

**WEB SCRAPING SITUS BERITA
MENGUNAKAN BAHASA
PEMOGRAMANPYTHON**

TUGAS AKHIR

Diajukan untuk memenuhi sebagian persyaratan menyelesaikan program Pendidikan Diploma Tiga (D-3)



Oleh:

MUHAMMAD RAMADAN FIKRI

41187005170004

**JURUSAN TEKNIK KOMPUTER D3
FAKULTAS TEKNIK
UNIVERSITAS ISLAM "45"
BEKASI
2022**

PERSETUJUAN TUGAS AKHIR

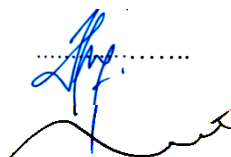
Tugas Akhir ini diajukan oleh:

Nama : Muhammad Ramadan Fikri
NPM : 41187005170004
Program Studi : Teknik Komputer
Fakultas : Teknik
Judul : Web Scraping Situs Berita Menggunakan Bahasa Pemrograman Python

Telah dipertahankan di depan tim penguji sidang Tugas Akhir dan diterima sebagai bagian persyaratan untuk memperoleh Diploma pada Program Studi Teknik Komputer Fakultas Universitas Islam “45” Bekasi.

Bekasi, 29 Juli 2022

Tim Penguji

	Nama	Tanda Tangan
Penguji 1	: Inna Ekawati, S.T., MMSI. NIK. 45116092019002	
Penguji 2	: Malikus Sumadyo, S.Si., MT NIK. 45101051995061

PENGESAHAN TUGAS AKHIR

Tugas Akhir ini diajukan oleh:

Nama : Muhammad Ramadan Fikri
NPM : 41187005170004
Program Studi : Teknik Komputer
Fakultas : Teknik
Judul : Web Scraping Situs Berita Menggunakan Bahasa Pemrograman Python

Telah dipertahankan di depan tim penguji sidang Tugas Akhir dan diterima sebagai bagian persyaratan untuk memperoleh Diploma pada Program Studi Teknik Komputer Fakultas Universitas Islam “45” Bekasi.

Bekasi, 29 Juli 2022

Disetujui oleh

Pembimbing 1



(Rahamdy Trias H.,S.Kom.,M.Kom.,P.hD.)

Pembimbing 2



(Dadan Irwan S.T.,M.Kom.)

Mengetahui,
Ketua Program Studi



(Rahmadya Trias H.,S.Kom.,M.Kom.,P.hD.)

PERNYATAAN KEASLIAN

Saya yang bertandatangan dibawah ini:

Nama : Muhammad Ramadan Fikri
NPM : 41187005170004
Program Studi : Teknik Komputer
Fakultas : Teknik
Judul : Web Scraping Situs Berita Menggunakan Bahasa Pemograman Python

Menyatakan dengan sesungguhnya bahwa tugas akhir ini benar-benar saya kerjakan sendiri. Tugas akhir ini bukan merupakan plagiarisme, pencurian hasil karya milik orang lain, hasil kerja orang lain untuk kepentingan saya karena hubungan material maupun non-material ataupun segala kemungkinan lain yang pada hakekatnya bukan merupakan karya tulis tugas akhir saya secara orisinil dan otentik.

Bila kemudian hari diduga kuat ada ketidak sesuaian antara fakta dengan kenyataan ini, saya bersedia diproses oleh tim Fakultas yang dibentuk untuk melakukan verifikasi, dengan sanksi terberat berupa pembatalan kelulusan/kesarjanaan.

Pernyataan ini saya buat dengan kesadaran sendiri dan tidak atas tekanan ataupun paksaan dari pihak maupun demi menegakan integritas akademik di institusi ini.

Bekasi, 29 Juli 2022

Saya yang menyatakan



Muhammad Ramadan Fikri

MOTTO DAN PERSEMBAHAN

MOTTO

“Waktu bagaikan pedang. Jika kamu tidak memanfaatkannya dengan baik, maka ia akan memanfaatkanmu ”

PERSEMBAHAN

Penulis mempersembahkan Tugas Akhir ini untuk:

1. Allah *Subhanahu waTa'ala* atas karunia dan Rahmat-Nya serta Junjungan Nabi Besar Muhammad *Shallahu'alaihiwasallam* atas perjuangan menegakkan ajaran Islam.
2. Keluarga dan saudara/i tercinta yang selalu senantiasa mendoakan, serta memberikan motivasi dan semangat untuk tetap melakukan terbaik.
3. Teman-teman satu almameter yang telah memberikan pengalaman yang berharga.

KATA PENGANTAR

Bismillaahirrahmaanirrahiim.

Assalamu'alaikum Warahmatullahi Wabarakatuh.

Puji syukur mari kita panjatkan kehadiran Allah SWT, atas segala rahmat, karunia, dan hidayah-Nya sehingga penulis dapat menyelesaikan Laporan Tugas Akhir ini dengan lancar dan baik.

Shalawat serta salam, tak lupa kita curah limpahkan kepada bagin dalam, yakni *habibanawa nabiyyanawa rasulanawa maulana* Muhammad *Shalallahu 'Alaihi Wasalam*, dan kepada keluarganya, para sahabatnya, para *tabi'indan tabi'atnya*, dan kepadakita yang semoga menjadi salah satu dari umatnya yang selaluta'at dan mendapatkan syafa'at hingga *yaumulakhir. Amiinn.*

Laporan Tugas Akhir ini diajukan untuk memenuhi sebagian dari pada menyelesaikan salah satu persyaratan kelulusan program studi Teknik Komputer D-3 dan juga sebagai bentuk pertanggung jawaban kepada pihak Program Studi Teknik Komputer D-3 Fakultas Teknik Universitas Islam "45" Bekasi. Penulis menyadari bahwa butuh waktu banyak dalam menyelesaikan laporan ini karena banyak kendala yang harus diselesaikan. Alhamdulillah, kendala-kendala tersebut dapat dilalui dengan baik berkat peran dan dukungan dari berbagai pihak yang terlibat. Maka dari itu, penulis mengucapkan terimakasih kepada:

1. Allah Subhanahu WaTa'ala;
2. Bapak Rahmadya Trias H., S.T., M.Kom., Ph.D. selaku Ketua Program Studi Teknik Komputer D-3 Fakultas Teknik Universitas Islam "45" Bekasi;
3. Bapak Rahmadya Trias H., S.T., M.Kom., Ph.D. selaku dosen pembimbing pertama yang telah memberikan arahan dan bimbingan selama kegiatan penyusunan laporan ini;

4. Dadan Irwan S.T.,M.Kom. selaku dosen pembimbing kedua yang telah memberikan arahan dan bimbingan selama kegiatan penyusunan laporan ini;
5. Kedua orang tua yang selalu memberikan semangat, motivasi, dan dukungan yang penuh kasih sayang, sehingga penulis dapat menyelesaikan kegiatan tugas akhir dengan baik;
6. Teman-teman seperjuangan dari Teknik Komputer, mulai dari angkatan 2017 hingga 2018 yang telah membantu dikala penulis mengalami kesusahan;
7. Segenap pihak yang terkait yang penulis tidak dapat sebutkan satu persatu, namun tidak menutup rasa hormat dan terimakasih kepada mereka semua, yang telah membantu dalam proses penyusunan laporan tugas akhir.

Dalam penyusunan laporan ini, penulis menyadari bahwa masih banyak kekurangan, baik dalam penulisan maupun isi yang disampaikan. Penulis berharap, tulisan ini dapat bermanfaat.

Demikian yang dapat kami sampaikan, semoga bermanfaat bagi pembaca.

Wassalamu'alaikum Warahmatullahi Wabarakatuh.

Bekasi, 29 Juli 2022



Penulis

ABSTRAK

Perkembangan teknologi saat ini banyak kemajuan dan memberikan inovasi, salah satunya memperoleh informasi dari situs web, salah satu contoh perkembangan teknologi Website portal berita Scraper yang menyediakan kebutuhan data berupa informasi, yang dimana proses pengambilan data informasi dari situs-situs nantinya diambil untuk diamati perilaku dan persepsi untuk mendapatkan segmentasi pasar yang tepat. Pengambilan data untuk saat ini masih banyak dilakukan secara manual, dapat ditemukan memiliki beberapa keterbatasan system yaitu lamanya mengumpulkan data sehingga memperlambat kinerja pengguna. Adapun resiko yang bisa terjadi yaitu tidak mendapatkannya segementasi pasar yang tepat. Di dalam mengatasi masalah tersebut maka diperlukan web scarping situs berita. Pada penelitian *web scraping* situs berita dengan menggunakan bahasa pemrograman python dan bantuan Flask untuk tampilan *web scraping*, *library Selenium* untuk mempermudah dan mengembangkan dalam berinteraksi dengan Web dan menyediakan fasilitas untuk mengontrol suatu peramban *Web (web browser)*. Program ini dapat mengambil data meliputi kata kunci yang diinginkan, judul (jika tertera), tanggal postingnya pada *Website* Berita dan rangkuman, lalu memasukkan data yang telah di ambil sebelumnya ke file berekstensi. CSV yang telah dibuat secara otomatis oleh program.

ABSTRACT

The development of technology is currently a lot of progress and provides innovation, one of which is obtaining information from websites, one example of technological developments is the Scraper news portal website that provides data needs in the form of information, where the process of retrieving information data from sites will be taken to observe behavior and perception to get the right market segmentation. Data retrieval for now is still mostly done manually, it can be found that it has several system limitations, namely the length of data collection so that it slows down user performance. The risk that can occur is not getting the right market segmentation. In overcoming these problems, it is necessary to web scarping news sites. On *web scarping* news sites using the python programming language and the help of Flask for *web scarping*, *Selenium library* to simplify and develop in interacting with the Web and providing facilities to control a *Web (web browser)*. This program can retrieve data including the desired keywords, title (if listed), the date of posting on *Website* and a summary, then enter the data that has been taken previously into a file with the extension . CSV that has been generated automatically by the program.

DAFTAR ISI

PERSETUJUAN TUGAS AKHIR.....	i
PENGESAHAN TUGAS AKHIR.....	ii
PERNYATAAN KEASLIAN	iii
MOTTO DAN PERSEMBAHAN.....	iv
KATA PENGANTAR	v
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xi
DAFTAR TABEL	xii
DAFTAR LAMPIRAN.....	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah.....	2
1.4 Tujuan Penelitian.....	2
1.5 Manfaat Penelitian.....	3
1.6 Sistematika Penulisan	4
BAB II LANDASAN TEORI.....	5
2.1 Penelitian Terdahulu	5
2.2 Konsep Dasar	5
2.2.1 <i>Web Scraper</i>	5
2.2.2 Selenium	6
2.2.3 <i>Selenium WebDriver</i>	8
2.2.4 <i>ChromeDriver</i>	9
2.2.5 Website Portal Berita.....	9
2.2.6 Python	10
2.2.7 CSV	11
2.2.8 DOM(<i>Document Object Model</i>).....	11

2.2.9	Flowchart	12
2.2.10	FLASK	15
BAB III	METODE PENELITIAN	16
3.1	Rancangan Penelitian.....	16
3.2	Analisis, Perancangan dan Implementasi	17
3.3	Implementasi	17
3.4	Identifikasi Kebutuhan <i>Hardware</i> dan <i>Software</i>	17
3.5	Pembuatan <i>web scraping</i>	18
3.6	Pengujiandan Evaluasi	19
BAB IV	HASIL DAN PEMBAHASAN	20
4.1	Hasil.....	20
4.1.1	Analisis masalah.....	20
4.1.2	Solusi masalah.....	20
4.1.3	Perancangan sistem	20
4.1.4	Flowchart <i>Web Scraping</i>	21
4.1.5	Implementasi Pemrograman	21
4.2	Pembahasan.....	27
4.2.1	Hasil <i>Scraping</i> web portal berita.....	27
BAB V	PENUTUP.....	36
5.1	Kesimpulan	36
5.2	Saran	36
	DAFTAR PUSTAKA	37
	LAMPIRAN	39

DAFTAR GAMBAR

Gambar 2. 1 <i>Web Scraping</i>	6
Gambar 2. 2 Logo Selenium	8
Gambar 2. 3 Logo Kompas.com.....	9
Gambar 2. 4 Logo Python	10
Gambar 2. 5 <i>Document Object Model</i>	12
Gambar 2. 6 Logo Flask.....	15
Gambar 3. 1 Rancangan Penelitian.....	16
Gambar 3. 2 Front End.....	18
Gambar 4. 1 <i>Flowchart Web Scraping</i> Portal Berita.....	21
Gambar 4. 2 <i>Install Python</i>	22
Gambar 4. 3 <i>Testing Python</i>	22
Gambar 4. 4 Script Python Web Scraping Portal Berita.....	23
Gambar 4. 5 <i>Import Flask</i>	24
Gambar 4. 6 Kata Kunci Covid-19.....	24
Gambar 4. 7 Proses <i>Scan</i>	25
Gambar 4. 8 Sedang Mencari Data.....	25
Gambar 4. 9 <i>Internet Error</i>	25
Gambar 4. 10 <i>Script Error</i>	26
Gambar 4. 11 <i>File Scraping</i>	26
Gambar 4. 12 Hasil <i>Web Scraping</i>	27

DAFTAR TABEL

Tabel 1. 1 Jenis Selenium	7
Tabel 2. 1 Simbol Flowchart	13
Tabel 4. 1 Hasil Web Scraping.....	27

DAFTAR LAMPIRAN

Lampiran 1: <i>Syntax</i> Program.....	39
--	----