

Web Scraper Dan Graphql API Untuk Data Perguruan Tinggi Di Indonesia Berdasarkan Website Kementerian Ristekdikti

(Studi Kasus: Website Kementerian Ristekdikti)

Lingga Tiara, Hadi Syaputra, Widya Cholil & A.Haidar Mirza

Fakultas Teknik Ilmu Komputer, Universitas Bina Darma

Email: linggatiara23@gmail.com

Abstract

The urge to obtain and transmit information is one of the reasons for current technological developments, such as the need for students to access higher education reporting forums or Forlap Dikti. Forlap Dikti is a website page developed by the Ministry of Research, Technology and Higher Education that contains data on student academic activities based on reporting data from universities in Indonesia. Therefore we need a support application to be able to facilitate and expedite the need for access to information on the Forlap Dikti website page using the android application. This study aims to: (1) Help collect student data as material for presenting the required information, (2) Design an application as a medium for finding information on student data at Bina Darma University, (3) With an efficient web scraping function it also helps in data analysis. This application uses puppeteer for scraping data from websites, App Inventor and GraphQL API to display student data.

Keywords: Graphql API, Scraper, Webserver, Web-Scraping, puppeteer

1. Pendahuluan

Internet memiliki keunggulan karena dapat menyediakan informasi yang mudah diakses. Terutama sebagai alat komunikasi yang lebih praktis. Dimana data yang diakses dari internet biasanya berupa halaman web yang berformat HTML (HyperText Markup language). Beberapa diantaranya berisi informasi bagi pengguna seperti tombol navigasi, penelusuran dan tata letak untuk memperindah, serta informasi lebih mudah dibaca. Salah satu pendorong perkembangan Internet adalah mesin pencari.

Mesin pencari secara teratur mengumpulkan informasi dari halaman website melalui bot yang menelusuri web. Mengambil informasi dari halaman web yang ada disebut “web scraping”. Web scraping digunakan untuk mengambil sejumlah besar data dari halaman web, data yang diekstraksi disimpan dalam database dalam format tabel (spreadsheet). Dimana isu-isu yang dikutip dalam buletin pendidikan pada disitus berita online jpnn.com (diakses pada senin 7 february 2021). Dalam berita online tersebut menyajikan berita peningkatan pemalsuan ijazah dan bagi universitas wajib untuk memastikan PIN (Penomoratan Ijazah Nasional). Dikarenakan, jika seseorang dikatakan mahasiswa atau lulusan dan jika datanya tidak tercantum pada halaman Forlap maka keberadaan mahasiswa atau lulusan akan dicurigai.

Disini penulis menggunakan Web Scraper dan GraphQL untuk pendidikan dengan mengekstraksi data dari situs website Forum Pelaporan Perguruan Tinggi, juga dikenal sebagai Forlap Dikti yang biasa digunakan sebagai cara untuk memvalidasi data tentang aktivitas belajar mahasiswa, riwayat profil dosen dan informasi universitas yang ada di

Indonesia. Forlap Dikti sangat penting bagi kelanjutan perguruan tinggi. Hal ini terlihat banyaknya mahasiswa yang mendaftar untuk penelitian, beasiswa, kontes maupun lulusan baru menggunakan semua data pada Forlap. Data akademik mahasiswa tersebut akan dijadikan acuan legalitas untuk data mahasiswa atau ijazah.

Forlap Dikti dapat diakses oleh seluruh lapisan masyarakat dengan mengakses web browser pada halaman (forlap.kemdikbud.go.id) yang dioperasikan oleh Pusat Data dan Informasi Iptek. Berdasarkan Surat Edaran Penerimaan dan kemahasiswaan No700/B/SE/2017 terkait penggunaan Ijazah Nasional, Forlap Dikti menyimpan data mahasiswa dari awal pendaftaran hingga penerbitan nomor seri Penomoran Ijazah Nasional (PIN) dan Sistem Verifikasi Ijazah secara Elektronik (SIVIL) pada 14 Desember 2017, Mahasiswa dan institusi pendidikan akan membutuhkannya dimasa depan. Namun dengan perkembangan teknologi yang semakin pesat, diperlukan terobosan-terobosan baru untuk mengirim dan menerima informasi dengan lebih cepat dan efisien. Kehadiran teknologi mobile semakin memudahkan mahasiswa dan institusi mencari informasi dan komunikasi baik secara verbal, non verbal maupun visual antara jarak dan waktu. Sistem operasi android digunakan sebagai pembaharuan teknologi informasi yang disesuaikan untuk memenuhi kebutuhan informasi suatu instansi. Dimana sering kali terkendala pada saat mengakses web yang membutuhkan Internet dan terkadang terkendala jika laman tersebut akan melakukan maintenance pada server.

Banyak penelitian telah dilakukan pada pengumpuan data dari situs web. Sebagai contoh (Riyadi, 2013) melakukan penelitian “Rancang Bangun REST Web Service untuk Perbandingan Harga Pengiriman dengan Metode Web Scrapping dan Pemanfaatan API”. Pada penelitian ini dibuat aplikasi perbandingan biaya pengiriman TIKI, JNE dan POS Indonesia. Pengguna tidak perlu membuka situs masing-masing penyedia layanan pengiriman atau memasukkan data berulang kali. Pengembangan aplikasi dalam penelitian ini menggunakan teknologi web service REST, web scraping dan penggunaan API. Dalam penelitian ini penggunaan API digunakan untuk memperoleh data harga pengiriman dan harga jasa JNE express. Metode web scraping digunakan dalam penelitian ini untuk menentukan perbandingan harga pengiriman TIKI dan POS Indonesia. Setelah anda memiliki informasi pengiriman kemudian impor ke database. Layanan REST web service digunakan untuk mengakses data pengiriman dalam database. Selanjutnya penelitian yang lebih detail dilakukan oleh (Josi, 2014) dengan judul “Penerapan Teknik Web Scrapping pada Mesin Pencari Artikel Ilmiah”. Aplikasi ini dibuat sebagai database web yang didistribusikan melalui web scraping pada aplikasi yang dibuat dan hasil dari pencarian disimpan dalam tabel menggunakan database MySQL.

Berdasarkan uraian pembahasan diatas, penulis akan mengeksplorasi, merancang dan mengimplementasi aplikasi web scraping pada platform Android. Aplikasi ini mengumpulkan data mahasiswa Univesitas Bina Darma Palembang dari situs Forlap Dikti yang tersimpan dalam database di web server dan membutuhkan android sebagai client untuk melihat data mahasiswa kampus tersebut melalui GraphQL API. Dengan menggunakan aplikasi ini, penulis ingin dengan mudah membantu mahasiswa maupun instansi yang menginginkan atau melihat data kegiatan akademik maupun sebagai legalitas data diri dan ijazah mahasiswa Universitas Bina Darma Palembang dengan lebih mudah.

2. Tinjauan Literatur

Menurut Turland (dalam Josi, 2014) Web Scraping adalah teknik untuk memperoleh informasi dari website yang disalin secara otomatis tanpa harus menyalin secara manual. Tujuan dari web scraper adalah untuk mencari informasi dan kemudian dikumpulkan dalam situs web yang baru. Web scraping berfokus untuk mendapatkan data dengan cara mengambil dan mengekstrak data tersebut. Manfaat dari proses web scraping ialah agar informasi yang diekstraksi lebih terkonsentrasi sehingga membuat pencarian lebih mudah. Aplikasi web scraping berfokus pada cara mendapatkan data mengambil dan ekstrak data berdasarkan ukuran data bervariasi.

Proses web scraping dilakukan untuk mengambil data tertentu secara semi-terstruktur pada halaman web biasanya dalam bahasa markup seperti HTML atau XHTML. Selanjutnya proses akan memulai menganalisis data tersebut dan kemudian data yang dibutuhkan diambil untuk digunakan bagi kepentingan lain.

GraphQL merupakan suatu metode untuk mengakses API dengan menggunakan deskripsi data sesuai dengan permintaan pemakai yang mudah dipahami (<https://graphql.org/>). Ini merupakan bahasa yang bisa diajarkan pada pemakai aplikasi. Aplikasi ini kemudian dapat berkomunikasi dengan layanan backed. Selain itu, GraphQL juga dapat meminta data. GraphQL ini sama dengan JSON dimana sama-sama memiliki operasi untuk membaca kueri dan penulisan data. Buna (dalam Arief Permana sastra, 2020)

Selain itu, GraphQL memberi para pengelola API fleksibilitas untuk menambahkan atau menghambat permintaan yang sudah ada. Pengembang dapat membangun APIs dengan metode apapun yang mereka inginkan dan spesifikasi GraphQL akan memastikan mereka berfungsi dengan cara yang dapat diprediksi bagi klien (Antoni, Fikari & Akbar, 2018; Antoni & Akbar, 2019; Antoni, Jie & Abareshi, 2020; Antoni, Herdiansyah, Akbar & Sumitro, 2021).

Menurut Cahyo Widianarko (2018) Puppeteer merupakan node library yang menyediakan fungsi API untuk bisa mengontrol browser web chromium. Selain itu, Puppeteer dapat dikonfigurasi untuk menjalankan mode non-headless. Berikut cara menggunakan puppeteer sebagai berikut:

1. Melakukan Tangkapan layar dan PDF berdasarkan tampilan pada halaman website.
2. Menjelajahi situs dalam SPA website untuk membuat prerendered content.
3. Melakukan Scraping pada situs website.
4. Mengotomatisasi element situs website, menjalankan pengujian UI testing dll
5. Merekam Runtime Performance Analysis untuk membantu mendiagnosis masalah dalam browser web.

Pada penelitian ini mengacu terhadap penelitian lain yang digunakan sebagai acuan referensi untuk mendukung penelitian yang akan diteliti. Penelitian mengenai pengambilan data dari website banyak dilakukan, Sebagai contoh (Riyadi, 2013) melakukan penelitian “Rancang Bangun REST Web Service untuk Perbandingan Harga Pengiriman dengan Metode Web Scraping dan Pemanfaatan API”. Pada penelitian ini dibuat aplikasi perbandingan biaya pengiriman TIKI, JNE dan POS Indonesia. Pengguna

tidak perlu membuka situs masing-masing penyedia layanan pengiriman atau memasukkan data berulang kali. Pengembangan aplikasi dalam penelitian ini menggunakan teknologi web service REST, web scraping dan penggunaan API. Dalam penelitian ini penggunaan API digunakan untuk memperoleh data harga pengiriman dan harga jasa JNE express. Metode web scraping digunakan dalam penelitian ini untuk menentukan perbandingan harga pengiriman TIKI dan POS Indonesia. Setelah anda memiliki informasi pengiriman kemudian impor ke database. Layanan REST web service digunakan untuk mengakses data pengiriman dalam database. Selanjutnya penelitian dilakukan oleh (Josi, 2014) dengan judul “Penerapan Teknik Web Scrapping pada Mesin Pencari Artikel Ilmiah”. Aplikasi ini dibuat sebagai database web yang didistribusikan melalui web scraping pada aplikasi yang dibuat dan hasil dari pencarian disimpan dalam tabel menggunakan database MySQL.

Penelitian lain mengenai metode ini dilakukan oleh B. Handaga (2019) forlap scraper menggunakan aplikasi android. Aplikasi ini merupakan aplikasi mobile berbasis android untuk sistem informasi yang memungkinkan menggunakan metode web scraping untuk mempermudah dan mempercepat dalam akses data pada website forlap ristekdikti menggunakan metode web scraping, android studio adalah bahasa pemrograman java yang menggunakan IDE dan REST API untuk pertukaran data.

Dalam penelitian ini penulis juga mengacu pada penelitian dilakukan oleh Edy Santoso (2019) mengenai Pengembangan aplikasi lelang ternak burung Lovebird berbasis Android. Sistem pada aplikasi ini beroperasi dengan menggunakan platform android dan proses akuisisi data menggunakan query pada server yang tersedia untuk memudahkan proses akuisisi data. Aplikasi ini terhubung melalui server menggunakan GraphQL API, yang membentuk skema sesuai dengan kebutuhan sistem.

3. Metode Penelitian

3.1 Metode Penelitian

Berdasarkan pada permasalahan yang diteliti, metode yang digunakan dalam penelitian ini adalah metode deskriptif kualitatif (Fauzi, Dencik & Asiati, 2019). Menurut Nana Syaodih Sukmadinata (2011: 73), Penelitian deskriptif kualitatif ditujukan untuk mendeskripsikan dan menggambarkan fenomena-fenomena yang ada, baik bersifat alamiah maupun rekayasa manusia, yang lebih memperhatikan mengenai karakteristik, kualitas, keterkaitan antar kegiatan.

3.2 Metode Pengumpulan Data

Adapun metode pengumpulan data yang dilakukan sebagai berikut:

1) Studi Literatur

Studi literatur dilakukan berdasarkan teori-teori yang mendukung penelitian dari berbagai referensi termasuk buku, paper atau jurnal penelitian, karya ilmiah, dan situs-situs penunjang lainnya untuk membantu dalam menyelesaikan tugas akhir terkait dengan materi penulisan skripsi ini.

2) Observasi

Tahap awal observasi yaitu pengumpulan data dengan melakukan suatu pengamatan secara langsung pada website kementerian ristekdikti yang akan menjadi objek penelitian.

3) Dokumentasi

Dokumentasi bisa catatan maupun berupa foto selama proses penelitian berlangsung. Dokumentasi merupakan teknik pengumpulan data dengan menghimpun dan menganalisis baik dokumen tertulis, gambar maupun elektronik.(Nana Syaodih Sukmadinata, 2009: 221).

3.3 Metode Pengembangan Aplikasi

Metode pengembangan aplikasi yang digunakan pada penelitian ini menggunakan metode Rational Unified Process memiliki empat tahapan yaitu:

- 1) Inception, Tahap ini lebih pada memodelkan proses bisnis yang dibutuhkan (business modeling), mendefinisikan kebutuhan akan sistem yang akan dibuat (requirement) serta analisis dan desain.
- 2) Elaboration, Tahap ini lebih pada analisis dan desain sistem serta implementasi sistem yang fokus pada purwarupa sistem (prototype).
- 3) Construction, Tahap construction berfokus pada pengembangan komponen dan fitur-fitur sistem. Melakukan implementasi dan pengujian sistem yang berfokus pada implementasi perangkat lunak pada kode program (coding).
- 4) Transition, Tahap transition lebih mengarah pada deployment atau instalasi sistem agar dapat dimengerti oleh pengguna. Aktivitas pada tahap ini termasuk pelatihan pengguna, pemeliharaan dan pengujian sistem apakah sudah memenuhi harapan pengguna.

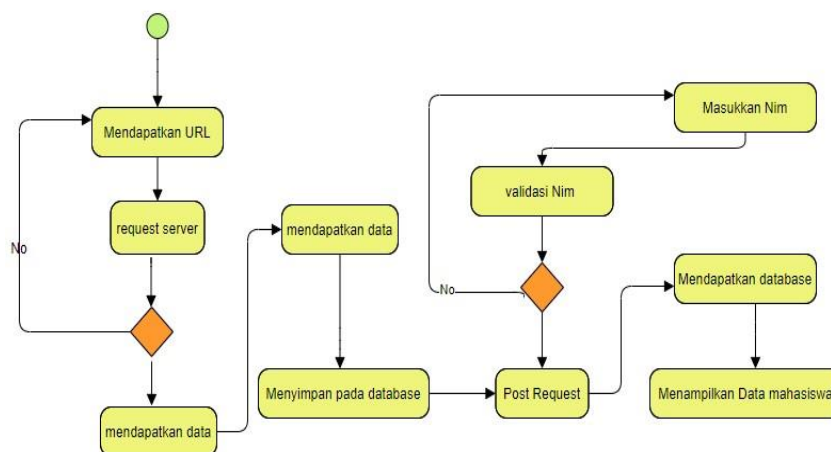
3.4 Analisis dan Perancangan

3.4.1 Inception

Tahap ini lebih pada memodelkan proses bisnis yang dibutuhkan (business modelling) dan mendefinisikan kebutuhan akan sistem yang akan dibuat (requirement).

1) Business Modelling

Business Modelling digunakan untuk menggambarkan model bisnis untuk meninjau, meningkatkan dan membuat sebuah bisnis. Model yang digunakan yaitu *Business Process Model* yang menggambarkan interaksi proses sistem tersebut dengan pihak luar.

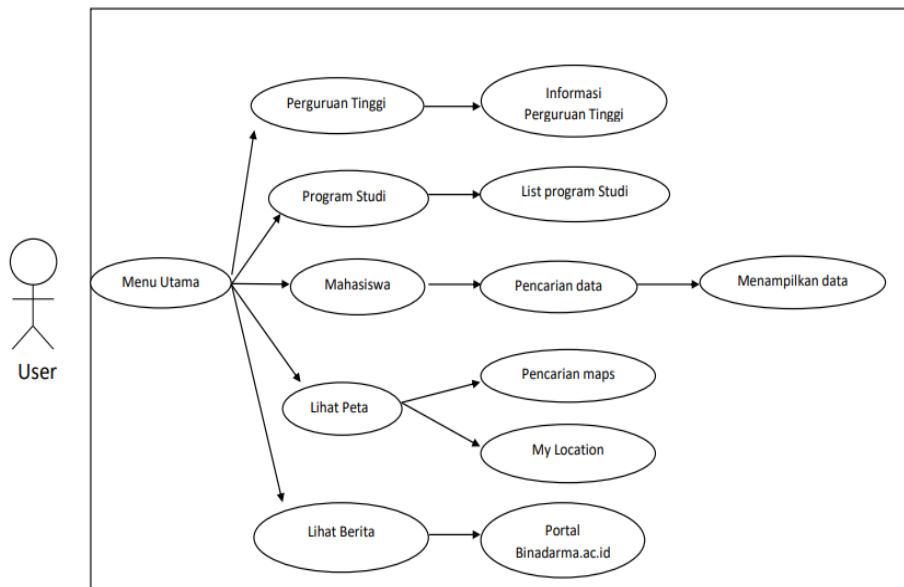


Gambar 1. Business Process Model

Proses ini diawali dengan mendapatkan alamat URL yang diinginkan atau target yang akan diambil informasinya. Setelah mendapat URL server melakukan request untuk mendapatkan data yang kita inginkan yang kemudian menyimpan data tersebut kedalam database. Pada android kita memasukkan nim kemudian server memvalidasi nim tersebut jika ada sesuai maka lanjut ke post request kemudian mendapatkan data kemudian melihat data tersebut, jika tidak ada maka akan kembali untuk memasukkan nim.

2) Requirement

Kegiatan yang dilakukan yaitu menganalisa kebutuhan fungsionalitas sesuai dengan aktor (pengguna). Kebutuhan fungsionalitas disusun berdasarkan pada hasil otomatisasi business case pada gambar 2 menunjukkan terdapat aktor yaitu user atau mahasiswa. Berikut kebutuhan fungsional yang berhasil diperoleh user dapat mengakses lima menu utama yaitu perguruan tinggi, program studi, mahasiswa, lihat peta dan lihat berita. dimana sistem ini berintegrasi pada sistem data mahasiswa Universitas Bina Darma.



Gambar 2. Use Case Diagram

3.4.2 Elaboration

Berdasarkan proses bisnis pada tahap inception kemudian dilanjutkan pada tahap elaboration. Tahap ini lebih difokuskan pada perencanaan arsitektur sistem dengan menggunakan diagram UML yaitu diagram alir (flowchart), activity diagram, dan sequence diagram.

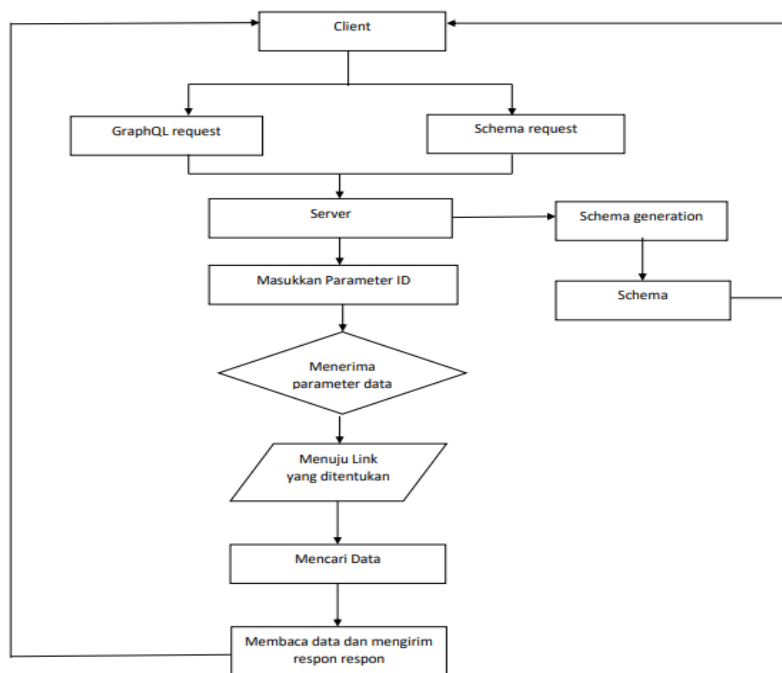
1) Analisis

Dari hasil penelitian yang dilakukan penulis dengan menggunakan aplikasi ini, penulis ingin membantu melihat data diri dan ijazah mahasiswa Universitas Bina Darma Palembang dengan lebih mudah dan efisien.

2) Desain

a. Flowchart GraphQL API

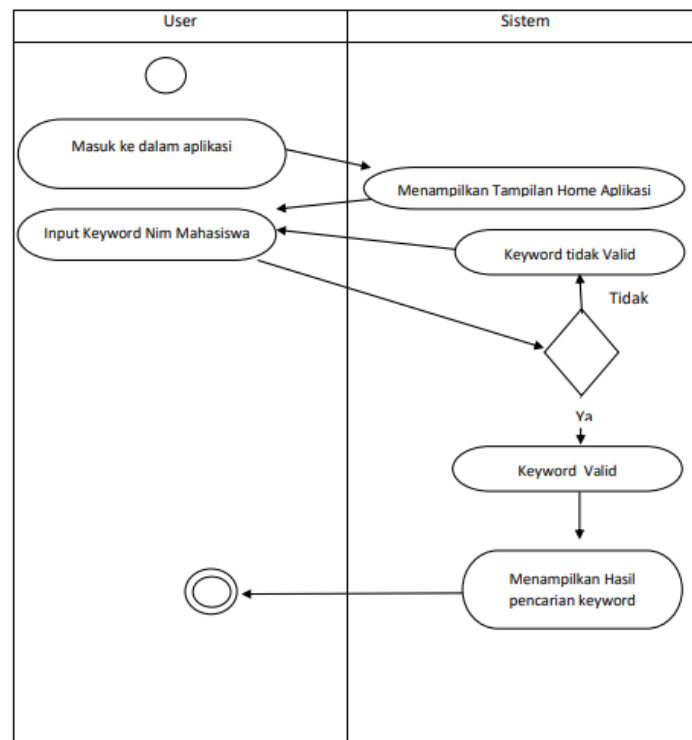
Pada gambar 3 flowchart GraphQL API menjelaskan alur dimulai dari client membuat schema yang dikoding berdasarkan data apa saja yang akan ditampilkan kemudian pada GraphQL request digunakan untuk mengetahui keyword yang akan kita gunakan untuk mencari data yang kita inginkan lalu pada server bisa melihat schema yang kita buat mulai dari nama, nim, jenis kelamin, kampus, jurusan, angkatan, status, tanggal lulus, ijazah dan jumlah sks, kemudian masukan ID Parameter data untuk mengetahui keyword suatu universitas yang ada di Indonesia, lalu parameter menerima data yang akan menuju link yang sudah ditentukan otomatis akan menuju halaman website forlap dikti yang berisi tentang data dari mahasiswa terkait. Mencari data kemudian akan tampil data dan kemudian di mengirimkan data respon tersebut kepada client.



Gambar 3. Flowchart GraphQL API

b. Diagram Activity

Pada gambar 4 adalah diagram activity dalam pencarian data mahasiswa. Proses dimulai ketika pengguna masuk kedalam menu tampilan awal pada aplikasi kemudian menekan tombol mahasiswa selanjutnya memasukkan input nim mahasiswa yang diinginkan dan menekan tombol cek. Maka sistem akan melakukan pengecekan data. Jika data ada didatabase, Sistem akan menampilkan hasil pencarian. Jika tidak maka sistem akan menampilkan notifier data tidak ada cek kembali nim anda.



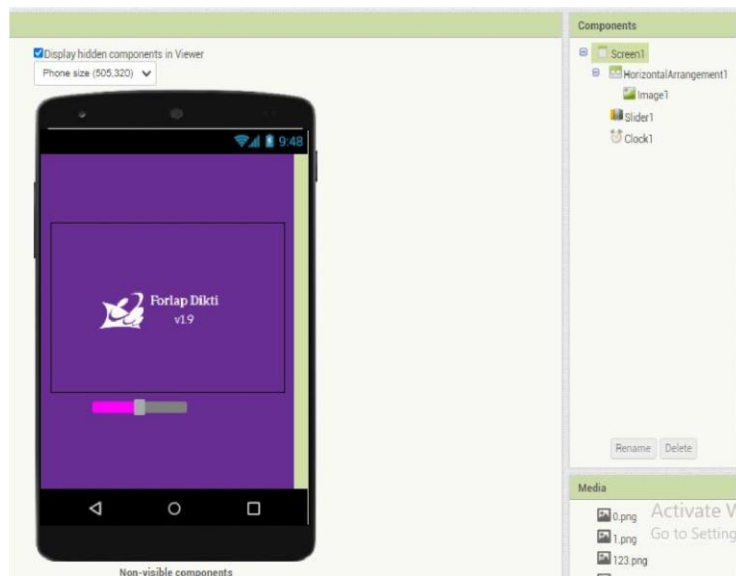
Gambar 4. *Diagram Activity*

3) Desain Struktur Menu

Perancangan antarmuka pengguna merupakan bagian penting dari proses desain aplikasi, karena ini berkaitan dengan tampilan dan interaksi antara pengguna dan aplikasi.

a. Rancangan Tampilan SplashScreen

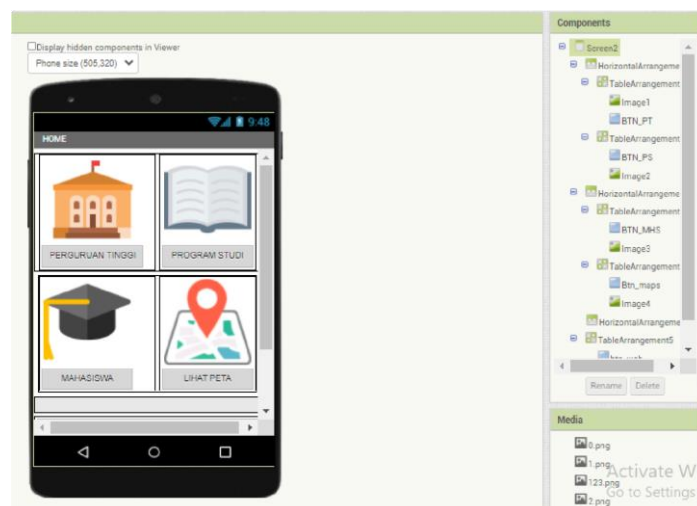
Tampilan SplashScreen merupakan tampilan yang pertama muncul pada saat user akan aplikasi dijalankan. Tampilan ini terdiri dari screen yang memiliki background berupa gambar splashscreen. Pada tampilan splashscreen akan tampil selama 3 detik kemudian tampilan ini akan tertutup dan secara otomatis akan membuka tampilan menu utama.



Gambar 5. Rancangan Tampilan SplashScreen

b. Rancangan Menu Utama

Pada tahap ini dilakukan pembangunan menu utama. Tampilan ini akan muncul pada saat user masuk kedalam sistem aplikasi. Pada tampilan menu utama berisi beberapa icon yang merupakan shortcut untuk membuka tampilan aplikasi yang lainya seperti berbagai fitur yang tersedia yaitu perguruan tinggi, program studi, mahasiswa, lihat peta serta lihat berita yang dapat dipilih oleh pengguna dengan mengklik button yang diinginkan. Berikut rancangan menu utama tersebut dapat dilihat pada gambar 3.8 dibawah ini

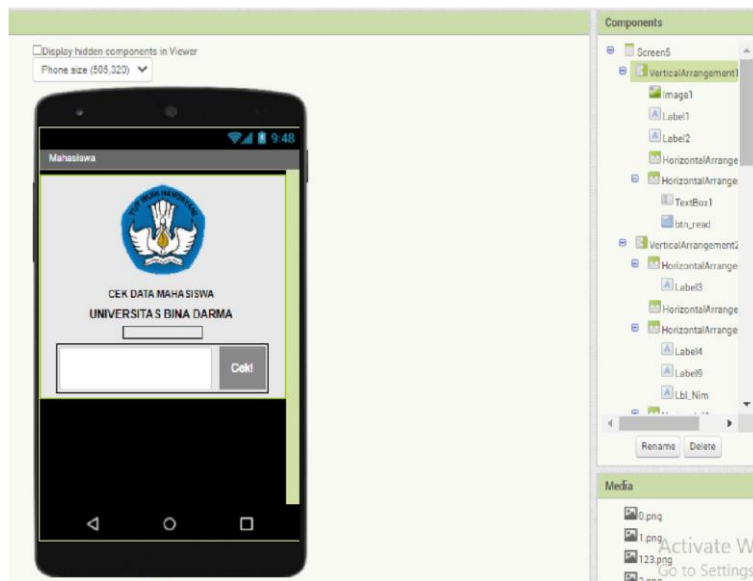


Gambar 6. Rancangan Menu Utama

c. Rancangan Menu Pencarian

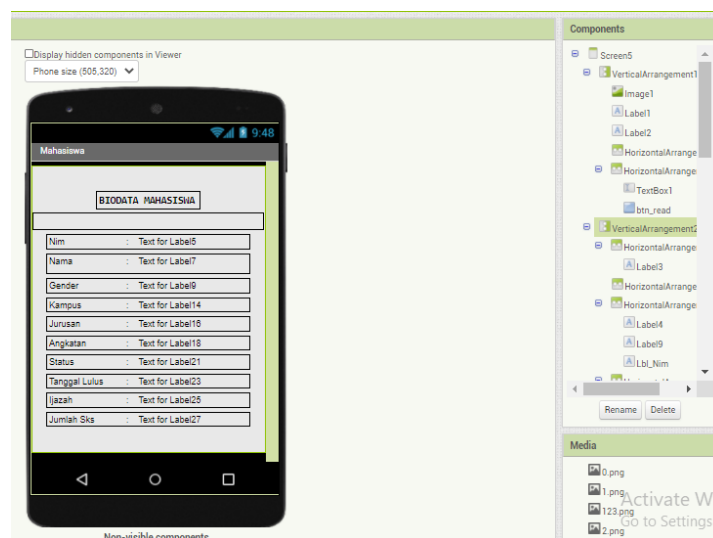
Menu pencarian mahasiswa akan memberikan informasi dan bagaimana pentingnya data mahasiswa untuk terdaftar didalamnya. Sedangkan button Pencarian Data akan menampilkan halaman form pencarian data, form tersebut terdiri dari field text input NIM dan dan button cek. User yang telah melakukan input pada form tersebut akan diproses melalui hosting web server dan akan ditampilkan data mahasiswa dari input yang telah dimasukkan oleh user.

Sedangkan pada button biodata mahasiswa yang terdiri 10 label akan menampilkan informasi tentang nim, nama, gender, kampus, jurusan, angkatan, status, tanggal lulus, ijazah dan jumlah sks pada tampilan aplikasi tersebut. Berikut adalah rancangan interface dari menu pencarian mahasiswa ditunjukkan pada gambar 3.11



Gambar 7. Rancangan Menu Pencarian

d. Rancangan Hasil Biodata Mahasiswa



Gambar 8 Rancangan Hasil Pencarian

4. Hasil dan Pembahasan

4.1 Construction

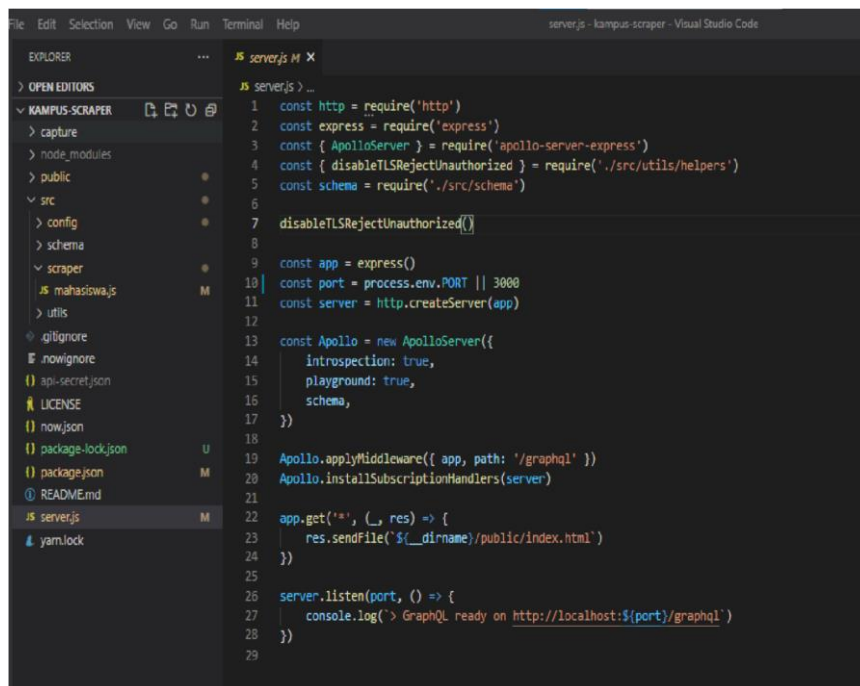
4.1.1 Implementation

Metode yang digunakan adalah metode web scraping dengan menggunakan NodeJs. Proses ini dibagi menjadi beberapa langkah sederhana sebagai berikut:

1) Load HTML

Buat credentials untuk gsheets di GCP Console. Buat file dengan nama apisecret.json kemudian buat 1 file dokumen google spreadsheet, kemudian kasih akses penuh untuk client_email yang didapatkan dari credentials tadi. Copy credentials yang sudah dibuat tadi kedalam api-secret.json, untuk key spreadsheet_id bisa ditambahkan berdasarkan id spreadsheet yang dibuat.

Untuk membuat http server menggunakan Node.js, maka kita perlu mengimport module bawaan Node.js yang bernama http dan express, cara mengimport module adalah dengan mengetikkan kode program require ('namamodule') dan menaruhnya kedalam variabel atau konstanta, contohnya seperti `const http = require('http')` load module http maksud dari potongan kode program itu adalah kita memuat module http kedalam konstanta bernama http. Kemudian pada express framework ini merupakan pengembangan dari NodeJS yang bertugas untuk menangani kebutuhan dalam server-side.



```

1  const http = require('http')
2  const express = require('express')
3  const { ApolloServer } = require('apollo-server-express')
4  const { disableTLSRejectUnauthorized } = require('./src/utlis/helpers')
5  const schema = require('./src/schema')
6
7  disableTLSRejectUnauthorized()
8
9  const app = express()
10 const port = process.env.PORT || 3000
11 const server = http.createServer(app)
12
13 const Apollo = new ApolloServer({
14   introspection: true,
15   playground: true,
16   schema,
17 })
18
19 Apollo.applyMiddleware({ app, path: '/graphql' })
20 Apollo.installSubscriptionHandlers(server)
21
22 app.get('*', (_, res) => {
23   res.sendFile(`${__dirname}/public/index.html`)
24 })
25
26 server.listen(port, () => {
27   console.log(` GraphQL ready on http://localhost:${port}/graphql`)
28 })
29

```

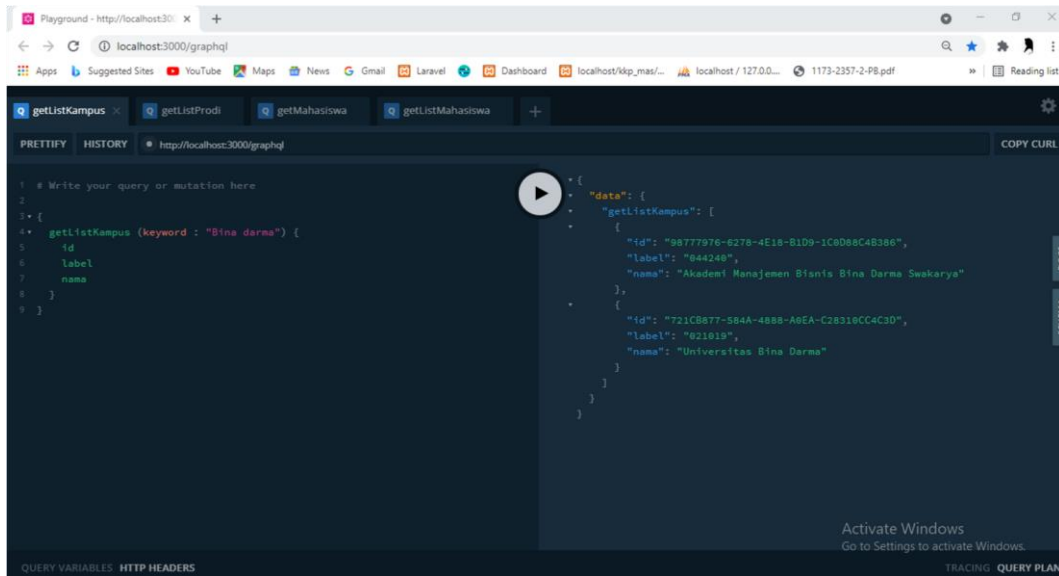
Gambar 9. Server.js

Setelah membuat HTTP Server kita juga harus menambahkan pada port beberapa backend kita akan mendengarkan dan akan memberikan respon, caranya adalah menambahkan listen (). Terakhir pada server.listen (port) akan dijalankan dan selalu dalam “mode stand-by” untuk mendapatkan request yang dikirim oleh user kita menjalankan HTTP Server pada port 3000 atau anda dapat mengakses server dengan mengetikkan pada web browser “http://localhost:3000”.

2) Parsing HTML

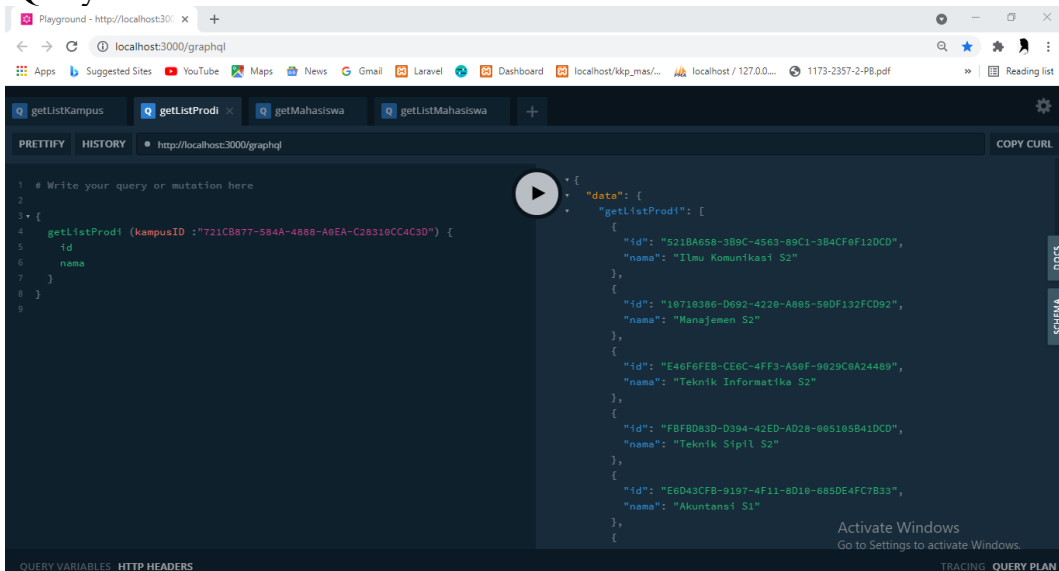
Setelah mendapatkan HTML dari request yang kita scraping lalu parsing untuk mengetahui keyword menggunakan Playground GraphQL API. Parameter kampusID bisa didapatkan dari request getListKampus (keyword) terlebih dahulu dari server GraphQL yang sudah dijalankan, contohnya dapat dilihat pada gambar berikut:

a. Query Get List Kampus



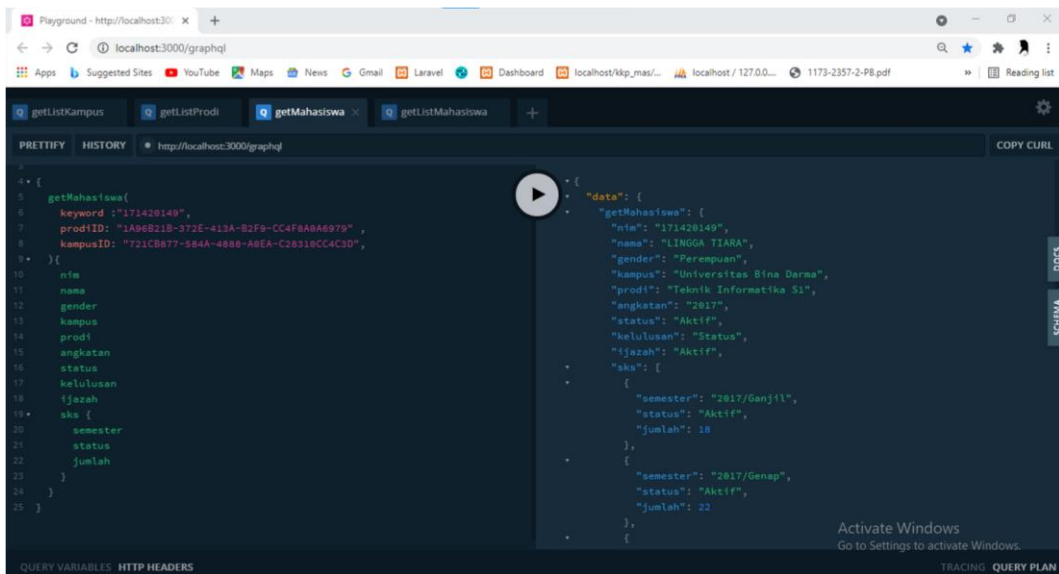
Gambar 10. Query Get List Kampus

b. Query Get List Prodi



Gambar 11. Query Get List Prodi

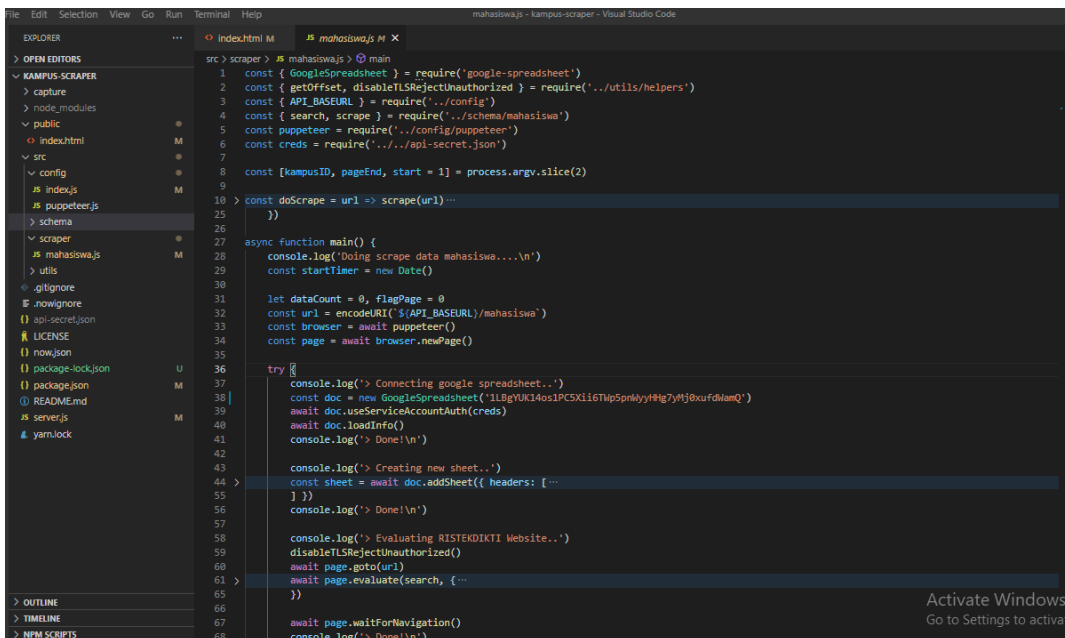
c. Query Get Detail Mahasiswa



Gambar 12. Query Get Detail Mahasiswa

3) Extract HTML

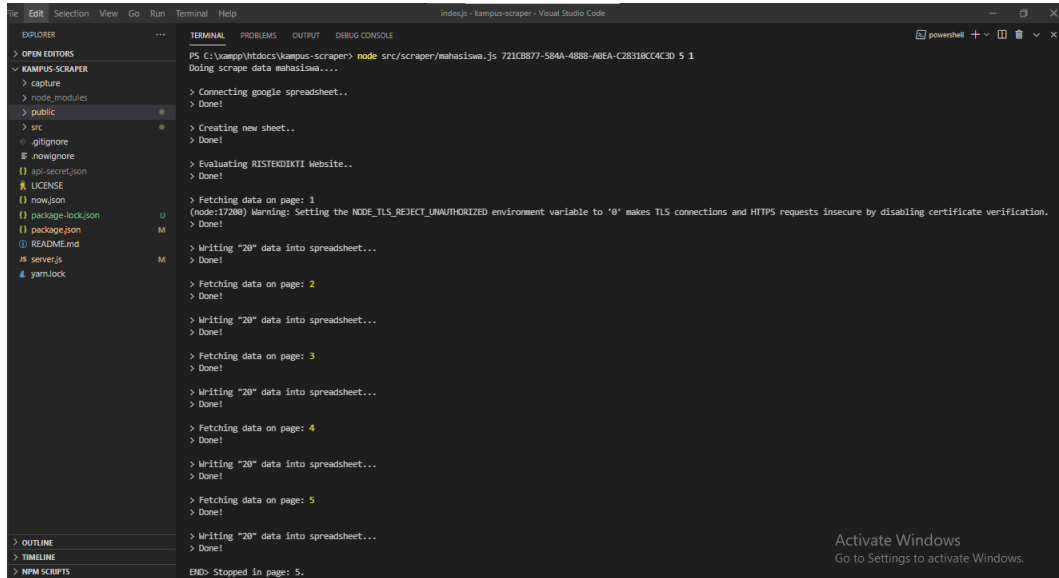
Sebelum melakukan ekstraksi data yang dibutuhkan kita perlu untuk memeriksa struktur HTML dari halaman tersebut yang kita koding kemudian membuat schema dari data yang akan kita ambil dari page tersebut. Data tersebut dapat berupa tampilan yang akan muncul jika kita melakukan proses web scraping.



Gambar 13. Schema Mahasiswa

Jika sudah menemukan Keyword kemudian kita memulai collecting data dari target kampus dengan parameter KampusID yang didapatkan dari server GraphQL yang sudah dijalankan. Contoh query perintah menjalankan proses scraping: `node src/scrapper/mahasiswa.js {kampusID} {batas_halaman} {start_halaman}`.

Data diambil dari situs berdasarkan halaman page yang ada pada website. 1 halaman page berisi data 20 mahasiswa yang langsung terhubung kedalam database, data tersebut dimasukkan kedalam bentuk tabel spreadsheet berdasarkan nama tabel yang tertera. Data ini tidak akan membuat terjadi duplikat karena dilakukan proses scraping dengan per halaman.



Gambar 14. query scraping

Berikut spreadsheet yang sudah collect ratusan ribu mahasiswa Universitas Bina Dharma yang sudah dibuat melalui web scraping dapat dilihat pada link atau gambar berikut: (<https://docs.google.com/spreadsheets/d/1LBgYUK14os1PC5Xii6TWp5pnWyyHHg7yMj0xufdWamQ>)

1	nim	nama	jenisKelamin	kampus	jurusan	angkatan	status	tanggal	Lulus	ijazah	jumlahSKS
2	01141221	A. ROMANI EKA SAPUTRA	Laki-laki	Universitas Bina Dharma	Sistem Informasi S1	2001	Mengundurkan c	1 Januari 2010	Status		115
3	11132339	A DORAJATUN SUPRIADI	Laki-laki	Universitas Bina Dharma	Pendidikan Olahraga S1	2011	Lulus	18 Agustus 2015	0411/IS-IPO/201		149
4	10142301	A FERI ANDRIAN	Laki-laki	Universitas Bina Dharma	Teknik Informatika S1	2010	Lulus	23 Agustus 2014	0401/IS-I/IT/2014		172
5	161460049	A HENDRA ALFIANSYAH	Laki-laki	Universitas Bina Dharma	Teknik Komputer D3	2016	Lulus	7 September 20	5640120190001		112
6	15142018P	A IKHSAN HARAHAP	Laki-laki	Universitas Bina Dharma	Teknik Informatika S1	2015	Universitas Sriwi		Lulus		98
7	06141054	A INO KERTAPATI	Laki-laki	Universitas Bina Dharma	Sistem Informasi S1	2006	Lulus	25 September 21	Status		158
8	04146026	A R. JOHAN	Laki-laki	Universitas Bina Dharma	Teknik Komputer D3	2004	Mengundurkan c	1 September 20	Status		44
9	05171004	A RACHMAN ANGGARA SAI	Laki-laki	Universitas Bina Dharma	Teknik Sipil S1	2005	Lulus	3 Oktober 2009	Status		164
10	13142119	A RIZKY GHAZALI	Laki-laki	Universitas Bina Dharma	Teknik Informatika S1	2013	Lulus	17 Februari 201	034/IS-1/IT/II/201		152
11	13142178	A YOGI	Laki-laki	Universitas Bina Dharma	Teknik Informatika S1	2013	Aktif	Status	Aktif		190
12	141510110	A YUNI ULFAH	Perempuan	Universitas Bina Dharma	Manajemen S1	2014	Lulus	1 September 20	117/S-1/MN/IX/2		148
13	171330006	A'ANG OKTAPIANUS	Laki-laki	Universitas Bina Dharma	Pendidikan Olahraga S1	2017	Aktif	Status	Aktif		148
14	201510084	A. AFRIANSYAH	Laki-laki	Universitas Bina Dharma	Manajemen S1	2020	Aktif	Status	Aktif		26
15	07142129	A. ALFATH	Laki-laki	Universitas Bina Dharma	Teknik Informatika S1	2007	Lulus	24 Maret 2012	Status		176
16	05146112	A. ANDRIFERIUS	Laki-laki	Universitas Bina Dharma	Teknik Komputer D3	2005	Mengundurkan c	25 September 21	Status		88
17	07171035P	A. ARDIANSYAH WIDODO	Laki-laki	Universitas Bina Dharma	Teknik Sipil S1	2007	Tidak diketahui		Mengundurkan c		44
18	05242030	A. ARI GUNAWAN SEPRIAN	Laki-laki	Universitas Bina Dharma	Teknik Informatika S2	2009	Lulus	14 Maret 2012	Status		30
19	052420301	A. ARI GUNAWAN SEPRIAN	Laki-laki	Universitas Bina Dharma	Teknik Informatika S2	2009	Lainnya	13 Mei 2019	Status		39
20	02151018	A. BASID	Laki-laki	Universitas Bina Dharma	Manajemen S1	2002	Mengundurkan c	1 April 2009	Status		24
21	10133111	A. BILAL AL-DI	Laki-laki	Universitas Bina Dharma	Pendidikan Olahraga S1	2010	Mengundurkan c	1 Februari 2016	Status		74

Gambar 15. Hasil Proses Web Scraping

4.1.2 Tampilan Menu

1) Tampilan SplashScreen

Aplikasi ini diawali dengan tampilan splashscreen ini muncul pada saat user akan membuka aplikasi yang akan menampilkan logo dari ristekdikti selama 3 detik kemudian tampilan ini akan tertutup dan secara otomatis akan membuka tampilan menu utama seperti pada gambar berikut:



Gambar 16. Tampilan SplashScreen

2) Tampilan Menu Utama

Tampilan ini akan muncul pada saat user masuk kedalam sistem aplikasi. Aplikasi ini akan menampilkan lima button yang dapat diakses user seperti berbagai fitur yang tersedia yaitu perguruan tinggi, program studi, mahasiswa, lihat peta serta lihat berita yang dapat dipilih oleh pengguna dengan mengklik button yang diinginkan. Tampilan menu utama dapat dilihat pada gambar berikut:



Gambar 17. Tampilan Menu Utama

Pada Gambar diatas menampilkan tampilan menu pencarian data mahasiswa, formulir ini

berisikan field text input NIM dan dan button cek. Pengguna dapat menggunakan kata kunci (keyword) untuk melakukan pencarian data yang diinginkan pada kotak teks yang ada pada menu pencarian.

3) Tampilan Pencarian Data Mahasiswa

Menu pencarian mahasiswa akan memberikan informasi dan bagaimana pentingnya data mahasiswa untuk terdaftar didalamnya. Sedangkan button Pencarian Data akan menampilkan halaman form pencarian data, form tersebut terdiri dari field text input NIM dan dan button cek. User yang telah melakukan input pada form tersebut akan diproses melalui hosting web server dan akan ditampilkan data mahasiswa dari input yang telah dimasukkan oleh user.



Gambar 18. Tampilan Form Pencarian mahasiswa

Pada gambar diatas menampilkan tampilan menu pencarian data mahasiswa, formulir ini berisikan field text input NIM dan dan button cek. Pengguna dapat menggunakan kata kunci (keyword) untuk melakukan pencarian data yang diinginkan pada kotak teks yang ada pada menu pencarian.

4) Tampilan Hasil Biodata Mahasiswa

Pada gambar 4.7 merupakan hasil proses dari inputan keyword berupa nim pengguna pada formulir pencarian mahasiswa kemudian akan diproses dan akan ditampilkan data mahasiswa yang diinputkan oleh pengguna, kemudian akan menampilkan informasi yang berisikan tentang nim, nama, gender, kampus, jurusan, angkatan, status, tanggal lulus, ijazah dan jumlah sks pada tampilan aplikasi tersebut dapat dilihat pada gambar berikut ini:



Gambar 19. Hasil Pencarian Mahasiswa

4.2 Transition

Pengujian sistem merupakan bagian penting dari proses pengembangan perangkat lunak. Pengujian dilakukan untuk memastikan kualitas dan mendeteksi kelemahan perangkat lunak. Tujuan dari pengujian ini adalah untuk memastikan bahwa perangkat lunak yang berkualitas tinggi dan berjalan dengan benar seperti yang diharapkan atau tidak. Hasil pengujian dapat ditunjukkan pada tabel berikut:

Tabel 1. Hasil Pengujian Black Box Testing

Button	Test	Hasil	Kesimpulan
Perguruan Tinggi	Mengklik menu Perguruan Tinggi	Menampilkan informasi atau pengenalan tentang PT	Valid
Program Studi	Mengklik menu Program Studi	Menampilkan list program studi yang ada di PT	Valid
Mahasiswa	Mengklik menu pencarian data	Menampilkan pencarian form data	Valid
	Mengisi form pencarian data	Menampilkan pencarian hasil data	Valid
Maps	Mengklik menu view map	Menampilkan form address yang terintegrasi ke google maps	Valid
	Mengklik menu My Location	Menampilkan form address yang terintegrasi ke lokasi awal	Valid
Lihat Berita	Mengklik menu Bina Darma	Menampilkan integrasi ke link binadarma.ac.id	Valid

5. Kesimpulan

Berdasarkan hasil implementasi aplikasi web scraping ini menggunakan sistem mobile android dengan metode web scraping, maka disini penulis dapat menarik beberapa kesimpulan sebagai berikut:

1. Berdasarkan kesimpulan data aplikasi yang dihasilkan dengan menerapkan teknik web scraping untuk mengekstrak informasi berupa data mahasiswa dari website kementerian ristekdikti berhasil dilakukan.

2. Aplikasi ini berhasil menyimpan otomatis data hasil dari ekstrak data kedalam database bentuk tabel spreadsheet.
3. Hasil pencarian informasi mahasiswa tersebut sesuai dengan keyword yang dimasukkan. Tidak hanya itu aplikasi ini juga bermanfaat untuk menampilkan sejumlah informasi mengenai universitas bina darma Palembang.

Referensi

Antoni, D., Herdiansyah, M. I., Akbar, M., & Sumitro, A. (2021). Pengembangan Infrastruktur Jaringan Untuk Meningkatkan Pelayanan Publik di Kota Palembang. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5(4), 1652-1659.

Antoni, D., Jie, F., & Abareshi, A. (2020). Critical factors in information technology capability for enhancing firm's environmental performance: case of Indonesian ICT sector. *International Journal of Agile Systems and Management*, 13(2), 159-181.

Antoni, D., & Akbar, M. (2019). E-supply chain management value concept for the palm oil industry. *Jurnal Sistem Informasi*, 15(2), 15-29.

Antoni, D., Fikari, D., & Akbar, M. (2018). The readiness of palm oil industry in enterprise resource planning. *Telkomnika*, 16(6), 2692-2702.

Fauzi, F., Dencik, A. B., & Asiati, D. I. (2019). Metodologi Penelitian untuk manajemen dan akuntansi. *Jakarta: Salemba Empat*.

Riyadi. 2013. *Rancang Bangun REST Web Service untuk Perbandingan Harga Pengiriman dengan Metode Web Scrapping dan Pemanfaatan API*. Yogyakarta: Sekolah Tinggi Manajemen Informatika dan Komputer Amikom Yogyakarta

Josi, Ahmat, and Leon Andretti Abdillah. "PENERAPAN TEKNIK WEB SCRAPING PADA MESIN Pencari ARTIKEL ILMIAH," n.d., 6.

Prayitno Agus, Yulia safitri. 2015. *Pemanfaatan Sistem Informasi Perpustakaan Digital Berbasis Website Untuk Para Penulis*. Volume 1 No.1. Jakarta timur: AMIKBSI

Sukmadinata, Nana Syaodih. 2009. *Metode Penelitian Pendidikan*. Bandung: PT. Remaja Rosdakarya.

H. Hesti. 2020. Rancang bangun aplikasi pelaporan kecelakaan lalu lintas berbasis online di Polres Kota Palopo. Skripsi thesis, UNIVERSITAS COKROAMINOTO PALOPO.

Sastra, Arief Permana. (2020) ANALISIS PERBANDINGAN METODE GRAPHQL DAN METODE REST API PADA TEKNOLOGI NODEJS. Skripsi thesis, STMIK

C, widianarko. (2018). Studi Perbandingan website view framework (vue js) dengan vanilla js. Yogyakarta: STMIK AKAKOM

A.S Rosa dan Salahuddin M, 2011. Modul Pembelajaran Rekayasa Perangkat Lunak (Terstruktur dan Berorientasi Objek), Modula, Bandung.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>)