

**PENGEMBANGAN SISTEM EKSTRAKSI INFORMASI
UNTUK DOKUMEN LEGAL INDONESIA:
STUDI KASUS DOKUMEN UNDANG-UNDANG
REPUBLIK INDONESIA**

SKRIPSI

**SUSY VIOLINA
120400084X**



**UNIVERSITAS INDONESIA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI ILMU KOMPUTER
DEPOK
JANUARI 2009**

**PENGEMBANGAN SISTEM EKSTRAKSI INFORMASI
UNTUK DOKUMEN LEGAL INDONESIA:
STUDI KASUS DOKUMEN UNDANG-UNDANG
REPUBLIK INDONESIA**

SKRIPSI

**Diajukan sebagai salah satu syarat untuk memperoleh gelar
Sarjana**

**SUSY VIOLINA
120400084X**



**UNIVERSITAS INDONESIA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI ILMU KOMPUTER
DEPOK
JANUARI 2009**

HALAMAN PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya saya sendiri,
dan semua sumber baik yang dikutip maupun dirujuk
telah saya nyatakan dengan benar.



Nama : Susy Violina

NPM : 120400084X

Tanda Tangan :

Tanggal : 31 Desember 2008

HALAMAN PENGESAHAN

Skripsi ini diajukan oleh:

Nama : Susy Violina
NPM : 120400084X
Program Studi : Ilmu Komputer
Judul Skripsi : Pengembangan Sistem Ekstraksi Informasi untuk Dokumen Legal
Indonesia: Studi Kasus Dokumen Undang-Undang Republik
Indonesia

Telah berhasil dipertahankan di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi Ilmu Komputer Fakultas Ilmu Komputer Universitas Indonesia.

DEWAN PENGUJI

Pembimbing : Dr. Indra Budi S.Kom, M.Kom (.....)

Penguji : Harry Budi Santoso S.Kom, M.Kom (.....)

Penguji : Rizal Fathoni Aji S.Kom, M.Kom (.....)

Ditetapkan di : Fakultas Ilmu Komputer UI Depok
Tanggal : 31 Desember 2008

KATA PENGANTAR/UCAPAN TERIMA KASIH

Puji syukur saya panjatkan kepada Allah SWT, karena atas berkat dan rahmat-Nya, saya dapat menyelesaikan skripsi ini. Penulisan skripsi ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer Program Studi Ilmu Komputer Fakultas Ilmu Komputer Universitas Indonesia. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan skripsi ini. Oleh karena itu, saya mengucapkan terima kasih pada:

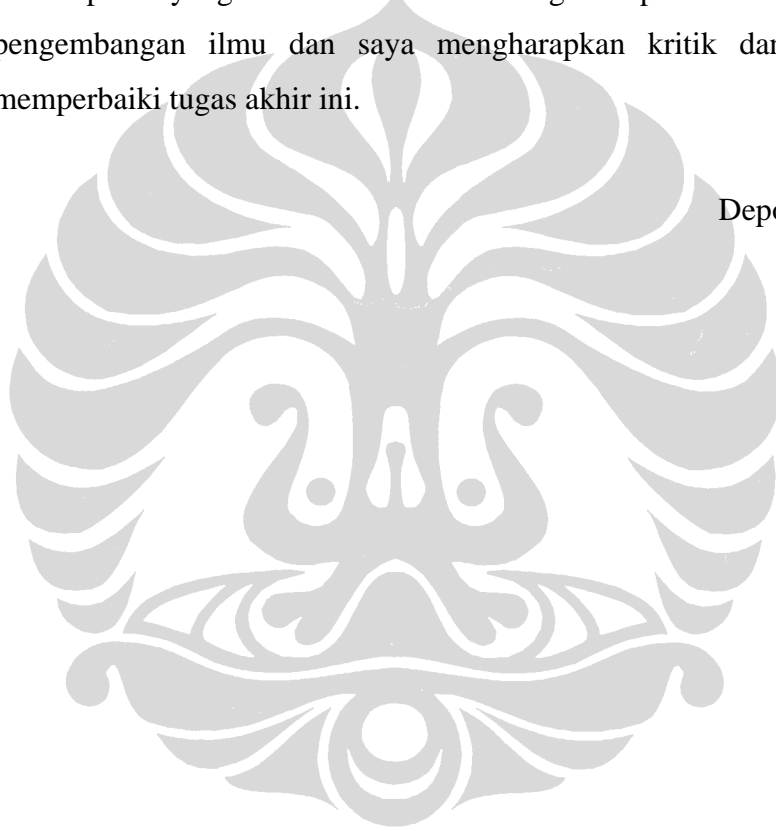
1. Papa(Alm.) dan Mama yang selalu memberikan cinta, dukungan, semangat, dan doa kepada saya;
2. Bapak Dr. Indra Budi S.Kom, M.Kom, selaku dosen pembimbing yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan saya dalam penyusunan skripsi ini;
3. Bapak Ir. Dana Indra Sensuse MLIS, PhD selaku pembimbing akademik yang telah memberikan dukungan, saran, dan kritik kepada saya;
4. Kakak dan Uda yang telah memberikan dukungan, semangat, dan doa kepada saya;
5. Seluruh keluarga besar yang selalu menyemangati dan mendukung saya;
6. Mba Siti Mawaddah dan Mba Ahlijati Nuraminah atas bantuannya kepada penulis dalam mengerjakan skripsi ini;
7. Teman-teman csui04: Icha (Almh.), Eva, Mea, Elly, Mega, Ipro, Ame, Sawie, Hendra, Mahen, Smel, dkk yang telah memberikan perhatian dan semangat;
8. Mba Nova dan Andreas atas segala bantuannya;
9. Teman-teman di ruangan 1233 dan 1235 yang selalu berbagi keceriaan dan semangat dalam perjuangan menyelesaikan tugas akhir;
10. Teman-teman di SMAN 1 Bukittinggi yang masih tetap berbagi cerita dan memberi semangat;
11. Pak Acep, Mas Wiryo, Bu Rita, Bu Yanti, dan seluruh staf Sekretariat Fasilkom UI yang telah membantu dalam proses administrasi skripsi ini;

12. Ibu-Ibu pustakawati Fasilkom UI yang telah membantu dalam mencari referensi;
13. Segenap sivitas akademika Fasilkom UI, atas segala bantuannya baik langsung maupun tidak langsung;
14. Semua orang yang telah membantu penulis selama ini, terimakasih atas semuanya.

Akhir kata, saya berharap Allah SWT berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga skripsi ini membawa manfaat bagi pengembangan ilmu dan saya mengharapkan kritik dan saran yang dapat memperbaiki tugas akhir ini.

Depok, 1 Desember 2008

Penulis



**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI
TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademik Universitas Indonesia, saya yang bertandatangan di bawah ini:

Nama : Susy Violina

NPM : 120400084X

Program Studi : Ilmu Komputer

Fakultas : Ilmu Komputer

Jenis Karya : Skripsi

demikian demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Indonesia **Hak Bebas Royalti Noneksklusif (*Non-exclusive Royalty-Free Right*)** atas karya ilmiah saya yang berjudul:

Pengembangan Sistem Ekstraksi Informasi untuk Dokumen Legal Indonesia: Studi Kasus Dokumen Undang-Undang Republik Indonesia beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Universitas Indonesia berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan memublikasikan tugas akhir saya tanpa meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Depok

Pada Tanggal : 31 Desember 2008

Yang Menyatakan

(Susy Violina)

DAFTAR ISI

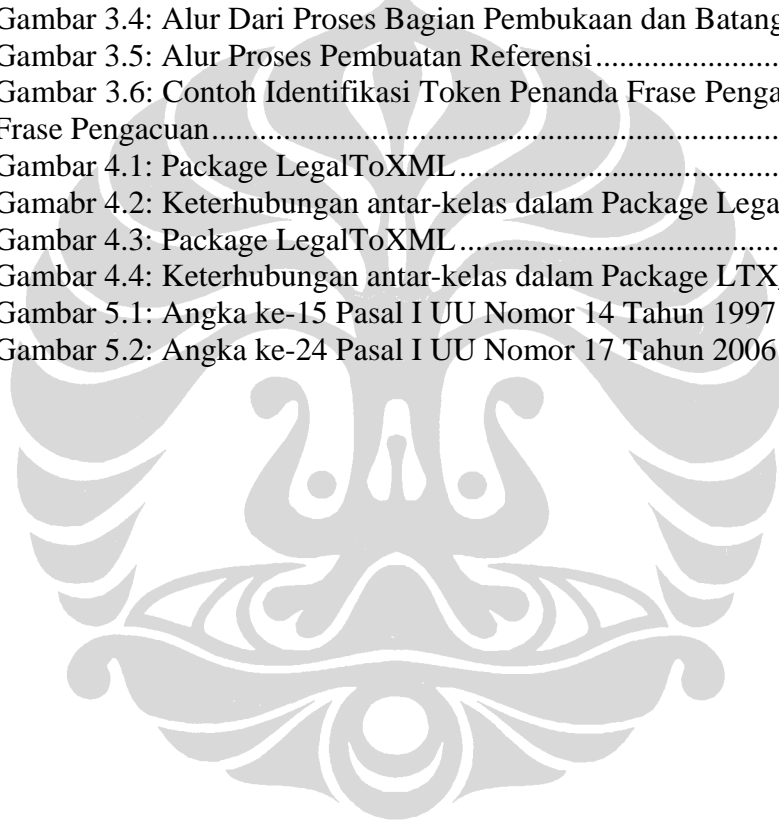
HALAMAN PERNYATAAN ORISINALITAS.....	ii
HALAMAN PENGESAHAN.....	iii
KATA PENGANTAR/UCAPAN TERIMAKASIH	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI.....	vi
ABSTRAK	vii
DAFTAR ISI.....	viii
DAFTAR GAMBAR	x
DAFTAR TABEL.....	xii
DAFTAR LAMPIRAN.....	xiii
PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	9
1.3 Tujuan Penelitian	9
1.4 Ruang Lingkup Penelitian.....	9
1.5 Tahapan Penelitian	10
1.6 Sistematika Penulisan	11
LANDASAN TEORI.....	13
2.1 Sistem Ekstraksi Informasi	13
2.1.1 Sejarah Sistem Ekstraksi Informasi	15
2.1.2 Tugas Sistem Ekstraksi Informasi.....	16
2.1.3 Pendekatan Sistem Ekstraksi Informasi.....	17
2.2 Extensible Markup Language (XML).....	20
2.3 Bahasa DOT	22
2.4 Dokumen Legal Indonesia	26
2.5 Undang-Undang Republik Indonesia.....	27
2.5.1 Sistematika Penyusunan Undang-Undang	27
2.5.2 Undang-Undang Perubahan	38
2.5.3 Teknik Pengacuan dalam UU.....	41
2.6 Evaluasi Sistem Ekstraksi Informasi.....	43
ANALISIS DAN PERANCANGAN SISTEM.....	47
3.1 Arsitektur Sistem.....	47
3.2 Dokumen Masukan dan Dokumen Keluaran	49
3.3 Kamus	50
3.4 Aturan.....	51
3.5 Tokenisasi	53
3.6 Identifikasi Bagian UU	53
3.7 Proses Bagian UU	54
3.8 Proses Bagian Pembukaan	56
3.9 Identifikasi Jenis UU.....	59
3.10 Proses Bagian Batang Tubuh UU Non-Perubahan	60
3.11 Proses Bagian Batang Tubuh UU Perubahan.....	63
3.12 Proses Bagian Penutup.....	68
3.13 Pembuatan Referensi.....	70
3.13.1 Sub Proses Identifikasi Token Penanda Frase Pengacuan	70

3.13.2 Sub Proses Identifikasi Frase Pengacuan.....	72
3.13.3 Sub Proses Pemberian Elemen Kosong REF.....	75
3.14 Pembuatan Dokumen dalam Format XML.....	77
3.15 Pencarian Referensi.....	78
3.16 Pembuatan Dokumen DOT.....	79
3.17 Evaluasi Sistem.....	80
IMPLEMENTASI.....	83
4.1 Struktur Implementasi SEI.....	83
4.1.1 Kelas LTX_Start.....	84
4.1.2 Kelas LTX_Judul.....	84
4.1.3 Kelas LTX_Pembukaan.....	85
4.1.4 Kelas LTX_BatangTubuh.....	85
4.1.5 Kelas LTX_BatangTubuh_Perubahan.....	85
4.1.6 Kelas LTX_Penutup.....	86
4.1.7 Kelas LTX_Referensi.....	86
4.1.8 Kelas LTX_Elemen.....	86
4.1.9 Kelas Referensi.....	87
4.2 Struktur Implementasi Program Evaluasi.....	87
UJI COBA DAN ANALISIS.....	90
5.1 Lingkungan Uji Coba.....	90
5.2 Karakteristik Dokumen.....	90
5.3 Hasil Uji Coba.....	91
5.3.1 Hasil Uji Coba Dokumen UU Non-Perubahan.....	92
5.3.2 Hasil Uji Coba Dokumen UU Perubahan.....	93
5.4 Analisis Hasil Uji Coba.....	93
5.4.1 Analisis Hasil Uji Coba UU Non-Perubahan.....	93
5.4.2 Analisis Hasil Uji Coba UU Perubahan.....	99
PENUTUP.....	104
6.1 Kesimpulan.....	104
6.2 Saran.....	104
DAFTAR REFERENSI.....	106

DAFTAR GAMBAR

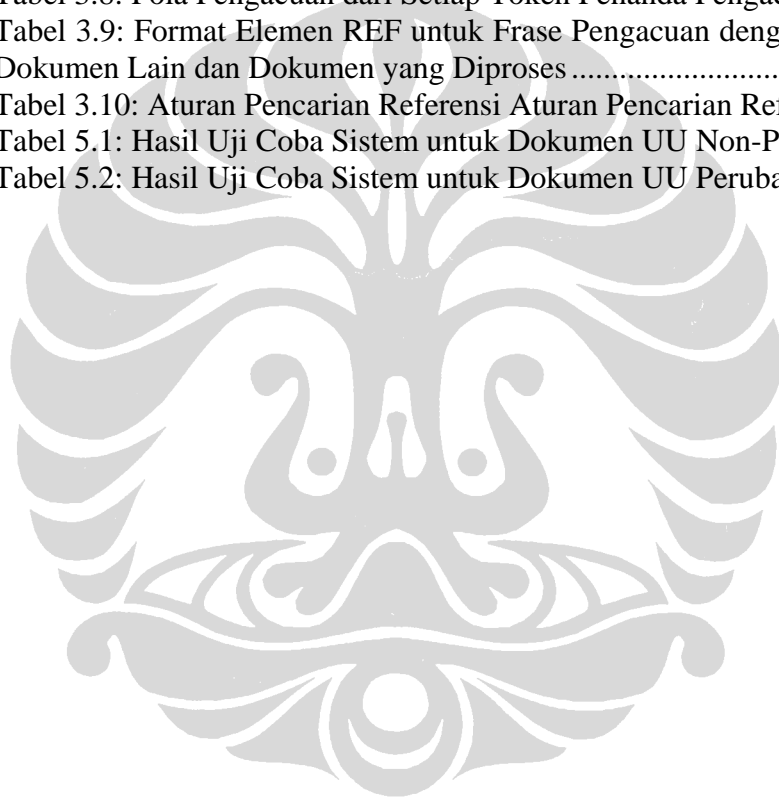
Gambar 1.1: Pasal 39 UU Nomor 16 Tahun 2004 tentang Kejaksaan Republik Indonesia	2
Gambar 1.2: Pasal 22 ayat (1) UU Nomor 30 Tahun 2007 tentang Energi	2
Gambar 1.3: Dasar Hukum UU Nomor 17 Tahun 2006	8
Gambar 1.4: Graf Berarah Dasar Hukum UU Nomor 17 Tahun 2006	8
Gambar 2.1: Perbedaan Arsitektur IR dan IE	14
Gambar 2.2: Contoh Penulisan Format XML	21
Gambar 2.3: Contoh Tampilan Dokumen XML dalam Browser.....	21
Gambar 2.4: Contoh Pendeklarasian Graf Tidak Berarah dengan Bahasa DOT ..	23
Gambar 2.5: Contoh Hasil Pemrosesan Graf Tidak Berarah Dengan Graphviz.....	23
Gambar 2.6: Contoh Pendeklarasian Graf Berarah dengan Bahasa DOT	24
Gambar 2.7: Contoh Hasil Pemrosesan Graf Berarah Dengan Graphviz	24
Gambar 2.8: Contoh Penulisan Komentar dalam bahasa DOT.....	24
Gambar 2.9: Contoh Penggunaan Attribute dalam bahasa DOT.....	25
Gambar 2.10: Contoh Graf Tidak Berarah dengan Menggunakan Attribute dalam bahasa DOT.....	25
Gambar 2.11: Sistematika Peraturan Perundang-Undangan	28
Gambar 2.12: Contoh Penulisan Judul UU pada Umumnya.....	28
Gambar 2.13: Contoh Penulisan Judul UU Perubahan Undang-Undang (perubahan pertama).....	29
Gambar 2.14: Contoh Penulisan Judul UU Perubahan Undang-Undang (perubahan ketiga).....	29
Gambar 2.15: Contoh Penulisan Judul UU Perubahan Undang-Undang dengan Menggunakan Nama Singkat UU yang Diubah.....	29
Gambar 2.16: Contoh Penulisan Judul UU Pencabutan Undang-Undang.....	29
Gambar 2.17: Contoh Penulisan Judul UU Penetapan Peraturan Pemerintah Pengganti Undang-Undang menjadi Undang-Undang.....	30
Gambar 2.18: Contoh Penulisan Judul UU Pengesahan Perjanjian Internasional dengan Penggunaan Bahasa Indonesia sebagai Teks Resmi	30
Gambar 2.19: Contoh Penulisan Judul UU Pengesahan Perjanjian Internasional Tanpa Penggunaan Bahasa Indonesia sebagai Teks Resmi	30
Gambar 2.20: Contoh Penulisan Frase ‘Dengan Rahmat Tuhan Yang Maha Esa’ dan Jabatan Pembentuk UU	31
Gambar 2.21: Format Penulisan Konsiderans.....	31
Gambar 2.22: Contoh Penulisan Dasar Hukum	32
Gambar 2.23: Contoh Penulisan Diktum	33
Gambar 2.24: Contoh Penulisan Buku, Bab, Bagian, dan Paragraf dalam UU	34
Gambar 2.25: Urutan Pengelompokan Materi UU Non-Perubahan	35
Gambar 2.26: Contoh Penulisan Pasal dan Ayat	35
Gambar 2.27: Contoh Penulisan Rincian Level 1	36
Gambar 2.28: Contoh Penulisan Rincian Sampai Level 2.....	36
Gambar 2.29: Contoh Penulisan Rincian Sampai Level 3.....	363
Gambar 2.30: Contoh Penulisan Rincian Sampai Level 4.....	37
Gambar 2.31: Contoh Penulisan Penutup	38
Gambar 2.31: Contoh Penulisan Penutup	38

Gambar 2.32: Format Penulisan Pasal I UU Perubahan	38
Gambar 2.33: Format Penulisan Rincian UU yang Telah Mengubah Suatu UU Pada Pasal I	39
Gambar 2.34: Contoh Penyisipan Bab Baru pada UU Perubahan	39
Gambar 2.35: Contoh Penyisipan Pasal Baru pada UU Perubahan.....	40
Gambar 2.36: Contoh Penyisipan Ayat Baru pada UU Perubahan.....	40
Gambar 2.37: Contoh Penghapusan Pasal dan Ayat pada UU Perubahan.....	40
Gambar 2.38: Urutan Pengelompokan Materi UU Perubahan.....	41
Gambar 2.39: Proses Evaluasi SEI.....	44
Gambar 2.40: Contoh Penghitungan Nilai Recall, Precision, dan F-Measure	46
Gambar 3.1: Arsitektur Sistem.....	47
Gambar 3.2: Proses Pembuatan Aturan	52
Gambar 3.3: Contoh Bagian Judul UU	54
Gambar 3.4: Alur Dari Proses Bagian Pembukaan dan Batang Tubuh	57
Gambar 3.5: Alur Proses Pembuatan Referensi	70
Gambar 3.6: Contoh Identifikasi Token Penanda Frase Pengacuan dan Identifikasi Frase Pengacuan.....	75
Gambar 4.1: Package LegalToXML	83
Gambar 4.2: Keterhubungan antar-kelas dalam Package LegalToXML	83
Gambar 4.3: Package LegalToXML	87
Gambar 4.4: Keterhubungan antar-kelas dalam Package LTX_Scorer	89
Gambar 5.1: Angka ke-15 Pasal I UU Nomor 14 Tahun 1997.....	100
Gambar 5.2: Angka ke-24 Pasal I UU Nomor 17 Tahun 2006.....	101



DAFTAR TABEL

Tabel 3.1: Aturan Pembuatan Elemen XML Bagian Judul.....	55
Tabel 3.2: Daftar Format ID Subbagian pada Bagian Pembukaan	58
Tabel 3.3: Daftar Format ID Subbagian pada Bagian Batang Tubuh UU Non-Perubahan.....	61
Tabel 3.4: Daftar Subbagian pada Bagian Batang Tubuh UU Perubahan	63
Tabel 3.5: Daftar Format ID Subbagian pada Bagian Batang Tubuh UU Perubahan.....	66
Tabel 3.6: Daftar Subbagian pada Bagian Penutup	68
Tabel 3.7: Jenis-Jenis Frase Pengacuan Diidentifikasi Sistem	71
Tabel 3.8: Pola Pengacuan dari Setiap Token Penanda Pengacuan.....	73
Tabel 3.9: Format Elemen REF untuk Frase Pengacuan dengan Target Pengacuan Dokumen Lain dan Dokumen yang Diproses	76
Tabel 3.10: Aturan Pencarian Referensi Aturan Pencarian Referensi	79
Tabel 5.1: Hasil Uji Coba Sistem untuk Dokumen UU Non-Perubahan.....	92
Tabel 5.2: Hasil Uji Coba Sistem untuk Dokumen UU Perubahan	93



DAFTAR LAMPIRAN

LAMPIRAN 1: Tabel Attribute untuk <i>Node</i> dalam Bahasa DOT	109
LAMPIRAN 2: Tabel Attribute untuk <i>Edge</i> dalam Bahasa DOT	110
LAMPIRAN 3: Tabel Attribute untuk Graf dalam Bahasa DOT	111
LAMPIRAN 4: Contoh Undang-Undang Non-Perubahan	113
LAMPIRAN 5: Contoh Undang-Undang Perubahan	115
LAMPIRAN 6: Contoh Dokumen Keluaran untuk Pemrosesan UU Non-Perubahan.....	118
LAMPIRAN 7: Contoh Dokumen Keluaran untuk Pemrosesan UU Perubahan	120
LAMPIRAN 8: Contoh Dokumen DOT	124
LAMPIRAN 9: Aturan Identifikasi Subbagian pada Bagian Batang Tubuh UU Perubahan.....	125
LAMPIRAN 10: Aturan Identifikasi ID Subbagian pada Bagian Batang Tubuh UU Perubahan	134
LAMPIRAN 11: Aturan Pembuatan Elemen XML Bagian Batang Tubuh UU Perubahan.....	140
LAMPIRAN 12: Aturan Identifikasi Subbagian pada Bagian Penutup.....	150
LAMPIRAN 13: Aturan Pembuatan Elemen XML Bagian Penutup	153
LAMPIRAN 14: Aturan Identifikasi Frase Pengacuan	159
LAMPIRAN 15: Aturan Pemberian Elemen Kosong REF.....	167
LAMPIRAN 16: Document Type Declaration	177
LAMPIRAN 17: Contoh Visualisasi Graf dengan Aplikasi Graphviz	178
LAMPIRAN 18: Deskripsi Kelas LTX_Start	179
LAMPIRAN 19: Deskripsi Kelas LTX_Judul.....	181
LAMPIRAN 20: Deskripsi Kelas LTX_Pembukaan.....	182
LAMPIRAN 21: Deskripsi Kelas LTX_BatangTubuh.....	183
LAMPIRAN 22: Deskripsi Kelas LTX_BatangTubuh_Perubahan.....	185
LAMPIRAN 23: Deskripsi Kelas LTX_Penutup	187
LAMPIRAN 24: Deskripsi Kelas LTX_Referensi	188
LAMPIRAN 25: Deskripsi Kelas LTX_Elemen	190
LAMPIRAN 26: Deskripsi Kelas Referensi	191