

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN SKRIPSI	ii
HALAMAN PENGESAHAN SKRIPSI	iii
HALAMAN PERNYATAAN SKRIPSI	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI	v
ABSTRAK	vi
ABSTRACT	vii
KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR GAMBAR	xii
DAFTAR TABEL	xiv
BAB I. PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Metodologi Penelitian	3
1.7 Sistematika Penulisan	5
BAB II. LANDASAN TEORI	6
2.1 <i>Question Answering System</i>	6
2.2 Kelas Kata	8
2.3 <i>Regular Expression</i>	10
2.4 <i>JavaScript Object Notation (JSON)</i>	12
2.5 <i>Web Service</i>	13
2.6 <i>Web Scraping</i>	14
2.7 <i>CSS Selector</i>	15
2.8 Android	16
2.9 PhoneGap	16
2.10 Studi Literatur	17
BAB III. ANALISIS DAN PERANCANGAN SISTEM	19
3.1 Analisis Sistem	20
3.2 Analisis Situs 21cineplex.com	20
3.3 Analisis <i>Web Scraping</i>	26
3.4 Analisis Pertanyaan	28
3.5 Analisis <i>Tokenization</i>	30
3.6 Analisis <i>Part of Speech</i>	31
3.7 Analisis <i>Named Entity Recognition</i>	34
3.8 Analisis Validasi Pertanyaan	38
3.9 Analisis <i>Answer Processing</i>	40

3.10	Perancangan Sistem	42
3.10.1.	Arsitektur Sistem	42
3.10.2.	<i>Data Flow Diagram</i> (DFD)	43
3.10.2.1.	DFD Level 0	43
3.10.2.2.	DFD Level 1	43
3.10.2.3.	DFD Level 2 <i>Question Processing</i>	44
3.10.2.4.	DFD Level 2 <i>Answer Processing</i>	45
3.10.3.	<i>Entity Relationship Diagram</i> (ERD)	45
3.10.3.1.	Tabel Actor	46
3.10.3.2.	Tabel Film	47
3.10.3.3.	Tabel ActorFilm	47
3.10.3.4.	Tabel POS	47
3.10.3.5.	Tabel WhiteListPOS	48
3.10.4.	Rancangan Antarmuka	48
BAB IV.	IMPLEMENTASI DAN PENGUJIAN	50
4.1	Implementasi	50
4.1.1.	Lingkungan Implementasi	50
4.1.1.1.	Spesifikasi <i>Hardware</i>	50
4.1.1.2.	Spesifikasi <i>Software</i>	50
4.2	Implementasi Antar Muka	51
4.3	Pengujian	54
4.3.1.	Pengujian <i>Black Box</i>	54
4.3.2.	Pengujian <i>White Box</i>	56
4.3.2.1.	Pengujian <i>White Box</i> Pada Modul <i>Tokenization</i>	57
4.3.2.2.	Pengujian <i>White Box</i> Pada Modul POS <i>Tagging</i>	58
4.3.2.3.	Pengujian <i>White Box</i> Pada Modul <i>Named Entity Recognition</i>	59
4.3.2.4.	Pengujian <i>White Box</i> Pada Modul <i>Answer Processing</i>	66
4.3.3.	Pengujian NLP	72
BAB V.	PENUTUP	76
5.1	Kesimpulan	76
5.2	Saran	76

DAFTAR PUSTAKA
LAMPIRAN

DAFTAR GAMBAR

Gambar 1.1 Siklus Model Sekuensial Linear.....	5
Gambar 2.1 <i>Tokenization</i> Pada Teks.....	6
Gambar 2.2 Pemberian Label POS	7
Gambar 3.1 Arsip Film Dikelompokkan Berdasarkan Tahun dan Bulan	8
Gambar 3.1 Arsip Film Berdasarkan Tahun dan Bulan	20
Gambar 3.2 Kontrol <i>Pagination</i> Dari Situs 21cinplex	21
Gambar 3.3 Judul Film, URL, Gambar, Dan Link <i>More</i> Dari Daftar Film	21
Gambar 3.4 <i>Flowchart</i> Mengekstrak Informasi Dari Arsip Film	22
Gambar 3.5 <i>Flowchart</i> Mengekstrak Informasi Film Yang Akan Tayang	24
Gambar 3.6 <i>Flowchart</i> Untuk Mengekstrak Informasi Film Yang Sedang Tayang	25
Gambar 3.7 <i>Flowchart</i> Teknik <i>Web Scraping</i> Pada <i>Domain</i> 21cinplex ..	26
Gambar 3.8 <i>Flowchart</i> <i>Tokenization</i>	30
Gambar 3.9 <i>Flowchart</i> <i>POS Tagging</i>	33
Gambar 3.10 <i>Flowchart</i> <i>Named Entity Recognition</i>	36
Gambar 3.11 <i>Flowchart</i> <i>Answer Processing</i>	40
Gambar 3.12 Arsitektur Sistem	41
Gambar 3.13 DFD Level 0	42
Gambar 3.14 DFD Level 1	43
Gambar 3.15 DFD Level 2 <i>Question Processing</i>	43
Gambar 3.16 DFD Level 2 <i>Answer Processing</i>	44
Gambar 3.17 ERD	45
Gambar 3.18 Rancangan Antarmuka Menyerupai Aplikasi <i>Messenger</i>	48
Gambar 3.19 Rancangan Antarmuka Daftar Film dan Gambar Poster Film	48
Gambar 3.20 Rancangan Antarmuka Saat Menampilkan Daftar Film, Gambar Poster Film, dan Jam Tayang	49
Gambar 4.1 Halaman Awal Program	50
Gambar 4.2 Pemberitahuan Bahwa <i>Server</i> Tidak Dapat Dihubungkan	51

Gambar 4.3 Konfigurasi Alamat <i>Server</i>	52
Gambar 4.4 Konfigurasi <i>Port</i> Pada <i>Server</i>	52
Gambar 4.5 Pemberitahuan Bahwa <i>Setting</i> Berhasil Diubah	53
Gambar 4.6 <i>Flowgraph Tokenization</i>	57
Gambar 4.7 <i>Flowgraph POS Tagging</i>	58
Gambar 4.8 <i>Flowgraph Named Entity Recognition</i>	60
Gambar 4.9 <i>Flowgraph Answer Processing</i>	67

DAFTAR TABEL

Tabel 2.1 Daftar <i>Tagset</i> POS	7
Tabel 2.2 Pola Umum Pada Regex	11
Tabel 2.3 Perbandingan Sistem QA.....	18
Tabel 3.1 Informasi Seputar Daftar Film Beserta <i>CSS Selectornya</i>	22
Tabel 3.2 Detail Film Yang Dicari Beserta <i>CSS Selectornya</i>	23
Tabel 3.3 Daftar Film Yang Akan Datang Beserta <i>CSS Selectornya</i>	23
Tabel 3.4 Daftar Film Yang Sedang Datang Beserta <i>CSS Selectornya</i>	25
Tabel 3.5 Sampel Pertanyaan Yang Dijadikan Ujicoba Penelitian	27
Tabel 3.6 Daftar Kategori POS Yang Diambil Dari Website	31
Tabel 3.7 Daftar Kategori POS Yang Digunakan Pada Penelitian Ini	32
Tabel 3.8 Fitur Kontekstual Yang Digunakan	34
Tabel 3.9 Hasil <i>Preprocessing</i> Pada Pertanyaan	34
Tabel 3.10 Daftar Variabel Yang Digunakan	35
Tabel 3.11 Daftar Aturan Yang Dipenuhi Untuk Mendeteksi Entitas Film	35
Tabel 3.12 Daftar Aturan Yang Dipenuhi Untuk Mendeteksi Entitas Aktor	35
Tabel 3.13 <i>White List</i> POS Hasil Dari Ekstrak Pada Sampel Pertanyaan ..	37
Tabel 3.14 Spesifikasi Tabel Actor	45
Tabel 3.15 Spesifikasi Tabel Film	46
Tabel 3.16 Spesifikasi Tabel Actorfilm	46
Tabel 3.17 Spesifikasi Tabel POS	46
Tabel 3.18 Spesifikasi Tabel WhiteListPOS	47
Tabel 4.1 Hasil Pengujian Pada Halaman Utama	53
Tabel 4.2 Hasil Pengujian Saat Melakukan Konfigurasi <i>Server</i>	54
Tabel 4.3 Bagian <i>Flowchart</i> Yang Akan Diuji Melalui <i>White Box</i>	55
Tabel 4.4 Hubungan Antara <i>Cyclomatic Complexity</i> dan Resiko	56
Tabel 4.5 Persentase Pertanyaan Yang Diterima	73