



Hak cipta dan penggunaan kembali:

Lisensi ini mengizinkan setiap orang untuk menggubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

Copyright and reuse:

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

**PENERAPAN TEKNIK *WEB SCRAPING* DALAM
APLIKASI KOMPARASI HARGA KOMPONEN
DAN PERAKITAN KOMPUTER BERBASIS *WEB***

SKRIPSI



Diajukan Guna Memenuhi Persyaratan Memperoleh
Gelar Sarjana Komputer (S.Kom.)

Leo Rizky Julian

11110310018

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNOLOGI INFORMASI DAN KOMUNIKASI
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG**

2015

PERNYATAAN

Dengan ini saya menyatakan bahwa penulisan tugas akhir ini merupakan karya ilmiah saya sendiri, bukan hasil plagiat dari karya ilmiah yang ditulis oleh orang lain atau lembaga lain. Semua karya ilmiah yang dituliskan oleh orang lain maupun lembaga lain yang dirujuk maupun dikutip yang terdapat di dalam penulisan skripsi ini telah disebutkan sumber kutipannya serta dicantumkan pada halaman daftar pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan atau penyimpangan, baik dalam pelaksanaan skripsi maupun dalam penulisan laporan skripsi, saya bersedia menerima konsekuensi dinyatakan tidak lulus untuk mata kuliah Skripsi yang telah saya tempuh.

Tangerang, 1 Juni 2015

UMMN

Leo Rizky Julian

HALAMAN PERSETUJUAN

Skripsi dengan judul

Penerapan Teknik *Web Scraping* dalam Aplikasi Komparasi
Harga Komponen dan Perakitan Komputer Berbasis *Web*

Oleh

Leo Rizky Julian

Telah disetujui untuk diajukan pada Sidang Ujian Skripsi

Universitas Multimedia Nusantara

Tangerang, 1 Juni 2015

Menyetujui,

Dosen Pembimbing

Friska Natalia, Ph.D.

Mengetahui,

Ketua Program Studi Sistem Informasi

Wira Munggana, S.Si., M.Sc.

HALAMAN PENGESAHAN

Skripsi dengan judul

“Penerapan Teknik *Web Scraping* dalam Aplikasi Komparasi
Harga Komponen dan Perakitan Komputer Berbasis *Web*”

Oleh
Leo Rizky Julian

Telah diujikan pada hari Jumat, tanggal 19 Juni 2015, pukul 10.30 s.d.12.00 dan
dinyatakan lulus dengan susunan penguji sebagai berikut

Ketua Sidang

Penguji Ahli

Ir. Raymond Sunardi Oetama, MCIS

Johan Setiawan, S.Kom, MM, MBA

Dosen Pembimbing

Friska Natalia, Ph.D.

Disahkan Oleh
Ketua Program Studi Sistem Informasi

Wira Munggana, S.Si., M.Sc

KATA PENGANTAR

Setiap awal dari setiap perjalanan tentu pada akhirnya nanti juga akan mencapai tujuan akhirnya. Begitu pula dengan perjalanan penulis dalam penulisan tugas akhirnya ini akhirnya telah mencapai tujuan akhirnya yang berarti telah terselesaikannya tugas akhir yang dijalani oleh penulis. Rasa syukur penulis panjatkan yang setinggi – tingginya kepada Tuhan Yang Maha Esa karena atas berkat rahmat dan karunia-Nya penulis dapat menyelesaikan skripsinya yang berjudul “Penerapan Teknik *Web Scraping* dalam Aplikasi Komparasi Harga dan Komponen Komputer Berbasis *Web*”.

Skripsi ini diajukan oleh penulis kepada Program Strata 1, Program Studi Sistem Informasi, Fakultas Teknologi Informasi dan Komunikasi, Universitas Multimedia Nusantara. Penulisan skripsi ini dilakukan penulis untuk dapat membantu para konsumen di Indonesia khususnya konsumen dalam hal pembelian dan perakitan komponen komputer untuk mendapatkan harga yang paling rendah berdasarkan perbandingan dari lima toko.

Namun, seperti halnya dalam perjalanan seseorang dalam mencapai tujuannya tentunya tidak selalu menemui jalan yang mulus. Oleh karena itu, penulis ingin berterima kasih kepada banyak pihak yang telah membantu dalam keadaan jalan yang tidak mulus tersebut sehingga dapat terselesaikannya tugas akhir yang dikerjakan oleh penulis ini. Penulis ingin mengucapkan banyak terima kasih kepada:

1. Ibu Dr. Friska Natalia, S.Kom., M.T. selaku pembimbing penulis dalam menyelesaikan tugas akhir ini yang telah memberikan pengarahan, bimbingan, motivasi, dan juga banyak masukkan ide sehingga penulis dapat menyelesaikan tugas akhir ini dengan baik.
2. Bapak Wira Mungguna, S.Si, M.Sc., Bapak Ir. Raymond Sunardi Oetama, MCIS, Bapak Yustinus Eko Soelistio, S.Kom, MM., Bapak Johan Setiawan, S.Kom, MM, MBA., Ibu Marcelli Indriana, S.Kom, M.Sc. selaku para dosen – dosen hebat program studi sistem informasi yang

telah memberikan masukan serta motivasi dalam penyelesaian tugas akhir yang dilakukan oleh penulis.

3. Bapak Feris Thia dan rekan - rekan dalam *Indonesian Business Intelligence Club Community* yang telah memberikan pengetahuan baru tentang *business intelligence* yang menjadi metode penting dalam perancangan aplikasi yang dibuat.
4. Ibu, Ayah, dan juga Adik yang selalu memberikan doa, motivasi, dan juga semangat kepada penulis agar tugas akhir ini dapat diselesaikan dengan baik.
5. Teman-teman seperjuangan dari program studi sistem informasi Universitas Multimedia Nusantara 2011 yang juga banyak membantu dalam pengisian kuisioner dari penelitian yang dijalani penulis serta memberikan motivasi dan semangat untuk dapat menyelesaikan skripsi dengan baik.
6. Raditya Praba, Basirudin Syarifan Radifan, serta Wimbo Adimas sebagai sahabat dari penulis yang juga turut membantu penulis dalam berbagi pengetahuan dan juga melakukan diskusi dan memberikan masukan untuk terciptanya hasil akhir skripsi penulis menjadi lebih baik.

Penulis berharap perancangan aplikasi yang telah dibuat ini dapat membantu para penggunanya sesuai dengan tujuan penelitian. Dan juga semoga perancangan aplikasi ini dapat menginspirasi serta dapat berguna untuk pengembang selanjutnya.

Tangerang, 1 Juni 2015

Leo Rizky Julian

DAFTAR ISI

PERNYATAAN.....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN.....	iii
KATA PENGANTAR.....	iv
DAFTAR ISI.....	vi
DAFTAR GAMBAR.....	ix
DAFTAR TABEL.....	xiv
ABSTRAK.....	xv
ABSTRACT.....	xvi
BAB I.....	1
1.1. Latar Belakang Masalah.....	1
1.2. Rumusan Masalah.....	2
1.3. Batasan Masalah.....	3
1.4. Tujuan dan Manfaat Penelitian.....	3
1.4.1. Tujuan Penelitian.....	3
1.4.2. Manfaat Penelitian.....	3
1.5. Sistematika Penulisan.....	4
BAB II.....	5
2.1. Teori Pengumpulan Data.....	5
2.2. Teori Pengolahan Data.....	6
2.2.1. <i>Data Mining</i>	6
2.2.2. <i>Data Cleaning dan Data Transformation</i>	7
2.3. Teori <i>Web Based Application</i>	7
2.3.1. Pengertian <i>Web</i>	7
2.3.2. Pengertian Internet.....	8

2.3.3.	<i>Web Based Application</i>	8
2.3.4.	<i>Website Dinamis</i>	9
2.3.5.	<i>Database</i>	10
2.3.6.	<i>Arsitektur Client – Server</i>	11
2.3.7.	<i>PHP Hypertext Preprocessor</i>	11
2.4.	<i>Tools yang Digunakan</i>	12
2.4.1.	<i>Adobe Dreamweaver</i>	12
2.4.2.	<i>Pentaho Data Integration</i>	13
2.4.3.	<i>Yed Graphic Editor</i>	13
2.4.4.	<i>Microsoft Visio</i>	13
2.4.5.	<i>MySql</i>	14
2.5.	<i>Unified Modeling Language Diagram</i>	15
2.5.1.	<i>Sequence Diagram</i>	15
2.5.2.	<i>Activity Diagram</i>	17
2.5.3.	<i>Use Case Diagram</i>	18
2.6.	<i>Flowchart Diagram</i>	19
2.7.	<i>Entity Relationship Diagram</i>	20
2.8.	<i>Waterfall Model</i>	21
BAB III	22
3.1.	<i>Gambaran Umum Objek Penelitian</i>	22
3.2.	<i>Penelitian Sebelumnya</i>	24
3.3.	<i>Metode Penelitian</i>	25
3.3.1.	<i>Requirement Analysis</i>	26
3.3.2.	<i>Design</i>	27
3.3.3.	<i>Implementation / Code</i>	27
3.3.4.	<i>Verification / Testing</i>	28
3.3.5.	<i>Maintenance</i>	28
3.4.	<i>Pengolahan dan Analisis Data</i>	29
3.4.1.	<i>Web Crawling</i>	29
3.4.2.	<i>Web Scraping</i>	30

3.4.3.	<i>Data Cleaning dan Data Transformation</i>	31
3.5.	<i>Unified Modeling Language Diagram</i>	32
3.6.	Spesifikasi Teknik	33
BAB IV		34
4.1.	<i>Requirement Analysis</i>	34
4.1.1.	Referensi Jurnal dan Observasi	34
4.1.2.	Pembahasan Hasil Penyebaran Kuisisioner	36
4.1.3.	Kesimpulan dari Hasil Observasi dan Kuisisioner	43
4.2.	<i>Design</i>	43
4.2.1.	<i>Use Case Diagram</i>	44
4.2.2.	<i>Activity Diagram</i>	48
4.2.3.	<i>Sequence Diagram</i>	57
4.2.4.	<i>Entity Relationship Diagram</i>	62
4.3.	<i>Implementation / Code</i>	65
4.3.1.	Tahap Pengambilan Data	65
4.3.2.	<i>User Interface</i>	112
4.3.3.	Algoritma Penentuan Harga Termurah	120
4.3.4.	Fitur-fitur di dalam Aplikasi	124
4.4.	<i>Verification / Testing</i>	124
4.5.	<i>Maintenance</i>	124
BAB V		127
5.1.	Kesimpulan	127
5.2.	Saran	128
Daftar Pustaka		129
Lampiran		131

DAFTAR GAMBAR

Gambar 2.1 Proses <i>request</i> dan <i>response</i> pada <i>website</i> dinamis	9
Gambar 2.2 Proses <i>request</i> dan <i>response</i> dasar pada arsitektur <i>client – server</i>	11
Gambar 2.3 Logo MySQL	14
Gambar 2.4 Contoh <i>Sequence Diagram</i>	16
Gambar 2.5 Contoh <i>Activity Diagram</i>	17
Gambar 2.6 Contoh <i>Use Case Diagram</i>	18
Gambar 2.7 <i>Waterfall Model</i>	21
Gambar 3.1 Objek Pengambilan Data Harga Komponen Komputer	23
Gambar 3.2 <i>Waterfall Model</i>	25
Gambar 3.3 <i>Flowchart Diagram</i> Proses <i>Web Crawling</i>	29
Gambar 3.4 <i>Flowchart Diagram</i> dari Metode <i>Web Scraping</i>	30
Gambar 3.5 Contoh penggunaan <i>regular expresion</i> untuk <i>data cleaning</i>	32
Gambar 4.1 www.webcrawler.com	35
Gambar 4.2 grafik jenis kelamin responden	36
Gambar 4.3 pekerjaan dari responden.....	37
Gambar 4.4 grafik kepemilikan komputer / laptop oleh responden.....	38
Gambar 4.5 grafik penggunaan komputer / laptop oleh responden	38
Gambar 4.6 grafik pengalaman transaksi <i>online</i> oleh responden	39
Gambar 4.7 <i>pie chart</i> pembelian komponen komputer secara <i>online</i> oleh responden	39

Gambar 4.8 alasan responden melakukan pembelian secara <i>online</i>	40
Gambar 4.9 grafik tingkat pengenalan terhadap <i>website</i> yang digunakan sebagai sumber penarikan data oleh peneliti.....	41
Gambar 4.10 pembelian komponen komputer di suatu <i>website</i> oleh responden ...	41
Gambar 4.11 fitur yang diinginkan oleh konsumen dari aplikasi yang dibuat	42
Gambar 4.12 <i>use case diagram</i> dari aplikasi yang dibuat	44
Gambar 4.13 <i>activity diagram</i> dari <i>use case</i> mencari informasi spesifikasi untuk perakitan komputer.....	48
Gambar 4.14 <i>activity diagram</i> dari <i>use case</i> mencari informasi harga komponen komputer secara individual	50
Gambar 4.15 <i>activity diagram</i> dari <i>use case</i> mencari informasi harga perakitan komputer	52
Gambar 4.16 <i>activity diagram</i> proses jual – beli	54
Gambar 4.17 <i>activity diagram</i> pengelolaan <i>website</i> oleh <i>web administrator</i>	56
Gambar 4.18 <i>sequence diagram user</i>	58
Gambar 4.19 <i>sequence diagram web administrator</i>	60
Gambar 4.20 <i>sequence diagram</i> penjual	61
Gambar 4.21 <i>entity relationship diagram database</i> rakitkomputer	63
Gambar 4.22 <i>homepage</i> www.enterkomputer.com	65
Gambar 4.23 <i>grabbing data</i> pada <i>stage</i> 1 dari <i>website</i> www.enterkomputer.com	66
Gambar 4.24 <i>steps split field to row</i>	67
Gambar 4.25 hasil dari <i>steps split field to row</i>	67
Gambar 4.26 <i>steps regular expression</i>	68
Gambar 4.27 hasil dari <i>steps regular expression</i>	69

Gambar 4.28 <i>steps modified java script value</i>	70
Gambar 4.29 hasil dari <i>steps modified java script value</i>	71
Gambar 4.30 hasil <i>filtering</i> terakhir pada <i>stage 1</i>	71
Gambar 4.31 proses penarikan data menggunakan <i>pentaho</i> pada <i>stage 2</i> dari <i>website</i> <i>www.enterkomputer.com</i>	72
Gambar 4.32 <i>steps regular expression</i> pada <i>stage 2</i>	73
Gambar 4.33 hasil dari <i>steps filter row</i> pada <i>stage 2</i>	74
Gambar 4.34 hasil akhir dari <i>stage 2</i>	76
Gambar 4.35 proses penarikan data menggunakan <i>pentaho</i> pada <i>stage 3</i> dari <i>website</i> <i>www.enterkomputer.com</i>	76
Gambar 4.36 data berhasil tersimpan di dalam <i>database</i>	77
Gambar 4.37 halaman <i>homepage</i> bengkel pc	77
Gambar 4.38 <i>design proses stage 1</i> penarikan data bengkel pc	78
Gambar 4.39 pola <i>regex</i> pada <i>stage 1</i> bengkel pc	79
Gambar 4.40 hasil dari <i>stage 1</i>	80
Gambar 4.41 <i>stage 2</i> penarikan data bengkel pc	80
Gambar 4.42 <i>stage 3</i> penarikan data bengkel pc	82
Gambar 4.43 <i>stage 4</i> penarikan data bengkel pc	83
Gambar 4.44 <i>stage 5</i> penarikan data bengkel pc	85
Gambar 4.45 <i>stage 6</i> penarikan data bengkel pc	86
Gambar 4.46 <i>stage 7</i> penarikan data bengkel pc	87
Gambar 4.47 <i>database</i> bengkel pc	87
Gambar 4.48 <i>homepage</i> dari <i>website</i> <i>quantum</i>	88
Gambar 4.49 <i>stage 1</i> penarikan data <i>website</i> <i>quantum</i>	89
Gambar 4.50 hasil dari <i>stage</i> pertama dalam <i>grabbing website</i> <i>quantum</i>	90
Gambar 4.51 <i>stage 2</i> penarikan data <i>website</i> <i>quantum</i>	91

Gambar 4.52 hasil dari <i>regex</i> pada <i>stage 2</i>	92
Gambar 4.53 <i>stage 3</i> penarikan data <i>website quantum</i>	93
Gambar 4.54 <i>database</i> tabel dari hasil <i>grabbing website quantum</i>	93
Gambar 4.55 tampilan <i>homepage website viraindo</i>	94
Gambar 4.56 <i>stage 1</i> proses penarikan data <i>website viraindo</i>	95
Gambar 4.57 hasil <i>grabbing</i> dari <i>stage 1</i>	96
Gambar 4.58 <i>stage 2</i> proses penarikan data <i>website viraindo</i>	97
Gambar 4.59 tampilan harga produk pada <i>website viraindo</i>	98
Gambar 4.60 <i>stage 3</i> proses penarikan data <i>website viraindo</i>	99
Gambar 4.61 <i>steps removing newlines</i>	100
Gambar 4.62 <i>stage 4</i> proses penarikan data <i>website viraindo</i>	100
Gambar 4.63 <i>homepage website nano komputer</i>	101
Gambar 4.64 <i>stage 1</i> proses penarikan data <i>website nano komputer</i>	102
Gambar 4.65 hasil dari <i>regex</i> pada <i>stage 1 website nano komputer</i>	103
Gambar 4.66 <i>stage 2</i> proses penarikan data <i>website nano komputer</i>	104
Gambar 4.67 hasil <i>regex</i> pada <i>stage 2</i>	105
Gambar 4.68 <i>stage 3</i> proses penarikan data <i>website nano komputer</i>	106
Gambar 4.69 hasil dari <i>regex</i> pada <i>stage 3</i>	107
gambar 4.70 <i>stage 4</i> proses penarikan data <i>website nano komputer</i>	108
Gambar 4.71 <i>stage 1</i> proses penarikan data <i>website nano komputer</i>	110
Gambar 4.72 <i>database nano komputer</i>	110
Gambar 4.73 contoh <i>match table</i> yang dibuat	111
Gambar 4.74 halaman <i>homepage</i> dari aplikasi yang dibuat	112
Gambar 4.75 halaman <i>homepage</i> bagian bawah	112
Gambar 4.76 Navigasi pada bagian bawah halaman <i>homepage</i>	113

Gambar 4.77 menu <i>build guides</i>	114
Gambar 4.78 spesifikasi komponen yang menyusun <i>guides</i>	115
Gambar 4.79 komparasi harga dari <i>guides</i> yang diberikan.....	115
Gambar 4.80 menu simulasi perakitan.....	116
Gambar 4.81 pengecekan komponen yang telah di- <i>input</i> pada <i>field</i> sebelumnya.....	117
Gambar 4.82 rekomendasi harga dari simulasi perakitan	118
Gambar 4.83 tampilan detail produk.....	119
Gambar 4.84 kolom rekomendasi harga	119
Gambar 4.85 menu <i>contact us</i>	120
Gambar 4.86 algoritma rekomendasi harga	122
Gambar 4.87 <i>Flowchart</i> algoritma perbandingan	123
Gambar 4.88 aktifitas perubahan pola <i>scraping</i>	125

UMMN

DAFTAR TABEL

Tabel 4.1 <i>use case scenario</i> informasi spesifikasi untuk perakitan komputer.....	45
Tabel 4.2 <i>use case scenario</i> mencari informasi spesifikasi perakitan komputer ...	45
Tabel 4.3 <i>use case scenario</i> mencari informasi harga perakitan komputer	46
Tabel 4.4 <i>use case scenario</i> proses jual – beli	46
Tabel 4.5 <i>use case scenario</i> pengelolaan <i>website</i> oleh web <i>administrator</i>	46

UMMN

Penerapan Teknik *Web Scraping* dalam Aplikasi Komparasi Harga Komponen dan Perakitan Komputer Berbasis *Web*

ABSTRAK

Oleh: Leo Rizky Julian

Jika pada awalnya komputer hanya digunakan sebagai alat bantu untuk melakukan perhitungan saja, berbeda hal dengan penggunaan komputer pada era modern ini dimana komputer mempunyai fungsi yang berbeda – beda sehingga seiring dengan beragamnya fungsionalitas dari komputer tersebut, maka dibutuhkan pula spesifikasi yang berbeda – beda dari tiap komputer untuk dapat berjalan sesuai dengan fungsinya. Maka dari itulah aplikasi ini dibuat dengan tujuan untuk memberikan solusi agar dalam melakukan perakitan komputer dapat sesuai dengan spesifikasi yang tepat sasaran dan juga dapat menghemat biaya pembelian maupun perakitan karena aplikasi ini juga memberikan komparasi harga berdasarkan lima toko yang dijadikan sumber penarikan data. Hal ini tentunya sesuai dengan prinsip dasar konsumen yang pada dasarnya ingin mendapatkan harga yang terendah akan tetapi juga ingin mendapatkan barang yang kualitasnya paling baik.

Dalam perancangan aplikasi komparasi harga komponen komputer ini diawali dengan melakukan penyebaran kuisioner untuk memastikan bahwa fitur – fitur yan sebelumnya sudah ditentukan oleh peneliti tepat sasaran sesuai dengan kebutuhan para penggunanya. Sementara itu, untuk melakukan penarikan data yang akan dipakai di dalam aplikasi peneliti menerapkan metode *web scraping* dan *web grabbing* sebagai metode dalam melakukan penarikan data. Metode *web scraping* ini dilakukan dengan menggunakan *software* pentaho. Metode ini memungkinkan aplikasi dapat memanfaatkan penarikan data yang dilakukan dari lima toko sumber penarikan data. Hasil dari perancangan aplikasi yang dilakukan ini adalah sebuah *website* yang dikembangkan dengan menggunakan PHP dan juga Javascript dengan menggunakan MySQL sebagai *database*.

Kata kunci: perakitan komputer, komparasi harga, *web scraping*, *website*, pentaho

THE USE OF WEB SCRAPING IN A WEB BASED APPLICATION OF COMPUTER PARTS AND ASSEMBLY PRICE COMPARISON

ABSTRACT

By: Leo Rizky Julian

Originally computers are used only as a tool to perform some calculations. Nowadays computers have a lot of functions to help people finishing their tasks in almost every aspect of human life. Despite computers have a lot of various functions, users also need some different specifications for each computer so they can do their tasks according to computer functionalities. Therefore, the author builds this application with a purpose to recommend a solution to its users in assembling computers that suit their needs. This application also has price comparison feature based on data sources retrieved from five computer shops. So the users can minimize the costs of purchasing PC parts and the computer easier. This comparison feature is based on a basic consumer's principle which is basically consumers want to buy items not only at the better price but also expecting the best quality as possible.

The research starts with the deployment of questionnaires to some respondents who already purchased computer parts online or assembled a computer. The questionnaires are made to assure that all features which have been specified by the author previously is appropriate to the user needs. Afterwards, in order to obtain required data from five computer shops, the author use Pentaho software as a tool to do web scraping and web grabbing method. The result of this research is a web- based application built in PHP and Javascript with MySQL as its database.

Keywords: Computer assembly, price comparison, web scraping, website, pentaho



UMN