

BAB III PELAKSANAAN KERJA MAGANG

3.1 Kedudukan dan Koordinasi

Selama program kerja magang di Cube Asia, penulis bekerja pada tim *product & tech* pada divisi data engineer. Penulis berada dibawah koordinasi CEO Sarabjit Singh, dan diatur juga dibimbing oleh proyek manager Sangeet Chandalika. Tim data extraction memiliki fungsi untuk mengambil data, membuat dan mengembangkan script untuk mengambil data secara otomatis pada environment AWS Lambda. Tim data extraction berkolaborasi dengan beberapa tim lain seperti tim backend database, dan tim data analyst.



Gambar 3.1 Kedudukan Penulis

Data extraction merupakan salah satu kunci penting dalam berjalannya perusahaan, karena perusahaan Cube Asia berjalan pada data analisis dan memberikan insight ke client. Sehingga proses data extraction sangat lah penting, karena untuk menganalisis sesuatu data diperlukan data aslinya terlebih dahulu, dan tugas tim data extraction yang menyiapkan data tersebut.

Penulis diberikan tugas untuk mengambil data dari tiktok, tokopedia, shopee, dan lazada. Penulis juga diberikan tugas untuk membuat dan mengembangkan script pengambilan data dari berbagai platform tersebut pada environment AWS Lambda, dan pengambilan data dilakukan secara otomatis. Seluruh script yang dibuat menggunakan bahasa pemograman python, untuk pengambilan data nya penulis menggunakan API dari TikAPI dan Zyte. Penulis juga berkoordinasi dengan tim lain untuk menentukan data apa saja yang perlu diambil. Selain itu juga penulis bertanggung jawab untuk membuat dokumentasi pada setiap script yang dibuat dan dikembangkan.

Penulis berkoordinasi dengan tim lain pada divisi data engineer dengan metode sprint. Setiap sprint dirancang agar tim dapat berkerja dengan cepat dan efisien. Setiap development life cycle terdapat beberapa pertemuan:

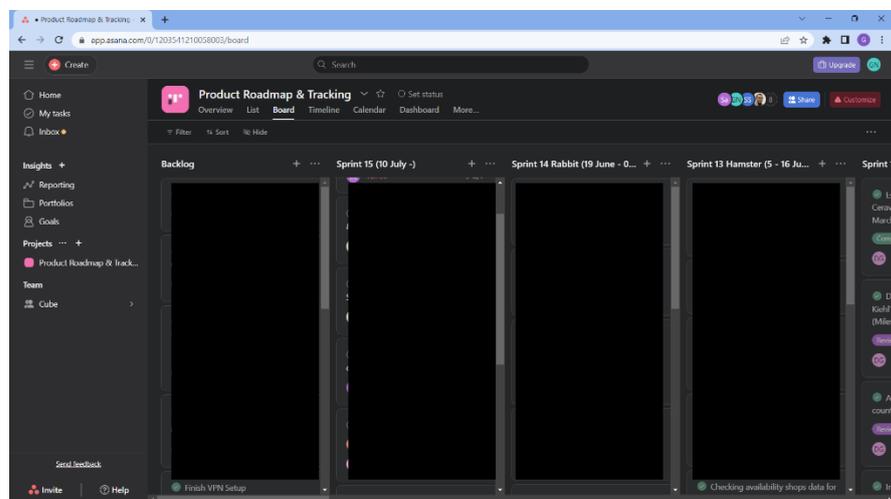
1. Daily check in : diskusi singkat mengenai apa yang sudah dilakukan kemarin dan apa yang akan dilakukan hari ini.
2. Sprint planning : diskusi membahas apa yang akan dilakukan pada sprint tersebut.
3. Sprint review : diskusi membahas apa yang sudah dikerjakan pada sprint tersebut.
4. Adhoc meeting : diskusi membahas hal yang urgent atau harus dibahas saat itu juga.

Untuk beberapa project khusus, project manager akan membuatkan tiket untuk penulis, namun biasanya penulis yang akan membuat tiket nya sendiri, pembuatan tiket tersebut dilakukan saat sprint planning, lalu penulis akan mengerjakan tugas-tugas, dan setiap hari pada daily check in penulis akan

memberitahukan tugas apa yang sudah dikerjakan. Untuk estimasi waktu pengerjaan tiket selalu didiskusikan antara penulis, project manager dan CEO. Pada sprint weekly, seluruh tim akan berkumpul dan membahas tiket-tiket yang sudah dibuat pada sprint planning bersama CEO, seperti tiket apa yang sudah selesai, dan tiket mana yang harus diperpanjang ke sprint selanjutnya. Semua tiket dan task yang sudah dikerjakan semua dirangkum dan disimpan pada aplikasi bernama Asana sebagai alat monitoring dan nantinya akan disampaikan ke CEO.



Gambar 3.2 Logo Asana



Gambar 3.3 Dashboard Asana

3.2 Tugas dan Uraian Dalam Kerja Magang

Beberapa tugas, uraian, kendala dan solusi yang dikerjakan penulis selama kegiatan magang dijelaskan pada sub bab dibawah ini

3.2.1 Tugas yang Dilakukan

Tugas-tugas yang dilakukan penulis selama kegiatan magang di Cube Asia, sebagai berikut:

Tabel 3.1. Tugas Kerja Magang

Week	Onboarding	Data Extraction Manual	Perancangan & Pembuatan Script Data Extraction secara otomatis	Testing & Bug Fixing	Documentation	Weekly Meeting
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						

Dari linimasa tugas kerja magang, dapat dijabarkan tugas yang dikerjakan penulis, yaitu:

1. Onboarding
Tahap ini, penulis belajar bagaimana *script data extraction* berjalan, dan penulis diminta untuk mempelajari dokumentasi dari API TikAPI.
2. Data Extraction Manual
Penulis diminta untuk mengambil data dari API TikApi, menggunakan script yang sudah diberikan kepada penulis secara manual.
3. Perancangan & Pembuatan Script Data Extraction secara otomatis
Penulis diminta untuk merancang script manual yang diberikan saat *onboarding* agar dapat berjalan pada *environment* AWS Lambda
4. Testing & Bug Fixing

Penulis teentunya akan melakukan kegiatan testing dan bug fixing pada scripts yang dibuatnya

5. Documentation

Setelah scripts selesai dan berjalan dengan baik, penulis diwajibkan untuk membuat dokumentasi mengenai scripts tersebut, dokumentasi akan dimasukkan ke gitlab.

6. Weekly Meeting

Penulis akan menghadiri weekly meeting di setiap hari selasa dan jumat untuk menceritakan progress pekerjaannya

Tabel 3.2 Ringkasan Kerja Magang per Minggu

Minggu	Tugas yang Dilakukan
1-2	<ol style="list-style-type: none"> 1. Pengenalan dengan Tim dan Mentor 2. Mempelajari Script python yang diberikan 3. Mempelajari cara kerja TikAPI 4. Menghadiri Weekly Meeting
3-4	<ol style="list-style-type: none"> 1. Mengambil data account dan videos untuk Tiktok 2. Mengirimkan data yang sudah diambil kepada data analyst dan database manager 3. Menghadiri Weekly Meeting
5-6	<ol style="list-style-type: none"> 1. Mengambil data account dan videos untuk Tiktok 2. Mengirimkan data yang sudah diambil kepada data analyst dan database manager 3. Menghadiri Weekly Meeting
7-8	<ol style="list-style-type: none"> 1. Mengambil data account dan videos untuk Tiktok 2. Mengirimkan data yang sudah diambil kepada data analyst 3. Menghadiri Weekly Meeting 4. Mempelajari layanan-layanan AWS khususnya Lambda dan S3
9-10	<ol style="list-style-type: none"> 1. Mencoba menggunakan layanan AWS Lambda dan S3

	<ol style="list-style-type: none"> 2. Membuat scripts yang dapat mengambil data dari TikAPI secara otomatis menggunakan Lambda, dan menyimpan data tersebut ke S3 3. Mengambil data Produk Tokopedia secara manual 4. Menghadiri weekly meeting
11-12	<ol style="list-style-type: none"> 1. Testing dan Bug fixing scripts data extraction untuk tiktok pada AWS 2. Mengambil data produk Tokopedia dan Lazada secara manual 3. Mengirimkan data yang sudah diambil kepada data analyst 4. Menghadiri weekly meeting
13-14	<ol style="list-style-type: none"> 1. Membuat dokumentasi untuk scripts data extraction tiktok di gitlab 2. Mengambil data produk Tokopedia, Lazada, dan shopee secara manual 3. Mengirimkan data yang sudah diambil kepada data analyst 4. Menghadiri weekly meeting
15-16	<ol style="list-style-type: none"> 1. Mengambil data sesuai request dari data analyst 2. Membuat scripts data extraction secara otomatis dari Tokopedia, Lazada, dan Shopee dari Zyte API yang berjalan di AWS Lambda 3. Testing dan Bug fixing scripts data extraction untuk Tokopedia, Lazada, dan shopee pada AWS 4. Menghadiri weekly meeting
17	<ol style="list-style-type: none"> 1. Testing dan Bug fixing scripts data extraction untuk Tokopedia, Lazada, dan shopee pada AWS 2. Membuat dokumentasi untuk scripts data extraction Tokopedia, Lazada, dan Shopee di gitlab 3. Menghadiri weekly meeting

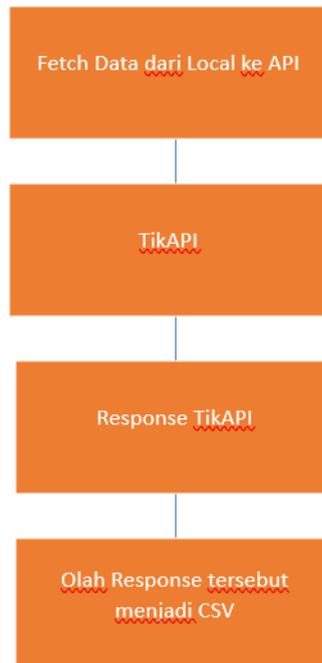
3.2.2 Uraian Pelaksanaan Kerja Magang

Uraian dan gambaran proyek yang dijelaskan dibawah merupakan ilustrasi yang dibuat penulis agar pembaca mendapatkan gambaran hasil proyek. Hal ini dilakukan untuk menjaga kerahasiaan data perusahaan Cube Asia.

Sistem *data extraction* pada Cube Asia sebelum dilakukan automasi adalah, *engineer* akan mengambil data secara manual dengan menjalankan script lalu hasil data tersebut akan dikirimkan ke data analyst dan database manager, yang nantinya dari database manager akan memasukkan ke database. Yang dimana hal tersebut akan membuang *resource* milik *engineer*. Oleh karena itu perusahaan mengusulkan untuk membuat proses *data extraction* ini menjadi otomatis, agar *resource* yang dimiliki *engineer* tidak terpakai, dan bisa dialokasikan mengerjakan tugas lainnya. Berikut ini rincian dari proses penulis membuat automasi dari proses *data extraction* pada Cube Asia:

1. Onboarding dan *Data Extraction* secara manual

Penulis diberikan scripts oleh mentor untuk dipelajari terlebih dahulu, scripts yang diberikan oleh mentor berisi bagaimana kita mengambil data dari TikAPI, kita akan hit API TikAPI untuk meminta data yang kita inginkan.



Gambar 3.4 Alur Pengambilan Data Tiktok

Berikut merupakan *code* yang kita gunakan untuk hit API nya, *code* ini diambil dari dokumentasi tikAPI.

```

from tikapi import TikAPI, ValidationException, ResponseException

api = TikAPI("myAPIKey")

try:
    response = api.public.check(
        username="lilyachty"
    )
    print(response.json())

except ValidationException as e:
    print(e, e.field)

except ResponseException as e:
    print(e, e.response.status_code)
  
```

Gambar 3.5 TikAPI Code

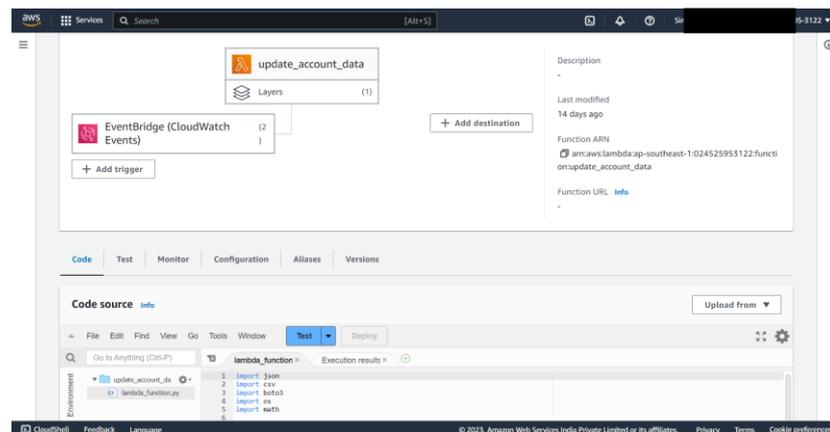
Setelah itu penulis diminta untuk mengambil data secara manual berdasarkan data yang di *request client*. Data yang diambil akan

disimpan ke file CSV, file CSV ini nantinya akan dikirim ke data analyst atau ke database manager.

2. Perancangan dan Pembuatan Script Data Extraction secara otomatis

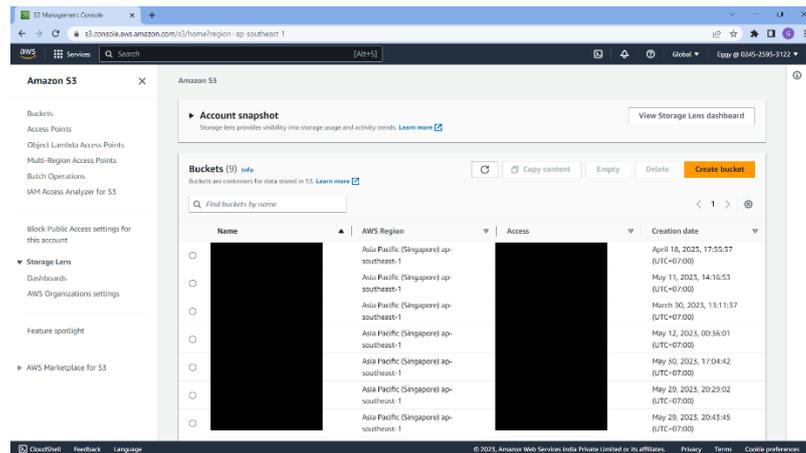
Penulis diminta mempelajari layanan-layanan AWS, khususnya Lambda dan S3. Lambda sendiri merupakan layanan AWS yang memungkinkan kita untuk menjalankan sebuah scripts dengan *serverless*. Sehingga kita tidak perlu menggunakan *resource* milik *engineer* dan resource tersebut dapat digunakan untuk keperluan yang lain. Beberapa hal yang ditemukan penulis mengenai Lambda:

1. Lambda merupakan layanan komputasi yang membuat kita dapat menjalankan scripts tanpa harus memiliki server
2. Lambda hanya memiliki waktu timeout selama 15 menit, setelah 15 menit itu scripts akan di *terminate* secara otomatis
3. Lambda memiliki temporary file untuk melakukan *file processing*



Gambar 3.6 AWS Lambda Dashboard

Penulis juga diminta untuk mempelajari S3 S3 sendiri merupakan layanan database yang diberikan oleh AWS. Penulis mendapatkan pelajaran bagaimana scripts pada Lambda dan mengambil dan memasukkan data ke S3.



Gambar 3.7 AWS S3 Dashboard

Untuk mengkoneksikan Lambda dengan S3 kita akan menggunakan library boto3 yang sudah disiapkan oleh AWS.

```

bucket_name = os.environ['BUCKET_NAME']
client = boto3.client('s3')
s3 = boto3.resource('s3')
bucket = s3.Bucket(bucket_name)

s3_object = s3.Object(bucket_name, key3)
s3data = s3_object.get()['Body'].read().decode('utf-8').splitlines()

lines = csv.reader(s3data)
headers = next(lines)

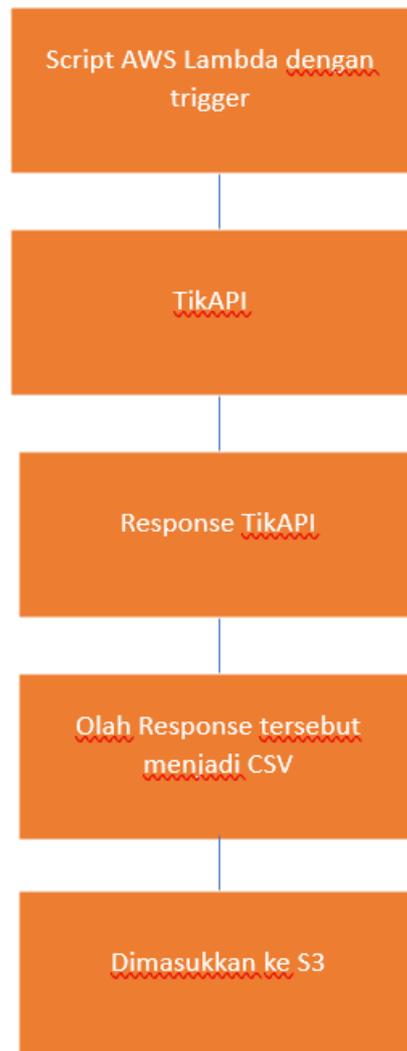
datas = []
for line in lines:
    datas.append(line[1])

```

Gambar 3.8 Code Mengambil Data dari S3

Code pada gambar 3.8 merupakan code untuk kita dapat mengambil data dari s3 menggunakan boto3, karena file yang akan kita proses adalah fil csv, yang dimana format dari csv selalu ada headernya, maka kita harus handle untuk kita tidak mengambil data headernya.

api TikAPI, dan hasil response dari TikAPI, kita olah response tersebut agar dapat dimasukkan ke S3.

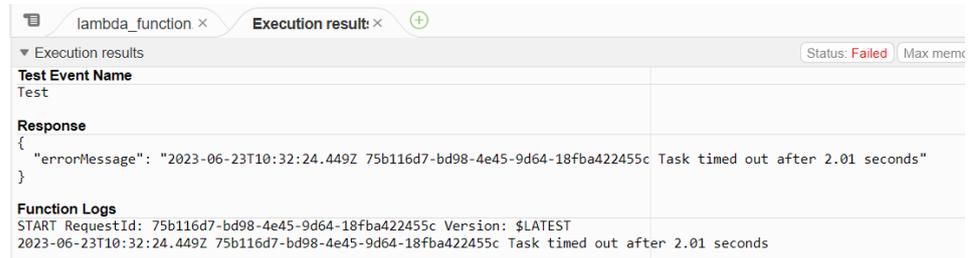


Gambar 3.12 Alur Automasi

3. Testing & Bug Fixing

Setelah penulis membuat scripts, penulis akan melakukan testing terhadap scripts tersebut, testing yang dilakukan adalah penulis akan mencoba mengambil data untuk 1 data terlebih dahulu baik account, video atau product. Lalu setelah data tersebut berhasil tersimpan di S3,

barulah. Testing dimudahkan karena terdapat log dalam AWS Lambda yang menunjukkan penulis dimana *error* pada scripts berada.



Gambar 3.13 Lambda Log

Berikut contoh log pada AWS Lambda yang memudahkan penulis untuk mengetahui dimana error pada scripts tersebut.

Setelah penulis menemukan *error* pada *scripts*, penulis akan langsung mencari cara untuk memperbaiki *error* tersebut, lalu akan melakukan testing kembali sampai tidak ada *error* lagi.

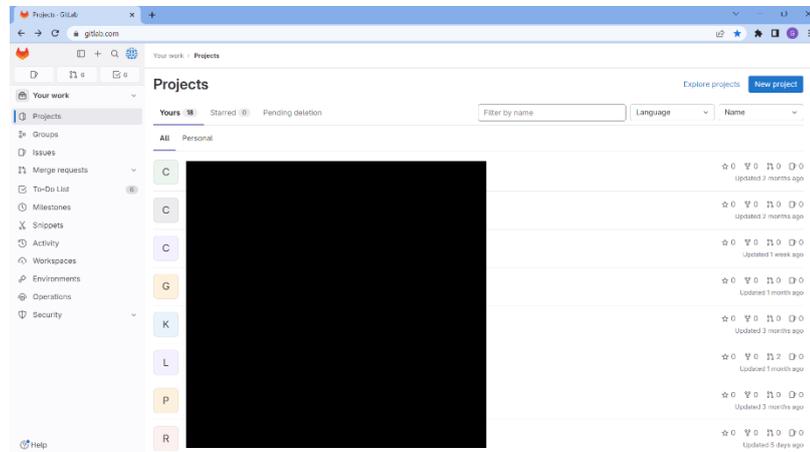
4. Weekly Meting

Penulis akan menghadiri weekly meeting yang diadakan secara *online* di setiap hari selasa dan jumat. Pada weekly meeting ini penulis akan menceritakan tugas-tugas apa saja yang sudah dikerjakan, atau akan dikerjakan untuk 1 minggu kedepan. Selain itu weekly meeting ini juga sebagai tempat penulis untuk demonstrasi scripts automasi *data extraction* buatan penulis kepada tim dan CEO Cube Asia. Selain itu penulis juga berkesempatan untuk menanyakan apakah ada *requirement* lain yang diinginkan CEO dalam *scripts* automasi tersebut.

5. Dokumentasi

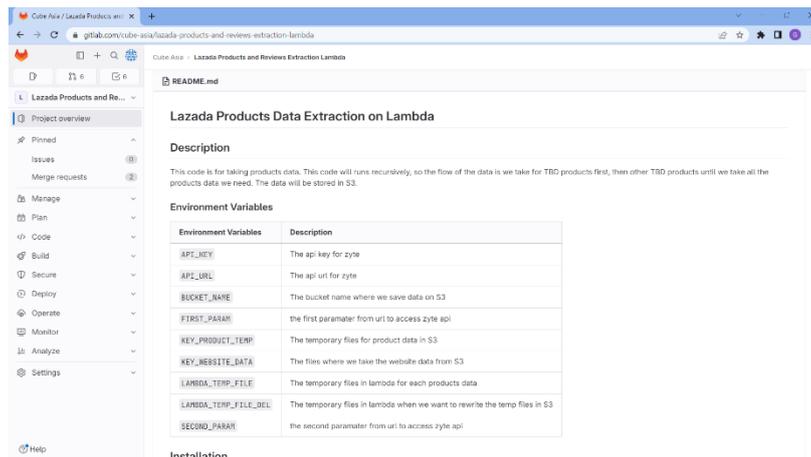
Setelah scripts automasi di terima oleh tim dan CEO Cube Asia, penulis bertanggung jawab untuk membuat dokumentasi yang menjelaskan secara umum bagaimana scripts itu bekerja, kapan scripts berjalan, dan *requirement* apa saja yang dibutuhkan. Dokumentasi

tersebut akan dibuat di repository gitlab, hal ini memudahkan penulis karena gitlab memiliki *web IDE* sehingga penulis tidak perlu *repot-repot* untuk *push* dan *pull* repository github tersebut.



Gambar 3.14 Dashboard GitLab

Dokumentasi yang saya buat merupakan deskripsi tentang script tersebut, lalu bagaimana cara instalasi nya, *library* apa yang digunakan untuk menjalankan script tersebut, dan pendekatan secara umum mengenai script tersebut.



Gambar 3.15 Dokumentasi Scripts

3.2.3 Kendala dan Solusi yang Ditemukan

Ada beberapa kendala yang dialami penulis selama kerja magang di Cube Asia, seperti kurangnya pengetahuan penulis mengenai

teknologi-teknologi yang dipakai, sehingga menghambat proses pengembangan *scripts*.

Solusi yang penulis lakukan adalah penulis akan mencari informasi-informasi terkait, dan melakukan *trial and error* pada *scripts* menggunakan teknologi yang dipakai. Apabila penulis tetap tidak menemukan solusi, penulis akan menanyakan kepada tim atau mentor untuk meminta arahan menyelesaikan masalah tersebut.

Salah satu kendala yang penulis temukan adalah waktu *timeout* lambda yang hanya 15 menit, sedangkan untuk mengambil list data yang banyak akan lebih dari 15 menit. Solusi yang penulis lakukan adalah dengan membuat *scripts* itu menjadi *recursive*, dimana penulis harus mencari informasi mengenai *recursive* dan apakah Teknik *recursive* dapat di implementasikan pada lambda.

Kendala lainnya adalah kurangnya pengetahuan penulis untuk membuat dokumentasi yang baik. Sehingga penulis harus melakukan revisi dokumentasi, dan untuk mengatasi masalah tersebut, penulis mencari informasi di internet, dan melakukan meeting pribadi dengan mentor untuk membahas bagaimana cara membuat dokumentasi yang baik.