



**Politechnika
Śląska**

PRACA MAGISTERSKA

Moduł do obsługi importów danych finansowych z plików PDF

Mateusz MARCZEWSKI

Nr albumu: 282700

Kierunek: Informatyka

Specjalność: internet i technologie sieciowe

PROWADZĄCY PRACĘ

Dr hab. inż. Arkadiusz Biernacki

KATEDRA Katedra Sieci i Systemów Komputerowych

Wydział Automatyki, Elektroniki i Informatyki

Gliwice 2023

Tytuł pracy

Moduł do obsługi importów danych finansowych z plików PDF

Streszczenie

(Streszczenie pracy – odpowiednie pole w systemie APD powinno zawierać kopię tego streszczenia.)

Słowa kluczowe

(2-5 słów (fraz) kluczowych, oddzielonych przecinkami)

Thesis title

Module for importing inventory data from PDF files

Abstract

(Thesis abstract – to be copied into an appropriate field during an electronic submission – in English.)

Key words

(2-5 keywords, separated by commas)

Spis treści

1	Wstęp	1
1.1	wprowadzenie w zagadnienie	1
1.2	cel pracy	1
1.3	charakterystyka rozdziałów	1
2	Definiowanie wymagań	3
2.1	Przegląd typów danych do importowania	3
2.2	Pola w plikach, które należy przechwycić	3
2.3	Formaty plików, z którymi pracujemy	3
3	Badanie istniejących rozwiązań	5
3.1	Przegląd istniejących modułów i bibliotek do importowania danych z PDF	5
3.2	Funkcje istniejących rozwiązań	5
3.3	Oceny użytkowników i opinie na temat istniejących rozwiązań	5
4	Ocena dostępnych opcji	7
4.1	Porównanie istniejących rozwiązań na podstawie zdefiniowanych wymagań	7
5	Rozwój istniejących rozwiązań	9
5.1	Rozważania dotyczące budowy niestandardowego modułu	9
5.2	Zasoby i umiejętności wymagane do rozwoju i utrzymania modułu	9
6	Rozwój wybranego rozwiązania	11
6.1	Testowanie wybranego rozwiązania	11
6.2	Udoskonalenie wybranego rozwiązania na podstawie napotkanych problemów	11
7	Dokumentacja i udostępnianie wybranego rozwiązania	13
7.1	Dokumentacja rozwiązania	13
7.2	Udostępnianie rozwiązania jako projektu open-source	13
8	Podsumowanie	15
8.1	Podsumowanie badania	15
8.2	Implikacje dla przyszłej pracy	15

Bibliografia	17
Spis skrótów i symboli	19
Lista dodatkowych plików, uzupełniających tekst pracy (jeżeli dotyczy)	21
Spis rysunków	23
Spis tabel	25

Rozdział 1

Wstęp

1.1 wprowadzenie w zagadnienie

Zarządzanie dostępnymi zasobami jest niezbędnym procesem dla dużych biznesów. Proces ten składa się z czynności takich jak, śledzenie ilości, lokalizacji i wartości towarów posiadanych przez firmę. Jest to niezbędna czynność dla utrzymania wydajności finansowej firmy oraz maksymalizacji zysków. Zważając na opisane operacje można wy-szczególnić duże wyzwanie stojące przed firmami chcącymi optymalnie zarządzać swoimi zasobami. Jest nim importowanie danych z różnych źródeł, takich jak faktury zamówień, potwierdzenia sprzedaży, dane o załadunkach i tym podobne, a następnie przechowywa-nie ich w bazie danych, dzięki czemu można łatwiej nimi zarządzać. Zazwyczaj takie dane będą dostępne w formie plików PDF, lub będzie możliwe ich do tej postaci zeskanowa-nie, co ułatwia ich przechowywanie i podgląd jednak utrudnia ekstrakcje danych z takich plików.

W celu poradzenia sobie z tym problemem

1.2 cel pracy

1.3 charakterystyka rozdziałów

Rozdział 2

Definiowanie wymagań

2.1 Przegląd typów danych do importowania

2.2 Pola w plikach, które należy przechwycić

2.3 Formaty plików, z którymi pracujemy

Odwołania do literatury: książek [bib:ksiazka], artykułów w czasopismach [bib:artykul], materiałów konferencyjnych [bib:konferencja] i stron www [bib:internet].

Równania powinny być numerowane

$$y = \frac{\partial x}{\partial t} \tag{2.1}$$

Rozdział 3

Badanie istniejących rozwiązań

tekst

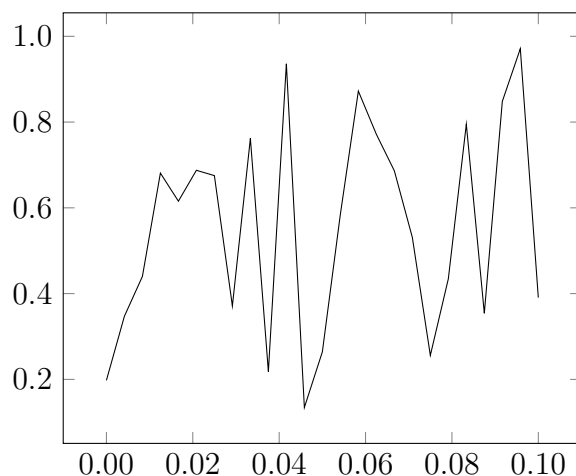
3.1 Przegląd istniejących modułów i bibliotek do importowania danych z PDF

3.2 Funkcje istniejących rozwiązań

3.3 Oceny użytkowników i opinie na temat istniejących rozwiązań

W całym dokumencie powinny znajdować się odniesienia do zawartych w nim ilustracji (rys. 3.1).

Tekst dokumentu powinien również zawierać odniesienia do tabel (tab. 3.1).



Rysunek 3.1: Wykres przebiegu funkcji.

Tabela 3.1: Opis tabeli nad nią.

ζ	metoda						
	alg. 1	alg. 2	alg. 3			alg. 4, $\gamma = 2$	
			$\alpha = 1.5$	$\alpha = 2$	$\alpha = 3$	$\beta = 0.1$	$\beta = -0.1$
0	8.3250	1.45305	7.5791	14.8517	20.0028	1.16396	1.1365
5	0.6111	2.27126	6.9952	13.8560	18.6064	1.18659	1.1630
10	11.6126	2.69218	6.2520	12.5202	16.8278	1.23180	1.2045
15	0.5665	2.95046	5.7753	11.4588	15.4837	1.25131	1.2614
20	15.8728	3.07225	5.3071	10.3935	13.8738	1.25307	1.2217
25	0.9791	3.19034	5.4575	9.9533	13.0721	1.27104	1.2640
30	2.0228	3.27474	5.7461	9.7164	12.2637	1.33404	1.3209
35	13.4210	3.36086	6.6735	10.0442	12.0270	1.35385	1.3059
40	13.2226	3.36420	7.7248	10.4495	12.0379	1.34919	1.2768
45	12.8445	3.47436	8.5539	10.8552	12.2773	1.42303	1.4362
50	12.9245	3.58228	9.2702	11.2183	12.3990	1.40922	1.3724

Rozdział 4

Ocena dostępnych opcji

4.1 Porównanie istniejących rozwiązań na podstawie zdefiniowanych wymagań

Rozdział 5

Rozwój istniejących rozwiązań

- 5.1 Rozważania dotyczące budowy niestandardowego modułu
- 5.2 Zasoby i umiejętności wymagane do rozwoju i utrzymania modułu

Rozdział 6

Rozwój wybranego rozwiązania

6.1 Testowanie wybranego rozwiązania

6.2 Udoskonalenie wybranego rozwiązania na podstawie napotkanych problemów

Rozdział 7

Dokumentacja i udostępnianie wybranego rozwiązania

7.1 Dokumentacja rozwiązania

7.2 Udostępnianie rozwiązania jako projektu open-source

Rozdział 8

Podsumowanie

8.1 Podsumowanie badania

8.2 Implikacje dla przyszłej pracy

- syntetyczny opis wykonanych prac
- wnioski
- możliwość rozwoju, kontynuacji prac, potencjalne nowe kierunki
- Czy cel pracy zrealizowany?

Dodatki

Spis skrótów i symboli

DNA kwas deoksyrybonukleinowy (ang. *deoxyribonucleic acid*)

MVC model – widok – kontroler (ang. *model-view-controller*)

N liczebność zbioru danych

μ stopnień przyleżności do zbioru

\mathbb{E} zbiór krawędzi grafu

\mathcal{L} transformata Laplace’a

Lista dodatkowych plików, uzupełniających tekst pracy (jeżeli dotyczy)

W systemie do pracy dołączono dodatkowe pliki zawierające:

- źródła programu,
- zbiory danych użyte w eksperymentach,
- film pokazujący działanie opracowanego oprogramowania lub zaprojektowanego i wykonanego urządzenia,
- itp.

Spis rysunków

3.1 Wykres przebiegu funkcji.	5
---------------------------------------	---

Spis tabel

3.1	Opis tabeli nad nią.	6
-----	------------------------------	---