

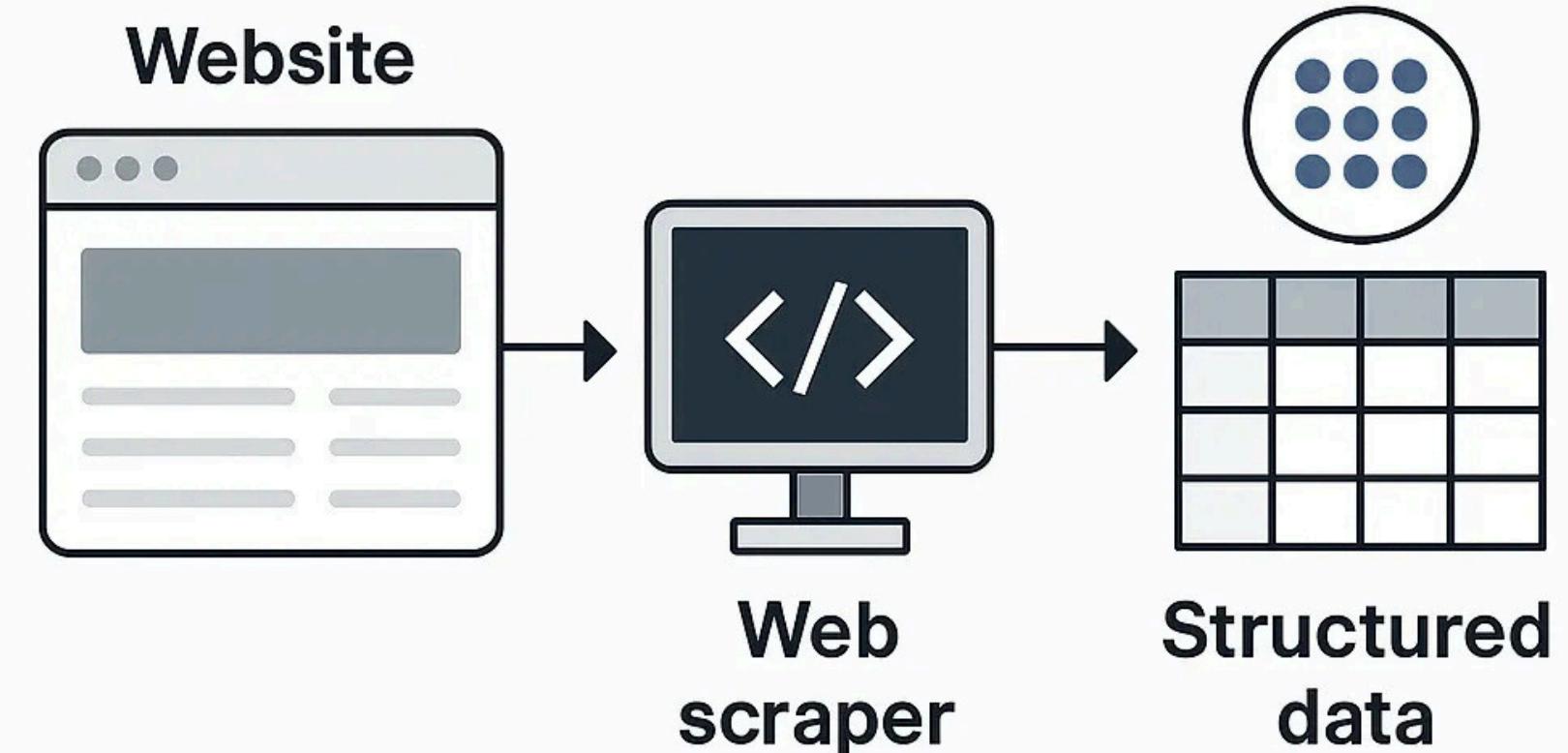
# **WEB SCRAPING Z PORTALU OTODOM.PL**

**ZUZANNA JASKUŁA**

# CZYM JEST WEB SCRAPING?

**Web scraping** jest techniką pozyskiwania danych z internetu w sposób automatyczny. Polega na przeglądaniu stron internetowych i wyciąganiu z nich informacji w formie, która nadaje się do dalszej analizy.

Jest wykorzystywany tam, gdzie dostęp do danych nie jest bezpośrednio udostępniony w formie bazy danych lub API.



# **DLACZEGO WARTO?**

- dostęp do wielu nowych źródeł danych,
- automatyczne pobieranie dużych ilości danych,
- szybsza analiza,
- dostęp do danych w czasie rzeczywistym.



# **PRZYKŁADY UŻYCIA**

- analiza ofert nieruchomości z portali ogłoszeniowych,
- zbieranie opinii i recenzji użytkowników,
- analiza ofert pracy,
- pobieranie informacji o produktach.

## ASPEKTY PRAWNE

Web scraping podlega ograniczeniom prawnym wynikającym z ochrony praw autorskich i ustawy o ochronie baz danych. Pobieranie danych do użytku własnego jest zazwyczaj dozwolone, lecz ich dalsze rozpowszechnianie może naruszać prawo. Efektywne scrapowanie wymaga przestrzegania wytycznych dla botów i scraperów zawartych w plikach *robots.txt* umieszczonych na stronach, unikania nadmiernego obciążania serwera i stosowania praktyk zgodnych z regulaminem witryny.



# NARZĘDZIA DO SCRAPINGU W PYTHONIE

Beautifulsoup

## BEAUTIFUL SOUP

- Biblioteka do parsowania HTML,
- Do statycznych stron,
- Licencja MIT (wolnego opr.).



## SELENIUM

- Platforma do automatyzacji,
- Do “poruszania się” po stronie,
- Obsługuje JavaScript,
- Licencja Apache (wolnego opr.).



## SCRAPY

- Framework do scrapingu,
- Do dużych projektów,
- Wydajny,
- Licencja BSD (wolnego opr.).

# **WYBÓR PORTALU OGŁOSZENIOWEGO**

Według badań **Otodom** jest w Polsce najskuteczniejszym portalem ogłoszeń nieruchomości. Korzysta z niego zdecydowana większość pośredników. Z tego powodu założyłam, że oferty znajdujące się na nim będą najbardziej szczegółowe i będą reprezentatywnym źródłem danych do analizy cech nieruchomości w Trójmieście.



# **SPOSÓBY POZYSKANIA DANYCH**

- Otodom Analytics (dla agencji nieruchomości i deweloperów),
- Firmy oferujące odpłatny scraping (Data Miners, ZenRows),
- Zbieranie danych ręcznie,
- Web scraping.



**ETYKA**

Plik *robots.txt* nie zabrania scrapowania danych z ogłoszeń. Jednocześnie regulamin portalu zakazuje agregowania i przetwarzania danych w celu dalszego udostępniania i w celach komercyjnych.

# SPECYFIKA STRONY

Otodom > Mieszkania na sprzedaż > pomorskie > Gdańsk

Mieszkania Na sprzedaż Gdańsk

Cena od zł do zł Powierzchnia od m<sup>2</sup> do m<sup>2</sup> Liczba pokoi 1 X 2 3 4 5 6+

Zapisz wyszukiwanie Wyczyszczyć kryteria Więcej filtrów Wyszukaj

**Mieszkania na sprzedaż: Gdańsk**

Sortuj Data dodania: najnowsze

1-36 ogłoszeń z 7486 [Jak pozycjonowane są ogłoszenia?](#)

**Wszystkie ogłoszenia**

**1 062 900 zł** 22 500 zł/m<sup>2</sup>  
2 pokoje - Gdańsk Śródmieście, Nadmotławie  
ul. Sienna Grobla, Śródmieście, Gdańsk, pomorskie  
2 pokoje 47.24 m<sup>2</sup> 2 piętro  
[Zobacz opis ogłoszenia](#)

**Ex Novo - Inwestycje i Nieruchomości**  
Biuro nieruchomości

**640 000 zł** 13 913 zł/m<sup>2</sup>  
Blisko Morza | Dwustronne | Balkon  
ul. Ksawerego Dunikowskiego, Brzeźno, Gdańsk, pomorskie

Dodane dzisiaj ZAKUP BEZPIECZNY ZAKUP

**Zobacz na mapie**

**Wyróżnione inwestycje**



**Opis**  
**Nadmotławie**  
Nadmotławie to prestiżowa inwestycja w centrum Gdańskiego, położona nad urokliwym brzegiem Motławy, z pięknym widokiem na rzekę i Stare Miasto. Jest to miejsce, w którym niezależnie od pory roku można poczuć wakacyjną atmosferę, sprzyjającą wypoczynkowi.

Poproś o rzut



EXNOVO

2 pokoje - Gdańsk Śródmieście, Nadmotławie

**1 062 900 zł** 22 500 zł/m<sup>2</sup>

Reklama Oblicz ratę dla tej nieruchomości RRSO 6,71% na dzień 24.11.2025.

ul. Sienna Grobla, Śródmieście, Gdańsk, pomorskie

**Mieszkanie na sprzedaż**

Powierzchnia: 47.24 m<sup>2</sup>  
Liczba pokoi: 2  
Ogrzewanie: miejskie  
Piętro: 2/3  
Czynsz: 800 zł  
Stan wykończenia: do zamieszkania  
Rynek: wtórny  
Forma własności: pełna własność  
Dostępne od: 2024-02-28  
Typ ogłoszeniodawcy: biuro nieruchomości  
Informacje dodatkowe: taras, garaż/miejsce parkingowe, oddzielna kuchnia

**Budynki i materiały**

**Wypożyczenie**

Reklama Finansowanie: Oblicz ratę dla tej nieruchomości z ING. RRSO 6,71% na dzień 24.11.2025.

**Opis**  
**Nadmotławie**  
Nadmotławie to prestiżowa inwestycja w centrum Gdańskiego, położona nad urokliwym brzegiem Motławy, z pięknym widokiem na rzekę i Stare Miasto. Jest to miejsce, w którym niezależnie od pory roku można poczuć wakacyjną atmosferę, sprzyjającą wypoczynkowi.

Pokaż więcej

ID: 67520976  
Nr w biurze: 200289

**Zgłoś**

Marcin Falek  
Ex Novo - Inwestycje i Nieruchomości  
Pokaż numer

Imię\*  
jasulkuzia@gmail.com  
+48 Numer telefonu\*  
Z chęcią dowiem się więcej na temat tego mieszkania na sprzedaż. Proszę o kontakt.

Administrator Twoich danych osobowych jest Grupa OLX Sp. z o.o. [więcej](#)

Wyślij wiadomość

# **PLAN DZIAŁANIA**

1. Pobranie adresów URL do konkretnych ogłoszeń (z użyciem bibliotek *requests* i *BeautifulSoup*),
2. Pobranie całego kodu html każdego z ogłoszeń i wyciągnięcie interesujących danych (*Selenium* i *BeautifulSoup*),
3. Zapisanie danych w postaci pliku JSON.



# EFEKT

W wyniku scrapowania dostaję poniższe informacje o ogłoszeniu:

|                |                      |                         |
|----------------|----------------------|-------------------------|
| ID ogłoszenia  | Piętro               | Obecność windy          |
| URL            | Czynsz               | Rodzaj zabudowy         |
| Cena           | Stan wykończenia     | Zabezpieczenia          |
| Cena za metr   | Rynek                | Bezpieczeństwo          |
| Adres          | Forma własności      | Wyposażenie             |
| Opis           | Dostępne od          | Materiał budynku        |
| Powierzchnia   | Typ ogłoszeniodawcy  | Typ okien               |
| Liczba pokoi   | Informacje dodatkowe | Media                   |
| Typ ogrzewania | Rok budowy           | Certyfikat energetyczny |

# ŹRÓDŁA:

- *Intro to Web Scraping: Build Your First Scraper in 5 Minutes*, medium.com [\[link\]](#)
- *The Complete Beginner’s Path to Web Scraping: From BeautifulSoup to Selenium and Scrapy*, medium.com [\[link\]](#)
- *Web scraping w Pythonie – jak pobierać dane z internetu legalnie i skutecznie*, cognity.pl [\[link\]](#)
- Maślankowski, J. (2019). *Pozyskiwanie i analiza danych na temat ofert pracy z wykorzystaniem big data*. Wiadomości Statystyczne. The Polish Statistician, 64(09), 60-74.
- Otodom, wikipedia.pl [\[link\]](#)

