

# Web Scraper Project Report

## 1. Project Overview

This project is a simple yet effective web scraping application built using Java and Jsoup library.

The primary goal of the project is to extract quotes along with their respective authors from the website 'https://quotes.toscrape.com/'

and store them into a text file named 'scraped\_quotes.txt'. The program fetches the HTML of the page, parses the required elements using CSS selectors, and saves the data in a human-readable format.

## 2. Duration

The project took a total of 3 days to complete. The initial setup faced some challenges due to encoding issues and Jsoup library configuration, but they were resolved with careful debugging and classpath setup.

## 3. Outcome

The application successfully scraped quotes and their respective authors from the given website and saved the data into 'scraped\_quotes.txt' file.

This demonstrates the capability of Java with Jsoup for real-world web scraping tasks.

## 4. Challenges Faced

- Initial issues in configuring the Jsoup JAR file in the classpath.
- Encoding issues while writing to and displaying content from the console.
- Confusion in locating the generated text file as it gets saved in the project root directory.
- Parsing the HTML structure properly using CSS selectors.

## 5. Conclusion

This project gave a solid understanding of how web scraping works and how to implement it in Java using Jsoup.

# Web Scraper Project Report

Despite facing minor challenges, the final output is successful and the project objectives were met. It sets the foundation for more complex scraping tasks in the future.