

Instrukcja obsługi aplikacji:

Instrukcja uruchomienia aplikacji:

Uruchomienie części front-end'owej:

Aby uruchomić wersję deweloperską aplikacji front-end'owej najlepiej skorzystać z wiersza poleceń (domyślny dostępny w systemie operacyjnym Windows). Powinniśmy znajdować się w katalogu aplikacji front-end'owej:

```
~\WebScrapper\webscrapper-frontend
```

Z tego miejsca uruchamiamy polecenie:

```
npm start
```

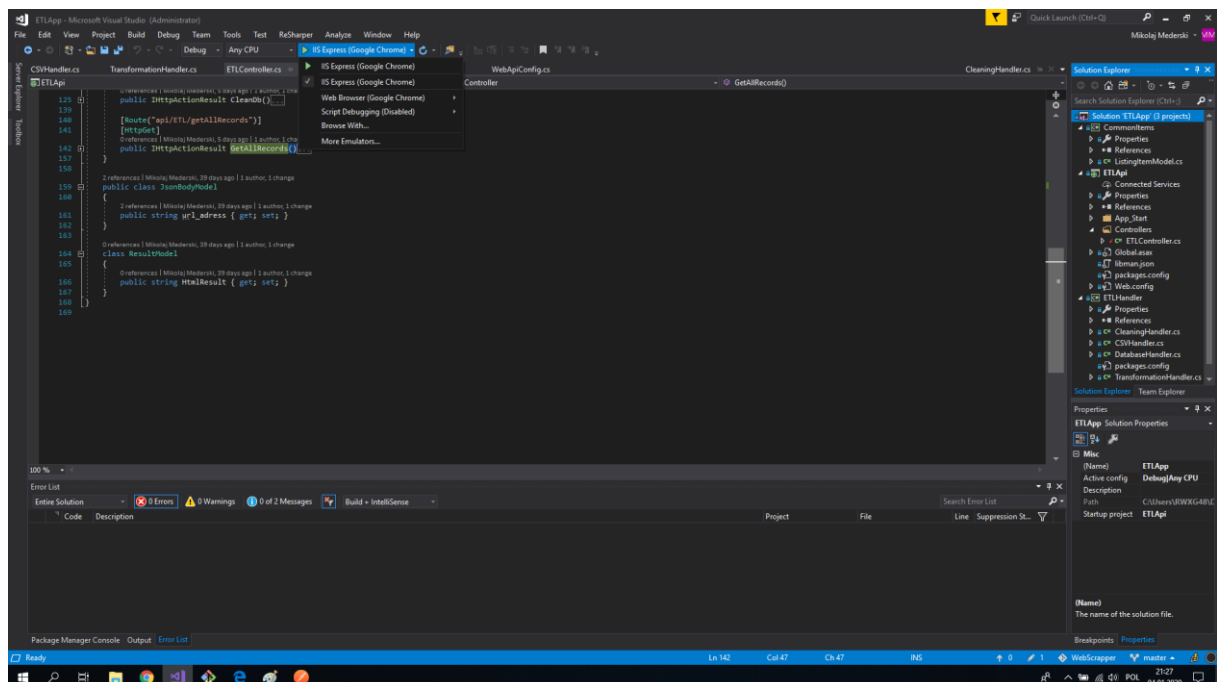
Powyższa komenda automatycznie otworzy domyślną przeglądarkę i uruchomi aplikację pod portem:

<http://localhost:3000/>

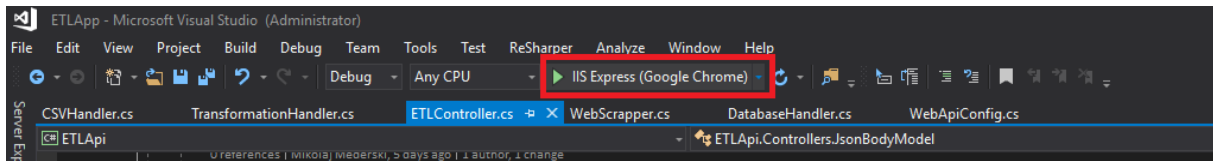
Uruchomienie części back-end'owej:

Aby uruchomić API obsługujące front-end aplikacji należy:

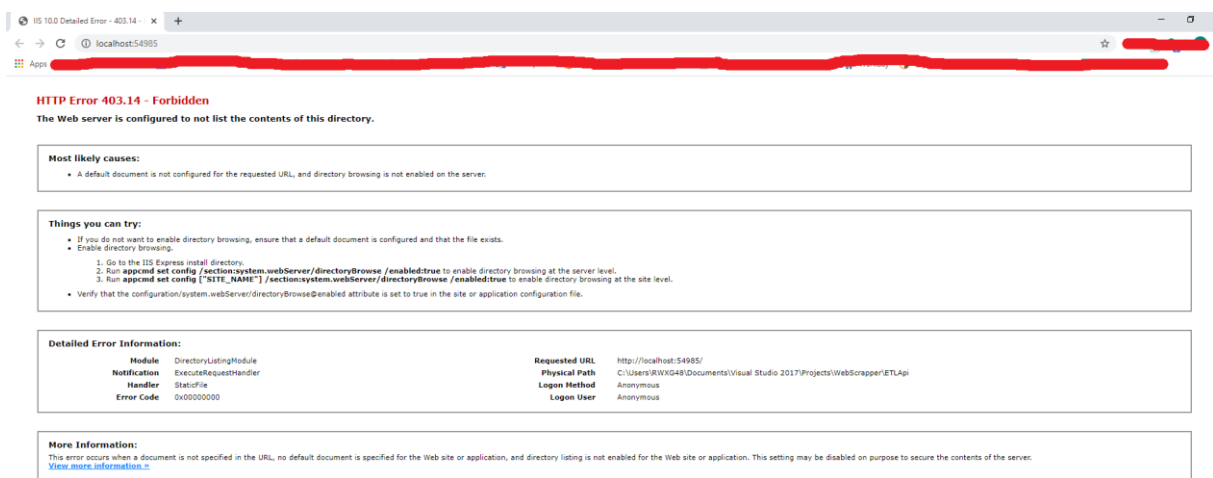
- Uruchomić program Visual Studio 2017
- Otworzyć plik "ETLApp.sln" w programie Visual Studio 2017 wybierając kolejno w lewym górnym rogu ekranu File -> Open -> Project/Solution
- W menu podświetlonym na zdjęciu poniżej wybrać opcję IIS Express



- Kliknąć w przycisk Startu



Po kliknięciu w przycisk Start, otworzy się nowe okno przeglądarki. Komunikat „**403 Forbidden**” wskazuje, na poprawne działanie serwisu.



Uruchomienie bazy danych

Opis funkcjonalności aplikacji:

Aplikacja ma na celu pobranie danych ze źródła (otodom.pl), odpowiedniego ich przekształcenia i zasilenia nimi bazy danych. Są to składowe procesy ETL, który główną funkcjonalnością naszej aplikacji. Ponadto istnieje możliwość exportu danych do pliku .csv lub .txt. Dane wyświetlane w tabeli możemy filtrować, sortować w sposób umożliwiając użytkownikowi jak najlepsze ich przetwarzanie.

Opis scenariuszy użycia aplikacji:

Aplikacja posiada dwa scenariusze użycia:

- W pierwszym musimy uruchomić wszystkie etapy manualnie pamiętając, że do uruchomienia kolejnego konieczne jest pozytywne zakończenie etapu poprzedniego.
- W drugim scenariuszu cały proces wykonywany jest automatycznie przy użyciu przycisku „ETL”.

Opis menu i widoków okna aplikacji:

The screenshot shows the 'WebScraper' application interface. At the top, there's a dark header with the title 'WebScraper'. Below it, a prompt says 'Type URL address from Otodom.pl'. A text input field is labeled 'URL address *'. Below the input field are five buttons: 'EXTRACT' (blue), 'TRANSFORM' (grey), 'LOAD' (grey), 'ETL' (blue), and 'EXPORT TO .CSV' (grey). Below the buttons, there are three sections: 'Extract detail:', 'Transform detail:', and 'Data:', each followed by a text input field. A small text 'Please type URL address' is visible on the left side of the 'Extract detail:' section.

Aplikacja podzielona została na kilka sekcji:

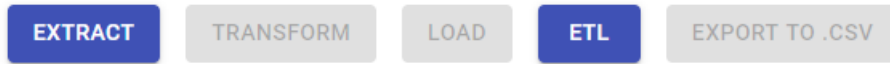
1. Menu funkcyjne:

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

This screenshot shows a close-up of the 'URL address *' input field. The field is empty. Below it, a tooltip with an orange exclamation mark icon and the text 'Wypełnij to pole.' (Fill this field) is visible. Below the tooltip, the input field now contains the text 'test'. At the bottom, a red error message reads '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 adress

Extract detail:

Invalid URL adress - please type correct adress

Extract detail:

Success - Data extracted !!!

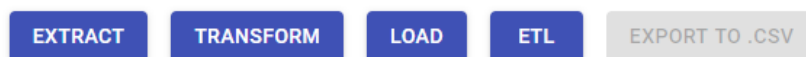
Po pozywanym zakończeniu etapu *EXTRACT* przycisk *TRANSFORM* staje się aktywny i możemy przejść do kolejnego procesu.



- b. *TRANSFORM* – informuje nas 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									Search	×
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 wykonywania 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

[Extract detail:](#)



[Transform detail:](#)

[Data:](#)

Please type URL adress