

HALAMAN PENGESAHAN .....	ii
PERNYATAAN BEBAS PLAGIASI .....	iii
HALAMAN PERSEMBAHAN .....	iv
KATA PENGANTAR .....	v
DAFTAR ISI .....	vi
DAFTAR TABEL .....	ix
DAFTAR GAMBAR .....	x
DAFTAR SINGKATAN .....	xii
INTISARI .....	xiii
ABSTRACT .....	xiv
BAB I Pendahuluan .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	2
1.3 Tujuan Penelitian .....	3
1.4 Batasan Masalah .....	3
1.5 Manfaat Penelitian .....	4
1.6 Sistematika Penulisan .....	4
BAB II Tinjauan Pustaka dan Dasar Teori .....	5
2.1 Tinjauan Pustaka .....	5
2.1.1 <i>Reverse Engineering</i> API Privat Instagram melalui <i>DevTools Network Interception</i> .....	5
2.1.2 <i>Scraping</i> Data Instagram melalui Otomasi Peramban .....	5
2.1.3 Pemanfaatan Layanan <i>Third-Party</i> sebagai <i>Instagram Data Provider</i> .....	6
2.1.4 Rangkuman Studi Literatur .....	7
2.2 Dasar Teori .....	11
2.2.1 Python .....	11
2.2.2 JavaScript .....	11
2.2.2.1 Node.js .....	11
2.2.3 Instagram .....	12
2.2.4 <i>Scraping</i> Data .....	12
2.2.5 <i>Reverse Engineering</i> .....	13
2.2.5.1 <i>Reverse Engineering API Private</i> .....	14
2.2.6 <i>Web Scraping</i> .....	14
2.2.7 API ( <i>Application Programming Interface</i> ) .....	14
2.2.7.1 Instagrapi .....	15

2.2.8	Docker .....	15
2.2.9	Prometheus .....	15
2.2.10	Grafana .....	16
2.2.11	Playwright .....	16
2.2.11.1	Patchright .....	16
2.2.12	Puppeteer .....	17
2.2.12.1	Puppeteer Stealth .....	17
2.2.13	Elasticsearch.....	17
2.2.14	Analisis Regresi .....	18
2.2.14.1	Regresi Linier Sederhana .....	18
2.2.15	Uji Friedman .....	18
2.2.16	Uji Peringkat Bertanda Wilcoxon (Wilcoxon Signed-Rank Test) ..	19
2.2.17	Koreksi Bonferroni untuk Perbandingan Ganda .....	20
BAB III Metode Penelitian.....		21
3.1	Alat dan Bahan Tugas akhir .....	21
3.1.1	Alat Tugas akhir.....	21
3.1.2	Bahan Tugas Akhir.....	23
3.2	Metode yang Digunakan.....	23
3.3	Alur Tugas Akhir .....	24
3.3.1	Identifikasi Masalah.....	26
3.3.2	Studi Literatur .....	26
3.3.3	Pemilihan Solusi dan Metode .....	26
3.3.3.1	Pemilihan Metode <i>Reverse Engineering</i> .....	27
3.3.3.2	Pemilihan Metode <i>Web Scraping</i> .....	27
3.3.4	Perancangan Arsitektur .....	29
3.3.5	Pengembangan Metode .....	30
3.3.5.1	Persiapan Data Target .....	30
3.3.5.2	Persiapan Data Akun Buatan.....	31
3.3.5.3	Instalasi Dependensi.....	31
3.3.5.4	Pengembangan Metode.....	31
3.3.5.5	Instalasi dan Konfigurasi Docker .....	32
3.3.6	Eksperimentasi dan Pengujian.....	32
3.3.7	Analisis Data Hasil Pengujian.....	33
3.3.8	Kesimpulan Metode Terbaik.....	34
3.3.9	Pembuatan Laporan .....	34
3.4	Etika dan Batasan Penelitian .....	34
3.4.1	Etika Penelitian.....	34
3.4.2	Batasan Penelitian.....	35
BAB IV Hasil dan Pembahasan.....		36

4.1	Hasil Pengembangan Metode .....	36
4.1.1	Pengembangan Metode Pendekatan <i>Reverse Engineering</i> Menggunakan Instagrapi .....	36
4.1.2	Pengembangan Metode <i>Web Scraping</i> .....	39
4.1.2.1	Pengembangan <i>Scraping</i> Pengambilan Data Postingan..	39
4.1.2.2	Pengembangan <i>Scraping</i> Pengambilan Data Komentar .	42
4.2	Hasil Perbandingan <i>Perfomance</i> .....	44
4.2.1	Hasil CPU .....	44
4.2.2	Hasil RAM .....	48
4.2.3	Hasil <i>Bandwidth</i> .....	51
4.2.3.1	<i>Bandwidth</i> yang Dikirim ( <i>Sent</i> ) .....	51
4.2.3.2	<i>Bandwidth</i> yang Diterima ( <i>Received</i> ).....	52
4.3	Hasil Perbandingan Pengambilan Data .....	53
4.3.1	Kecepatan Pengambilan Data .....	53
4.3.2	<i>Completeness</i> Data .....	58
4.3.3	<i>Coverage</i> Data.....	60
4.4	Hasil Uji Statistik .....	61
4.4.1	Hipotesis Penelitian .....	62
4.4.1.1	Uji Friedman .....	62
4.4.1.2	Uji Lanjutan Wilcoxon dengan Koreksi Bonferroni .....	62
4.4.2	Hasil Analisis .....	62
4.4.3	Interpretasi Hasil .....	63
4.5	Analisis Pengambilan Data dengan Studi Kasus <i>Big Data</i> .....	64
4.5.1	Metodologi Estimasi Kebutuhan Sumber Daya .....	66
4.5.2	Studi Kasus: 5.000.000 Komentar Instagram dalam 24 Jam.....	66
4.5.3	Studi Kasus: 100.000 Postingan Instagram dengan 5 <i>Worker</i> .....	67
4.5.4	Keterbatasan dan Rekomendasi Lanjutan .....	68
BAB V	Kesimpulan dan Saran.....	69
5.1	Kesimpulan.....	69
5.2	Saran.....	69
DAFTAR PUSTAKA	.....	71

Tabel 2.1	Ringkasan Penelitian tentang Teknik <i>Scraping</i> Instagram .....	8
Tabel 3.1	Variabel Monitoring dan Kueri PromQL Sederhana.....	33
Tabel 4.1	Rata-rata Penggunaan CPU (%) menurut Metode, Jenis Data, dan Tingkat Paralelisme .....	47
Tabel 4.2	Rata-rata Penggunaan RAM (GiB) menurut Metode, Jenis Data, dan Tingkat Paralelisme .....	50
Tabel 4.3	Total <i>Bandwidth</i> Terkirim (GB) berdasarkan Metode, Jenis Data, dan Paralelisme .....	51
Tabel 4.4	Kecepatan Rata-rata <i>Bandwidth</i> Terkirim (MB/min) .....	52
Tabel 4.5	Total <i>Bandwidth</i> Diterima (GB) berdasarkan Metode, Jenis Data, dan Paralelisme .....	52
Tabel 4.6	Kecepatan Rata-rata <i>Bandwidth</i> Diterima (MB/min) .....	53
Tabel 4.7	Rata-rata Kecepatan dan Total Data yang Diambil per Metode, Jenis Data, dan Paralelisme .....	57
Tabel 4.8	Perbandingan Kelengkapan Atribut Data Postingan Berdasarkan Metode .....	59
Tabel 4.9	Perbandingan Kelengkapan Atribut Data Komentar Berdasarkan Metode .....	59
Tabel 4.10	Perbandingan Cakupan Relasi dan Kedalaman Konten antar Metode	61
Tabel 4.11	Ringkasan Hasil Uji Statistik Perbandingan Kinerja Metode Scraping	63
Tabel 4.12	Model Regresi Linier ( $y = mx + b$ ) untuk Setiap Metrik, Metode, dan Jenis Data .....	64
Tabel 4.13	Estimasi Kebutuhan <i>Node</i> dan Total Klaster untuk Studi Kasus 5.000.000 Komentar dalam 24 Jam Menggunakan Playwright. ....	67
Tabel 4.14	Estimasi Konsumsi Sumber Daya untuk 100.000 Postingan dengan 5 <i>Worker</i> .....	68

Gambar 3.1	Alur Tugas Akhir .....	25
Gambar 3.2	Alur Rancangan Arsitektur Extract Data .....	29
Gambar 3.3	Alur Pengembangan Metode .....	30
Gambar 4.4	Skema Reverse Engineering melalui MITM Proxy pada Aplikasi Instagram .....	36
Gambar 4.5	Alur Utama Proses <i>Scraping</i> Menggunakan Pendekatan Instagrapi	37
Gambar 4.6	Alur Lanjutan Instagrapi: <i>Login</i> dan Pengambilan Daftar Target .	38
Gambar 4.7	Alur Lanjutan Instagrapi: Pengambilan Postingan dan Komentar.	39
Gambar 4.8	Alur Utama Proses <i>Scraping</i> Postingan Instagram .....	40
Gambar 4.9	Alur Lanjutan Proses Pengambilan Data Postingan .....	41
Gambar 4.10	Alur Utama Proses <i>Scraping</i> Komentar Instagram .....	42
Gambar 4.11	Alur Lanjutan Proses Pengambilan Data Komentar .....	44
Gambar 4.12	Penggunaan CPU pada Proses Pengambilan Komentar dengan Playwright .....	45
Gambar 4.13	Penggunaan CPU pada Proses Pengambilan Post dengan Playwright .....	45
Gambar 4.14	Penggunaan CPU pada Proses Pengambilan Komentar dengan Puppeteer .....	46
Gambar 4.15	Penggunaan CPU pada Proses Pengambilan Post dengan Puppeteer .....	46
Gambar 4.16	Penggunaan CPU pada Proses Pengambilan Data dengan Instagrapi .....	47
Gambar 4.17	Penggunaan Memori pada Proses Pengambilan Post dengan Puppeteer .....	48
Gambar 4.18	Penggunaan Memori pada Proses Pengambilan Komentar dengan Puppeteer .....	48
Gambar 4.19	Penggunaan Memori pada Proses Pengambilan Post dengan Playwright .....	49
Gambar 4.20	Penggunaan Memori pada Proses Pengambilan Komentar dengan Playwright .....	49
Gambar 4.21	Penggunaan Memori pada Proses Pengambilan Data dengan Instagrapi .....	50
Gambar 4.22	Kecepatan Pengambilan Komentar dengan Instagrapi (item/menit)	54
Gambar 4.23	Kecepatan Pengambilan Post dengan Instagrapi (item/menit) ....	54
Gambar 4.24	Kecepatan Pengambilan Post dengan Playwright (item/menit) ....	55
Gambar 4.25	Kecepatan Pengambilan Komentar dengan Playwright (item/menit) .....	55
Gambar 4.26	Kecepatan Pengambilan Komentar dengan Puppeteer (item/menit)	56
Gambar 4.27	Kecepatan Pengambilan Post dengan Puppeteer (item/menit) ....	56
Gambar L.1	Tampilan <i>Network DevTools</i> saat Pengambilan Data Post .....	L-1
Gambar L.2	Tampilan <i>Network DevTools</i> saat Pengambilan Data Komentar ...	L-1
Gambar L.3	<i>Web User Interface</i> Hasil Olah Data Instagram (Versi Pertama) ..	L-2
Gambar L.4	<i>Web User Interface</i> Hasil Olah Data Instagram (Versi Kedua) ....	L-2
Gambar L.5	Tampilan Data Hasil <i>Scraping</i> di dalam Elasticsearch .....	L-3
Gambar L.6	<i>Web User Interface</i> Portainer untuk Manajemen Container .....	L-3



UNIVERSITAS  
GADJAH MADA

**Benchmarking Pendekatan API Reverse-Engineered dan Web Scraping Berbasis Browser untuk Pengambilan**

**Postingan dan Komentar di Instagram**

Haikal Filmi, Widyawan, S.T., M.Sc., Ph.D.; Dr. Indriana Hidayah, S.T., M.T.

Universitas Gadjah Mada, 2025 | Diunduh dari <http://etd.repository.ugm.ac.id/>

Gambar L.7 Tampilan Antarmuka Pengguna Prometheus untuk Monitoring... L-4