

# Projekt i implementacja aplikacji webowej do zarządzania usługą OCR dla faktur oraz usługami pomocniczymi

Raport końcowy

Opracowali:

Martyna Bużek

Paulina Chachuła

Marta Fabrowska

Daniel Rakoczy

Joanna Skolarska

Tomasz Wiśniewski

## Spis treści

1.	Wstęp .....	2
2.	Lista zrealizowanych zadań przez każdego z członków .....	2
3.	Cel projektu.....	2
4.	Założenia i zakres projektu.....	2
5.	Definicja zadań .....	3
5.1	Usługa OCR .....	3
5.2	Walidator poprawności danych.....	3
5.3	Front-end.....	3
5.4	Usługa wyliczająca podatek .....	3
5.5	Tagowanie .....	4
5.6	Sprawdzanie i zwracanie informacji o firmie .....	4
5.7	Pobieranie i przekazywanie plików .....	4
5.8	Wysyłanie maila .....	4
6.	Opis opracowanego systemu webowego.....	5
6.1	Zadania i wymagania .....	5
6.2	Potrzebne klasy (?) .....	5
6.3	Diagramy UML .....	5
6.4	Diagram komponentów .....	5
6.4.1	Diagram klas .....	5
7.	Podsumowanie .....	6

## 1. Wstęp

Księgowi często spędzają bardzo dużo czasu na wprowadzaniu faktur do systemu, wyliczaniu różnych kwot czy podatków. Usługa OCR czyli optyczne rozpoznawanie znaków ułatwia im tę pracę. Wystarczy, że prześlą skan faktury, a OCR rozpozna tekst, który znajduje się na dokumencie i zwróci wartości możliwe do skopiowania, a w większym systemie automatycznie wypełni wartości w programie księgowym.

## 2. Lista zrealizowanych zadań przez każdego z członków

Grupa składa się z 6 osób. Każdej osobie przyporządkowane zostały odpowiednie zadania.

Imię i nazwisko	Zadanie
Martyna Bużek	OCR, tagowanie, wyliczanie podatku, zwracanie danych wystawiającego fakturę
Paulina Chachuła	Raport końcowy, pobieranie i przekazywanie plików
Marta Fabrowska	Aplikacja webowa spajająca wszystkie serwisy Front-end
Daniel Rakoczy	OCR, tagowanie, wyliczanie podatku, zwracanie danych wystawiającego fakturę
Joanna Skolarska	Raport końcowy, pobieranie i przekazywanie plików
Tomasz Wiśniewski	Sprawdzenie danych wystawiającego w sieci Walidator poprawności danych Wysyłanie emaili

## 3. Cel projektu

Celem projektu jest opracowanie i udostępnienie usługi OCR, która po przesłaniu skanu dokumentu rozpoznaje znaki znajdujące się na niej. Następnie nastąpi przeliczenie odpowiednich kwot, na skutek czego usługa wyliczy podatek, który dana firma musi zapłacić. Odpowiednie wartości liczbowe zostaną przesłane na maila konkretnej osoby.

## 4. Założenia i zakres projektu

Projekt obejmuje:

# Projekt i implementacja aplikacji webowej do zarządzania usługą OCR dla faktur oraz usługami pomocniczymi

© © ©

- przygotowanie projektu systemu,
- opracowanie modułu OCR,
- wykonanie usługi wysyłającej maile do odpowiednich osób oraz wyliczającej podatek,
- tagowanie, zwracanie danych wystawiającego fakturę i wystawiającego w sieci, a także pobieranie i przekazywanie plików
- sporządzenie narzędzia porównującego i sprawdzającego dane wystawiających faktury i wystawiających w sieci

## 5. Definicja zadań

### 5.1 Usługa OCR

Usługa OCR służy do usprawnienia i skrócenia czasu pracy pracowników biurowych. Usługa ma za zadanie pobrać url z plikiem zapisanym w formacie pdf (np. skan faktury), a po optycznym rozpoznaniu znaków program wygeneruje tekst, możliwy do skopiowania.

### 5.2 Walidator poprawności danych

Walidator jest narzędziem, które sprawdza poprawność danych kontaktowych osób wystawiających faktury oraz dane firmy z sieci. Danymi wejściowymi są konkretne dane kontaktowe (takie jak nip, numer KRS, nazwa firmy – z zaznaczeniem, że dane muszą być podane z obu źródeł), a następnie po ich wprowadzeniu dostajemy odpowiedź w formie logicznego typu danych (0-1, czyli *True* lub *False*), czy wprowadzone informacje są jednakowe czy nie. Mikroserwis wykonany jest w języku *Python*. Do sporządzenia narzędzia wykorzystano również *Docker*. Zapytanie wysyłane jest metodą POST.

### 5.3 Front-end

Aplikację webową uruchamia się w przeglądarce, a użytkownik wchodzi z nią w interakcję. Do pobrania danych od użytkownika służy Front-end.

### 5.4 Usługa wyliczająca podatek

Jedną z ważniejszych części tego systemu, jest usługa wyliczająca podatek. Aplikacja obliczy podatek za pomocą rozpoznanych znaków z faktury. Po wprowadzeniu tekstu, system oblicza kwotę konieczną do zapłacenia.

## 5.5 Tagowanie

Tagi umożliwiają utrzymanie porządku w aplikacji. Powstają one poprzez wprowadzenie tekstu. System tworzy z nich tagi.

## 5.6 Sprawdzanie i zwracanie informacji o firmie

Ta usługa powstała w celu otrzymania potrzebnych informacji o danym przedsiębiorstwie. Wpisując NIP firmy, dostaje się numer REGON, numer KRS oraz pełną nazwę jednostki. Zapytanie wysyłane jest metodą GET. Do stworzenia tego mikroservisu wykorzystano język *Python* oraz użyto *Dockera*.

## 5.7 Pobieranie i przekazywanie plików

Przekazanie skanu faktury powinno odbyć się na początku całego procesu, a pobranie pliku na jego końcu, opcjonalnie lub równolegle do przesłania. Zapytania wykonane są za pomocą metod POST oraz GET. Do stworzenia mikroservisów wykorzystano język C# oraz wykorzystano *Dockera*.

## 5.8 Wysyłanie maila

Po procesie rozpoznania znaków na fakturze, wyliczeniu wartości podatku oraz porównaniu i sprawdzeniu danych o firmie, zostanie wysłana wiadomość do odpowiednich podmiotów z obliczoną kwotą podatku. Mikroservis powstał w języku *Python*, zastosowano także oprogramowanie *Docker*. Zapytanie wysyłane jest metodą POST z elementami takimi jak: adres adresata, tytuł wiadomości oraz treścią maila.

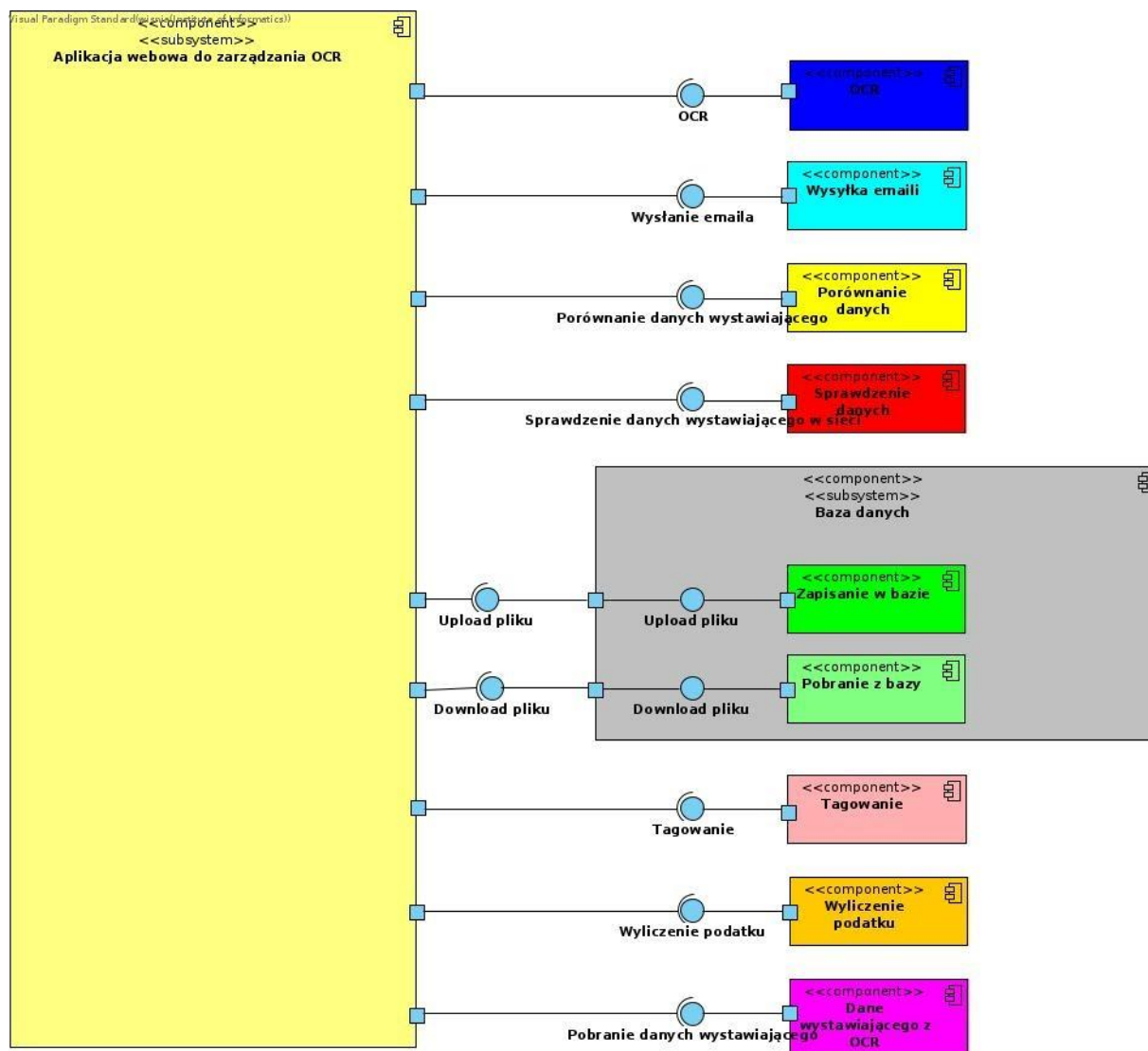
## 6. Opis opracowanego systemu webowego

### 6.1 Zadania i wymagania

Zadaniem aplikacji jest odczytywanie znaków, znajdujących się na fakturze, a następnie wyliczanie podatku koniecznego do zapłacenia. Ponadto system porównuje informacje o firmie z dwóch źródeł i sprawdza, czy są jednakowe, pobiera informacje o działalności a także wysyła maila zawierającego odpowiednią kwotę podatku lub pozwala je pobrać.

### 6.2 Diagram komponentów

Przed przystąpieniem do wykonywania aplikacji został ustalony diagram komponentów, w celu ułatwienia podziału pracy. Pozwolił on również poznać szczegóły budowy programu. Diagram komponentów został wzbogacony o kilka komponentów więcej, ze względu na dużą grupę.



Rys. 1 Projekt systemu OCR

Na przedstawionym diagramie komponentów można zobaczyć, co wchodzi w skład aplikacji. Jest to implementacja OCR, wysyłanie maili, porównywanie danych wystawiającego, sprawdzenie informacji o wystawiającym w sieci, pobieranie i przekazywanie plików, co mieści się w osobnym podsystemie. Kolejnymi komponentami są tagowanie, wyliczenie podatku i pobieranie danych wystawiającego.

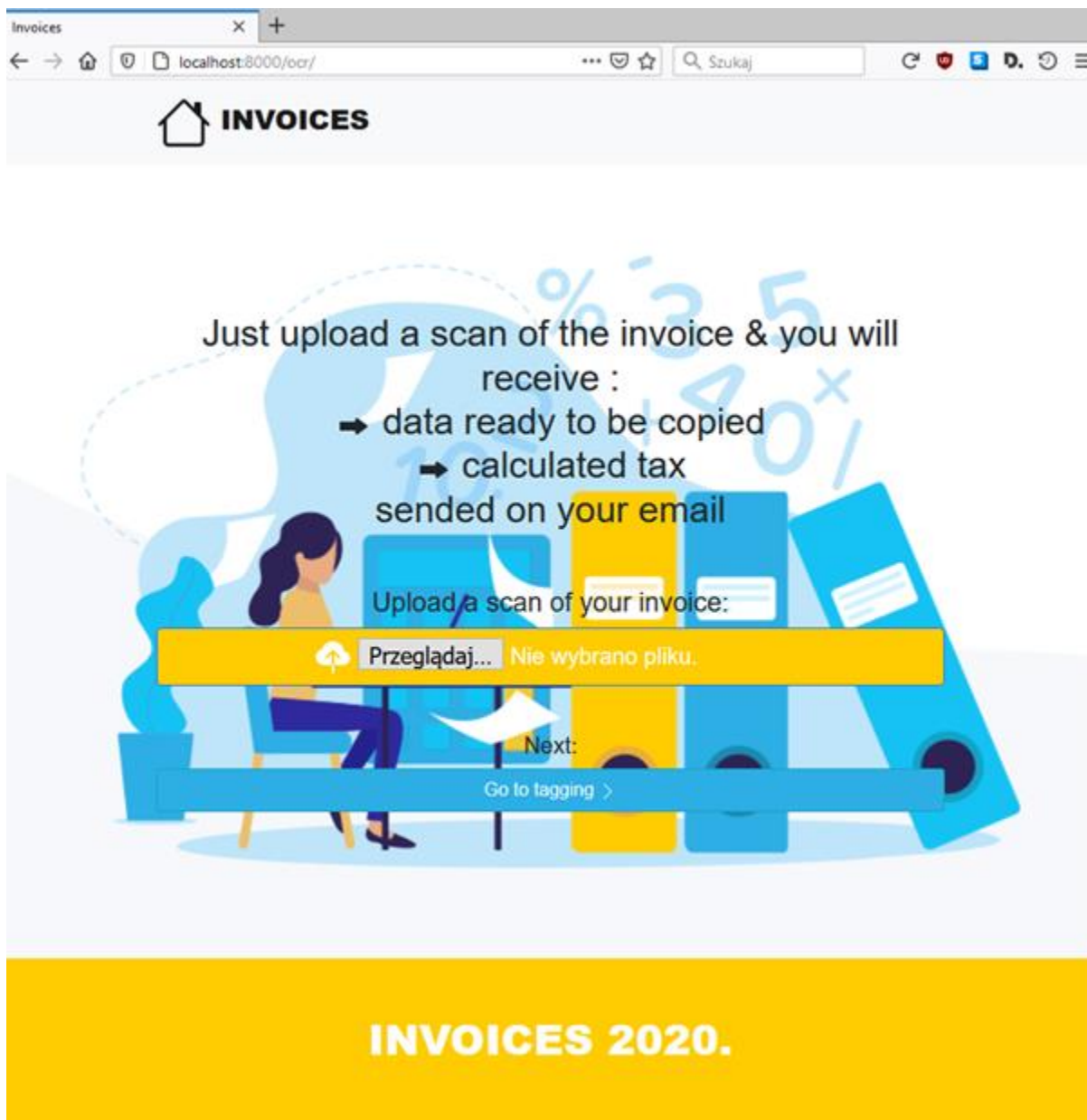
## 7. Opis funkcjonalności udostępnianej użytkownikowi

Osoba korzystająca z aplikacji ma szeroki wachlarz możliwości – oprócz wykrycia znaków z faktury przez system, użytkownik może sprawdzić pobrać i przekazać pliki do systemu, a także utworzyć tagi.

## Projekt i implementacja aplikacji webowej do zarządzania usługą OCR dla faktur oraz usługami pomocniczymi

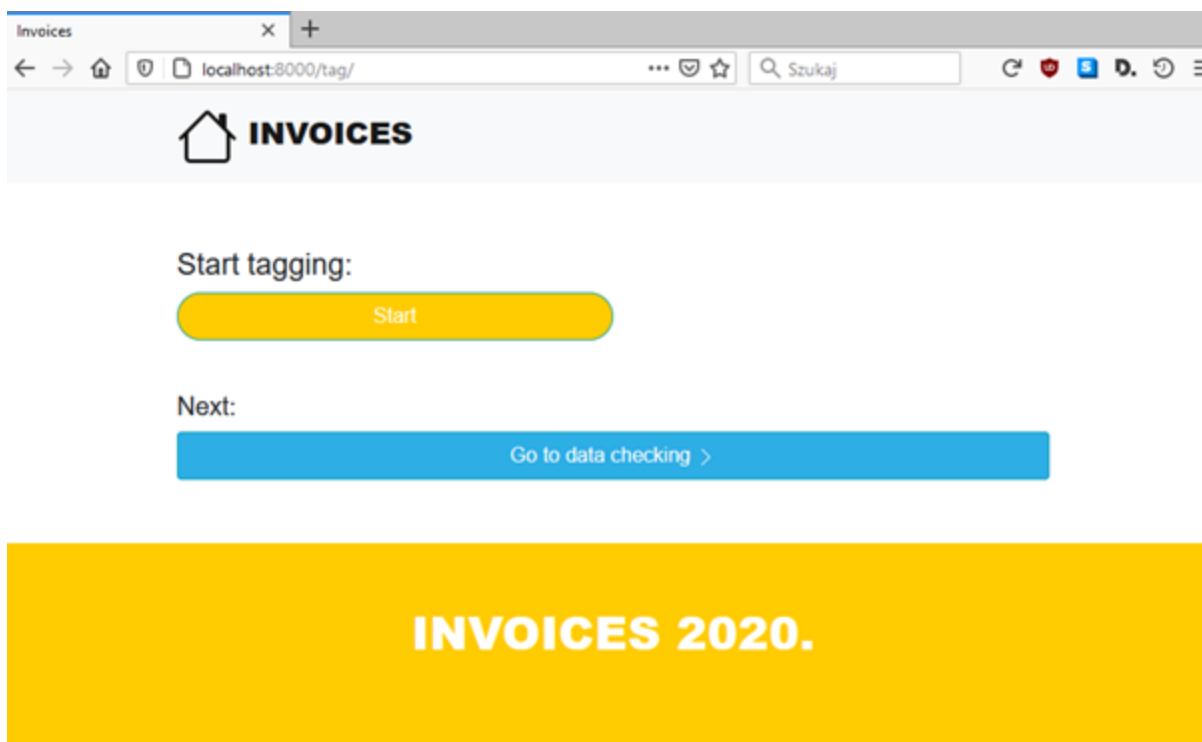
© © ©

Dużym ułatwieniem jest wysłanie maila z kwotą podatku – użytkownik nie musi sam wykonywać tych czynności. Co więcej korzystający z aplikacji ma możliwość uzyskać dane o wystawiających z sieci.

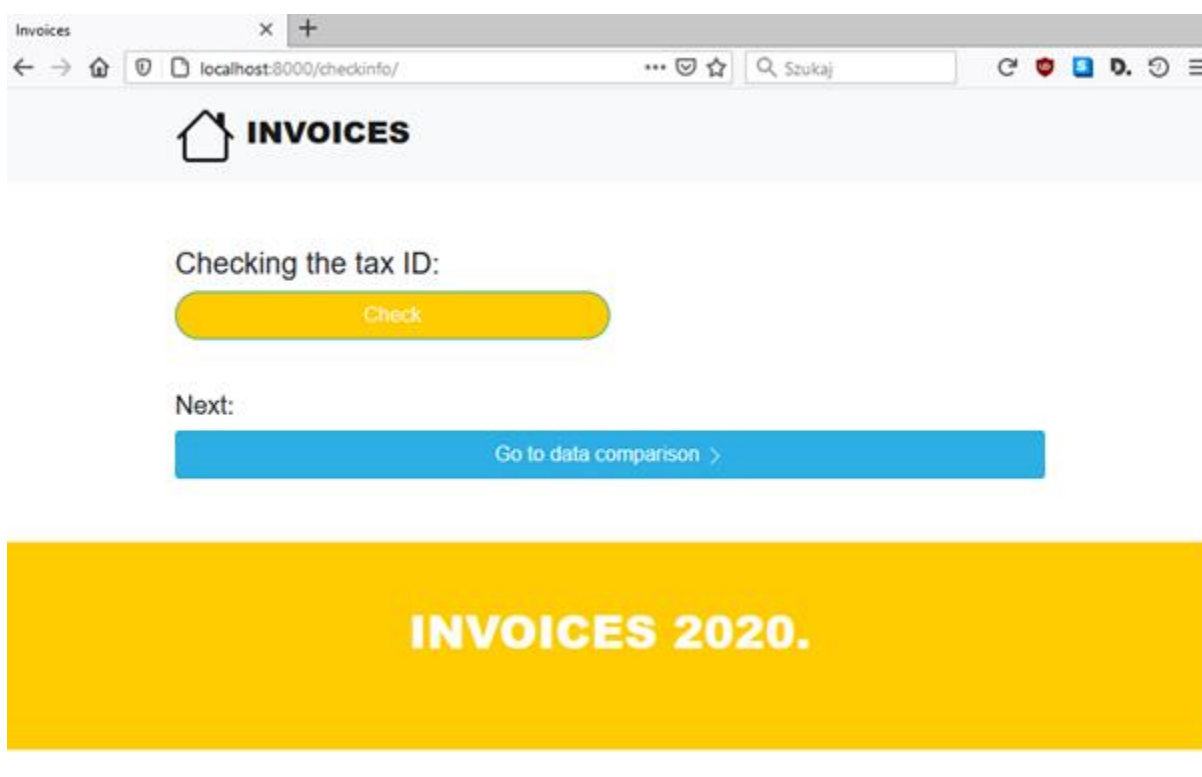


Rys. 2 Widok z polem przekazywania pliku





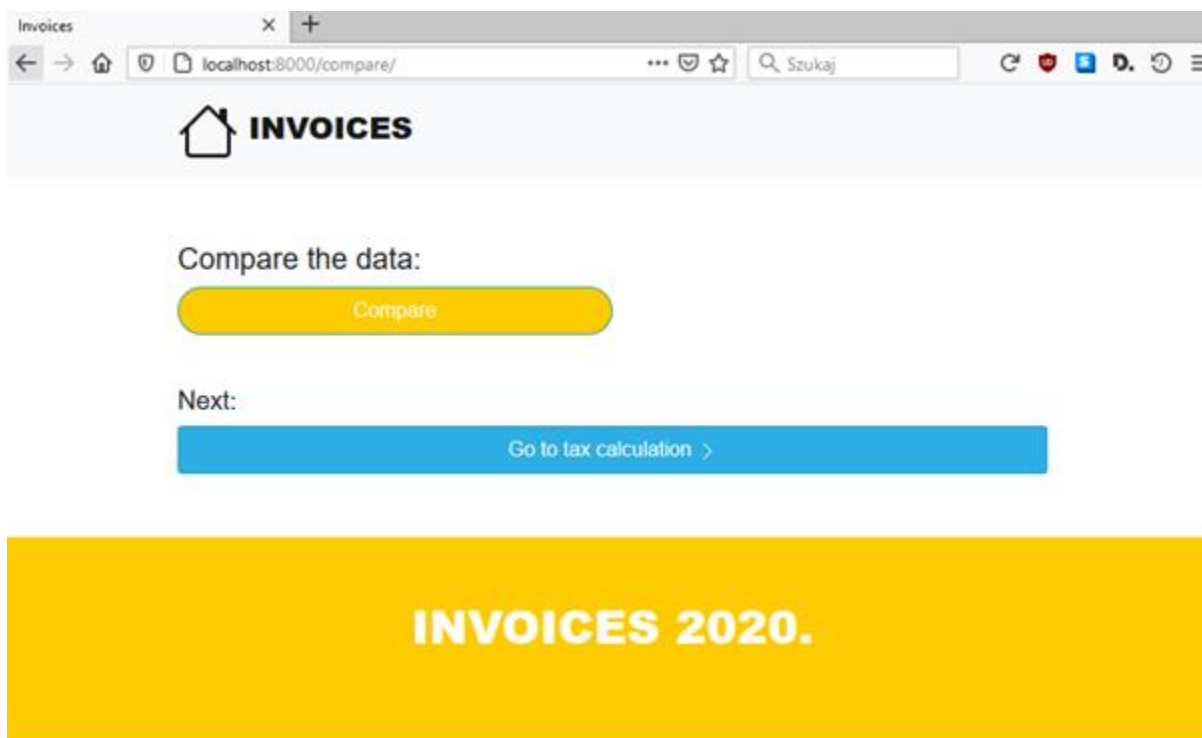
Rys. 3 Widok strony pozwalającej na tagowanie



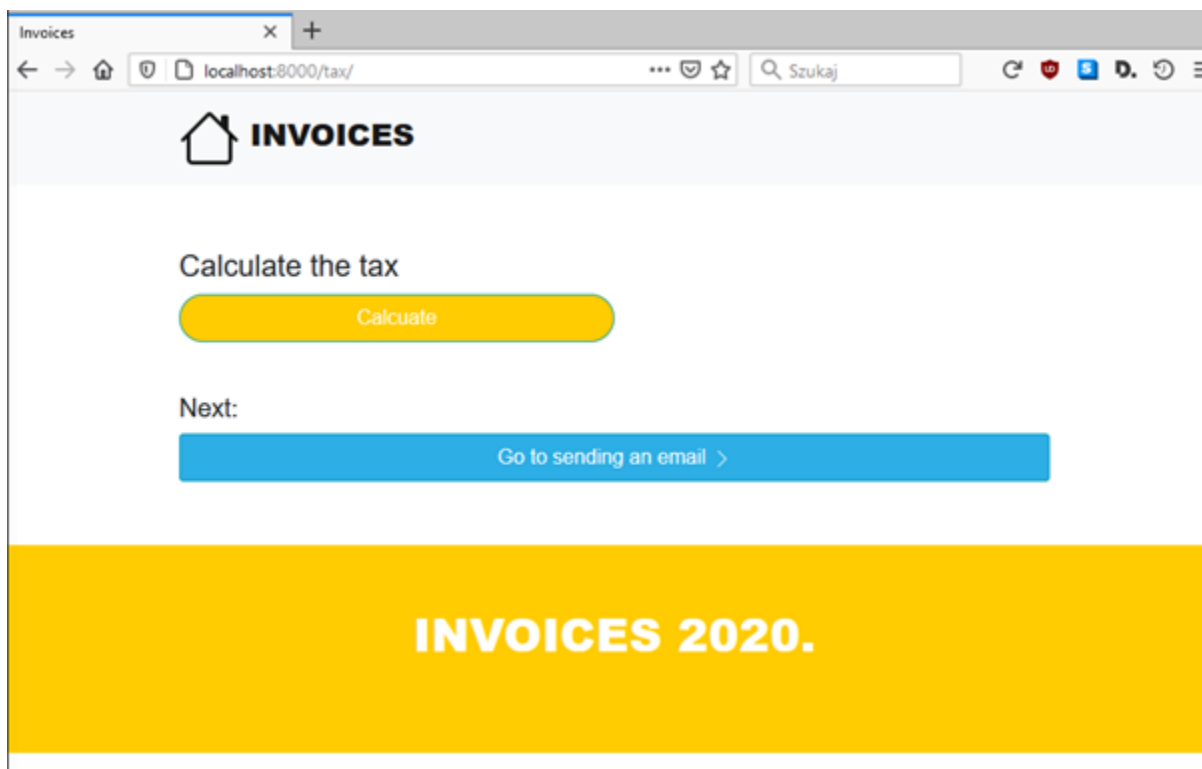
Rys. 4 Widok strony pozwalającej sprawdzanie danych wystawiającego

# Projekt i implementacja aplikacji webowej do zarządzania usługą OCR dla faktur oraz usługami pomocniczymi

© © ©



Rys. 5 Widok strony umożliwiającej porównanie danych firmy z dwóch źródeł - walidacja



Rys. 6 Widok strony do wyliczania podatku

The screenshot shows a web browser window with the title 'Invoices'. The address bar displays 'localhost:8000/sendemail/'. The page features a header with a house icon and the word 'INVOICES'. Below the header, there is a form with the label 'Enter your e-mail address:'. The form includes a text input field labeled 'Recipient:' and a yellow 'Send' button. At the bottom of the form area, there is a blue button that says 'Click if you want to go to the beginning >'. A large yellow banner at the bottom of the page contains the text 'INVOICES 2020.'.

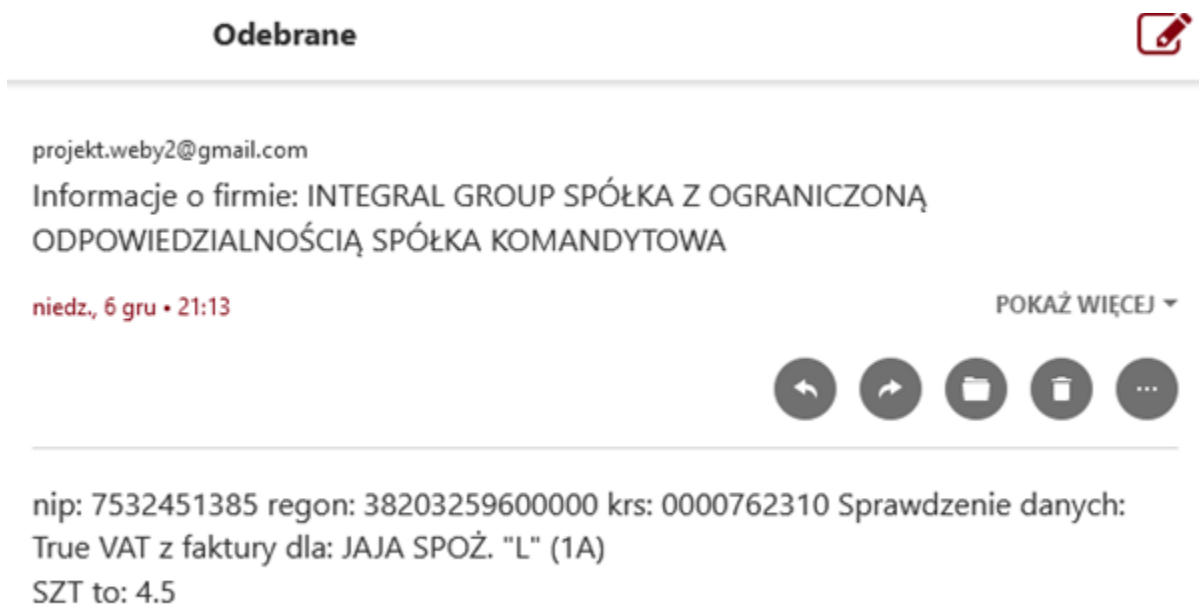
Rys. 7 Widok strony do wysyłania maila z danymi zwrotnymi

This screenshot shows the same 'Invoices' web application as in Rys. 6, but with the email input field filled with 'martfab@tutanota.com'. A confirmation dialog box is displayed in the center of the screen, with the message 'Your data has been sent, check your email' and an 'OK' button. The background of the page is dimmed, and the yellow banner at the bottom still displays 'INVOICES 2020.'.

## Projekt i implementacja aplikacji webowej do zarządzania usługą OCR dla faktur oraz usługami pomocniczymi

© © ©

Rys. 8 Potwierdzenie wysłania maila z danymi zwrotnymi



Rys. 9 Odebrany mail z serwisu w skrzynce odbiorczej

### 8. Podsumowanie

Systemy z usługą OCR są coraz bardziej potrzebne w dzisiejszym świecie, ponieważ oszczędzają ludziom dużo czasu. Pracownicy nie muszą wykonywać prostych i powtarzalnych czynności, ponieważ wykona to za nich aplikacja.