian Cari Dengan VSM Dengan Metode *Black Box* 122
4.2.1.9 Pengujian Dokumen Hyperteks Dengan Metode *Black Box* 122
4.2.2 Pengujian Kadar Relevansi Dokumen …… 123
4.2.2.1 Pengujian Ini Dilakukan Dengan Tidak Terdapatnya,

Sampel Dokumen Dengan Tidak Mengunakan
Kata Kunci Dalam Pencarian ……. 123
4.2.2.2 Pengujian ini dilakukan dengan tidak terdapatnya, sampel
dokumen dengan mengunakan kata kunci dalam pencarian ..……. 124
4.2.2.3 Pengujian Ini Dilakukan Dengan Terdapatnya,
Satu Sampel Dokumen Kosong Dengan Tidak
Mengunakan Kata Kunci …… 125
4.2.2.4 Pengujian Ini Dilakukan Dengan Terdapatnya,
Satu Sampel Dokumen Kosong Dengan Mengunakan
Kata Kunci Lebih Dari Satu ……..… 125
4.2.2.5 Pengujian Ini Dilakukan Dengan Terdapatnya,
Satu Sampel Dokumen Yang Terdapat Satu Istilah
“tolong” Di Dalamnya. Dengan Mengunakan
Kata Kunci “tolong” ……. .....……... 126
4.2.2.6 Pengujian Ini Dilakukan Dengan Terdapatnya,
Satu Sampel Dokumen Yang Terdapat Tiga Istilah
“tolong” Di Dalamnya. Dengan Mengunakan
Kata Kunci “tolong” ……. .....…... 127
4.2.2.7 Pengujian Ini Dilakukan Dengan Terdapatnya,
Sebelas Sampel Dokumen Yang Terdapat
Satu Istilah “tolong” Setiap Dokumen Di Dalamnya.
Dengan Mengunakan Kata Kunci “tolong” ……. .....…. 128
4.2.2.8 Pengujian Ini Dilakukan Dengan Terdapatnya, Sebelas
Sampel Dokumen Dengan Istilah “tolong, menolong,
tertolong, ditolong, tertolong, pertolongan “
Yang Secara Acak Penempatannya Dan Jumlahnya.
Dengan Mengunakan Kata Kunci “tolong” …….....……. 129
4.2.2.9 Pengujian Ini Dilakukan Dengan Terdapatnya,

Sebelas Sampel Dokumen Dengan Banyak Istilah
Dengan Salah Satu Dokumen Hanya Mengandung Kata
"tolong" Dengan Kata Kunci "tolong" ……..... 130
4.2.2.10 Pengujian Ini Dilakukan Dengan Terdapatnya,
Sebelas Sampel Dokumen Dengan Banyak Istilah
Dengan Kata Kunci Yang Berbeda Lebih Dari Satu …… 131
4.2.2.11 Pengujian Ini Dilakukan Dengan Terdapatnya,
Sebelas Sampel Dokumen Dengan Salah Satu
Dokumen Terdapat Istilah Asing "information"
Dengan Kata Kunci "information" ……..…….. 132
4.2.2.12 Pengujian Ini Dilakukan Dengan Terdapatnya,
Sembilan Sampel Dokumen Dengan Salah Satu
Dokumen Terdapat Istilah Asing "information"
Dengan Kata Kunci "informations" ……... 134
4.2.2.13 Pengujian Ini Dilakukan Dengan Terdapatnya,
Sembilan Sampel Dokumen Dengan Kata Kunci
Menggunakan Kata-Kata Buang …… .....…… 135
4.2.2.14 Hasil Dari Semua Pengujian Kadar Relevansi ….....…… 135
4.2.3 Pengujian Betha …….. 137
BAB V KESIMPULAN DAN SARAN ……..... 141
5.1 Kesimpulan …… .....……. 141
5.2 Saran …… .....…… 141
DAFTAR PUSTAKA 142