

Raport z negocjacji

Marek Kacprzak

10 października 2024

Contents

1	Podsumowanie negocjacji	3
2	Opis działania aplikacji	3
2.1	Główne widoki	3
2.2	Logika biznesowa aplikacji	4
2.3	Mechanizmy	4
2.4	Możliwe wyzwania i rozwiązania	5
2.5	Pytania do zamawiającego	5
3	Sekcja techniczna	6
3.1	Scraper	6
3.2	Baza danych	6
3.3	API	6
3.4	Frontend	6
3.5	Hosting	6
3.6	Docker	6
3.7	Dokumentacja	7
3.8	Testowanie	7
4	Rozszerzenie projektu na przyszłość	8

1 Podsumowanie negocjacji

Przed zespołem programistów zostało postawione zadanie stworzenia aplikacji webowej, która pozwoli na wgląd w osiągnięcia naukowe pracowników akademickich. Aplikacja będzie prezentować różnorodne dane, takie jak obszary badawcze, publikacje oraz wskaźniki bibliometryczne poszczególnych naukowców. Dodatkowo, system umożliwi porównywanie dorobku między naukowcami.

2 Opis działania aplikacji

Aplikacja zostanie podzielona na widoki oraz mechanizmy, które będą realizować określone funkcje.

2.1 Główne widoki

- **Strona główna** - Zawiera wyszukiwarkę pozwalającą na znalezienie pracowników naukowych według różnych kryteriów. Na stronie tej wyświetlani będą też naukowcy o najwyższej liczbie publikacji oraz najlepiej wypadający pod względem wskaźników bibliometrycznych.
- **Widok pracownika** - Prezentuje szczegółowe informacje o naukowcu, w tym jego obszary badań, wskaźniki bibliometryczne oraz publikacje. Użytkownik może również przejść do porównania z innymi naukowcami.
- **Widok porównania** - Umożliwia zestawienie dorobku naukowców według różnych filtrów, takich jak stopień naukowy czy dziedzina badań. Aplikacja wygeneruje graficzną wizualizację wyników porównania.
- **Widok dziedzin** - Prezentuje listę dziedzin naukowych oraz naukowców z danej dziedziny. Użytkownik może wybrać dziedzinę, aby zobaczyć listę naukowców.
- **Widok publikacji** - Wyświetla listę publikacji naukowca wraz z informacjami takimi jak liczba cytowań czy czasopismo naukowe.
- **Widok statystyk** - Prezentuje ogólne statystyki dotyczące naukowców, takie jak średnia liczba publikacji czy średni wskaźnik cytowań.
- **Konto przeglądającego** - Umożliwia zalogowanie się do aplikacji, co pozwala na zapisywanie ulubionych naukowców oraz dostęp do dodatkowych funkcji.
- **Widok ustawień** - Pozwala na dostosowanie ustawień aplikacji, takich jak język czy powiadomienia.
- **Widok pomocy** - Zawiera informacje o aplikacji, w tym instrukcje obsługi i odpowiedzi na najczęściej zadawane pytania.

- **Widok o aplikacji** - Zawiera informacje o autorach aplikacji, źródłach danych oraz sposobie działania aplikacji.
- **Widok kontaktu** - Pozwala na kontakt z zespołem odpowiedzialnym za aplikację.

2.2 Logika biznesowa aplikacji

- Aplikacja będzie regularnie 'zdzierać' dane ze strony bw.sggw.edu.pl, oczyszczać je, a następnie aktualizować swoją bazę danych.
- Pobierane będą informacje o profesorach, adiunktach, doktorach oraz innych pracownikach naukowych SGGW.
- Będziemy pobierać dane takie jak imię, nazwisko, stopień naukowy, dziedzina badań, liczba publikacji, wskaźniki bibliometryczne, publikacje (w czasopismach oraz po konferencjach) oraz cytowania.
- Dodatkowym źródłem informacji o pracach naukowych będzie portal <https://punktoza.pl/>.
- Po wpisaniu odpowiednich informacji i ustawieniu filtrów, użytkownik będzie mógł wyszukać pracownika SGGW, a dane zostaną pobrane z bazy i wyświetlone w widoku naukowca.
- Filtrowanie będzie możliwe po instytutach, pracowniach, dziedzinach badań, stopniach naukowych, wskaźnikach bibliometrycznych oraz liczbie publikacji.
- Najważniejszymi wskaźnikami będzie IF, liczba cytowań oraz publikacji, punktacja ministerialna oraz H-index.
- Aplikacja pozwoli także na porównanie dorobku kilku naukowców, z odpowiednią wizualizacją wyników.
- Użytkownik będzie mógł zalogować się do aplikacji, co pozwoli na zapisywanie ulubionych naukowców oraz dostęp do dodatkowych funkcji.

2.3 Mechanizmy

- **Mechanizm pobierania danych** - Odpowiada za regularne zbieranie danych ze strony bw.sggw.edu.pl.
- **Mechanizm aktualizacji danych** - Umożliwia aktualizację bazy danych w chmurze po każdej operacji pobierania danych.
- **Mechanizm wyszukiwania** - Pozwala na wyszukiwanie pracowników według filtrów z tolerancją na błędy użytkownika.
- **Mechanizm wizualizacji** - Przetwarza dane z bazy i przedstawia je w formie graficznych podsumowań i porównań.

- **Mechanizm zapisywania** - Pozwala na zapisywanie ulubionych naukowców i ustawień użytkownika.

2.4 Możliwe wyzwania i rozwiązania

- **Problemy z pobieraniem danych** - Zmiana struktury strony może spowodować, że skrypty przestaną działać. Możliwe rozwiązanie to wprowadzenie systemu powiadomień dla osób odpowiedzialnych za utrzymanie aplikacji.
- **Trudności w porównywaniu dorobku** - Porównanie naukowców z różnych dziedzin może być trudne ze względu na różnice w ich wskaźnikach bibliometrycznych. Rozwiązaniem może być dodanie sekcji wyjaśniającej metodę obliczania wskaźników i specyfikę różnych dziedzin naukowych.
- **Wydażność przy dużych zbiorach danych** - Aplikacja może działać wolno przy dużej liczbie danych. Konieczne jest wdrożenie optymalnych algorytmów wyszukiwania i sortowania.
- **Bezpieczeństwo danych** - Dane naukowców mogą być wrażliwe, dlatego konieczne jest zabezpieczenie bazy danych i zapewnienie, że tylko uprawnione osoby mają do nich dostęp.

2.5 Pytania do zamawiającego

1. Kim będą główni użytkownicy aplikacji?
2. Z jakich urządzeń będą korzystać użytkownicy?
3. Czy aplikacja będzie dostępna dla ogółu, czy tylko dla klienta?
4. Czy aplikacja ma być dostępna tylko w języku polskim?
5. Czy interfejs powinien obsługiwać różne wersje językowe?
6. Czy aplikacja powinna umożliwiać zapisywanie wyników wyszukiwań?
7. Czy aplikacja powinna umożliwiać zapisywanie ustawień użytkownika?
8. Który wskaźnik bibliometryczny jest najważniejszy dla użytkownika?
9. Jak często baza danych powinna być aktualizowana?

3 Sekcja techniczna

Na podstawie wymagań projektowych podzielono projekt na poniższe komponenty:

3.1 Scraper

- **Python:** **Beautiful Soup** lub **Scrapy**
- **JavaScript:** **Puppeteer**
- Scraper będzie odpowiadał za przeglądanie strony i zbieranie danych, które następnie zostaną zapisane w bazie danych.

3.2 Baza danych

- **PostgreSQL** - Baza danych o lepszej wydajności niż **MySQL**.

3.3 API

- API będzie pośredniczyć między bazą danych a frontendem użytkownika.
- API zostanie zaimplementowane w języku **Go** z użyciem frameworka **Gin** lub **Rust** z użyciem frameworka **Axum**.

3.4 Frontend

- Możliwe są dwa podejścia: tradycyjne z użyciem **HTML**, **CSS** i **JS** lub nowoczesne z wykorzystaniem frameworka, np. **React**, **Angular** czy **Vue**.
- Frontend będzie odpowiedzialny za prezentację danych z bazy i umożliwi użytkownikowi filtrowanie informacji.

3.5 Hosting

- **DigitalOcean** - Dostawca chmury, który zapewni nam serwer do hostowania aplikacji.

3.6 Docker

- **Docker** - Umożliwi nam łatwe zarządzanie kontenerami i wdrożenie aplikacji.
- **Docker Compose** - Umożliwi nam zarządzanie wieloma kontenerami jednocześnie.

3.7 Dokumentacja

- Dokumentację można generować automatycznie tam, gdzie jest to obsługiwane, np. za pomocą **OpenAPI**, lub przygotować ręcznie w plikach `.yaml`.

3.8 Testowanie

- **Testowanie API** - Testy jednostkowe, zostaną ustalone z drużyną testerów.
- **Testowanie frontendu** - Testy integracyjne i jednostkowe, zostaną ustalone z drużyną testerów.
- **Testowanie scrapera** - Sposób testowania będzie ustalony później z drużyną testerów.

4 Rozszerzenie projektu na przyszłość

- **Kolejne platformy** - Możliwe jest rozszerzenie projektu o dodatkowe platformy, z których można pobierać dane o pracownikach naukowych.