

Instrukcja korzystania z aplikacji:

WebScrapper

Type URL address from Otodom.pl

URL address *

EXTRACT

TRANSFORM

LOAD

ETL

EXPORT TO .CSV

Extract detail:

Please type URL address

Transform detail:

Data:

Opis elementów aplikacji:

Aplikacja podzielona została na kilka sekcji:

1. Menu funkcyjne zawierające:

- Pole typu input – pozwalające wprowadzić adres URL strony, którą użytkownik będzie Scrapował wraz z walidacją wymagającą wpisania poprawnego adresu ze strony otodom.pl

URL address *

URL address *

!

Wypełnij to pole.

URL address *

test

Url have to be from otodom.pl

- b. Przycisku funkcyjne odpowiedzialne za uruchamianie kolejnych etapów procesu ETL osobno - 'EXTRACT', 'TRANSFORM', 'LOAD', przycisk 'ETL' wykonujący cały proces automatycznie oraz 'EXPORT TO .CSV' generujący plik z wszystkimi danymi znajdującymi się w tabeli w formacie .csv. W początkowej fazie działania aplikacji nie ma możliwości użycia przycisków wyszarzonych, ponieważ wymagają one ukończenia poprzedniego etapu procesu.



2. Informacje o postępie i efektach procesu

- a. *EXTRACT* – sekcja mówiąca o postępie etapu *EXTRACT*. Informuje nas czy etap zakończył się sukcesem lub czy podany adres URL jest błędny, ponieważ nie zawiera żadnych ogłoszeń możliwych do pobrania (np. strona startowa domeny otodom.pl).

Extract detail:

Please type URL address

Extract detail:

Invalid URL address - please type correct address

Extract detail:

Success - Data extracted !!!

Po pozywanym zakończeniu etapu *EXTRACT* przycisk *TRANSFORM* staje się aktywny i możemy przejść do kolejnego procesu.



- b. *TRANSFORM* – informuje nasz ile ofert zostało pobranych i pozywanie przeszło etap *TRANSFORM* po czym aktywuje kolejny etap *LOAD*:

Transform detail:

Transformed 26 Results



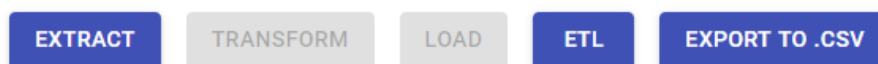
- c. *DATA* – sekcja w której załadowane zostają pobrane dane w formie tabeli. Tabela posiada sortowanie, wyszukiwarkę, możliwość przesuwania i decydowania o kolejności kolumn, możliwość zmiany ilości wierszy oraz w pierwszej kolumnie 'Action' przycisk do exportu pojedynczego wiersza i zapisania go w formacie .txt.

Data:

Adverts									
Q Search X									
Actions	Title	Rooms	Area	Price	Bond	BuildingType	FloorsInBuilding	Windows	HeatingType
	Rumia ul.Jeziorna, 2 pokoje, ogródek, Orlex Invest	2	38,46 m²	215 000 zł	0	blok	3	plastikowe	gazowe
	Inwestuj w najem Okazja *** 26m2, przy Stacji SKM!	2	26 m²	215 000 zł	0	blok	0	Brak Informacji	Brak Informacji
	komfortowe m3 parter duży balkon umeblowane !	2	36,60 m²	215 000 zł	0	Brak Informacji	10	Brak Informacji	Brak Informacji
	Mieszkanie, 33,50 m², Gorzów Wielkopolski	2	33,50 m²	215 000 zł	0	Brak Informacji	2	plastikowe	miejskie
	II Piętro- Super Lokalizacja!!!	2	37,70 m²	215 000 zł	0	blok	4	plastikowe	miejskie

5 rows |< < 1-5 of 26 > >|

Po zakończeniu etapu *LOAD* możemy wyeksportować całą tabelę do pliku .csv jednocześnie wyłączmy możliwość wykonania etapu *TRANSFORM* oraz *LOAD* umożliwiając jedynie rozpoczęcie procesu na nowo i kolejne zasilenie hurtowni danych.



Podczas działania każdego procesu uruchamiany jest loader w celu zablokowania interfejsu użytkownika na czas trwania procesu:

WebScrapper

Type URL adress from [Otodom.pl](#)

URL adress *

https://www.otodom.pl/sprzedaz/mieszkanie/?search%5Bfilter_float_c

EXTRACT

TRANSFORM

LOAD

ETL

EXPORT TO .CSV

Please type URL adress

Extract detail:

Transform detail:

Data: