Członkowie grupy:

1. Bartłomiej Ratajczak
2. Jakub Turek
3. Konrad Świercz
4. Michalina Mudlaff

Ceneo WebScrapper

Projekt zespołowy

Spis treści

[1. Opis projektu 2](#_Toc136672321)

[2. technologia wykorzystywana w ramach projektu (stos technologiczny) 2](#_Toc136672322)

[3. OPIS INFRASTRUKTURY TESTOWEJ 3](#_Toc136672323)

[4. WYMAGANIA SYSTEMOWE 4](#_Toc136672324)

[5. OPIS ZABEZPIECZEŃ SYSTEMU 4](#_Toc136672325)

[6. Przykładowe testy APLIKACJI 5](#_Toc136672326)

[7. dokumentacja techniczna (instalacja i uruchomienie aplikacji) 6](#_Toc136672327)

[8. dokumentacja administratora 9](#_Toc136672328)

[TWORZENIE NOWEGO KONTA ADMINISTRATORA 9](#_Toc136672329)

[ZARZĄDZANIE PROCESEM AKTUALIZACJI SUBSKRYPCJI 9](#_Toc136672330)

[AKTYWACJA HARMONOGRAMU ZADAŃ 9](#_Toc136672331)

[WSTRZYMANIE HARMONOGRAMU ZADAŃ 9](#_Toc136672332)

[JEDNORAZOWE URUCHOMIENIE PROCESU 9](#_Toc136672333)

[9. dokumentacja Użytkownika 11](#_Toc136672334)

[INSTALACJA I URUCHOMIENIE APLIKACJI 11](#_Toc136672335)

[REJESTRACJA I LOGOWANIE 11](#_Toc136672336)

[TWORZENIE NOWEGO KONTA 11](#_Toc136672337)

[LOGOWANIE SIĘ DO APLIKACJI 13](#_Toc136672338)

[PRZEGLĄDANIE PRODUKTÓW 13](#_Toc136672339)

[ZARZĄDZANIE SUBSKRYPCJAMI 13](#_Toc136672340)

[DODAWANIE NOWEJ SUBSKRYPCJI 13](#_Toc136672341)

[MODYFIKACJA  ISTNIEJĄCEJ SUBSKRYPCJI 14](#_Toc136672342)

[USUWANIE SUBSKRYPCJI 16](#_Toc136672343)

# Opis projektu

Tworzona aplikacja ma na celu pomóc użytkownikowi w wyszukaniu i zagregowaniu opinii dotyczących wybranego przez niego przedmiotu. Serwis Ceneo.pl, z racji swojej pozycji na rynku, dysponuje ogromną bazą produktów, przez co przeszukiwanie go może okazać się kłopotliwe i pochłaniające znaczne zasoby czasowe. Ceneo WebScrapper redukuje czas potrzebny na znalezienie informacji, ułatwia proces decyzyjny dotyczący zakupu, a jednocześnie systematyzuje informacje dostępne na wielu stronach i przedstawia je w jednym miejscu, czyniąc cały proces przyjemniejszym, a porównywanie bardziej rzetelnym.

Funkcjonalności WebScrappera obejmują:

1. Wyszukiwanie produktów, opinii i ceny o nich.
2. Generowanie powiadomień dla użytkownika o zmianie ceny, by umożliwić zakup wybranego produktu w najkorzystniejszych warunkach.

# technologia wykorzystywana w ramach projektu (stos technologiczny)

Kod aplikacji utworzony został w języku Python. Oprócz tego w kolejnych etapach pracy wykorzystaliśmy następujące technologie: Postresql, Docker, natomiast Frontend będzie zbudowany przy pomocy Jinja2 wraz z biblioteką TailwindCSS. Koordynacja i postęp prac zostały zaewidencjonowane w Jira.

# OPIS INFRASTRUKTURY TESTOWEJ

Obraz zawierający tekst, diagram, Plan, Rysunek techniczny

Opis wygenerowany automatycznie

Ryc. 1. Diagram UML komponentów dla aplikacji CeneoWebScrapper.

Obraz zawierający diagram, tekst, linia, szkic

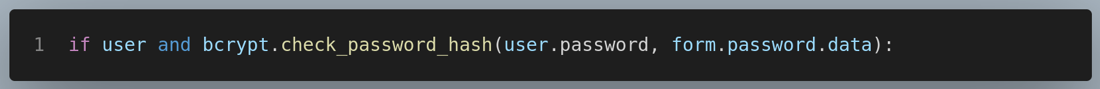
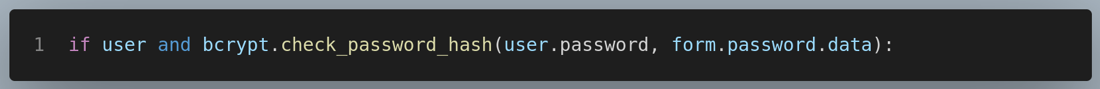
Opis wygenerowany automatycznie

Ryc. 2. Diagram UML wdrożenia dla aplikacji CeneoWebScrapper.

# WYMAGANIA SYSTEMOWE

Zainstalowany docker oraz docker-compose (najnowsze wersje).

# OPIS ZABEZPIECZEŃ SYSTEMU

1. **Przechowywanie haseł.** Aby zapewnić bezpieczeństwo przechowywania haseł, nie powinny one być przechowywane w postaci czytelnego tekstu. Zamiast tego, po zastosowaniu biblioteki bcrypt, hasła są zahaszowane i tak zapisane w bazie danych. Do obsługi hashowania w konfiguracji zapisane jest ziarno dla biblioteki bcrypt.
2.   
   
3. **Zabezpieczenie formularzy przed atakami CRSF (Cross-Site Request Forgery).** Do implementacji tego rozwiązania wykorzystany został pakiet Flask-WTF, implementując mechanizmy zabezpieczeń formularzy. Zabezpiecza to formularze przed atakami SQL injection, cross-site scripting i cross-site request forgery. Do wszystkich formularzy dołączony jest token-csrf.  
   Poniżej przykład z formularza logowania użytkownika,:  
     
     
   Obraz zawierający tekst, zrzut ekranu, Czcionka

   Opis wygenerowany automatycznie
4. **Autoryzacja, uwierzytelnianie i obsługiwanie sesji użytkownika.** Biblioteka Flask-Login obsługuje bezpieczne logowanie i rejestrację użytkownika, zarządza oraz generuje sesję dla przeglądarki użytkownika. Wszystkie endpointy poza logowaniem i rejestracją, wymagają od użytkownika zalogowania i potwierdzenia konta.  
   Obraz zawierający tekst, zrzut ekranu, Czcionka

   Opis wygenerowany automatycznie

# Przykładowe testy APLIKACJI

W pliku tests/services/ceneo/test\_ceneo\_item.py znajdują się przykładowe testy.

test\_fetch\_lowest\_price - test jednostkowy sprawdzający poprawność działania interfejsu modułu do pobierania informacji z ceneo. [Podział zawijania tekstu]Obraz zawierający tekst, zrzut ekranu, oprogramowanie

Opis wygenerowany automatycznie

test\_e2e\_fetch\_lowest\_price - test e2e sprawdzający poprawność całego procesu otwierania przeglądarki, strony ceneo.pl, wybierania produktu, scrapowania strony oraz przekazywania uzyskanych informacji w formie obiektu.

Obraz zawierający tekst, zrzut ekranu, Czcionka, oprogramowanie

Opis wygenerowany automatycznie

Raport uruchomienia testów

Obraz zawierający tekst, zrzut ekranu, oprogramowanie, Oprogramowanie multimedialne

Opis wygenerowany automatycznie

# dokumentacja techniczna (instalacja i uruchomienie aplikacji)

**Instalacja i uruchomienie aplikacji:**

1. W pierwszej kolejności należy utworzyć osobny folder dla projektu.

~$ mkdir ceneo\_scrapper  
 ~$ cd ceneo\_scrapper

1. Kod repozytorium można pobrać za pomocą następującego polecenia.  
     
   ~$ git clone <https://github.com/aeh-projekt-magia/APZ-CeneoScrapper.git>Alternatywnie można go również pobrać za pomocą przycisku “Download ZIP” ze strony projektu: <https://github.com/aeh-projekt-magia/APZ-CeneoScrapper>
2. Kolejny krok obejmuje przejście do pobranego folderu i uruchomienie skryptu startup.sh, który umożliwia instalację.  
   ~$ cd APZ-CeneoScrapper  
   ~$ ./startup.sh  
     
   > <1.> docker compose build  
     <2.> docker compose up -d  
    <3.> docker compose stop  
    <4.> docker compose down  
    <5.> docker compose up -d --scale app=5  
    <6.> docker exec python manage.py recreate\_db  
    <7.> docker exec python manage.py create\_admin  
    Choose option:  
     
   Program wypisuje możliwe do wykonania operacje. Wybór operacji polega na wpisaniu odpowiedniej cyfry i wciśnięciu klawisza Enter. Za utworzenie kontenera i pobranie potrzebnych zależności odpowiada opcja nr 1. Należy uruchomić ją w pierwszej kolejności.  
     
     
    Choose option: 1  
    Building app…  
    … Successfully build [container id]  
    Done!
3. Aby aplikacja działała poprawnie konieczne jest utworzenie konta Gmail i utworzenie hasła dla aplikacji zgodnie z instrukcją <https://support.google.com/accounts/answer/185833?hl=en>
4. Przed pierwszym uruchomieniem aplikacji należy uruchomić skrypt tworzący schemat bazy danych. W tym celu należy ponownie skorzystać ze skryptu startup.sh i wybrać opcję nr 6.  
   ~$ ./startup.sh  
     
   > …  
    Choose option: 6
5. Aplikację uruchamia się za pomocą opcji nr 2 w skrypcie startup.sh  
   ~$ ./startup.sh  
     
   > …  
    Choose option: 2

**Konfiguracja aplikacji**

1. Konfiguracja pliku config.ini  
   Plik konfiguracyjny dzieli się na trzy sekcje
   1. task\_scheduler  
        
      Wartości dla *day, hour, minute* powinny być wyrażeniami typu *crone*. Te trzy wartości ustalają częstotliwość uruchamiania głównego zadania aplikacji, które aktualizuje ceny subskrybowanych produktów i wysyła powiadomienia subskrybentom. Poniżej widnieje przykładowa konfiguracja, która nakazuje uruchomienie zadania codziennie o 12:30.   
      Obraz zawierający tekst, Czcionka, zrzut ekranu, design

      Opis wygenerowany automatycznie
   2. email  
        
      Domyślnie wpisane wartości dla *port* oraz *smtp\_server* powinny pozostać niezmienione, jako, że są to aktualnie zalecane ustawienia do wysyłania e-maili przez serwer smtp od Google. Wartości dla *email\_address* oraz *password* należy ustawić zgodnie z utworzonym adresem e-mail oraz hasłem wygenerowanym przez Google dla tej aplikacji.  
      Obraz zawierający tekst, Czcionka, zrzut ekranu

      Opis wygenerowany automatycznie
   3. subscription\_updates  
        
      Konfiguracja wartości dla *date\_format* oraz *currency* pozwala ustalić odpowiednio format daty oraz symbol / skrót waluty dla mailowych powiadomień subskrybentów.  
      Obraz zawierający tekst, Czcionka, zrzut ekranu, design

      Opis wygenerowany automatycznie
2. Konfiguracja pliku app/\_\_init\_\_.py  
     
   Aby wybrać domyślną konfigurację dla systemu produkcyjnego, testowego lub deweloperskiego należy ustawić domyślny argument *config\_class* odpowiednio jako *ProductionConfig*, *TestingConfig* lub *DevelopmentConfig*Obraz zawierający tekst, Czcionka, zrzut ekranu, design

   Opis wygenerowany automatycznie

# dokumentacja administratora

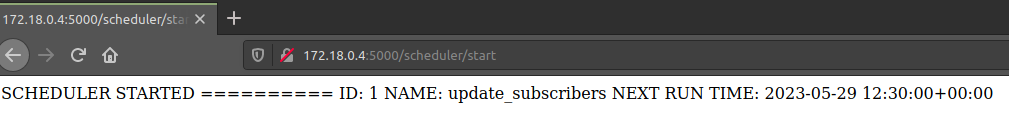
TWORZENIE NOWEGO KONTA ADMINISTRATORA

Nowe konto administratora można stworzyć uruchamiając skrypt startup.sh i wybierając opcję nr 7. Po wpisaniu i zatwierdzeniu (przycisk Enter) adresu e-mail i hasła konto administratora zostanie utworzone.  
~$ ./startup.sh  
 Choose option: 7  
 Using container id: [container id]  
 Enter email: [adminemail@gmail.com](mailto:adminemail@gmail.com) Enter password:   
 Done!  
Zalogowanie się na konto administratora jest możliwe po uruchomieniu aplikacji w przeglądarce po wpisaniu poprawnego loginu i hasła.

ZARZĄDZANIE PROCESEM AKTUALIZACJI SUBSKRYPCJI

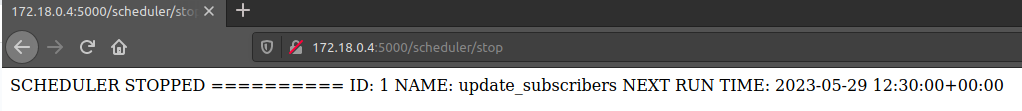
Przed użyciem dowolnego z poniższych narzędzi administratora konieczne jest zalogowanie się do aplikacji poprawnym adresem e-mail i hasłem przypisanym do konta administratora.

AKTYWACJA HARMONOGRAMU ZADAŃ

W celu aktywacji zaplanowanego zgodnie z konfiguracją zadania wysyłania aktualizacji dla subskrybentów należy wywołać następujący adres url:  
  
<http://172.18.0.4:5000/scheduler/start>Przy poprawnym wykonaniu całej operacji wyświetli się komunikat z listą wszystkich zaplanowanych zadań i terminami najbliższego uruchomienia.  


WSTRZYMANIE HARMONOGRAMU ZADAŃ

Aby anulować zaplanowane do wykonania zadania należy wywołać następujący adres url:

<http://172.18.0.4:5000/scheduler/stop>Przy poprawnym wykonaniu całej operacji wyświetli się komunikat z listą wszystkich wstrzymanych zadań i anulowanymi terminami najbliższego uruchomienia.  
  


JEDNORAZOWE URUCHOMIENIE PROCESU

Pojedyncze uruchomienie zadania wysyłania aktualizacji subskrypcji można wykonać otwierając poniższy adres url uprzednio logując się jako administrator:

<http://172.18.0.4:5000/scheduler/run-task>

Wyświetlenie komunikatu “OK” oznacza, że zadanie zostało wykonane i aktualizacje zostały wysłane na poczty e-mail subskrybentów. Zadanie można uruchomić jednorazowo niezależnie od aktualnego stanu harmonogramu zadań (może być zarówno wstrzymany jak i aktywny).

…  
Obraz zawierający tekst, zrzut ekranu, Oprogramowanie multimedialne

Opis wygenerowany automatycznie

# dokumentacja Użytkownika

# INSTALACJA I URUCHOMIENIE APLIKACJI

Instalacja i uruchomienie aplikacji zostały dokładnie wytłumaczone w “**Dokumentacji technicznej**”.

# REJESTRACJA I LOGOWANIE

Przed użyciem aplikacji należy utworzyć konto i się zalogować.

### TWORZENIE NOWEGO KONTA

Aby utworzyć nowe konto należy nacisnąć przycisk “**register**” znajdujący się na górze strony. Uzupełnić następujące pola z emailem, hasłem i potwierdzeniem hasła, a następnie nacisnąć przycisk “**Register**”.

Jeżeli użyjemy adresu mailowego, który już został użyty w aplikacji wyświetli się następujący komunikat:

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

Z kolei w przypadku niezgodności haseł będzie widoczna następująca informacja:

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

Po udanym zarejestrowaniu widać na stronie komunikat na zielonym tle „**You registered and are now logged in. Welcome**!”. Należy jeszcze potwierdzić konto klikając w czerwony przycisk “**Confirm Account**”.

Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie

Ostatnim krokiem rejestracji jest wklejenie kodu weryfikacyjnego (“token”), który wyświetla się powyżej:

Obraz zawierający tekst, zrzut ekranu

Opis wygenerowany automatycznie

### LOGOWANIE SIĘ DO APLIKACJI

Aby zalogować się do aplikacji należy kliknąć w “Login” znajdujący się po lewej stronie przy nazwie aplikacji.

Następnie uzupełnić pola “email” oraz “password”. Zatwierdzić przyciskiem “login”. Po poprawnym zalogowaniu widoczny jest następujący komunikat:

Obraz zawierający tekst, zrzut ekranu, Czcionka, design

Opis wygenerowany automatycznie

# PRZEGLĄDANIE PRODUKTÓW

W celu przeglądania interesujących cię produktów kliknij w zakładkę “products”. Możesz wyszukać produkty wykorzystując nazwę z naszej aplikacji lub ze strony Ceneo.

# ZARZĄDZANIE SUBSKRYPCJAMI

### DODAWANIE NOWEJ SUBSKRYPCJI

Aby zasubskrybować nowy produkt należy go wpierw wyszukać w zakładce Products. Jako że produkt nie był wcześniej wyszukiwany należy skorzystać z drugiego pola tekstowego “**Search for product in Ceneo**”.  Po wpisaniu nazwy produktu i kliknięciu przycisku “**Search**” wyszukany produkt pojawia się poniżej.

Aby zasubskrybować znaleziony produkt należy kliknąć “**Check product description**”, a następnie “**Subscribe product**”

Po kliknięciu przycisku pojawia się wiadomość z potwierdzeniem -  “**Product subscribed**”.

Nowo dodaną subskrypcję można znaleźć w zakładce “**Subscriptions**”.

Obraz zawierający tekst, Czcionka, zrzut ekranu

Opis wygenerowany automatycznie

### MODYFIKACJA  ISTNIEJĄCEJ SUBSKRYPCJI

Aby zmienić preferencje częstotliwości otrzymywania powiadomień należy zmodyfikować subskrypcję. W tym celu należy wyszukać subskrypcję w zakładce “**Subscriptions**” i kliknąć “**Check subscription description**” pod wybraną subskrypcją.

Obraz zawierający tekst, Czcionka, zrzut ekranu

Opis wygenerowany automatycznie

Na ekranie pojawia się podsumowanie danej subskrypcji. Aby rozpocząć modyfikację należy wcisnąć przycisk “**Edit subscription**”.

Obraz zawierający tekst, zrzut ekranu, Czcionka, oprogramowanie

Opis wygenerowany automatycznie

Pole “**Notification frequency**” wyznacza częstotliwość powiadomień w dniach, natomiast flaga “**Notify on price change**” ustawia filtr na otrzymywanie powiadomień tylko w przypadku, gdy cena zmieniła się od poprzedniego powiadomienia.

Obraz zawierający tekst, zrzut ekranu, oprogramowanie, Strona internetowa

Opis wygenerowany automatycznie

Po zmodyfikowaniu wartości należy zatwierdzić zmiany przyciskiem “**Update subscription**”.

### USUWANIE SUBSKRYPCJI

Aby usunąć subskrypcję danego produktu należy go wyszukać w zakładce **Products.**

Obraz zawierający tekst, zrzut ekranu

Opis wygenerowany automatycznie

Po znalezieniu produktu, którego subskrypcję zamierza się anulować, należy kliknąć ”**Check product description**”, po czym kliknąć przycisk “**Unsubscribe product**”.

Obraz zawierający tekst, Czcionka, zrzut ekranu

Opis wygenerowany automatycznie

Wiadomość o udanej anulacji subskrypcji pojawia się na górze ekranu, a przycisk “**Unsubscribe product**” zamienia się na “**Subscribe product**”.

Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie