

# Implementation of Web Scraping in Inventory Management System for Drop-Shipping

Lutfi Pratama Yogaswara, Yan Puspitarani  
Department of Informatics,  
Widyatama University Bandung, Indonesia

**Abstract:-** The drop-shipping is a simplified business model in order fulfillment, which allows retailers to sell products without keeping any physical inventory. The product order is sent to supplier, who ships the order directly to the customers. Therefore, currently, around 27% of online retailers are starting to turn to the drop-shipping business model as a method of fulfilling orders from customers. But despite its advantages, the drop-shipping business model still has disadvantages, such as inventory problems. This paper proposes an approach to solve the inventory problem in drop-shopping by using web scraper. The web scraper extracts product information from supplier website and stores it in the database. For managing the extracted product information and upload it to the marketplace, we propose an Inventory Management System, implemented as a web-based application. However, users can save time doing drop-shipping without extracting product information by visiting each product page, managing product stock, and uploading to the marketplace manually.

**Keywords:-** drop-shipping, web scraper, marketplace.

## I. INTRODUCTION

The drop-shipping is a simplified business model in order fulfillment, which allows retailers to sell products without keeping any physical inventory. The product order is sent to supplier, who ships the order directly to the customer. Therefore, currently, around 27% of online retailers are starting to turn to the drop-shipping business model as a method of fulfilling orders from customers [1].

Despite its advantages, drop-shipping still has disadvantages. Tracking inventory is quite easy when you keep stock on hand. With drop-shipping, keeping an eye on stock can be tough to do. Especially if there is a new product, users need to extract the product information from the product page and upload it manually to the marketplace.

This paper proposes an approach to solve the inventory problem in drop-shopping by using web scraper. The web scraper extracts product information from supplier website and stores it in the database. For managing the extracted product information and upload it to the marketplace, we propose an Inventory Management System, implemented as a web-based application.

This paper is divided into five sections. The first section is introduction. Related works is discussed in section 2. Section 3 explains the proposed system. Section

4 is the results of proposed system. And finally, section 5 is conclusion and future work.

## II. RELATED WORKS

This section discussed about previous works that related to the utilization of web scraping. In [2], the authors developed a platform that called NewsOne, which aggregates latest news that extracted from online news websites using web scraping technique and presented it in summarize.

Web scraping technique can be utilized in collecting of marketing promotion as well. The authors of [3] build a mobile platform that named ProCircle, which collect promotion news from websites using web scraping technique into one place, so users can easily search for promotional information.

Other related work is a platform named Smart Chef that presented in [4]. This paper describes the utilization of web scraping to extract recipes and ingredients information from a website called [www.foodnetwork.com](http://www.foodnetwork.com) and stored it in the recipe database. By using the recipe/ingredient algorithm, users can search recipe based on ingredient, nutritional details, or diet plan.

And there is similar work that using web scraping technique for drop-shipping, which is the authors of [5] that extract information from marketplace, compare the price of products, store information on the development of sales and prices of a product and provide recommendations for the best product choices for drop-shipping.

As for authors of [6] using web scraping to extract product information from several marketplaces in Indonesia into CSV (Comma Separated Values) files, and has a feature to upload the extracted product information into another marketplace.

## III. PROPOSED SYSTEM

To get all product information on the supplier website is a time-consuming activity. It will be difficult for users to visit each product page to get the product information, keeping track of the product stock, and upload it to the marketplace manually. Thus, users can use the web scraping technique to extract product information and utilize the marketplace API to upload the products.

The functionality of this proposed system includes scraping product information from supplier websites, storing extracted product information into a database, retrieving data from the product database, implementing

web applications to manage products, and uploading the products to the marketplace using API. Fig. 1 shows the architecture of the proposed approach.

This proposed system is divided into two main components: (1) Web Scraper, and (2) Web Application.

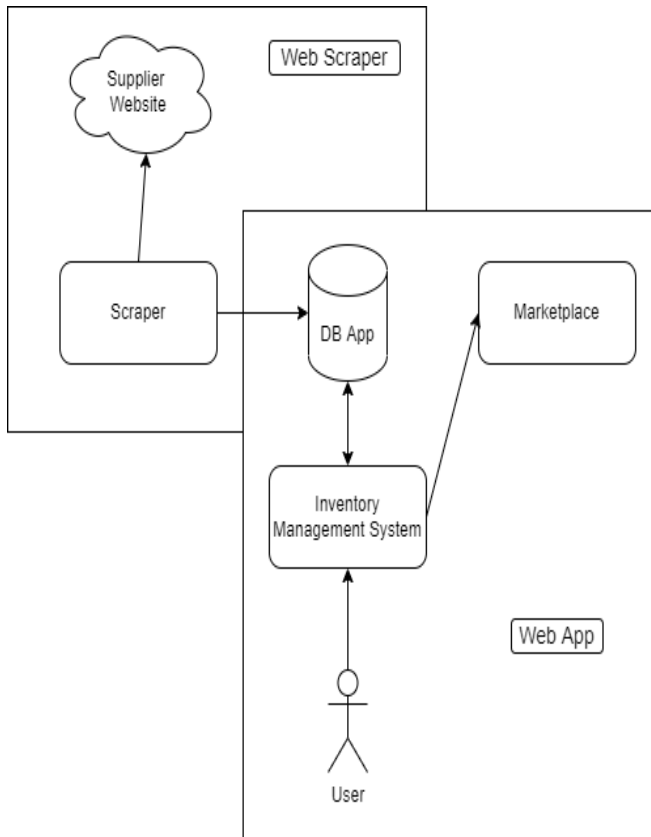


Fig. 1: Architecture of Proposed System

**A. Web Scraper**

Web scraping is a technique to extracting text information from web pages automatically. Web scraping can save more time, instead of manually copying the data from websites [2].

In this proposed system, a web scraper is built based on Python programming language with BeautifulSoup library. Python is a general-purpose programming language with a mature ecosystem of high quality third-party packages for almost any task [7], including web scraping. BeautifulSoup is a python library that used to parsing HTML/XML.

The supplier website that will be scraped is unomax.com, one of the suppliers in Indonesia that supply accessories for smartphones. Web scraping steps to get all product information effectively are based on website sitemap: (1) Frontpage, (2) List product by category page, (3) Product detail page. Fig. 2 depicts the web scraping steps.

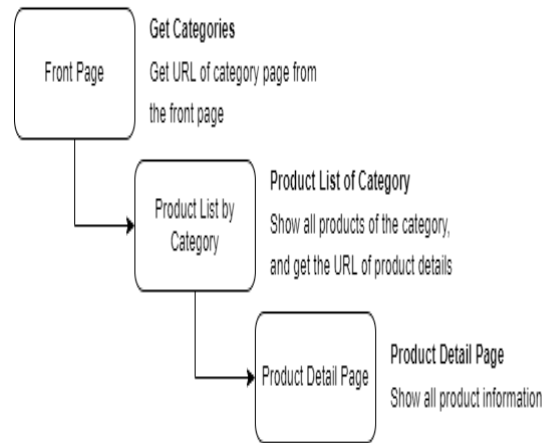


Fig. 2: Web Scraping Step based on Website Sitemap

The algorithm that is used for scraping product information from supplier websites:

- Import the libraries that needed for scraping: requests, bs4
- Go to the front page, parse the content using bs4
- Fetch all URLs of the category and save them in a variable called categories
- Iterate all URLs of category and go to every URL
- Parse the content of the list of products by category using bs4
- Get URL for every product
- Check if the URL product exists in the database, if exists, update the status stock, and continue to the next product, if not exist, get the URL product and go to the product detail page
- Parse the content of the product detail page using bs4
- Extract all product information and store it in the database

Web scraper will be scheduled by task scheduler to run every day at midnight. In Linux, task scheduler run by a command named cron job [8], while on Windows, scheduled tasks is used to schedule a job that is executed periodically. The product information extracted is product name, SKU (Stock Keeping Unit), brand, description, price, weight, stock status, product URL, and product images URL.

**B. Web Application**

Web application is an application software that is stored on a remote server and delivered over the internet through browser interface. Web application needs authentication [9]. A browser is an application that is used to browse the internet.

The web application in this proposed system is Inventory Management System that is built on PHP (PHP: HyperText Preprocessor) with Laravel as the framework and MySQL as the Database Engine. PHP is a scripting designed specifically for use on the web [10]. Laravel is a framework created to help developers create a emphasizing the simplicity and flexibility of the design [11]. MySQL is a fast, easy-to-use RDBMS (Relational Database Management System) used on many web sites [10]. Inventory Management System is used to manage products that have been extracted from the supplier website.

a) Mockup

This web application has six modules: login, register, add product, edit product, product list, and shop. By utilizing these modules users can manage the products.

• Login Page

Fig. 3 shows a mockup of the login page. Login is a set of credentials used to gain access to a web application requiring proper authentication.

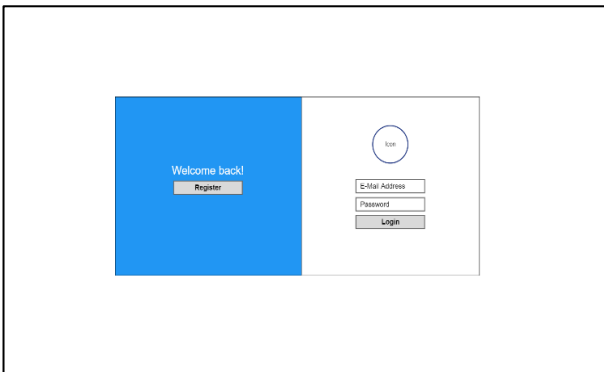


Fig. 3: Login Page

• Register Page

Fig. 4 shows a mockup of the register page. The register enables users to independently register and gain access to a web application.

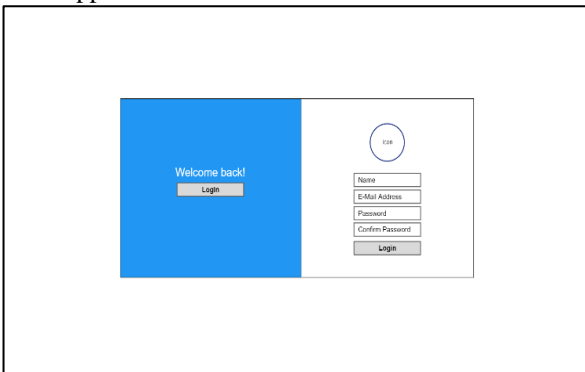


Fig. 4: Register Page

• Add Product Page

Fig. 5 shows a mockup of the add product page. Add product is a page that is used to add a new product to the database.

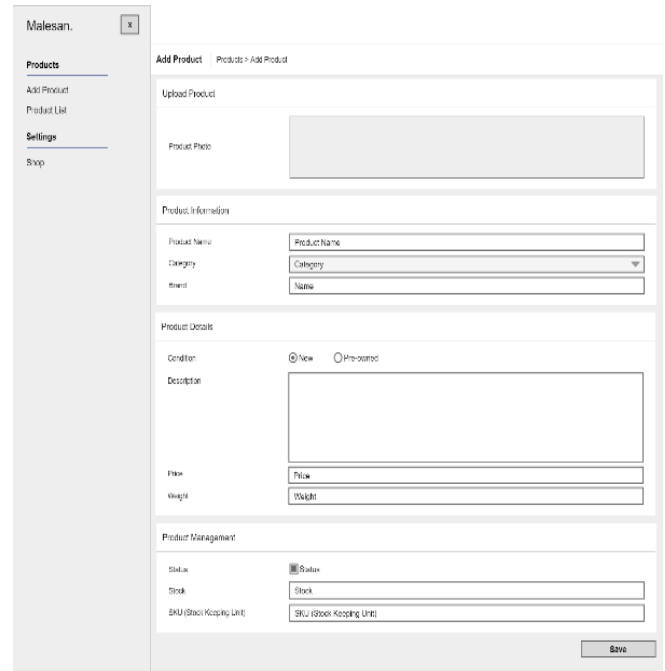


Fig. 5: Add Product Page

• Edit Product Page

Fig. 6 shows a mockup of the edit product page. Edit Product is a page that is used to edit the product in the database.

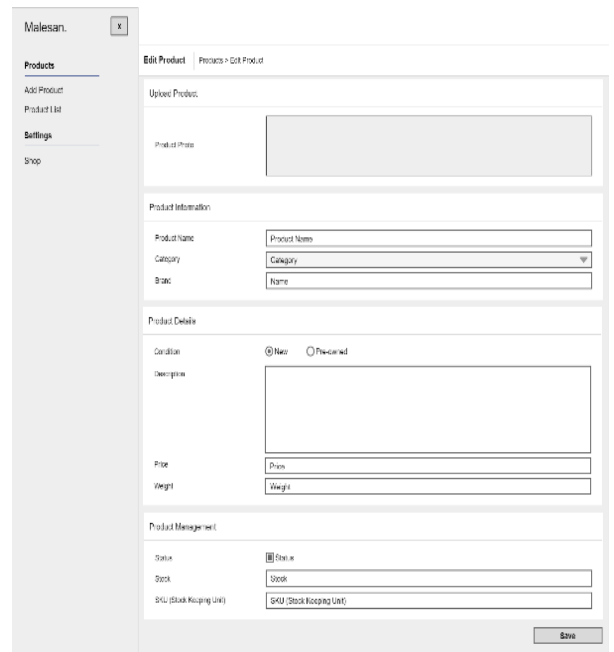


Fig. 6: Edit Product Page

• Product List Page

Fig. 7 shows a mockup of the product list page. Product List is a page that shows all products in the database and has a function to upload the product to the marketplace.

On this page, there are two tabs, local and online. Local is a tab for products that have not been uploaded to the marketplace, while online is a tab for products that have already been uploaded to the marketplace.

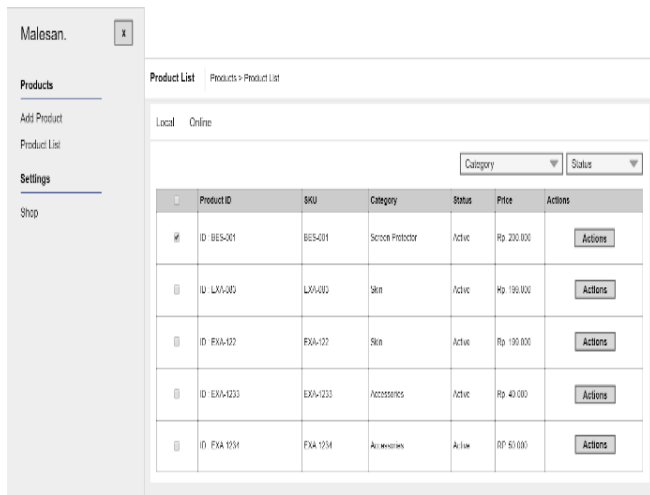


Fig. 7: Product List Page

• Shop Page

Fig. 8 shows a mockup of the shop page. A shop is a page that shows all information about the shop in the marketplace.

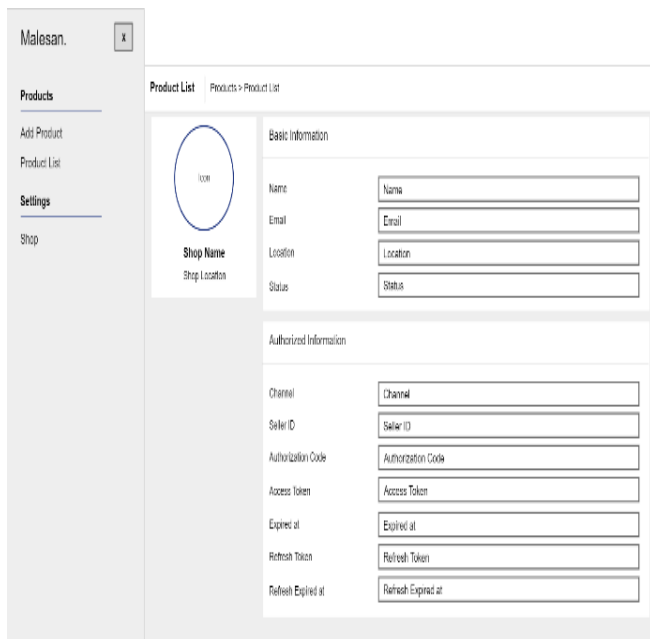


Fig. 8: Shop Page

b) Database Schema

The database schema that was designed to build this proposed system is using MySQL. MySQL is a Relational Database Management System (RDBMS) that is based on Structured Query Language (SQL). MySQL is used to store structured data which is suitable with product information data that has been extracted by web scraper.

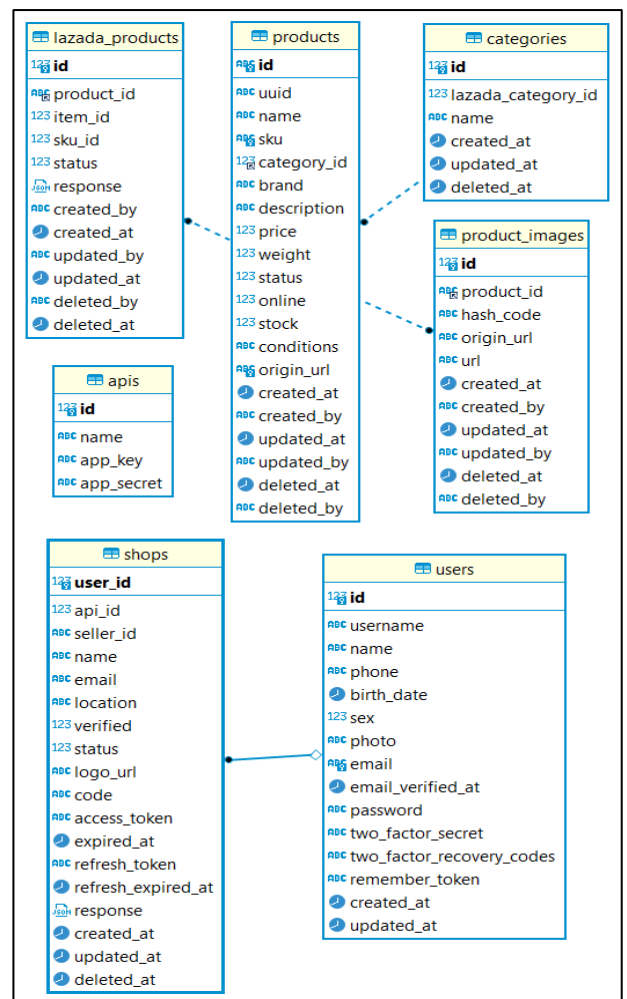


Fig. 9: Database Schema for Web Application

Fig. 9 describes the database schema in a form of Entity Relationship Diagram (ERD). Entity Relationship Diagram (ERD) is a type of structural diagram used in database design. The schema contains seven tables: users, products, categories, product\_images, lazada\_products, apis, and shops.

- users  
Users table is used to store user information and has a foreign key in table shops so that every user has their shop.
- products  
The products table is used for product information.
- categories  
Table categories are used to store master data categories and map from local product category to Lazada product category.
- product\_images  
Product\_images table is used to store product images of every product.
- lazada\_products  
Lazada\_products table is used to store products that have already been uploaded to the marketplace.

- **apis**  
Apis table is used to store credentials for Lazada API.
- **shops**  
The shops table is used to store shop information for every user.

**IV. RESULTS**

The product information will get extracted automatically by web scraper and shown on the product list page in the web application as shown in Fig. 10.

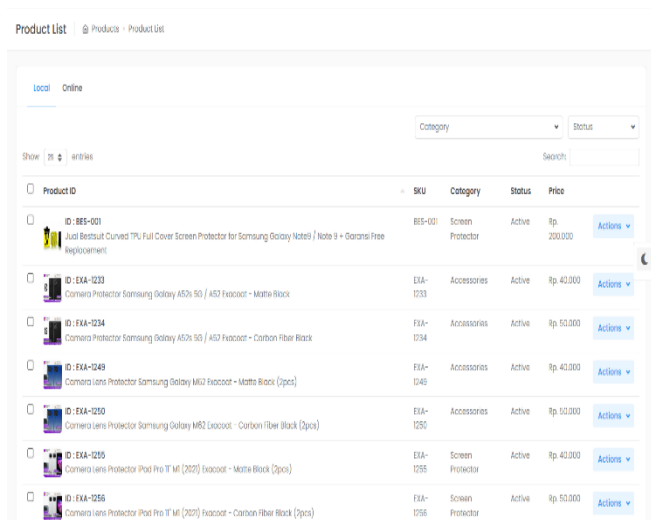


Fig. 10: Product List Page that shows extracted products

Users can upload the product to the marketplace by clicking on the actions button and choosing upload and a confirmation box will show up as shown in Fig. 11.

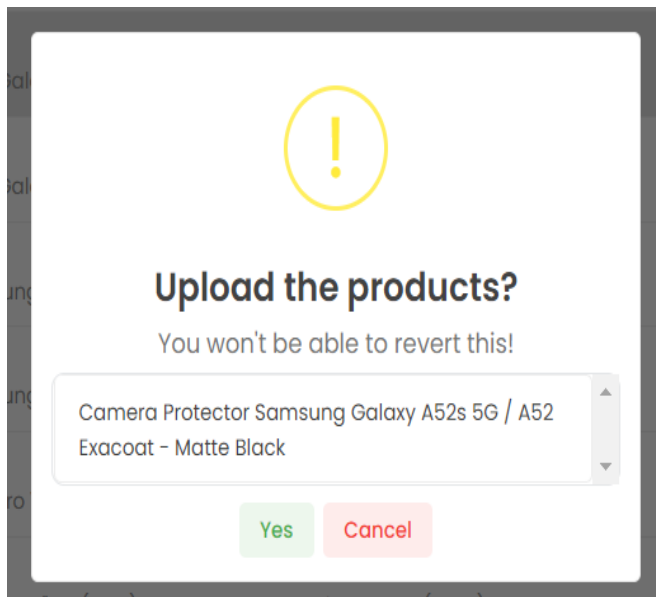


Fig. 11: Confirmation box when uploading product

After the user confirmed by clicking Yes, the product will be uploaded to the marketplace.

The product has been uploaded to the marketplace and shown in the product list in the marketplace as shown in Fig. 12.

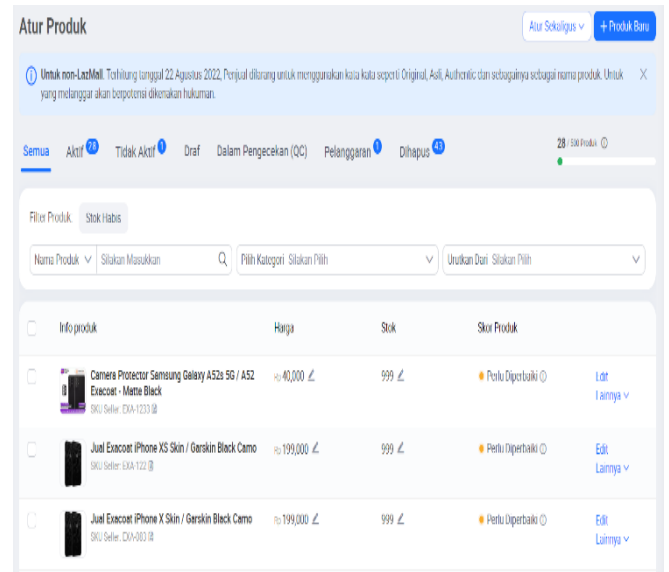


Fig. 12: List of products that have been uploaded to the marketplace

**V. CONCLUSION AND FUTURE WORK**

This system can help users save time doing dropshipping without extracting product information by visiting each product page, managing product stock, and uploading to the marketplace manually. Thus, users can focus on sales and managing orders without having to worry about the products. Even so, this system still needs to get some enhancement. In the future, this system can be enhanced by adding more marketplace channels, so users only need to upload once, and products will be uploaded to many marketplace channels instead of one.

**REFERENCES**

- [1.] “Dropshipping and ecommerce key statistics [Updated: August 2022] - AppScenic.” <https://appscenic.com/dropshipping-and-ecommerce-key-statistics/> (accessed Aug. 15, 2022).
- [2.] K. Sundaramoorthy, R. Durga, and S. Nagadarshini, “NewsOne - An Aggregation System for News Using Web Scraping Method,” *Proc. - 2017 Int. Conf. Tech. Adv. Comput. Commun. ICTACC 2017*, vol. 2017-Octob, pp. 136–140, 2017, doi: 10.1109/ICTACC.2017.43.
- [3.] L. Junjoewong, S. Sangnapachai, and T. Sunetnanta, “ProCircle: A promotion platform using crowdsourcing and web data scraping technique,” *Proceeding 2018 7th ICT Int. Student Proj. Conf. ICT-ISPC 2018*, pp. 1–5, 2018, doi: 10.1109/ICT-ISPC.2018.8524003.
- [4.] S. Chaudhari, R. Aparna, V. G. Tekkur, G. L. Pavan, and S. R. Karki, “Ingredient/Recipe Algorithm using Web Mining and Web Scraping for Smart Chef,” *Proc. CONECCT 2020 - 6th IEEE Int. Conf. Electron. Comput. Commun. Technol.*, no. 3, pp. 22–25, 2020, doi: 10.1109/CONECCT50063.2020.9198450.
- [5.] F. A. Pratama, F. Amalia, and E. Santoso, “Pengembangan Aplikasi Rekomendasi Produk Terbaik Pada Marketplace Bukalapak Bagi

- Dropshipper dan Reseller Menggunakan AHP-TOPSIS,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 4, no. 1, pp. 36–45, 2020.
- [6.] A. A. Maulana, A. Susanto, and D. P. Kusumaningrum, “Rancang Bangun Web Scraping Pada Marketplace di Indonesia,” *JOINS (Journal Inf. Syst.*, vol. 4, no. 1, pp. 41–53, 2019, doi: 10.33633/joins.v4i1.2544.
- [7.] Dogan Can, V. R. Martinez, P. Papadopoulos, and S. S. Narayanan, “PYKALDI: A PYTHON WRAPPER FOR KALDI Do ~ gan Can , Victor R . Martinez , Pavlos Papadopoulos , Shrikanth S . Narayanan Signal Analysis and Interpretation Lab University of Southern California Los Angeles , CA , USA,” pp. 5889–5893, 2018.
- [8.] “What is Cron Job? - Cron Jobs and Scheduled Tasks - Hivelocity Hosting.” <https://www.hivelocity.net/kb/what-is-cron-job/> (accessed Aug. 22, 2022).
- [9.] “Difference Between Web application and Website - GeeksforGeeks.” <https://www.geeksforgeeks.org/difference-between-web-application-and-website/> (accessed Aug. 12, 2022).
- [10.] Stephen R.Davis, *Get More and Do More at Dummies*. 2010.
- [11.] M. Anif, A. Denta, and H. W. S. Sindung, “Designing internship monitoring system web based with laravel framework,” *2017 IEEE Int. Conf. Commun. Networks Satell. COMNETSAT 2017 - Proc.*, vol. 2018-Janua, pp. 112–117, 2017, doi: 10.1109/COMNETSAT.2017.8263583.