

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi pembuatan *web* telah mengalami perubahan yang sangat pesat dalam 5 tahun terakhir. Sejak kemunculannya, *web* telah menjadi platform utama dalam penyebaran informasi, komunikasi, hingga transaksi digital. Seiring meningkatnya kebutuhan pengguna terhadap kecepatan, keamanan, dan interaktivitas, berbagai teknologi baru terus bermunculan untuk mendukung pengembangan *web* yang lebih efisien dan responsif.

Internet sendiri merupakan jaringan yang menghubungkan berbagai perangkat untuk memungkinkan pertukaran informasi secara cepat. Pertukaran informasi ini diatur oleh protokol utama TCP/IP (*Transmission Control Protocol/Internet Protocol*). Namun, informasi yang dikirimkan di internet harus mudah dipahami oleh pengguna, tidak hanya dalam bentuk teks tetapi juga melalui gambar, video, dan suara. Kebutuhan inilah yang mendorong berkembangnya layanan *web* (*World Wide web*), yang memungkinkan penyajian informasi secara lebih interaktif dengan memanfaatkan protokol HTTP (*HyperText Transfer Protocol*).

Teknologi pembuatan *web* semakin beragam dalam perkembangannya, baik dari sisi *front-end* maupun *back-end*. Beberapa teknologi utama yang mendukung pengembangan *web* di antaranya adalah JavaScript, PHP, dan MySQL. Munculnya berbagai *framework* dan pustaka seperti React, Vue.js, dan Node.js juga mempercepat adopsi teknologi baru dalam pengembangan *web* modern. Perubahan ini membuat pentingnya pemantauan tren teknologi *web* agar pengembang dapat memilih teknologi yang sesuai dengan kebutuhan pengguna dan standar industri.

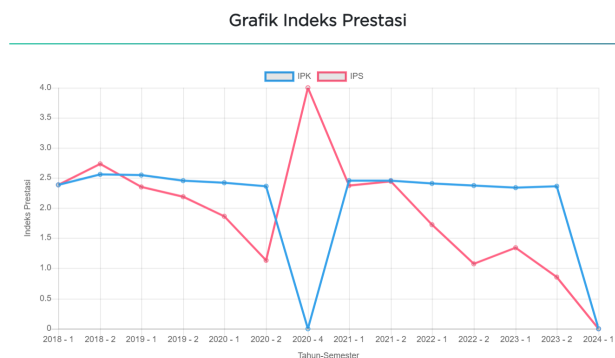
Penelitian ini bertujuan untuk mengetahui tren perkembangan penggunaan teknologi pembuatan *web*. Penelitian ini merupakan pengembangan dari penelitian yang telah dilakukan sebelumnya [1] yang menjawab pertanyaan berapa banyak *web* yang menggunakan teknologi yang jarang dipakai oleh kebanyakan *developer*.

Situs *Http Archive* menyediakan data tentang teknologi yang digunakan dalam pembuatan *web* untuk mencatat perkembangan teknologi ini. Situs ini mengumpulkan data berdasarkan berbagai aspek, seperti pengalaman pengguna dalam mengakses *web*, kecepatan pemuatan halaman, serta tingkat aksesibilitas. Salah satu aspek utama yang diamati dalam penelitian ini adalah *Chrome User Experience Report* (CrUX), yang mengukur tingkat interaktivitas dan kecepatan pemuatan *web* berdasarkan data nyata dari pengguna peramban Google Chrome.

Data dari *Http Archive* kemudian disimpan dalam *Google BigQuery*, layanan penyimpanan dan analisis data berbasis *cloud* yang memungkinkan pemrosesan data dalam skala besar menggunakan *query SQL*. Dengan adanya teknologi ini, analisis terhadap perkembangan teknologi pembuatan *web* dapat dilakukan secara lebih mendalam dan berbasis data yang akurat.

Untuk mempermudah pemahaman terhadap hasil analisis, penelitian ini akan menggunakan visualisasi data dalam bentuk grafik. Salah satu bentuk visualisasi yang digunakan adalah *line chart*, yang dapat menunjukkan tren perubahan teknologi dalam rentang waktu tertentu secara lebih jelas. Contoh *line chart* dapat dilihat pada gambar 1.1.

Perkembangan penggunaan berbagai teknologi *web* dapat divisualisasikan sehingga pola-pola perubahan dapat dikenali dengan lebih mudah dengan menggunakan *line chart*. Selain itu, bentuk



Gambar 1.1: Contoh *line chart*

visualisasi lainnya seperti *bar chart* dan *scatter plot* juga digunakan untuk memberikan perspektif tambahan terhadap data yang dianalisis.

Penelitian ini bertujuan untuk memahami bagaimana tren teknologi pembuatan *web* berkembang dalam 5 tahun terakhir, dari Oktober 2018 hingga Desember 2024. Dengan menggunakan data dari *Http Archive* dan *Google BigQuery*, penelitian ini akan mengeksplorasi perubahan signifikan dalam penggunaan teknologi *web* dan dampaknya terhadap pengalaman pengguna. Hasil analisis ini diharapkan dapat memberikan wawasan bagi pengembang *web* dan industri teknologi dalam memahami arah perkembangan *web* di masa depan.

1.2 Rumusan Masalah

Rumusan masalah yang akan diselesaikan dalam penelitian ini adalah:

1. Bagaimana perkembangan teknologi pembuatan *web* selama 5 tahun terakhir?
2. Bagaimana perkembangan teknologi pembuatan *web* yang banyak digunakan oleh pembuat *web*?
3. Bagaimana cara menyajikan perkembangan teknologi pembuatan *web* kepada pengguna?

1.3 Tujuan

Tujuan yang akan dicapai dari penelitian ini adalah:

1. Mengetahui perkembangan teknologi pembuatan *web* selama 5 tahun terakhir.
2. Mengetahui perkembangan teknologi pembuatan *web* yang banyak digunakan oleh pembuat *web*.
3. Membuat perangkat lunak untuk menyajikan perkembangan teknologi pembuatan *web*.

1.4 Batasan Masalah

Batasan masalah yang diterapkan pada penelitian ini adalah sebagai berikut:

1. Data yang digunakan berasal dari rentang 5 tahun terakhir. Hal ini dimaksudkan untuk membatasi ukuran data agar tidak besar.
2. Data yang akan dianalisis adalah data jumlah penggunaan dan persentase penggunaan. Hal ini dilakukan agar cakupan analisis tidak terlalu besar.

1.5 Metodologi

Metodologi yang digunakan dalam melakukan penelitian ini adalah sebagai berikut:

- Mengumpulkan data penggunaan teknologi pembuatan *web* selama 5 tahun terakhir.

- Membersihkan data dari kolom dan baris yang tidak digunakan.
- Melakukan analisis dengan menggunakan data dengan skala lebih kecil.
- Melakukan analisis dengan menggunakan data yang sebenarnya.
- Membuat perangkat lunak untuk menampilkan hasil analisis secara interaktif.

1.6 Sistematika Pembahasan

Sistematika pembahasan tugas akhir ini adalah:

1. Bab 1: Pendahuluan
Membahas latar belakang, tujuan, rumusan masalah, batasan masalah, dan metodologi penelitian yang digunakan.
2. Bab 2: Landasan Teori
Membahas *web*, *Http Archive*, bahasa SQL, *Google Big Query*, dan visualisasi data yang digunakan.
3. Bab 3: Analisis Masalah
Membahas tentang analisis masalah dan solusinya dan melakukan analisis dengan menggunakan data yang skalanya lebih kecil.
4. Bab 4: Penambangan Data
Membahas eksplorasi dan analisis data dengan menggunakan data *real*.
5. Bab 5: Pembuatan perangkat lunak dan Peluncuran Model
Membahas tentang pembuatan perangkat lunak dan pengujian fungsional perangkat lunak untuk menampilkan hasil analisis secara interaktif.
6. Bab 6 : Kesimpulan dan Saran
Membahas tentang kesimpulan dari hasil penelitian yang telah dilakukan dan saran agar penelitian ini lebih baik.

BAB 2

LANDASAN TEORI

Bab ini akan berisikan tentang beberapa teori dari metode atau hal-hal yang diperlukan dalam melakukan penelitian ini seperti apa itu *Http Archive*, SQL, statistika, dan visualisasi data.

2.1 *Http Archive* [2]

Http Archive merupakan sebuah situs *web* yang menyediakan catatan dari kurun waktu tertentu. Hal ini dilakukan agar pemilik halaman *web* dapat melihat kembali hal yang sudah terjadi, perilaku yang sedang terjadi, dan menemukan tren yang akan muncul.

2.1.1 *Reports*

Reports berisi informasi terperinci mengenai sumber daya yang diambil, API dan fitur platform yang digunakan, serta jejak eksekusi dari setiap halaman dari situs-situs teratas yang ada di *web*. Informasi yang telah didapatkan kemudian diolah dan dianalisis untuk melihat perkembangan tren. *Reports* yang dimiliki oleh situs *Http Archive* dibagi menjadi beberapa kategori. Kategori tersebut adalah sebagai berikut:

State of the Web

State of the Web berisi *report* yang menangkap perkembangan *web* secara jangka panjang termasuk teknik untuk efisiensi jaringan dan penggunaan standar seperti HTTPS. *Reports* ini mencakup beberapa hal yaitu:

- *Sample size* yang berisi perkembangan jumlah URLs yang digunakan untuk dianalisis.
- *Total Kilobytes* yang berisi jumlah dari ukuran perpindahan *kilobytes* dari semua sumber daya yang di *request* oleh halaman *web*.
- *Total Request* yang berisi jumlah *resource* yang di *request* oleh halaman *web*.
- *HTTPS Requests* yang berisi persentase dari semua *request* dari halaman *web* yang menggunakan URL dengan awalan HTTPS.
- *TCP Connections per Page* yang berisi jumlah koneksi TCP dari setiap halaman.
- *HTTP/2 Requests* yang berisi persentase dari semua *request* yang menggunakan HTTP/2.
- *HTTP/3 Support* yang berisi persentase dari semua *request* yang mendukung protokol HTTP/3.
- *Font Display* yang berisi persentase dari halaman yang menghindari munculnya teks tidak terlihat dengan sekejap sewaktu *web* memuat font dengan menggunakan properti CSS `font-display`. Matriks ini diukur dengan menggunakan *Lighthouse*.

State of JavaScript

JavaScript membuat halaman *web* dapat memiliki aplikasi yang kaya dan lebih interaktif. *Report* dalam kategori ini bertujuan untuk melihat penggunaan *JavaScript* dalam *web* dan adopsi serta trennya untuk perangkat *mobile*. Report ini akan menganalisis skrip eksternal. Skrip eksternal ini dimaksudkan untuk *resource file* yang menggunakan ekstensi `js` atau `json` atau sebuah tipe

MIME((*Multipurpose Internet Mail Extensions*)) yang mengandung **script** atau **json**. Beberapa hal yang dianalisis adalah sebagai berikut:

- *JavaScript Bytes* yang berisi jumlah ukuran perpindahan *kilobytes* dari skrip eksternal yang di *request*.
- *JavaScript Requests* yang berisi jumlah skrip eksternal yang di *request* oleh halaman *web*.
- *JavaScript Boot-Up Time* yang berisi jumlah dari waktu CPU yang dibutuhkan oleh setiap *script* di setiap halaman *web*. Matriks ini diukur menggunakan *Lighthouse*.

State of Images

Images atau gambar merupakan tipe *resource* yang populer digunakan dalam *web*. *Report* ini adalah hasil analisa penggunaan gambar eksternal di seluruh *web*. Gambar eksternal adalah *resource* yang memiliki ekstensi **png**, **gif**, **jpg**, **jpeg**, **webp**, **ico**, atau **svg** atau sebuah tipe MIME((*Multipurpose Internet Mail Extensions*)) yang mengandung **image**. *Report* yang masuk kategori ini adalah:

- *Image Bytes* yang berisi jumlah ukuran perpindahan *kilobytes* dari gambar eksternal yang di *request* oleh halaman *web*.
- *Image Request* yang berisi jumlah gambar eksternal yang di *request* oleh halaman *web*.
- *Offscreen Image Save* yang berisi jumlah *kilobytes* yang dapat dihemat oleh setiap halaman menggunakan *lazy-loading offscreen* dan gambar tersembunyi. Matriks ini berasal dari *Lighthouse*,
- *Optimize Image Savings* yang berisi jumlah *kilobytes* yang dapat dihemat oleh setiap halaman dengan mengatur kompresi JPEG ke 85 atau lebih kecil.
- *Native Image Lazy Loading* yang berisi persentase dari halaman yang memiliki atribut **loading = lazy** di dalam elemen **img**.

Loading Speed

Performa *web* dapat berpengaruh langsung terhadap bisnis seperti kepuasan pengguna. *Reports* ini adalah akan menganalisis berbagai matriks performa dalam siklus pemuatan halaman *web* termasuk yang digunakan oleh aplikasi *web* modern.

Progressive Web App

Report ini akan mengkaji status dari *Progressive Web App*. *Progressive Web App* adalah kelas baru dari aplikasi *web* yang disediakan oleh *Service Workers APIs*. *Service Workers* memungkinkan aplikasi untuk mendukung proses muat jaringan secara independen, menerima *push notifications* untuk menyinkronkan data di *background*.

Accessibility

Report ini menjelaskan tingkat aksesibilitas dari sebuah halaman *web*. Penilaian ini dilakukan oleh *Lighthouse*.

SEO

Report yang akan menelusuri penggunaan beberapa teknik agar halaman *web* dapat dikenali oleh mesin pencarian secara lebih baik.

Page Weight

Report ini menelusuri ukuran dan banyaknya *resource* dari banyak halaman *web* populer. Ukuran dalam hal ini merepresentasikan jumlah *byte* yang dikirimkan melalui jaringan.

CrUx

Report ini akan menelusuri tingkat interaktivitas dan proses muat dari pengguna *Chrome* di dunia nyata melalui berbagai kondisi perangkat keras dan jaringan.

Capabilities

Core Web Vitals Technology Report

2.1.2 2024 Web Almanac

Introduction

Web Almanac merupakan kombinasi dari statistik yang masih mentah dan tren yang ada di *Http Archive* dengan keahlian komunitas *web*. *Web Almanac* memiliki 20 bab yang mencakup banyak aspek seperti isi konten dalam halaman, pengalaman pengguna, distribusi, dan penerbitan.

Method

Metode yang digunakan oleh *Http Archive* dalam mengumpulkan data perkembangan pembuatan *web* sejak 2010 adalah dengan menggunakan *WebPageTest* dan *Lighthouse*. *Lighthouse* adalah sebuah alat *open-source* yang disediakan oleh *Google* untuk meningkatkan kualitas halaman *web* [3]. *Lighthouse* dapat melakukan pemeriksaan terhadap performa, aksesibilitas, dan SEO (*Search Engine Optimization*) dari sebuah halaman *web*

2.1.3 Public Dataset

Http Archive memberikan akses terhadap informasi yang lebih detail mengenai hal-hal yang ada di setiap *website*, seperti *metadata* dari *request* dan *response*, *response bodies*, jejak eksekusi, dan lain-lain.

2.2 Structure Query Language

Structure Query Language atau yang biasa disebut SQL merupakan bahasa pemrograman yang bertujuan untuk memanipulasi atau mengubah basis data. Ben Forta dalam bukunya menjelaskan bahwa bahasa ini didesain untuk mengerjakan sebuah perintah dengan tepat dan benar agar proses pembuatan atau pengambilan data berjalan lebih efisien [4]. Data disimpan dalam bentuk tabel ke dalam basis data, untuk mengakses data tersebut SQL menyediakan beberapa sintaks yang bisa dipakai. Sintaks yang dapat dipakai adalah sebagai berikut:

- **SELECT** dan **FROM** merupakan sintaks yang berguna untuk memilih bagian data yang dibutuhkan dari tabel tertentu.
- **WHERE** adalah sintaks yang berguna untuk memberikan kondisi tertentu dalam memilih data.
- **GROUP BY** adalah sintaks yang berguna untuk mengelompokkan data berdasarkan kelas tertentu yang terdapat dalam data.
- **JOIN** adalah sintaks yang berguna untuk menggabungkan dua buah tabel. Penggabungan data ini bisa dibagi kedalam beberapa cara yaitu *inner join*, *outer join*, *right join*, *self join*. sintaks-sintaks tersebut merupakan sebagian kecil dari sintaks yang dimiliki oleh bahasa SQL.

2.3 Statistika [5]

Statistika adalah ilmu yang mempelajari tentang eksplorasi, analisis, implementasi, dan pengumpulan data. Statistika memiliki beberapa properti untuk melihat *Central Tendency* dari data. *Central Tendency* adalah pusat kumpulan sebuah data. Properti yang dapat digunakan untuk melihat pusat kumpulan data adalah sebagai berikut:

- *Mean* atau rata-rata adalah properti untuk mengukur distribusi nilai dari sebuah data. Persamaan 2.1 digunakan untuk mencari nilai rata-rata dari sebuah data. N merupakan jumlah data yang sedang diamati sedangkan nilai x_N merupakan nilai-nilai yang akan dijumlahkan mulai dari x_1 hingga x_N .

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_N}{N} \quad (2.1)$$

- *Median* merupakan nilai tengah dari data yang sedang diamati. Nilai *median* dapat dicari dengan cara mengasumsikan bahwa data telah terurut, nilai N merupakan jumlah data kemudian jika N memiliki nilai yang ganjil maka letak nilai *median* nya terpadat pada posisi $\frac{N+1}{2}$, sedangkan jika nilai N nya genap maka nilai *median* nya terletak pada posisi $\frac{N}{2}$.
- *Mode* atau modus adalah nilai yang kemunculannya paling banyak pada sebuah data.

Properti lain yang dapat digunakan selain *Central Tendency* adalah pengukuran distribusi dan sebaran data, beberapa properti yang dapat digunakan adalah sebagai berikut:

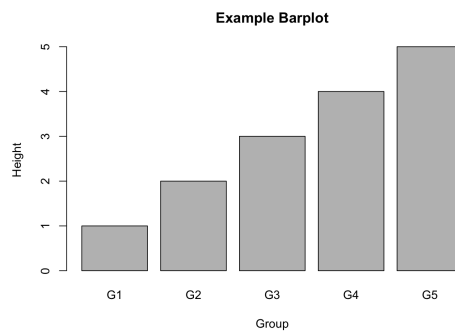
- *Max* merupakan nilai paling besar dari sebuah data.
- *Min* merupakan nilai paling kecil dari sebuah data.
- *Range* merupakan perbedaan dari nilai paling besar dengan nilai paling kecil
- *Variance* dan Standar Deviasi adalah metode untuk mengukur sebaran data. *Variance* didapatkan dengan cara mengkuadratkan perbedaan setiap titik pada data dengan rata-ratanya, sedangkan standar deviasi merupakan akar dari *variance*. *Variance* cenderung menghasilkan nilai yang lebih besar dari nilai-nilai yang terdapat pada data asli karena merupakan hasil kuadrat, sedangkan standar deviasi cenderung menghasilkan nilai yang hampir sama dengan nilai-nilai yang terdapat pada data asli. Standar deviasi dapat dicari dengan menggunakan Rumus 2.2, dimana nilai N adalah jumlah data, nilai X_i adalah nilai ke- i dari atribut X , dan \bar{X} merupakan nilai rata-rata dari atribut X .

$$\sigma = \sqrt{\left(\frac{1}{N} \sum_{i=1}^N X_i^2\right) - \bar{X}^2} \quad (2.2)$$

Semakin besar nilai standar deviasinya maka dapat dikatakan bahwa data semakin tersebar dari nilai rata-ratanya, sebaliknya semakin kecil nilai standar deviasinya maka dapat dikatakan bahwa data semakin dekat sebarannya dari nilai rata-ratanya.

2.4 Visualisasi Data

2.4.1 Bar Plot



Gambar 2.1: Contoh visualisasi dari tinggi beberapa grup dengan menggunakan *bar plot*

Bar plot merupakan teknik visualisasi data yang menggunakan batang vertikal atau horizontal untuk menunjukkan nilai-nilai dari data. Visualisasi ini berguna untuk menunjukkan pengukuran

-
- 1 statistik sebuah data secara terpisah. *Bar Plot* memiliki elemen utama yaitu sumbu x dan sumbu y .
 - 2 Gambar 2.1 [6] merupakan contoh penggunaan *bar plot* untuk memvisualisasikan data di mana pada
 - 3 contoh ini sumbu x nya menunjukkan grup yang dimiliki data sedangkan sumbu y nya menunjukkan
 - 4 tinggi dari masing-masing grup.

DAFTAR REFERENSI

- [1] Nugroho, P. A., Putra, R. C., Maulana, R. C., dan Tandra, V. (2024) Usage of unsupported technologies in websites worldwide. *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika*, **10**, 332–344.
- [2] Viscomia, R., Calvano, P., Chen, S., Hobbs, M., dan Demir, N. (2024) Http archive about. Technical report. Http Archive.
- [3] Version 3.0 (2016) *Introduction to Lighthouse*. Google for developers.
- [4] Forta, B. (2019) *SQL in 10 Minutes a Day, Sams Teach Yourself*, 5th edition. Sams Publishing.
- [5] Jiawei Han, H. T., Jian Pei (2022) *Data Mining: Concepts and Techniques*, 4 edition The Morgan Kaufmann Series in Data Management Systems. Morgan Kaufmann, Cambridge, MA.
- [6] Phillips, N. (2017) Yarrrr! the pirate’s guide to r. *The Observer* , ?