# Web Scraping: A Review of Methodologies Used

**Abstract.** The incorporation of information and communication technologies (ICT) has enabled the buying and selling of various products through web pages, which generate large volumes of data. Recently, online prices have generated significant interest in the empirical literature. Web scraping (WS) is a technique used to extract available information, which allows the data to be utilised. This article presents the results of a literature review of 20 articles published between 2018 and 2023, aiming to understand the methodologies employed for WS and the benefits of its use. The findings of the review indicate that there is no standard WS methodology. The most common phases of the proposed methodologies include collecting, processing, storing and analysing or predicting. The benefits of WS are diverse, benefiting both sellers and buyers. These advantages are primarily related to product price comparisons and time savings.

**Keywords:** web scraping; prices; review; methodology; product.

## 1 Introduction

In recent years, national statistical agencies have shown interest in using web-based consumer price index data, which could complement or replace manually collected price quotations [1]. One tool used to extract data from web pages is web scraping (WS) [2]. In the information age, WS has emerged as an essential tool across various fields for any company wishing to maintain its online presence a key factor for survival in today's market [3]. The collected data is available in real time, with no delays in accessing and processing the information [4]. WS can collect unstructured content and convert it into a structured type that can be used for further study [5].

Recently, online pricing has sparked significant interest in the empirical literature [6]. The introduction of new data collection methods like WS offers the ability to utilise highly disaggregated and timely data for analysis and decision-making [7]. Although data is scattered across hundreds of websites, advances in automated 'scraping' software now allow anyone to design and implement large-scale data collections on the web [7]. Some websites hold a considerable amount of invaluable data, such as product details, sports statistics, etc. Manual data extraction can be quite challenging when dealing with a large volume of information from a website [8].

WS has been used to compare online stores for displaying and analysing digital marketing strategies [9], select features to rank web pages and recommend specific pages [10], and acquire skills and demands for jobs in data science [11]. It has also been used to analyse the real estate market [12], estimate food inflation based on prices [13], determine housing or rental price levels [12] and in the pharmaceutical sector.

WS is an efficient technique for extracting data from various websites and it can be used for purposes including web mining, data mining and weather monitoring, among

others [14]. WS is a low-cost data collection method for food price research. Using WS to create customised data sets can help overcome common problems such as incomplete data, omitted variables and sample selection [15]. As price data becomes more widely available online, the applications for WS are expected to grow.

WS is characterised by its cost-effectiveness, speed and accuracy in online data collection, making it an ideal complement to traditional price collection methods, especially for product categories that are well-represented online [7]. WS enables quick and cost-effective access to information about all items available in online stores [6]. However, it's important to note that data collection from a website may require the appropriate license in certain cases [2].

Python is one of the languages commonly used to implement WS due to its ease of learning and straightforward syntax [3]. WS also requires a structured methodology. The purpose of this study is to comprehensively review and present the various WS methodologies proposed in the existing research. This research contributes to the WS knowledge base, particularly concerning methodological aspects.

The article is structured in five sections, including this introduction. Section 2 outlines the theoretical framework of WS, Section 3 describes the methodology used, Section 4 presents and discusses the results and Section 5 offers the conclusions.

## 2 Web Scraping (WS)

WS is a technique that allows organisations to analyse large amounts of data and gather new information [2]. It is a data collection method used to track changes in product prices over time and across different locations [4]. WS facilitates the extraction of large volumes of data from websites [11]. This process involves gathering extensive data from websites and storing it locally in different formats. WS is utilised for various purposes, including research, market analysis, price comparison, gathering public opinions about businesses, job advertisements and collecting contact details for businesses [16].

WS is technically feasible for all businesses. However, a site's susceptibility to scraping is determined by one or more of the following key factors: (1) impending competition from similar businesses, (2) the popularity of the website in terms of traffic or user engagement, (3) the uniqueness or necessity of the content being created or modified and (4) the website's security vulnerabilities [3].

WS tools include Selenium, Beautiful Soup and Scrapy. Selenium is a versatile web crawler due to its ability to perform simultaneous automated WS tests. Beautiful Soup, which is used for HTML and XML analysis, is easy to learn but can be inefficient due to the inherently slow nature of the WS process. Scrapy is asynchronous and can send multiple requests at once [14].

## 3 Methodology

This research is fundamental and seeks to understand the WS methodologies employed by various authors. To achieve this, a review of narrative literature has been

conducted. The sample consisted of 20 articles published in the period 2018–2023. The articles were selected based on the terms ‘web scraping’, ‘web mining’, ‘web scraper’, ‘price’ and ‘price online’. The databases selected were Scopus and Web of Science.

One criterion for inclusion in the study was a clear presentation of the methodology used, including the steps for performing WS. A total of 20 articles were selected. Relevant information, such as WS steps, tools used, benefits, study objectives and the focus of the WS, were extracted and recorded in an Excel sheet for subsequent analysis.

# 4 Results

## 4.1 Tools Used

Conducting WS requires several tools. The results of the literature review indicate that Python and its libraries are the most commonly used tools for WS. In contrast, software developed specifically for WS saw the lowest usage. Python is favoured for WS in part because it is an easy-to-learn language that does not require extensive programming experience [3]. The available Python libraries are used to extract the information from HTML pages [10].

## 4.2 WS Methodologies Used

The methodology used to perform web scraping varies depending on the author. We identified the main steps for web scraping from the selected studies describing the phases of the methodology used. Table 1 shows the phases of the methodology used by selected articles’ authors to develop their systems. If we analyse the steps in general, the most common methodological actions are collect, store, analyse and process. Specifically, we argue that the general action(s) for each step of the methodology is (are) as follows: Step 1: Collect; Step 2: Process; Step 3: Store;5); Step 4: Analyse and/or predict.

**Table 1.** Description of the steps of the methodology

| N | Author | Methodology used | Software |
|---|---|---|---|
| **1** | [17] | Develop and improve database: The database of product information is developed and improved. Capture and manage: Data and photos are uploaded from the website using Optical Character Recognition technology to extract lists of ingredients in photo format. Processing and categorisation: The products obtained are categorised into 24 categories and 172 sub-categories. | Retail websites. Python |
| **2** | [18] | Collect: User reviews and metadata from popular e-commerce websites are collected and cleaned by the script created for that function. Processing: Unstructured data (raw HTML) is converted into a structured format (relative table type structure) where tools can be used to define various aspects and corresponding sentiments for each product. Categorise: Indicate all E elements and A attributes commented on in the given message. One must | Online Stores. Octoparse. |

| N | Author | Methodology used | Software |
|---|---|---|---|
| | | choose E and A from the predefined Entity inventory. Attribute labels per domain E belong to A; each pair determines the message's image category. | |
| **3** | [4] | Scan: Online prices were collected using specialised software, which scans the websites of selected retailers displaying online prices. Find information: Relevant information is found on product websites. Store: The information found is stored in a database. | Online Retail Chain. Python. |
| **4** | [7] | Download: At a fixed daily time, software downloads a list of public websites displaying product and price information. Parse: The underlying code is parsed to locate each piece of relevant information by using special characters in the code that identify the start and end of each variable. Store: The software stores the extracted information in a database containing one record per product per day. These variables include the product price, date, category information and an indicator of whether the item was on sale. | Not mention. |
| **5** | [14] | Scrape: Recipe-related information, e.g. name, ingredients used, URL, is extracted from recipe websites. Store: The recipe extracted is stored in the MongoDB database functionality, which stores three fields: the name, the ingredients and the URL of the web scrapping results. Take: Retrieves data from the database. The algorithm takes two data elements – the name and ingredients. Provide list: The data are entered as a list, and a list of recipes for the mentioned ingredient is provided. | Websites. Python. |
| **6** | [19] | Extract cost: The cost of delivery is extracted, which reflects the amount of money charged for delivering the food from the supplier to the customer. Extract time: One obtains the expected delivery time: when the suppliers declare they will deliver orders to their customers. Obtain minimum charge: The minimum order is obtained, i.e. the minimum charge required for suppliers to deliver their orders to the customer. Get feedback: The number of comments that customers have registered for each supplier is collected. | Online Sites Agentv. |
| **7** | [20] | Web scraping: A database is created containing positive and negative reviews on products in the categories detected by scraping review sites according to attributes. Processing: the TextRank algorithm is used for keyword detection, and the Python NLTK library is used for empty word removal, tokenisation of sentences and words and POS tagging of products with positive and negative reviews based on some templates. Estimate: The approach is needed to auto-calculate subjective probabilities based on the TextRank algorithm. | Websites. Not mention. |
| **8** | [21] | Download: Content is downloaded from web pages. This | Websites. |

| N | Author | Methodology used | Software |
|---|---|---|---|
| | | can be any active website link where data extraction can be performed. Extract: The downloaded contents are HTML and mostly unstructured. Therefore, the Beautiful Soup library is used to parse the data and extract structured data from the downloaded content. Store: Extracted data are stored and saved in any structured format such as CSV, Database or JSON. Analyse: The web scraper will parse the stored data, which can be transformed by writing code using different Python libraries. | Python. |
| **9** | [22] | Collect: Information is collected from a detailed category. Processing: The product pages of each category are then used to create a list of all product names and URLs (website addresses). Store: Lists of product pairs and categories, unique product names, and URLs are stored. Analyse: Hierarchies are examined to select all potential products and then filtered using regular expressions to include or exclude specific sub-categories and products. | Websites. Python and Selenium. |
| **10** | [15] | Download: Once the script is written and successfully tested, its execution can be fully automated: a scheduler starts the download at defined time intervals. Send alerts: Alerts can be created to send an email if a script does not run completely or if the download size is unusually small.<br>Save: The accessed data can be saved in any format with defined elements, such as a text or CSV file. | Supermarkets. |
| **11** | [23] | Extract: Data are extracted from social networks, valid entities of product names from authentic sites. Process: All unnecessary columns are removed, and only the name and comments column is retained. Visualise: The dataset is visualised with the SeaBorn heat map, and null and duplicate values are discarded. Split: Divide the datasets into training, test and validation sets, ensuring each set contains at least 1000 valid data. | Website. Python. |
| **12** | [24] | Define chain: Out-of-home food chains are defined as any chain where food or beverages are prepared for immediate consumption by the purchaser. Define menu: All out-of-home menu items with nutrition information online were included in the data collection. Collect: The string name, menu item name, section, description, size, etc., are collected. The data are cleaned and standardised. | Website. Python. |
| **13** | [13] | Retrieve: Daily, at a fixed time, the software detects all individual product web pages. These pages are then retrieved. Parse: The underlying code of websites is analysed to locate each piece of relevant information. Special characters in the code identify the start and end of each variable placed by the website programmers. Store: The software stores the extracted information in a database containing one record per product per day. | Food. Not mention. |
| **14** | [25] | Connect: The application connects to the company's | Online |

| N | Author | Methodology used | Software |
|---|---|---|---|
| | | server. To do this, the company has provided access to its API. The application can access the web server using an HTTP request and the credentials provided. Obtain: You get a list of all the products with their corresponding attributes. Store All the information obtained is exported to a JSON file and imported into a local database. An automatic process is executed once a day to keep this information updated. Analyse: Another IT process analyses each product to determine whether it is possible to reduce the price based on the following parameters previously defined by the company manager | stores. Selenium. |
| **15** | [26] | Extract: Web scraping is the primary method of data collection. Store: The extracted data are stored in CSV format. The scraped data still contains many missing values and inconsistencies in the specifications. Clean and validate: The data must be cleaned and validated before further processing. | Sitio web. Python. |
| **16** | [6] | Web scraping: A programmer distributes daily tasks that execute Python scripts for specific web shops. Each script is tailored to a given webshop. Sorting: In contrast to the traditional survey-based collection method, web scraping enables using information on all items available in online shops quickly and cost-effectively. Data: The database contains almost 159 million quotes for around 650 thousand products. Forecast: The Inflation rate is defined as the monthly (not seasonally adjusted) price changes. | Food. Phyton. |
| **17** | [27] | Clean up: Starting with the discarded web files, the authors perform some basic operations to check for missing data and other basic validation operations. Merge: All files obtained from the data collection process for a given month are automatically merged. Restricting: The observations made for the variable “name” represent a guarantor of the invariance of the technical and qualitative characteristics of the items. This technique is used to code the entire sample. | Food Websites. Robot. |
| **18** | [28] | Search: Before the user searches for any product, the system will undergo maintenance every day for a few hours and store the entries for each product. Web crawler: Products will be crawled and processed in the system hourly for up-to-date results. Search and sort database: The system sorts each product according to the product's popularity that day and the price of that product instead of the web pages. Display: After the calculation, the filtered products will be displayed to the user. | Websites. Not mention. |
| **19** | [5] | Query URL: When the web-scraping application is run, a query will be sent to the listed URL. Data: The server sends the data in response to the request and allows you to access the HTML tab. Parse and delete: The application parses and deletes the data in the HTML file. Prediction: Prediction by supervised machine learning as | Websites. |

| N | Author | Methodology used | Software |
|---|---|---|---|
| | | a prediction of the decision tree algorithm is useful. | |
| **20** | [8] | Invoke: The initial phase of this model is to invoke the selenium library driver module. Reaching the URL: After the invocation, one hits the URL of the target website ('Hit' means browsing the URL that the machine will confirm the specified number of times. One must know about web page design to automate it). Prepare and analyse: The separated information is then prepared and parsed through the Beautifulsoup4 library and stored in a new object. Predict: Fure predictions can be made, and the results are displayed. | Python and other tools. |

## 4.3 Benefits of Using WS

There are many benefits to using WS. Based on the literature review, search accuracy varies between 85% [29], 86% [14] and 93% [30]. Additional benefits of WS are shown in Table 2.

**Table 2.** WS benefits

| Description | Author |
|---|---|
| Avoid spending extra time and money/Identify low prices/Suggest the best price/Save time on shopping | [30] , [31], [28], [32] |
| Compare prices from different online sites/Compare your prices with those of competitors to enable price regulation/Monitor competitors' prices and track increases | [21], [29], [25], [33], [34] |
| Report consumer prices in real time/Identify price differences in real time/Inform users of price trends | [15], [4] |
| Provide customers with a better shopping experience | [28] |
| Compare prices and experiences | [35] |
| Compare products | [14] |
| Select better products according to needs | [20] |
| Proper customer orientation | [36] |

## 4.4 Discussion

The results of the review show that the most widely used programming language is Python (including its libraries), potentially due to its easy-to-learn language and syntax. In recent years, publications related to WS have increased. This growth aligns with the rise of e-commerce during the COVID-19 period, when global health restrictions significantly accelerated online shopping. The results align with those of Benedetti et al. [6], who found that online prices have gained interest in the empirical literature.

Regarding the methodology, there appears to be no standardised approach. However, the steps most frequently mentioned by the authors include collecting, processing and

storing. Based on the commonalities found in the authors' methodologies and the analysed steps, the proposed methodology is as follows: Step 1: Collect, Step 2: Process, Step 3: Store and finally Step 4: Analyse and/or predict. Among the types of scraping are website, price, online stores and products.

In terms of benefits, the focus is primarily on price and time. This includes real-time price tracking, comparing your prices with competitors and adjusting your prices accordingly. These benefits are supported by tangible results, such as system performance metrics or user surveys, and are recognised as valuable to both consumer and retailer stakeholders.

## 5 Conclusions

This article presents a narrative literature review on WS in the purchase and sale of food and products. The selected databases were Web of Science and Scopus. Publications from the period 2018–2022 were reviewed and 20 articles were selected. The results reveal various types of WS, with a focus on websites, prices, online stores and products. Python is the most commonly used programming language, making WS accessible to both people with IT knowledge and those without. Although no standard methodology was identified, the most frequently cited steps were collecting, processing and storing data. The review also highlights key benefits of using WS for food and products, such as time savings and the ability for users to compare or visualise prices. While consumers reap the majority of these advantages, retailers also benefit. The limitations of this article include not incorporating all relevant scientific databases, such as IEEE Xplore, SpringerLink and Wiley, among others, which could have expanded the range of articles reviewed.

## References


1. Ayoubkhani, D., Thomas, H.: Estimating Weights for Web-Scraped Data in Consumer Price Indices. Journal of Official Statistics. 38, 5–21 (2022). https://doi.org/10.2478/jos-2022-0002
2. Dahiya, R., Nidhi, Kumari, K., Kumari, S., Agarwal, N.: Usage of Web Scraping in the Pharmaceutical Sector. EAI Endorsed Transactions on Pervasive Health and Technology. 9, 1–4 (2023). https://doi.org/10.4108/eetpht.9.4312
3. Khder, M.A.: Web scraping or web crawling: State of art, techniques, approaches and application. International Journal of Advances in Soft Computing and its Applications. 13, 144–168 (2021). https://doi.org/10.15849/ijasca.211128.11
4. Benedetti, I., Laureti, T., Palumbo, L., Rose, B.M.: Computation of High-Frequency Sub-National Spatial Consumer Price Indexes Using Web Scraping Techniques. Economies. 10, (2022). https://doi.org/10.3390/economies10040095
5. Saranya, G., Gopinath, N., Geetha, G., Meenakshi, K., Nithya, M.: Prediction of Customer Purchase Intention Using Linear Support Vector Machine in

Digital Marketing. Journal of Physics: Conference Series. 1712, (2020). https://doi.org/10.1088/1742-6596/1712/1/012024
6. Macias, P., Stelmasiak, D., Szafranek, K.: Nowcasting food inflation with a massive amount of online prices. International Journal of Forecasting. 39, 809–826 (2023). https://doi.org/10.1016/j.ijforecast.2022.02.007
7. Cavallo, A., Rigobon, R.: The billion prices project: Using online prices for measurement and research. Journal of Economic Perspectives. 30, 151–178 (2016). https://doi.org/10.1257/jep.30.2.151
8. Sidhu, A.S., Misra, N., Kaushik, V., Shankar, A., Joshi, K., Singh, R.: Analysis of Global Finance Using Web Scraping and Topic Modeling. Proceedings of 3rd International Conference on Intelligent Engineering and Management, ICIEM 2022. 747–753 (2022). https://doi.org/10.1109/ICIEM54221.2022.9853165
9. Addo, P.C., Akpatsa, S.K., Nukpe, P., Ohemeng, A.A., Kulbo, N.B.: Digital analytics approach to understanding short video advertising in digital marketing. Journal of Marketing Theory and Practice. 30, 405–420 (2022). https://doi.org/10.1080/10696679.2022.2056487
10. Chaithra, Malleshappa, L.G., Sreenivasaiah, J.: Classification of Web Pages Using the Machine Learning Algorithms with Web Page Recommendations. International Journal of Intelligent Engineering and Systems. 15, 637–650 (2022). https://doi.org/10.22266/ijies2022.0831.57
11. Zulkifli, N.D., Satari, S.Z.: Insight on Competency Skills from Job Advertisement for Data Scientists. Journal of Advanced Research in Applied Sciences and Engineering Technology. 44, 25–34 (2025). https://doi.org/10.37934/araset.44.2.2534
12. Bricongne, J.C., Meunier, B., Pouget, S.: Web-scraping housing prices in real-time: The Covid-19 crisis in the UK. Journal of Housing Economics. 59, 101906 (2023). https://doi.org/10.1016/j.jhe.2022.101906
13. Jaworski, K.: Measuring food inflation during the COVID-19 pandemic in real time using online data: a case study of Poland. British Food Journal. 123, 260–280 (2021). https://doi.org/10.1108/BFJ-06-2020-0532
14. Chaudhari, S., Aparna, R., Tekkur, V.G., Pavan, G.L., Karki, S.R.: Ingredient/Recipe Algorithm using Web Mining and Web Scraping for Smart Chef. Proceedings of CONECCT 2020 - 6th IEEE International Conference on Electronics, Computing and Communication Technologies. 22–25 (2020). https://doi.org/10.1109/CONECCT50063.2020.9198450
15. Hillen, J.: Web scraping for food price research. British Food Journal. 121, 3350–3361 (2019). https://doi.org/10.1108/BFJ-02-2019-0081
16. Ullah, H., Ullah, Z., Maqsood, S., Hafeez, A.: Web scraper revealing trends of target products and new insights in online shopping websites. International Journal of Advanced Computer Science and Applications. 9, 427–432 (2018). https://doi.org/10.14569/IJACSA.2018.090658
17. Ahmed, M., Schermel, A., Lee, J., Weippert, M., Franco-Arellano, B., L'Abbé, M.: Development of the Food Label Information Program: A Comprehensive Canadian Branded Food Composition Database. Frontiers in Nutrition. 8, 1–11 (2022). https://doi.org/10.3389/fnut.2021.825050

18. Bawiskar, M., Injamuri, V.: Aspect based mobile recommendation system. International Journal of Innovative Technology and Exploring Engineering. 8, 3840–3846 (2019). https://doi.org/10.35940/ijitee.J9996.0881019
19. Correa, J.C., Garzón, W., Brooker, P., Sakarkar, G., Carranza, S.A., Yunado, L., Rincón, A.: Evaluation of collaborative consumption of food delivery services through web mining techniques. Journal of Retailing and Consumer Services. 46, 45–50 (2019). https://doi.org/10.1016/j.jretconser.2018.05.002
20. Dyulicheva, Y.Y.: Auto-generation of the customer questions and their ranking in e-commerce system. Journal of Physics: Conference Series. 1399, (2019). https://doi.org/10.1088/1742-6596/1399/3/033081
21. Greeshma, N., Raghavendra, C., Prasad, K.R.: Visualization of optimal product pricing using e-commerce data. International Journal of Innovative Technology and Exploring Engineering. 9, 4441–4443 (2019). https://doi.org/10.35940/ijitee.A5262.119119
22. Harrington, R.A., Adhikari, V., Rayner, M., Scarborough, P.: Nutrient composition databases in the age of big data: FoodDB, a comprehensive, real-time database infrastructure. BMJ Open. 9, 1–10 (2019). https://doi.org/10.1136/bmjopen-2018-026652
23. Hossain, M.S., Nayla, N., Rassel, A.A.: Product Market Demand Analysis Using Nlp in Banglish Text With Sentiment Analysis and Named Entity Recognition. 2022 56th Annual Conference on Information Sciences and Systems, CISS 2022. 166–171 (2022). https://doi.org/10.1109/CISS53076.2022.9751188
24. Huang, Y., Burgoine, T., Essman, M., Theis, D.R.Z., Bishop, T.R.P., Adams, J.: Monitoring the Nutrient Composition of Food Prepared Out-of-Home in the United Kingdom: Database Development and Case Study, (2022)
25. Jorge, O., Pons, A., Rius, J., Vintró, C., Mateo, J., Vilaplana, J.: Increasing online shop revenues with web scraping: a case study for the wine sector. British Food Journal. 122, 3383–3401 (2020). https://doi.org/10.1108/BFJ-07-2019-0522
26. Listianingrum, T., Jayanti, D., Afendi, F.M.: Smartphone hedonic price study based on online retail price in Indonesia. Journal of Physics: Conference Series. 1863, (2021). https://doi.org/10.1088/1742-6596/1863/1/012032
27. Oancea, B.: Using Web scrapping techniques for price statistics-the Romanian experience. 1–10 (2021)
28. Rathod, U., Pavate, A., Patil, V.: Product Rank Based Search Engine for E-Commerce. 2018 3rd International Conference for Convergence in Technology, I2CT 2018. 1–5 (2018). https://doi.org/10.1109/I2CT.2018.8529503
29. Horch, A., Kett, H., Weisbecker, A.: Mining E-commerce Data from E-shop Websites. Proceedings - 14th IEEE International Conference on Trust, Security and Privacy in Computing and Communications, TrustCom 2015. 2, 153–160 (2015). https://doi.org/10.1109/Trustcom.2015.575
30. Alam, A., Anjum, A.A., Tasin, F.S., Reyad, M.R., Sinthee, S.A., Hossain, N.: Upoma: A Dynamic Online Price Comparison Tool for Bangladeshi E-commerce Websites. 2020 IEEE Region 10 Symposium, TENSYMP 2020. 194–197 (2020). https://doi.org/10.1109/TENSYMP50017.2020.9230862

31. Ma, S., Jiang, S., Ling, M., Chen, J., Shang, C.: Price Promotions of E-Liquid Products Sold in Online Stores. International Journal of Environmental Research and Public Health. 19, (2022). https://doi.org/10.3390/ijerph19148870
32. Mehak, S., Zafar, R., Aslam, S., Bhatti, S.M.: Exploiting filtering approach with web scrapping for smart online shopping : PPenny Wise: A wise tool for online shopping. 2019 2nd International Conference on Computing, Mathematics and Engineering Technologies, iCoMET 2019. (2019). https://doi.org/10.1109/ICOMET.2019.8673399
33. Sui, F.M., Chang, J.C., Hsiao, H.C., Su, S.C.: A study on entrepreneurial education regarding college students' creative tendency, entrepreneurship self-efficacy and entrepreneurial motivation. Presented at the (2018)
34. Lin, C.F., Yang, S.C.: Taiwan Stock Tape Reading Periodically Using Web Scraping Technology with GUI. Applied System Innovation. 5, (2022). https://doi.org/10.3390/asi5010028
35. Kwarteng, M.A., Jibril, A.B., Botha, E., Osakwe, C.N.: The Influence of Price Comparison Websites on Online Switching Behavior: A Consumer Empowerment Perspective. Springer International Publishing (2020)
36. Nurdin, Bustami, Hutomi, M., Elveny, M., Syah, R.: Implementation of the bfs algorithm and web scraping techniques for online shop detection in Indonesia. Journal of Theoretical and Applied Information Technology. 99, 2878–2889 (2021)