

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Zettabyte adalah unit ukuran untuk jumlah data yang tersedia di internet. Jumlah data yang sangat besar ini memberikan banyak kesempatan bagi para peneliti dan praktisi untuk mendapatkan wawasan langsung tentang organisasi, individu, atau bahkan fenomena sosio-teknis tingkat makro. Tidak mengherankan bahwa para peneliti Sistem Informasi semakin banyak bergantung pada internet untuk mendapatkan informasi yang dapat membantu mereka menemukan jawaban atas pertanyaan penelitian mereka (V. Krotov et al., 2021.).

Dengan bertambahnya jumlah data yang ada di internet, *web scraping* adalah salah satu metode yang dibutuhkan untuk mengumpulkan banyak data dari suatu website (Satriajati et al., 2020). *Web scraping* adalah solusi untuk membantu masalah ini. *Web scraping* adalah metode cepat dan efektif untuk otomatisasi ekstraksi data dari halaman web. Proses ini dapat dilakukan menggunakan layanan perangkat lunak atau program khusus yang dibuat sendiri, yang dapat mengakses dan menavigasi situs web pada berbagai tingkat. (Abodayeh et al., 2023). Teknik *web scraping* ini pernah diterapkan dalam penelitian Y. A. Hafiz dan Endah Sudarmilah pada tahun 2023 dengan melakukan *web scraping* pada portal berita online dengan tujuan melihat seberapa efektif penggunaan *web scraping* pada situs portal berita. Dibutuhkan library seperti BeautifulSoup untuk membantu teknik *web scraping* ini.

Di era modern ini, popularitas anime di kalangan masyarakat semakin meningkat seiring dengan berkembangnya media sosial yang kian masif, mendorong orang untuk memilih berbagai bentuk hiburan. Anime, jenis animasi yang berasal dari Jepang, mencakup beragam genre dan menjadi salah satu bentuk kartun Jepang yang paling diminati (Ramadhan & Ekawardhani,

2022). Seiring berjalannya waktu, anime semakin dikenal berkat beragam ceritanya yang berasal dari berbagai genre serta peningkatan kualitas grafisnya. Namun, dalam perkembangan anime, peran penting seorang seiyuu (声優), atau pengisi suara dalam bahasa Jepang, tidak bisa diabaikan.

*Seiyuu* memainkan peran yang penting dalam penggambaran sebuah karakter maupun *live action*. Penjiwaan, emosi dan intonasi merupakan faktor yang mempengaruhi *seiyuu* dalam sebuah karakter animasi yang diperankannya. *Seiyuu* juga berkontribusi dalam sejarah dan perkembangan budaya *anime*, *seiyuu* sendiri berkontribusi ke pasar *anime* dengan mempromosikan karakter yang mereka perankan dan serial *animenya*. Dalam dekade terakhir, *anime* telah mengalami pertumbuhan popularitas yang signifikan secara global. Data dari Association of Japanese Animations menunjukkan bahwa pasar *anime* global telah tumbuh sebesar 10% setiap tahun sejak 2010, mencapai nilai pasar lebih dari 19 miliar USD pada tahun 2023. Dalam menghadapi pertumbuhan ini, industri memerlukan alat analitis yang mampu mengelola dan menganalisis volume data yang besar, yang tersedia melalui berbagai platform online seperti MyAnimeList.

Dalam penelitian ini, peneliti memilih website [myanimelist.net](http://myanimelist.net), sebuah komunitas dan database anime serta manga online paling aktif di dunia, sebagai target untuk pengujian implementasi web scraping. Berdasarkan data dari [similarweb.com](http://similarweb.com) per 28 Maret 2024, [myanimelist.net](http://myanimelist.net) menduduki peringkat 978 dalam daftar website dengan traffic tertinggi di dunia dan peringkat 13 dalam kategori Seni & Hiburan. Database [myanimelist.net](http://myanimelist.net) sangat kaya akan informasi seputar anime, mencakup judul, tanggal rilis, status, pengisi suara, hingga rating. Dengan banyaknya data yang tersedia, peneliti akan mengambil dan memproses informasi terkait seiyuu dan anime untuk menarik kesimpulan yang relevan.

## 1.2 Identifikasi Masalah

Berdasarkan latar belakang penelitian yang sudah dijelaskan diatas, identifikasi masalah dapat diuraikan seperti berikut:

1. Ketidakmampuan untuk mencegah kesalahan pengambilan data yang berlebihan dan bercabang dalam web scraping.
2. Keterbatasan dalam pengolahan data *seiyuu* yang tersedia di MyAnimeList untuk analisis lebih lanjut.
3. Tidak adanya data yang jelas mengenai korelasi antara popularitas *seiyuu* (pengisi suara) dengan *anime* yang mereka perankan.

## 1.3 Tujuan

Tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut:

1. Mengatasi kesalahan pengambilan data secara banyak dan bercabang dalam web scraping melalui penggunaan metode time sleep.
2. Mengembangkan strategi pengolahan data yang efektif untuk data *seiyuu* dari MyAnimeList.
3. Menganalisis dan menjelaskan pengaruh hubungan popularitas *seiyuu* terhadap *anime* menggunakan data yang telah dikumpulkan.

## 1.4 Batasan Masalah

Adapun batasan masalah untuk penelitian ini sebagai berikut:

1. Penelitian ini hanya berdasarkan data yang ada dalam myanimelist.net.
2. Penelitian ini mencakup 350 orang dalam 7 halaman yang ada dalam daftar orang populer di myanimelist.net.
3. Penelitian hanya menggunakan uji korelasi untuk mencari tahu hubungan antara *seiyuu* dan *anime*.

## 1.5 Kontribusi

Pada penelitian analisis ini, ditunjukkan agar dapat memperkaya literatur dan pemahaman di bidang pengolahan data dan analisis informasi dalam konteks industri *anime*. Serta memberikan pemahaman yang lebih mendalam tentang efektivitas dan relevansi metode *web scraping*.