

PEMANFAATAN TEKNIK WEB SCRAPING PYTHON UNTUK SISTEM PENCARIAN PRODUK DI TOKO ONLINE

Adi Sopian¹⁾, Andy Dharmalau²⁾, Alpindo³⁾

¹⁾Prodi Sistem Informasi, Fakultas Teknologi, ITB Swadharma

^{2,3)}Prodi Teknik Informatika, Fakultas Teknologi, ITB Swadharma

Correspondence author: Adi Sopian, adisopian@swadharma.ac.id, Jakarta, Indonesia

Abstract

Sales and purchase transactions utilizing electronic systems with the internet through computer networks has grown rapidly. Buying goods in online stores sometimes the items you are looking for are not what you want, especially in price and item details, it can even happen that the item you are looking for is not found. In searching for goods using search engines such as google, yahoo, it will generate information on goods in the form of a list of website addresses. So that the information obtained is not as desired, people have to visit the online shop address one by one, causing people to waste a lot of time. The results of the research on the product search system using web scraping techniques to several e-commerce, namely tokopedia, shopee, evelania and blibli, have succeeded in obtaining data in the form of product names, number of stock items, url, rating and product images. So that people can easily find the desired item along with the data.

Keywords: e-commerce, python, search systems, web scraping

Abstrak

Transaksi penjualan dan pembelian memanfaatkan sistem elektronik dengan internet melalui jaringan komputer telah berkembang dengan pesat. Membeli barang di toko online terkadang barang yang dicari tidak sesuai yang diinginkan terutama dalam harga dan detail barang, bahkan bisa terjadi barang yang dicari tidak ditemukan. Dalam melakukan pencarian barang menggunakan search engine seperti google, yahoo maka akan menghasilkan informasi barang berupa daftar alamat website. Sehingga informasi yang didapat tidak sesuai yang diinginkan, masyarakat harus mengunjungi satu-persatu alamat toko online, menyebabkan masyarakat harus membuang banyak waktu. Tujuan dari penelitian ini membangun aplikasi berbasis web untuk menampung informasi dari berbagai toko online. Pengembangan aplikasi diperlukan teknik web scraping menggunakan fungsi dari Python. Hasil penelitian sistem pencarian produk dengan teknik web scraping ke beberapa e-commerce yaitu tokopedia, shopee, evelania dan blibli telah berhasil mendapatkan data berupa nama produk, jumlah stok barang, url, rating dan gambar produk. Sehingga masyarakat dengan mudah dapat mencari barang yang diinginkan berikut data-datanya.

Kata Kunci: toko online, python, pencarian produk, web scraping

A. PENDAHULUAN

Perkembangan teknologi informasi terus berkembang, kebanyakan masyarakat melakukan kegiatan sehari-hari tidak lepas dari teknologi informasi, mulai dari berinteraksi/berkomunikasi, berniaga, berpolitik, dan lain-lain. Kegiatan berniaga di masyarakat juga tidak lepas dari teknologi informasi, baik televisi maupun website (toko online). Perkembangan teknologi informasi juga melahirkan sistem baru dalam berjualan, dikenal dengan E commerce. Toko online atau dikenal e-commerce (Electronic Commerce) adalah suatu industri yang sudah melakukan transaksi penjualan dan pembelian memanfaatkan sistem elektronik seperti internet atau jaringan komputer (AYani, Pratiwi, & Muhandi, 2019).

Dalam transaksi di toko online terkadang barang yang dicari tidak sesuai yang diinginkan terutama dalam harga, bahkan bisa terjadi barang yang dicari tidak ditemukan (Setiono, Rostianingsih, & Purbowo, 2018). Untuk mengatasi masalah pencarian barang ini, masyarakat bisa menggunakan aplikasi mesin pencarian (search engine) untuk mencari barang yang diinginkan (Josi, Abdillah, & Suryayusra, 2014; Utomo, 2013).

Search engine merupakan program komputer yang dirancang sebagai alat bantu untuk mencari informasi di internet dengan cara mengetikkan kata kunci (keyword) yang dimaksud sehingga akan ditampilkan pada hasil pencarian yang berupa website asli yang berisi berbagai bentuk informasi seperti tulisan, gambar dan video (Rosario B, Pratama, & Fachruddin, 2017). Search engine memberikan pencarian content media dengan kriteria yang spesifik dan memperoleh daftar file yang memenuhi kriteria tersebut.

Search engine (mesin pencarian) yaitu kombinasi perangkat keras dan perangkat lunak komputer yang disediakan oleh perusahaan tertentu melalui website yang

telah ditentukan. Dalam melakukan pencarian barang menggunakan search engine seperti google, yahoo maka akan menghasilkan informasi barang berupa daftar alamat website bukan daftar barang untuk diperjual belikan ataupun hasil pencarian berupa artikel barang yang dicari, berita yang membahas mengenai kata kunci barang dicari. Sehingga dapat menyebabkan informasi yang dicari tidak sesuai yang diinginkan walaupun hasil pencarian mengenai daftar alamat website yang menjual barang yang dicari maka masyarakat harus mengunjungi satu-persatu alamat toko online untuk mencari dan membandingkan barang, yang dapat menyebabkan masyarakat harus membuang banyak waktu. Oleh karena itu diperlukan suatu aplikasi untuk menampung berbagai informasi barang dari berbagai toko online (Syarifudin, Syaifiandini, & Prasadana, 2018).

Aplikasi yang akan dibangun merupakan suatu aplikasi berbasis web untuk menampung informasi dari berbagai toko online Cara kerja aplikasi yang akan dibangun yaitu pertama aplikasi akan mengumpulkan informasi mengenai kata kunci barang yang dicari dari berbagai toko online yang sudah terdaftar di aplikasi tersebut. Untuk pengembangan aplikasi tersebut diperlukan teknik web scraping menggunakan fungsi dari Python.

Web Scraping adalah proses yang melibatkan pengambilan sebuah dokumen semi-terstruktur dari internet, umumnya halaman web dalam bahasa markup seperti HTML (Josi et al., 2014; Setiawan, Tristiyanto, & Hijriani, 2020). Untuk melakukan pengambilan data tersebut dapat memanfaatkan fungsi dari Python. Sehingga dengan memanfaatkan teknik web scraping dapat mengumpulkan informasi barang yang dicari dari berbagai toko online.

Web scraping proses yang memanfaatkan dokumen berbentuk semi terstruktur yang didapatkan dari internet, yang dimana dokumen tersebut berbentuk sebuah halaman website yang dibangun

oleh bahasa markup seperti HTML yang kemudian dianalisis untuk mendapatkan informasi yang berguna yang dapat dilakukan untuk konteks lain .

Implementasi web scraping dapat diterapkan untuk mendapatkan informasi tertentu dari sebuah website yang nantinya akan diolah untuk ditampilkan ke dalam bentuk lain yang menghasilkan informasi untuk para penggunaanya.

Pada perancangan aplikasi yang dilakukan peneliti, dilakukan dengan mengambil data dari berbagai toko online. Setelah berhasil memperoleh data dari website tersebut, aplikasi yang dibuat akan menampilkan barang yang di cari beserta harganya sehingga output yang di hasilkan nantinya adalah aplikasi untuk membandingkan harga produk yang di cari dari berbagai toko online.

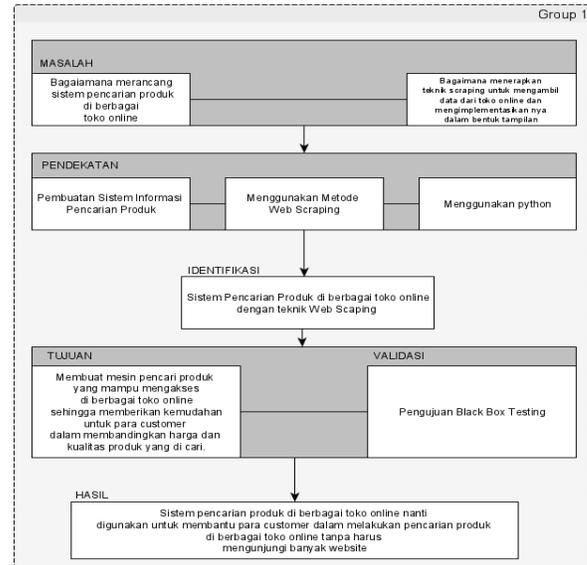
B. METODE PENELITIAN

Metode observasi merupakan pengamatan langsung ke lapangan untuk meneliti terhadap objek-objek dan dokumen-dokumen yang diamati, untuk mendapatkan informasi tentang dokumen yang digunakan, dan laporan yang diperlukan, serta data lain yang diperlukan untuk perancangan dan pengembangan sistem aplikasi yang akan dikembangkan. Selain dengan pengamatan ke lapangan juga mempelajari dari berbagai buku bacaan diperpustakaan, jurnal berita online yang erat kaitannya dengan tema penelitian, sehingga didapatkan dasar ilmiah yang kuat dan wawasan teori.

Selain itu dilakukan juga metode pengumpulan data dengan melakukan tanya jawab secara langsung dengan customer yang terkait dengan proses pencarian produk ke berbagai toko online.

Aplikasi yang dibangun merupakan suatu aplikasi berbasis web untuk menampung informasi dari berbagai toko online. Pengembangan aplikasi tersebut diperlukan teknik web scraping menggunakan fungsi dari Python.

Pengambilan data menggunakan metode web scraping dengan memanfaatkan fungsi dari Python, untuk mengumpulkan informasi barang yang dicari dari berbagai website toko online.



Gambar 1. Kerangka Penelitian

Penelitian ini dilakukan bertujuan agar sistem informasi ini dapat dikembangkan dan digunakan untuk membantu proses pencarian produk di berbagai toko online, dengan web scraping produk penjualan online pada e-marketplace seperti Tokopedia, Bukalapak, shopee, elevania, dan blibli yang dilakukan oleh pengguna.

C. HASIL DAN PEMBAHASAN

Perancangan web scraping dirancang dengan menggunakan pendekatan terstruktur, berikut ini adalah hasil analisa, perancangan dan implementasi sistem.

Analisa Kebutuhan Sistem

Analisa kebutuhan sistem yang dibuat yaitu:

1. Website
 - a. Tokopedia
 - b. Elevania
 - c. Shopee

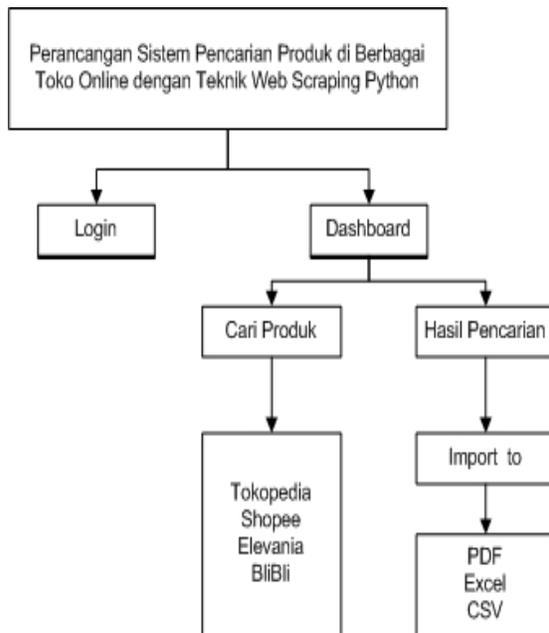
- d. Bibli
- 2. Input
 - a. Input login
 - b. Input keyword produk
- 3. Output
 - a. Hasil scraping produk
 - b. Import PDF, Excel dan CSV



Gambar 3. Struktur Web Sistem Pencarian Produk

Diagram Dekomposisi

Berikut ini adalah dekomposisi dari sistem wab scraping:



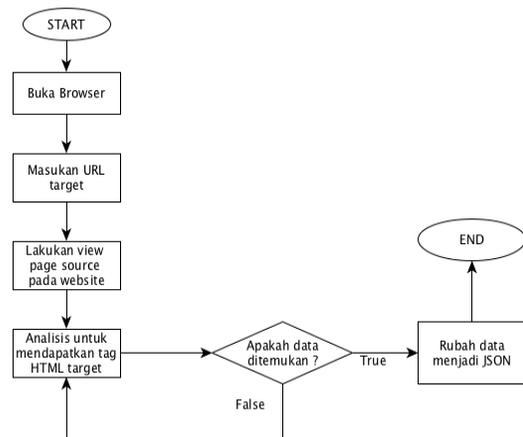
Gambar 2. Diagram Dekomposisi

Analisa Web Sistem Pencarian Produk

Website E-Commerce perlu dianalisa untuk mengetahui hubungan atau link antar halaman dan letak data yang akan diekstrasi. Hubungan link antara halaman diperlukan untuk melakukan proses scraping.

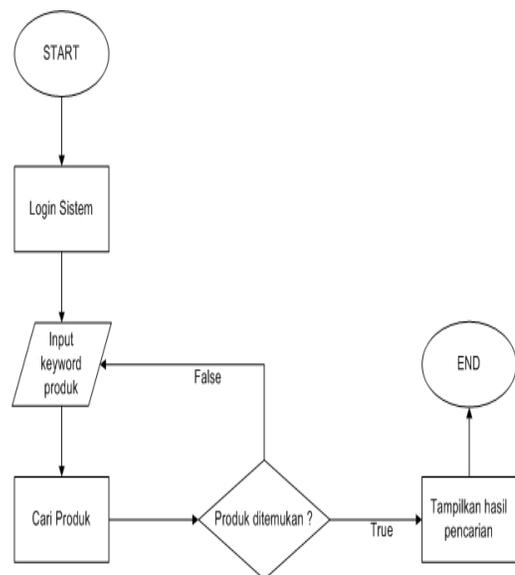
Ada 3 halaman utama yang menjadi perhatian dan letak sumber data. Pertama halaman home sebagai awal akses ke web E-Commerce tersebut, kedua halaman input keyword produk dan yang ketiga adalah halaman list produk dimana terdapat daftar semua data produk, Hubungan antar halaman tersebut dapat dilihat pada Gambar 3.

Flowchart Scraping Sumber Data



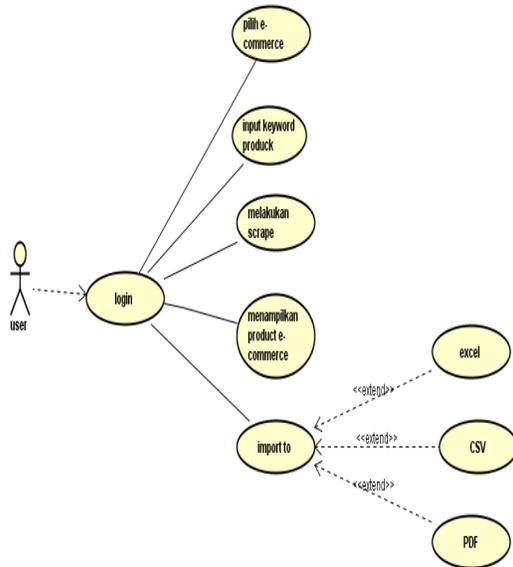
Gambar 4. Flowchart Scraping Sumber Data

Flowchart Sistem Pencarian Produk



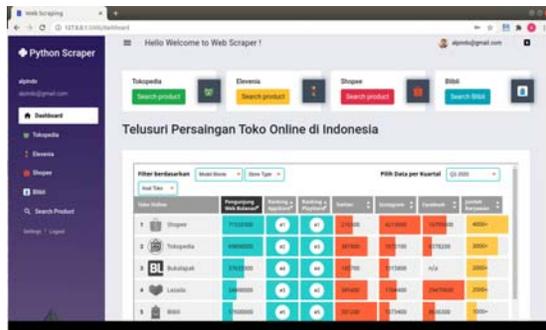
Gambar 5. Flowchart sistem pencarian produk

Use Case

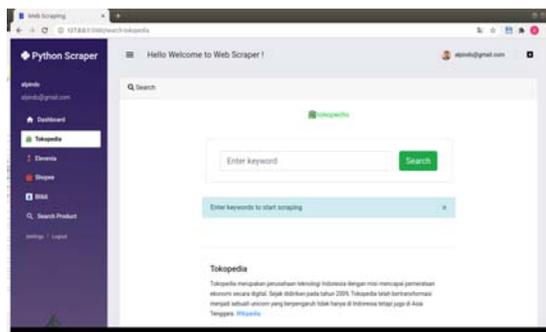


Gambar 6. Use Case Diagram

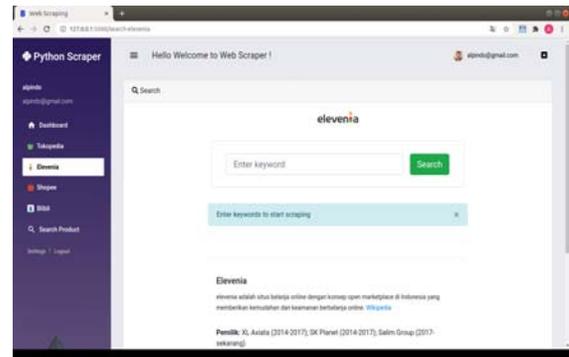
Rancangan Tampilan Antarmuka



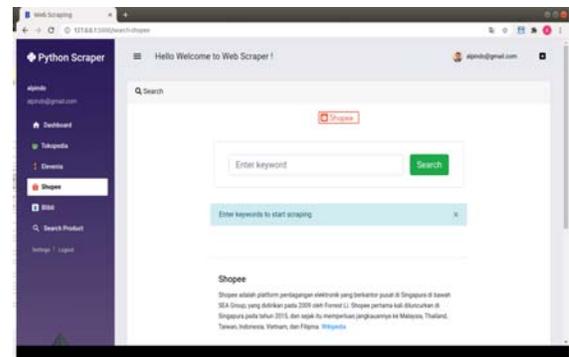
Gambar 7. Dashboard



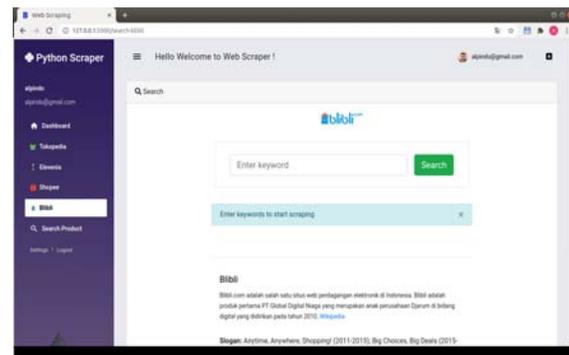
Gambar 8. Halaman Pencarian Tokopedia



Gambar 9. Halaman Pencarian Elexenia

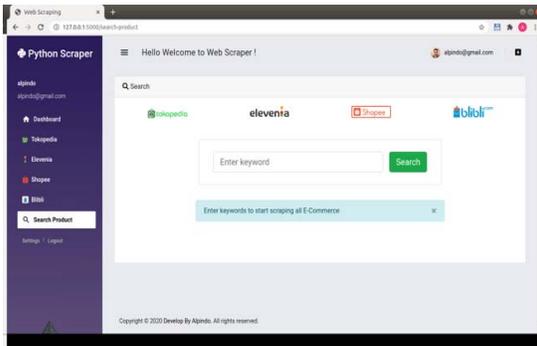


Gambar 10. Halaman Pencarian Shopee

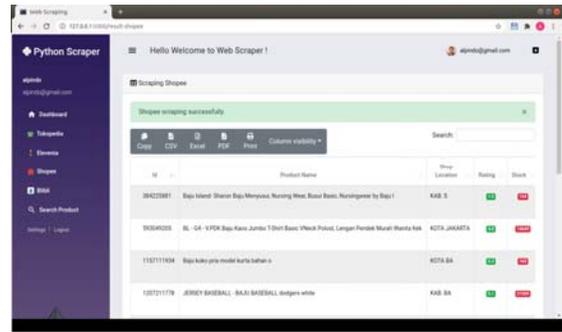


Gambar 11. Halaman Pencarian Blibli

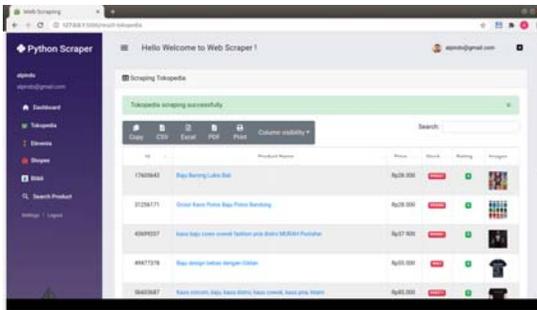
Pemanfaatan Teknik Web Scraping Python Untuk Sistem Pencarian Produk di Toko Online
 Adi Sopian, Andy Dharmalau, Alpindo



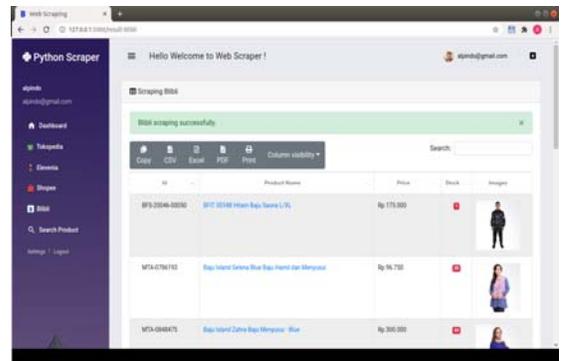
Gambar 12. Halaman Pencarian Tokopedia, Elevenia, Shopee dan Blibli



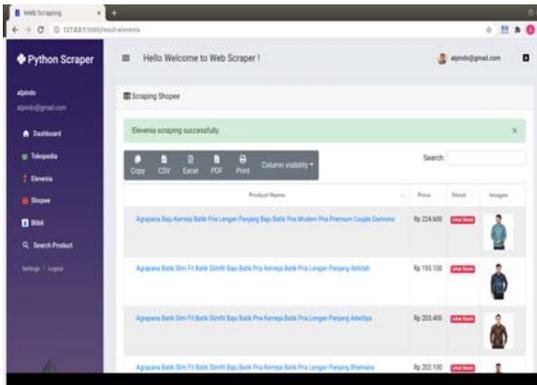
Gambar 15. Halaman Hasil Pencarian Produk Shopee



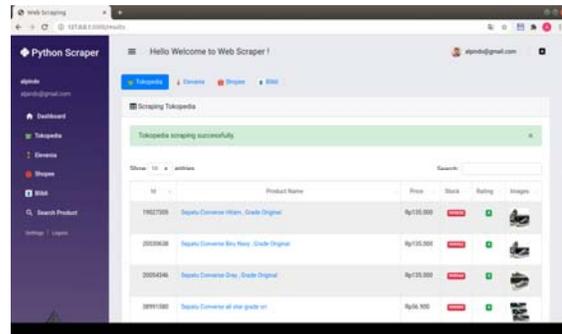
Gambar 13. Halaman Hasil Pencarian Produk Tokopedia



Gambar 16. Halaman Hasil Pencarian Produk Blibli



Gambar 14. Halaman Hasil Pencarian Produk Elevenia



Gambar 17. Hasil Pencarian Produk Tokopedia, Elevenia, Shopee, dan Blibli

Web Scraping



Id	Product Name	Price	Stock	Rating	Images
17605643	Baju Barong Lukis Bali	Rp28.000	999601	4	
31256171	Grosir Kaos Polos Baju Polos Bandung	Rp28.000	995085	5	
43696207	kaos baju cowo cowok fashion pria distro MURAH Punisher	Rp37.900	999081	5	
49477278	Baju design bebas dengan Gildan	Rp55.000	9957	5	
56603687	Kaos volcom, baju, kaos distro, kaos cowok, kaos pria, hitam	Rp85.000	999971	5	
64243638	Kaos Baju Gildan Softstyle Colour 63000 warna Original GROSIR S M L XL	Rp34.000	998192	5	
64244505	Kaos Baju Gildan Softstyle Colour 63000 warna Original GROSIR XXL	Rp40.000	999507	5	
64452111	Kaos Baju Gildan Custom Desain Bebas Suka-suka Gambar Fullcolour Uk A4	Rp68.750	9574	5	
64458010	Kaos Baju Gildan Custom Desain Bebas Suka-suka Tuisan 1 warna Uk A4	Rp58.500	9745	5	

Gambar 18. Export PDF Tokopedia

Web Scraping



Product Name	Price	Stock	Images
Agrapana Baju Kemeja Batik Pria Lengan Panjang Baju Batik Pria Modern Pria Premium Couple Dapinono	Rp 224.600	Lihat Stock	
Agrapana Batik Slim Fit Batik Slimfit Baju Batik Pria Kemeja Batik Pria Lengan Panjang Abhitah	Rp 193.100	Lihat Stock	
Agrapana Batik Slim Fit Batik Slimfit Baju Batik Pria Kemeja Batik Pria Lengan Panjang Adeniyah	Rp 203.400	Lihat Stock	
Agrapana Batik Slim Fit Batik Slimfit Baju Batik Pria Kemeja Batik Pria Lengan Panjang Bhamana	Rp 202.100	Lihat Stock	
Agrapana Batik Slim Fit Batik Slimfit Baju Batik Pria Kemeja Batik Pria Lengan Panjang Bhamana 2	Rp 202.100	Lihat Stock	
Agrapana Batik Slim Fit Batik Slimfit Baju Batik Pria Kemeja Batik Pria Lengan Panjang Damiano	Rp 224.600	Lihat Stock	
Agrapana Batik Slim Fit Batik Slimfit Baju Batik Pria Kemeja Batik Pria Lengan Panjang Hanasta	Rp 199.200	Lihat Stock	

Gambar 19. Export PDF Elevania

Web Scraping



Id	Product Name	Shop Location	Rating	Stock	Images
384225881	Baju Island- Sharon Baju Menyusui, Nursing Wear, Busui Basic, Nursingwear by Baju I	KAB. S	0.3	140	
593049205	Bl - G4 - VPOK Baju Kaos Jumbo T-Shirt Basic Vneck Polos, Lengan Pendek Murah Wanita Kek	KOTA JAKARTA	2.9	13083	
115711934	Baju koko pria model kurta bahan o	KOTA BA	2.8	275	
120721178	JERSEY BASEBALL - BAJU BASEBALL dodgers white	KAB. BA	3.0	26282	
1442695491	Promo Temurah Baju Manset Atasan Dalemah kaftan / inner All	KOTA JAKARTA	5.1	234327	
1457360109	PEPOY SHOP Stelan Piyama Baju Tidur Tangan Panjang (PP) Emboss Motif Wanita LD 1	KOTA JAKARTA	0.8	50	
1489344053	BAJU KOKO LENGAN PANJANG LIST HEDLONG FASHION MUSLIM MODEL KURTA KOKO	KOTA BA	4.9	162	
1553875023	Baju baseball NY sahur	KAB. BA	4.4	1363	

Gambar 20. Export PDF Shopee

Web Scraping



Id	Product Name	Price	Stock	Images
BFS-20046-00050	BFIT 30348 Himam Baju Sauna L/XL	Rp 175.000	1	
MTA-0786193	Baju Island Selena Blue Baju Hamil dan Menyusui	Rp 96.750	30	
MTA-0848475	Baju Island Zahra Baju Menyusui - Blue	Rp 300.000	44	
MTA-1163111	Baju Island Debby Dress Hamil & Menyusui	Rp 126.720	0	
MTA-1177514	Baju Island Aiko Baju Menyusui	Rp 275.000	0	
MTA-2259424	Baju Island Mala Nursingwear Baju Menyusui	Rp 650.000	2	
MTA-2461703	Baju Island Jacie Celana Hamil - Blue	Rp 87.200	2	
MTA-2466999	Baju Island Una Celana Hamil - Coklat	Rp 119.000	0	
MTA-3023192	Baju Island Ling Nursingwear Dress Menyusui	Rp 189.000	0	
MTA-3349079	Wellborn X Blibli Uncontrolled T-Shirt Pria - Burgundy	Rp 95.700	16	
MTA-3790299	Hijab 8313 Baju Gamis Wanita	Rp 75.000	14	

Gambar 21. Export PDF Blibli

D. PENUTUP

Berdasarkan dari hasil yang telah diuraikan sebelumnya, telah dibangun sistem pencarian produk dengan teknik *Web Scraping* menggunakan Python untuk mempermudah para *customer* dalam mencari barang dan mengisi deskripsi produk dengan mudah dan cepat. Penerapan teknik *Web Scraping* dapat membantu mengatasi permasalahan dalam melakukan pencarian produk. Karena Teknik ini mendapatkan data dari server toko online tersebut. sehingga data tersebut akurat dan valid.

Proses *scraping* telah berjalan dengan baik berdasarkan pengujian yang telah dilakukan. Sistem yang berjalan saat ini masih bergantung terhadap server toko online tersebut. Apabila server toko online terjadi kendala, maka sistem tersebut juga akan mengalami kendala. Apabila toko online tersebut merubah struktur data. Maka sistem yang di buat juga harus merubah sturktur kode untuk mendapatkan data terbaru dan menghindari error pada saat melakukan *scraping*.

Atas dasar permasalahan, analisis dan rancangan yang telah diuraikan sebelumnya, disarankan untuk menambahkan filter pada sistem pencarian sehingga hasil pencarian yang ditampilkan lebih spesifik lagi. Mengembangkan sistem informasi *web scraping* ke dalam bentuk *mobile* yang menggunakan Android dan iOS sehingga mempermudah dalam mengakses sistem informasi *web scraping*.

E. DAFTAR PUSTAKA

AYani, D. D., Pratiwi, H. S., & Muhardi, H. (2019). Implementasi web scraping untuk pengambilan data pada situs marketplace. *Jurnal Sistem Dan Teknologi Informasi (JUSTIN)*, 7(4), 257.

<https://doi.org/10.26418/justin.v7i4.309>
30

- Josi, A., Abdillah, L. A., & Suryayusra. (2014). *Penerapan teknik web scraping pada mesin pencari artikel ilmiah*. 5. Retrieved from <http://arxiv.org/abs/1410.5777>
- Rosario B, M., Pratama, Y., & Fachruddin. (2017). Penerapan web scraping pada website company profile. *Kntia*, 4(4), 37–43.
- Setiawan, D. F., Tristiyanto, T., & Hijriani, A. (2020). Aplikasi web scraping deskripsi produk. *Jurnal Teknoinfo*, 14(1), 41. <https://doi.org/10.33365/jti.v14i1.498>
- Setiono, A., Rostianingsih, S., & Purbowo, A. N. (2018). Website pencari produk terlaris untuk membantu penjualan pada toko online. *Jurnal Infra*, 3–7. Retrieved from <http://publication.petra.ac.id/index.php/teknik-informatika/article/view/7462>
- Syaifudin, Y. W., Syafiandini, A. F., & Prasadana, H. R. (2018). Aplikasi pencarian penjualan laptop menggunakan teknologi web scraping. *Jurnal Informatika Polinema*, 4(4), 246. <https://doi.org/10.33795/jip.v4i4.214>
- Utomo, M. S. (2013). Web scraping pada situs wikipedia menggunakan metode ekspresi regular. *Jurnal Teknologi Informasi DINAMIK*, 18(2), 153–160.