IMPLEMENTASI WEB SCRAPING PADA WEBSITE EDUESIA.COM UNTUK PENGUKUR KESENJANGAN JUMLAH MAHASISWA PERGURUAN TINGGI DI INDONESIA

SKRIPSI

Diajukan Guna Memenuhi Persyaratan Memperoleh

Gelar Sarjana Komputer (S.Kom.)

Yuri Pramana

11110310002

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS TEKNOLOGI INFORMASI DAN KOMUNIKASI

UNIVERSITAS MULTIMEDIA NUSANTARA

2015

HALAMAN PENGESAHAN SKRIPSI

IMPLEMENTASI WEB SCRAPING PADA WEBSITE DALAM KESENJANGAN PERGURUAN TINGGI DI INDONESIA

Oleh:

Yuri Pramana

Telah diujikan pada hari Senin, 10 Agustus 2015

Pukul 13.00 s/d 14.30 dan dinyatakan lulus

dengan susunan penguji sebagai berikut

Ketua Sidang **Penguji**

Johan Setiawan, S.Kom., M.M., M.B.A. Marcelli Indriana, S.Kom., M.Sc.

Pembimbing Skripsi

Wira Munggana, S.Si.,M.Sc

Ketua Program Studi Sistem Informasi

Wira Munggana, S.Si.,M.Sc

PERNYATAAN TIDAK MELAKUKAN PLAGIAT

Dengan ini saya:

Nama	: Yuri Pramana
NIM	: 11110310002
Program Studi	: Sistem Informasi

Dengan ini saya menyatakan bahwa skripsi ini adalah karya ilmiah saya sendiri, dan saya tidak melakukan plagiat. Semua kutipan karya ilmiah orang lain atau lembaga lain yang dirujuk dalam skripsi ini telah saya sebutkan sumber kutipannya serta saya cantumkan di Daftar Pustaka.

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

Tangerang, 7 Juli 2015

Yuri Pramana

IMPLEMENTASI WEB SCRAPING PADA WEBSITE EDUESIA.COM UNTUK PENGUKUR KESENJANGAN JUMLAH MAHASISWA PERGURUAN TINGGI DI INDONESIA

ABSTRAKSI

Oleh : Yuri Pramana

Pendidikan merupakan aset penting bagi masa depan masing-masing individu, masyarakat, negara, dan dunia. Namun, tidak semua orang bisa mendapatkan kesempatan yang sama karena kendala dibidang ekonomi dan geografis yang memicu kesenjangan dalam penilaian publik terhadap kualitas intelektual seseorang. Berdasarkan data yang diberikan oleh UNESCO, Indonesia tercatat diurutan ke 4 dari negara ASEAN lainnya, dalam *Gross Enrolment Ratio*. Proses penelitian ini ditujukan untuk membentuk gambaran kesenjangan edukasi pada tingkat perguruan tinggi setiap daerah yang ada di Indonesia. Penelitian ini dibuat dengan teknik *Web scraping*, dan menghasilkan sebuah website yang menyajikan informasi yang valid dan akurat serta analisis rasio *Gross Enrolment Ratio* untuk memperlihatkan tingkat partisipasi pada tingkat perguruan tinggi.

Kata kunci ─ **Data Collection, UMN, Web Crawling, Data Cleansing, Web Scraping**

IMPLEMENTATION OF WEB SCRAPING ON EDUESIA.COM FOR MEASURING THE GAP NUMBER OF COLLEGE STUDENTS IN INDONESIA

ABSTRACT

By: Yuri Pramana

Education is an important asset for the future of each individual, community, nation, and world. However, not everyone can have the same opportunity as the field of economic and geographic constraints that trigger gaps in the public assessment of the quality of one's intellectual. Based on data provided by UNESCO, Indonesia recorded no. 4 from other ASEAN countries, in the level of *Gross Enrolment Ratio*. The research process is intended to form a picture of educational inequality at the college level in every region in Indonesia. This study was made with Web scraping's techniques and produce a website that presents a valid and accurate information and analysis of the ratio of *Gross Enrolment Ratio* to show the level of participation at the college level.

Keywords ─ Data Collection, UMN, Web Crawling, Data Cleansing, Web Scraping

Kata Pengantar

Di setiap tarikan nafas, langkah kaki, dan desiran aliran darah kita, sudah seharusnya kita selalu mengucapkan syukur atas kemudahan dan kenikmatan dalam mencapai tujuan hidup. Begitu pula dengan penulis yang saat ini telah menyelesaikan skripsi yang berjudul "Implementasi Web Scraping Pada Website Dalam Kesenjangan Perguruan Tinggi Di Indonesia" yang di tujukan kepada Program Studi Sistem Informasi, Fakultas Teknologi Informasi dan Komunikasi, Universitas Multimedia Nusantara.

Penulis menyadari bahwa tanpa bantuan dan dukungan dari berbagai pihak. Oleh karena itu penulis ingin menyampaikan rasa terimakasih kepada:

1. Wira Munggana, S. Si., M.Sc. selaku Ketua Program Studi Sistem Informasi di Universitas Multimedia Nusantara dan sebagai Dosen Pembimbing yang selalu memberikan banyak dukungan dalam pembuatan skripsi dan penulisan laporan ini.
2. IR. Raymond Sunardi Oetama, M.C.I.S yang telah memberikan saran, dukungan, dan inspirasi dalam pembuatan dan penulisan laporan skripsi ini
3. Pak Feris Thia, dan tim IBICC yang turut memberikan wawasan mengenai *Pentaho Data Integration*.
4. Wisnu Satyagraha, Theodora Giovani dan Fededi selaku teman penulis dan rekan riset yang melewati suka duka bersama selama pembuatan riset dan penulisan laporan skripsi.
5. Yonathan Hadiputra, Gustave Lyman, Stefanus Hendri Jason Japutra, Aril Rulif, Agustyan Hidayat, Ivan Dermawan, Granodio Pratama, Reisha Pahlevi, dan teman-teman yang turut membantu memberikan saran, semangat, dan mengibur penulis selama pengerjaan laporan ini.

6. Orang Tua, saudara, dan keluarga yang selalu memberikan dukungan doa dan kepercayaan selama pengerjaan laporan skripsi ini.
7. Kepada semua pihak yang telah membantu memberikan dukungan dan tidak dapat disebutkan satu per satu

Penulis menyadari bahwa banyak kekuran pada laporan skripsi. Semoga isi dari laporan skripsi ini dapat memberikan wawasan dan inspirasi bagi pembaca dan pihak-pihak yang membutuhkan

Tanggerang 1 Juni 2015

Yuri Pramana

Daftar Isi

Daftar Gambar