

DAFTAR ISI

LEMBAR PENGESAHAN	i
LEMBAR PERNYATAAN ORISINALITAS	ii
ABSTRAK	iv
ABSTRACT	v
KATA PENGANTAR	vi
DAFTAR ISI	vii
DAFTAR GAMBAR DAN ILUSTRASI	x
DAFTAR TABEL	xiv
DAFTAR SINGKATAN DAN LAMBANG	xv
BAB I PENDAHULUAN	1
I.1 Latar Belakang	1
I.2 Rumusan Masalah	5
I.3 Tujuan Penelitian	5
I.4 Batasan Penelitian	5
I.5 Manfaat Penelitian	6
I.6 Sistematika Pelaporan	6
BAB II TINJAUAN PUSTAKA	8
II.1 Data Profiling	8
II.2 Proses Data Profiling	12
II.3 Penelitian Terkait	13
II.4 Perangkat Open Source Data Profiling	17
II.4.1 OpenRefine	17
II.4.2 SQL Power Group	17
II.4.3 Talend Open Studio	18
II.4.4 DataCleaner Community Edition	19
II.5 Pentaho Data Integration (Kettle)	19
II.6 Algoritma Clustering	21
II.7 Arsitektur Three-tier Aplikasi	23
II.8 Skenario Testing	24
BAB III METODE PENELITIAN	25

III.1	Konseptual Model	25
III.2	Sistematika Penelitian	26
III.2.1	Fase Inisiasi.....	26
III.2.2	Fase Data Profiling.....	27
III.2.3	Fase Pembuatan Aplikasi Website.....	27
III.2.4	Fase Data Cleansing.....	28
BAB IV	ANALISIS DAN DESAIN.....	29
IV.1	Analisis Pemilihan Aplikasi Sebagai <i>Benchmark</i>	29
IV.2	Analisis Desain Arsitektur Alternatif	33
IV.2.1.1	Arsitektur Aplikasi OpenRefine	33
IV.2.1.2	Arsitektur Aplikasi SQL Power Architect	35
IV.2.1.3	Arsitektur Aplikasi Talend Open Studio	36
IV.2.1.4	Arsitektur Aplikasi DataCleaner Community Edition.....	38
IV.2.1.5	Arsitektur Aplikasi Pentaho Data Integration (Kettle)	39
IV.3	Analisis Kebutuhan Arsitektur Aplikasi Usulan.....	40
IV.4	Arsitektur Tier Aplikasi Usulan.....	41
IV.5	Usecase Diagram	42
IV.5.1	Usecase form.....	44
IV.6	Activity Diagram	47
IV.6.1	Activity Diagram Login	47
IV.6.2	Activity Diagram untuk Admin	47
IV.6.3	Activity Diagram untuk Manajerial	48
IV.7	Class Diagram.....	49
IV.8	Sequence Diagram	50
IV.9	Desain Antarmuka	53
BAB V	IMPLEMENTASI DAN PENGUJIAN.....	58
V.1	Implementasi	58
V.1.1	Pembuatan Koneksi Database pada Kettle.....	58
V.1.2	Implementasi Logika Clustering	61
V.1.3	Implementasi Logika Null/Blank rows	70
V.1.4	Implementasi Job Scheduling	72
V.2	Aplikasi Website	74

V.2.1	Halaman login	74
V.2.2	Halaman dashboard admin	74
V.2.3	Halaman list data source	75
V.2.4	Halaman list clustering	75
V.2.5	Halaman list editing clustering	76
V.2.6	Halaman monitoring admin	78
V.2.7	Halaman list show Null	80
V.2.8	Halaman dashboard manager	80
V.2.9	Halaman monitoring manager	81
V.3	Pengujian	81
V.3.1	Pengujian fungsionalitas	81
V.4	Evaluasi	85
BAB VI KESIMPULAN DAN SARAN		91
VI.1	Kesimpulan	91
VI.2	Kesimpulan	91
DAFTAR PUSTAKA		92