

ABSTRACT

Java Search Tool is a search engine application built on Java programming language. Java Search Tool give simple approach for Inverted File Indexing and Vector Space Model algorithms for information retrieval core engine. Beside the fact that Java Search Tool is dekstop search application, Java Search Tool also provide native integration with host operating system indexing service and also provide search feature through GoogleBase services. Native integration only available for UNIX based operating system including Mac OS X. While for GoogleBase services search, Java Search Tool only provide GoogleBase directory search and video search though YouTube feeds service.

Java Search Tool tries to implement indexing process using multi threading instance. Some of data structures used on base Inverted File Indexing algorithm change into Hashtable implementation rather than LinkedList. Modification also apply to Vector Space Model so it can improve similarity measurement for phrase based search. To accomodate indexing process for all documents in hard drive, Java Search Tool also uses some of text or metadata extraction library including, Apache POI for Microsoft Office PowerPoint and Word 2000/XP compatible documents, PDFBox for pdf to text extraction and also JID3Tag for MP3 ID3Tag extraction. All library used by Java Search Tool is open source based and free to use with each GNU based licenses.

For such system such as Mac OS X 10.5, Java Seach Tool also provide QuickLook for previewing document and some cool GUI experience. Beside that, Java Search Tool also provide logging feature and embedded cool help system to provide user with easy to use and fully documented application.

Keywords: java, search, vector space model, pemodelan ruang vektor, inverted file indexing, indexing, native indexing.

DAFTAR ISI

LEMBAR PENGESAHAN	ii
PERNYATAAN ORISINALITAS LAPORAN	iii
KATA PENGANTAR	iv
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS	v
ABSTRACT	vi
DAFTAR ISI	vii
DAFTAR GAMBAR	xiv
Bab I : PENDAHULUAN	1
I.1. Latar Belakang	1
I.2. Tujuan	1
I.3. Gambaran Umum	1
I.4. Batasan Masalah	2
I.5. Detail Teknis	3
I.6. Definisi dan Singkatan	4
I.6.1. JRE(Java Runtime Environment) atau JVM	4
I.6.2. JDK(Java Development Kit)	4
I.6.3. Search Engine	4
I.6.4. Inverted File Indexing	4
I.6.5. <i>Vector Space Model</i> atau Pemodelan Ruang Vektor	5
I.6.6. Index	5
I.6.7. Google dan GoogleBase	5
I.6.8. Mime Type	5
I.6.9. ID3Tag, ID3Tag versi 1.0 dan ID3Tag versi 2.3	6
I.6.10. Runtime atau service	6
I.6.11. Native dan native indexing service	6

I.6.12.	Native Library	7
I.6.13.	AppleScript.....	7
I.6.14.	QuickLook.....	7
I.6.15.	Spotlight	7
I.6.16.	Locate	8
I.6.17.	YouTube.....	8
I.6.18.	DTrace	8
I.6.19.	Apple Instruments	9
I.6.20.	I/O dan I/O Usage.....	9
I.6.21.	CPU dan CPU Usage.....	9
I.6.22.	Object Allocation.....	9
I.6.23.	Memory Leak	10
I.6.24.	Garbage Collector.....	10
I.7.	Struktur Laporan	10
Bab II :	Spesifikasi Produk.....	12
II.1.	Gambaran Keseluruhan	12
II.1.1.	Perspektif Produk.....	12
II.1.1.1.	Antarmuka Sistem.....	12
II.1.1.2.	Antarmuka Pengguna.....	12
II.1.1.3.	Antarmuka Perangkat Keras	14
II.1.1.4.	Antarmuka Perangkat Lunak	14
II.1.1.5.	Antarmuka Jaringan.....	15
II.1.1.6.	Batasan Memori	15
II.1.1.7.	Persyaratan Adaptasi.....	15
II.1.2.	Fungsi Produk	15
II.1.3.	Karakteristik Pengguna	16
II.1.4.	Asumsi dan Ketergantungan	16
II.1.5.	Penundaan Persyaratan.....	16

II.2. Fitur Produk Perangkat Lunak.....	17
II.2.1. <i>Indexing</i>	17
II.2.1.1. Tujuan	17
II.2.1.2. Urutan Stimulus	17
II.2.1.3. Persyaratan Fungsional yang Berhubungan.....	18
II.2.1.3.1. Masukan	18
II.2.1.3.2. Proses.....	18
II.2.1.3.3. Keluaran	19
II.2.2. <i>Local Searching</i>	19
II.2.2.1. Tujuan	19
II.2.2.2. Urutan Stimulus	19
II.2.2.3. Persyaratan Fungsional yang Berhubungan.....	19
II.2.2.3.1. Masukan	19
II.2.2.3.2. Proses.....	19
II.2.2.3.3. Keluaran	20
II.2.3. <i>Native Searching</i>	20
II.2.3.1. Tujuan	20
II.2.3.2. Urutan Stimulus	20
II.2.3.3. Persyaratan Fungsional yang Berhubungan.....	21
II.2.3.3.1. Masukan	21
II.2.3.3.2. Proses.....	21
II.2.3.3.3. Keluaran	21
II.2.4. <i>Google Searching</i>	21
II.2.4.1. Tujuan	22
II.2.4.2. Urutan Stimulus	22
II.2.4.3. Persyaratan Fungsional yang Berhubungan.....	22
II.2.4.3.1. Masukan	22
II.2.4.3.2. Proses.....	22

II.2.4.3.3. Keluaran	23
II.2.5. <i>Bookmark</i>	23
II.2.5.1. Tujuan	23
II.2.5.2. Urutan Stimulus	23
II.2.5.3. Persyaratan Fungsional yang Berhubungan.....	23
II.2.5.3.1. Masukan	23
II.2.5.3.2. Proses.....	24
II.2.5.3.3. Keluaran	24
II.3. Batasan Performa.....	24
II.4. Atribut – Atribut Perangkat Lunak.....	24
II.4.1. Keandalan.....	25
II.4.2. Ketersediaan.....	25
II.4.3. Keamanan.....	25
II.4.4. Pemeliharaan	25
II.4.5. Perpindahan.....	25
Bab III : Desain Perangkat Lunak	26
III.1. Overview	26
III.1.1. Komponen Perangkat Lunak	26
III.2. Desain Arsitektur Aplikasi.....	29
III.2.1. Teori Penunjang.....	29
III.2.1.1. Inverted File Indexing	29
III.2.1.2. Vector Space Model	30
III.2.2. Use Case	32
III.2.3. Dependency Diagram	35
III.2.3.1. Model-Controller Components.....	36
III.2.3.2. View-Controller Components	38
III.2.4. Class Diagram.....	39
III.2.4.1. Package net.sf.jsearchtool.core.extractor	39

III.2.4.1.1.	Extractor.....	39
III.2.4.1.2.	AlienExtractor.....	40
III.2.4.1.3.	MP3Extractor.....	40
III.2.4.1.4.	OfficePowerPointExtractor.....	40
III.2.4.1.5.	OfficeWordExtractor.....	41
III.2.4.1.6.	PDFExtractor.....	41
III.2.4.2.	TextExtractor.....	42
III.2.4.3.	Package net.sf.jsearchtool.core.index.....	42
III.2.4.3.1.	FileIndexer.....	42
III.2.4.3.2.	InvertedFileEntry.....	43
III.2.4.4.	Package net.sf.jsearchtool.core.search.....	43
III.2.4.4.1.	VectorSpaceHandler.....	44
III.2.4.4.2.	GoogleBaseClient.....	44
III.2.4.4.3.	LocateHandler.....	45
III.2.4.4.4.	SpotlightHandler.....	45
III.2.4.4.5.	YouTubeClient.....	46
III.2.4.5.	Package net.sf.jsearchtool.io.....	46
III.2.4.5.1.	FileManager.....	46
III.2.4.5.2.	FileTypeDescriptor.....	47
III.2.4.5.3.	ID3TagDescriptor.....	47
III.2.4.5.4.	ImagesDescriptor.....	48
III.2.4.6.	Package net.sf.jsearchtool.launcher.....	48
III.2.4.6.1.	JSearchToolLauncher.....	48
III.2.4.6.2.	MacLauncher.....	49
III.2.4.7.	Package net.sf.jsearchtool.UIComponents.....	49
III.2.4.7.1.	AppForm.....	49
III.2.4.7.2.	AboutDialog.....	50
III.2.4.7.3.	BookmarkDialog.....	50

III.2.4.7.4.	BookmarkManager.....	50
III.2.4.7.5.	ExceptionDialog.....	51
III.2.4.7.6.	LogViewer.....	51
III.2.4.7.7.	PreferencesDialog	51
III.2.4.7.8.	TableFactory	52
III.2.4.8.	Package net.sf.jsearchtool.util	52
III.2.4.8.1.	Logger	52
III.2.4.8.2.	StopWord	53
III.2.4.8.3.	SystemUtilities	53
III.2.4.8.4.	VectorMath	54
III.2.4.9.	Package net.sf.jsearchtool.xml	54
III.2.4.9.1.	XMLReader.....	54
III.2.4.9.2.	XMLWriter	55
III.2.4.10.	Package uk.ac.comp.lancs	55
III.2.4.10.1.	Paice	55
Bab IV :	Implementasi Desain.....	56
IV.1.	Perencanaan Tahap Implementasi.....	56
IV.1.1.	Implementasi Kelas	56
IV.1.2.	Keterkaitan Antar Kelas	78
IV.2.	Perjalanan Tahap Implementasi.....	79
IV.2.1.	Implementasi Bottom Up.....	79
IV.2.2.	Debugging	80
IV.3.	Ulasan Realisasi Fungsionalitas.....	81
IV.3.1.	Fitur Indexing	81
IV.3.2.	Fitur Searching	83
IV.4.	Ulasan Realisasi Antarmuka Pengguna	84
Bab V :	Testing dan Evaluasi Sistem	93
V.1.	Rencana Pengujian Sistem Terimplementasi	93

V.1.1.	Test Case.....	93
V.1.2.	Uji Fungsionalitas	94
V.2.	Perjalanan Metodologi Pengujian	95
V.2.1.	White Box	95
V.2.1.1.	Fitur Indexing	97
V.2.1.2.	Fitur Pencarian Menggunakan Vector Space Model.....	101
V.2.2.	Black Box.....	107
V.2.2.1.	Fitur Pencarian Menggunakan Runtime Spotlight	107
V.2.2.2.	Fitur Pencarian Menggunakan Runtime Locate	107
V.2.2.3.	Fitur Pencarian Menggunakan Service Google	108
V.2.2.4.	Fitur Bookmark.....	108
V.2.2.5.	Fitur QuickLook	108
V.2.2.6.	Fitur Logging dan Log Viewer.....	109
V.2.3.	Ulasan Hasil Evaluasi	109
Bab VI :	Kesimpulan dan Saran	110
VI.1.	Saran Pengembangan	110
VI.2.	Keterkaitan antara Kesimpulan dan Hasil Evaluasi.....	110
VI.3.	Keterkaitan antara Saran dan Hasil Evaluasi	110
VI.4.	Rencana Perbaikan terhadap Saran yang Diberikan	111
Referensi	112
Lampiran	114
Algoritma Inverted File Indexing.....		114
Algoritma Vector Space Model.....		116
GNU General Public License		117
Apache License.....		125
PDFBox License.....		130
Mozilla Public License		131

DAFTAR GAMBAR

Gambar 3.1: Use Case JSearchTool.....	33
Gambar 3.2: Model-Controller Dependency Diagram	36
Gambar 3.3: View-Controller Dependency Diagram	38
Gambar 3.4: Kelas Diagram Extractor.....	39
Gambar 3.5: Kelas Diagram AlienExtractor.....	40
Gambar 3.6: Kelas Diagram MP3Extractor.....	40
Gambar 3.7: Kelas Diagram OfficePowerPointExtractor.....	41
Gambar 3.8: Kelas Diagram OfficeWordExtractor	41
Gambar 3.9: Kelas Diagram PDFExtractor	41
Gambar 3.10: Kelas Diagram TextExtractor	42
Gambar 3.11: Kelas Diagram FileIndexer	42
Gambar 3.12: Kelas Diagram InvertedFileEntry	43
Gambar 3.13: Kelas Diagram VectorSpaceHandler	44
Gambar 3.14: Kelas Diagram GoogleBaseClient	44
Gambar 3.15: Kelas Diagram LocateHandler.....	45
Gambar 3.16: Kelas Diagram SpotlightHandler.....	45
Gambar 3.17: Kelas Diagram YouTubeClient.....	46
Gambar 3.18: Kelas Diagram FileManager	46
Gambar 3.19: Kelas Diagram FileTypeDescriptor	47
Gambar 3.20: Kelas Diagram ID3TagDescriptor	47
Gambar 3.21: Kelas Diagram ImagesDescriptor	48
Gambar 3.22: Kelas Diagram JSearchTooLauncher	48
Gambar 3.23: Kelas Diagram MacLauncher	49
Gambar 3.24: Kelas Diagram AppForm.....	50
Gambar 3.25: Kelas Diagram AboutDialog.....	50
Gambar 3.26: Kelas Diagram BookmarkDialog.....	50
Gambar 3.27: Kelas Diagram BookmarkManager	51
Gambar 3.28: Kelas Diagram ExceptionDialog	51
Gambar 3.29: Kelas Diagram LogViewer	51
Gambar 3.30: Kelas Diagram PreferencesDialog.....	52
Gambar 3.31: Kelas Diagram TableFactory	52

Gambar 3.32: Kelas Diagram Logger.....	52
Gambar 3.33: Kelas Diagram StopWordWriter.....	53
Gambar 3.34: Kelas Diagram SystemUtilities.....	53
Gambar 3.35: Kelas Diagram VectorMath.....	54
Gambar 3.36: Kelas Diagram XMLReader.....	54
Gambar 3.37: Kelas Diagram XMLWriter.....	55
Gambar 3.38: Kelas Diagram Paice.....	55
Gambar 4.39: Sequence Diagram Proses Indexing.....	78
Gambar 4.40: Sequence Diagram Proses Pencarian.....	79
Gambar 4.41: Tampilan Window Utama Aplikasi.....	84
Gambar 4.42: Menu Tools.....	85
Gambar 4.43: Tampilan Window Utama Pencarian Google.....	86
Gambar 4.44: Menu JSearchTool Pada Mac OS X.....	86
Gambar 4.45: Tampilan <i>Window Preferences</i>	87
Gambar 4.46: Menu Bookmark.....	88
Gambar 5.47: Window Manajemen Bookmark.....	88
Gambar 5.48: Menu Help.....	89
Gambar 4.49: Tampilan Hasil Pencarian.....	90
Gambar 4.50: Tampilan Dock di Mac OS X.....	90
Gambar 4.51: Preview Dokumen pada Mac OS X 10.5.....	91
Gambar 4.52: Tampilan Window Log Viewer.....	92
Gambar 5.53 : Hirarki Dokumen Pengujian.....	97
Gambar 5.54: Tampilan Hasil DTrace Fitur Indexing.....	98
Gambar 5.55: Tampilan Hasil CPU Monitor.....	99
Gambar 5.56: Tabel Hasil Process Sampler Fitur Indexing.....	101
Gambar 5.57: Hasil Pencarian Menggunakan Satu Kata.....	103
Gambar 5.58: Hasil Pencarian Menggunakan Frasa.....	103
Gambar 5.59: Penggunaan Resources Fitur Pencarian.....	105
Gambar 5.60: Tampilan Hasil Process Sampler Proses Pencarian.....	106
Gambar 5.61: Tampilan Hasil Directory I/O Monitor.....	106