Politechnika Łódzka

Wydział Elektrotechniki, Elektroniki, Informatyki i Automatyki

Instytut Informatyki Stosowanej

**PRACA DYPLOMOWA MAGISTERSKA**

Ekstrakcja informacji z utworu literackiego

Extracting essential data from a literary work

Paweł Wolski

234014

Dr hab. Szymon Grabowski, prof. PŁ

Łódź, luty 2020

*Składam serdeczne podziękowania dr hab. Szymonowi Grabowskiemu za poświęcony czas i pomoc w przygotowaniu niniejszej pracy*

Streszczenie:

Słowa kluczowe:

Summary:

Keywords:

Spis treści

[1. Wstęp 9](#_Toc66549197)

[2. Cel i zakres pracy 9](#_Toc66549198)

[3. Algorytmy przetwarzające tekst 9](#_Toc66549199)

[3.1. Historia NLP 9](#_Toc66549200)

[3.2. Podstawowe operacje 9](#_Toc66549201)

[3.3. Historia Spacy 9](#_Toc66549202)

[3.4. Podstawowe operacje 9](#_Toc66549203)

[3.5. Inne podobne rozwiązania 9](#_Toc66549204)

[4. Opis środowiska i aplikacji 9](#_Toc66549205)

[4.1. Język Python 9](#_Toc66549206)

[4.2. Framework nklp 9](#_Toc66549207)

[4.3. Framework spacy 9](#_Toc66549208)

[4.4. Framework pyqt 9](#_Toc66549209)

[5. Opis metryk 9](#_Toc66549210)

[5.1. Opis stylów autorskich 9](#_Toc66549211)

[5.2. Opis gatunków literackich 9](#_Toc66549212)

[5.3. Różnice językowe 9](#_Toc66549213)

[6. Projekt aplikacji do ekstrakcji informacji z dzieła literackiego 9](#_Toc66549214)

[6.1. Założenia aplikacji 9](#_Toc66549215)

[6.2. Wymagania aplikacji 9](#_Toc66549216)

[6.3. Obsługa aplikacji 10](#_Toc66549217)

[6.4. Zarys implementacji 10](#_Toc66549218)

[6.5. Wzorce projektowe ? 10](#_Toc66549219)

[6.6. Opis aplikacji 10](#_Toc66549220)

[6.7. Testy aplikacji 10](#_Toc66549221)

[7. Wnioski 10](#_Toc66549222)

[7.1. Wnioski badawcze 10](#_Toc66549223)

[7.2. Drogi rozwoju aplikacji 10](#_Toc66549224)

[Literatura 11](#_Toc66549225)

[Spis ilustracji 12](#_Toc66549232)

# Wstęp

# Cel i zakres pracy

# Algorytmy przetwarzające tekst

## Historia NLP

## Podstawowe operacje

## Historia Spacy

## Podstawowe operacje

## Inne podobne rozwiązania

# Opis środowiska i aplikacji

## Język Python

## Framework nklp

## Framework spacy

## Framework pyqt

# Opis metryk

## Opis stylów autorskich

## Opis gatunków literackich

## Różnice językowe

# Projekt aplikacji do ekstrakcji informacji z dzieła literackiego

## Założenia aplikacji

## Wymagania aplikacji

## Obsługa aplikacji

## Zarys implementacji

## Wzorce projektowe ?

## Opis aplikacji

## Testy aplikacji

# Wnioski

## Wnioski badawcze

## Drogi rozwoju aplikacji

# Literatura

### Lech Pijanowski, „Przewodnik gier", Wydawnictwo Iskry, 1978

### Rob Miles, „Microsoft XNA Game Studio 4.0”, Wydawnictwo Helion, 2012

### Steven John Metsker, „C#. Wzorce projektowe”, Wydawnictwo Helion, 2005

### Stanisław Wrycza, Bartosz Marcinkowski, Krzysztof Wyrzykowski, „Język UML 2.0 w modelowaniu systemów informatycznych”, Wydawnictwo Helion, 2006

### Strony internetowe:

#### <https://www.pasjans.pl/historia-pasjansa>

#### <https://pl.wikipedia.org/wiki/Diagram_klas>

#### <https://pl.wikipedia.org/wiki/Pasjans>

#### <https://pl.wikipedia.org/wiki/Kompilator>

#### <https://pl.wikipedia.org/wiki/C_Sharp>

#### <https://en.wikipedia.org/wiki/MonoGame>

#### <https://en.wikipedia.org/wiki/Windows_Forms>

#### <https://pl.wikipedia.org/wiki/Py%C5%82ek_(wzorzec_projektowy)>

#### <https://pl.wikipedia.org/wiki/Budowniczy_(wzorzec_projektowy)>

#### <https://devcave.pl/effective-java/wzorzec-projektowy-builder>

#### <https://en.wikipedia.org/wiki/Interpreter_(computing)>

#### <https://docs.microsoft.com/pl-pl/visualstudio/designers/getting-started-with-wpf?view=vs-2019>

#### <https://pl.wikipedia.org/wiki/Extensible_Application_Markup_Language>

#### <https://pl.wikipedia.org/wiki/Data_Manipulation_Language>

#### <https://pl.wikipedia.org/wiki/JSON>

#### <https://en.wikipedia.org/wiki/NUnit>

### Gry karciane:

#### Open Solitaire Classic - https://github.com/SoundGoddess/OpenSolitaire/blob/master/LICENSE.md

#### 123 Solitaire firmy TreeCardGames, 2019

#### Microsoft Solitaire Collection firmy Microsoft, 2010

# Spis ilustracji

[Rys. 3.1 Pasjans Jan Gotard 1931 r. (https://commons.wikimedia.org/wiki/File:Jan\_Gotard\_-\_Pasjans.jpg) 10](#_Toc32004628)

[Rys. 3.2 Pasjans Grandfather's Clock (https://pl.panoidl.com/sol/fbpreview/grandfathersclock.png) 11](#_Toc32004629)

[Rys. 4.1 Cechy języka C# (https://www.javatpoint.com/csharp-features) 14](#_Toc32004630)

[Rys. 4.2 Przykładowe działanie programu na konsoli Xbox (Miles R. – „Microsoft XNA Game Studio 4.0”) 15](#_Toc32004631)

[Rys. 4.3 Schemat działania aplikacji w frameworku Monogame (Miles R. – „Microsoft XNA Game Studio 4.0”) 16](#_Toc32004632)

[Rys. 4.4 Schemat działania funkcji do tekstur (Miles R. – „Microsoft XNA Game Studio 4.0”) 16](#_Toc32004633)

[Rys. 4.5 Przykładowe komendy DML dla języka SQL (https://www.splessons.com/lesson/mysql-dml/) 18](#_Toc32004634)

[Rys. 4.6 Różnica zapisu danych w JSON i XML (https://pl.wikipedia.org/wiki/JSON) 19](#_Toc32004635)

[Rys. 4.7 Przykładowy dokument napisany w języku YAML (https://pl.wikipedia.org/wiki/YAML) 20](#_Toc32004636)

[Rys. 5.1 Kolory kart w standardowych grach typu Pasjans (https://pl.wikipedia.org/wiki/Karty) 22](#_Toc32004637)

[Rys. 5.2 Rozkładanie kart na wszystkie stosy pomocnicze w Pasjansie Spider w aplikacji Microsoft Windows Solitaire (https://www.wikiwand.com/en/Microsoft\_Spider\_Solitaire) 23](#_Toc32004638)

[Rys. 5.3 Rozkładanie kart na stos pomocniczy umiejscowiony przy talii w aplikacji 123 Solitaire 24](#_Toc32004639)

[Rys. 5.4 Pasjans Spider jednokolorowy (https://www.free-spider-solitaire.com/) 24](#_Toc32004640)

[Rys. 5.5 Siatka, na której mogą znajdować się modele w aplikacji Solitaire Plugins 30](#_Toc32004641)

[Rys. 5.6 Odwzorowanie komend znajdujących się w edytorze tekstowym w aplikacji Solitaire Plugins, które po uruchomieniu aplikacji zostaną zapisane w formacie JSON 34](#_Toc32004642)

[Rys. 5.7 Przedstawienie komend dla przykładowych stosów pomocniczych i zbierających znajdujących się w edytorze tekstowym w aplikacji Solitaire Plugins, które po uruchomieniu aplikacji zostaną zapisane w formacie JSON 35](#_Toc32004643)

[Rys. 5.8 Odwzorowanie komend dla zasad wygranej znajdujących się w edytorze tekstowym w aplikacji Solitaire Plugins, które po uruchomieniu aplikacji zostaną zapisane w formacie JSON 35](#_Toc32004644)

[Rys. 5.9 Wygląd pliku w formacie JSON dla pasjansa Klondike wraz z wskazaniem podstawowych przetworzonych właściwości dla talii, kolizji i przykładowego stosu pomocniczego w aplikacji Solitaire Plugins 36](#_Toc32004645)

[Rys. 5.10 Zapis pliku JSON dla pasjansa Klondike wraz z pokazaniem przetworzenia zasad dla stosów zbierających typu Play 36](#_Toc32004646)

[Rys. 5.11 Zapis pliku JSON dla pasjansa Klondike wraz z pokazaniem przetworzenia zasad dla opcji wygranej 37](#_Toc32004647)

[Rys. 6.1 Główne okno Main Menu aplikacji Solitaire Plugins 41](#_Toc32004648)

[Rys. 6.2 Okno wyboru opcji 42](#_Toc32004649)

[Rys. 6.3 Widok okna edytora aplikacji Solitaire Plugins wraz z kodem do utworzenia pasjansa Klondike napisanym w ScriptsCards 43](#_Toc32004650)

[Rys. 6.4 Widok okna debugera w przypadku niezapisania pliku przed uruchomieniem 43](#_Toc32004651)

[Rys. 6.5 Najprostszy pasjans możliwy do stworzenia w aplikacji Solitaire Plugins 44](#_Toc32004652)

[Rys. 6.6 Diagram klas UML dla modelu DragAndDrop 45](#_Toc32004653)

[Rys. 6.7 Diagram klas UML dla modelu obiektów w grze 46](#_Toc32004654)

[Rys. 6.8 Diagram klas UML dla modelu przetwarzania zasad napisanych w ScriptsCards 47](#_Toc32004655)

[Rys. 6.9 Diagram klas UML dla klas wyjątków w grze 47](#_Toc32004656)

[Rys. 6.10 Diagram klas UML dla klas modeli ustawionych przez użytkownika za pomocą ScriptsCards 48](#_Toc32004657)

[Rys. 6.11 Diagram klas UML dla klasy Game1 oraz powiązanych z nią obiektów odpowiedzialnych za zachowanie podczas gry 49](#_Toc32004658)

[Rys. 6.12 Diagram klas UML dla klas formularzy wraz z możliwością przechodzenia między oknami 50](#_Toc32004659)

[Rys. 6.13 Kod przetwarzający zasady opisane dla stosu zbierającego typu Play w aplikacji Solitaire Plugins 51](#_Toc32004660)

[Rys. 6.14 Diagram klas UML dla wzorca Flyweight w aplikacji Solitaire Plugins 52](#_Toc32004661)

[Rys. 6.15 Diagram sekwencji wzorca budowniczy (https://pl.wikipedia.org/wiki/Plik:Builder\_sequence\_pl.svg) 53](#_Toc32004662)

[Rys. 6.16 Widok projektu aplikacji Solitaire Plugins 55](#_Toc32004663)

Załącznik nr 1

do Procedury antyplagiatowej prac dyplomowych

w Politechnice Łódzkiej

Łódź, dn. ......……………….

**OŚWIADCZENIE**

Paweł Wolski

………………………………………………………….

*(Imię i nazwisko studenta)*

Orla 3 m. 10, 90-317 Łódź

………………………………………………………….

*(Adres)*

209509

………………………………………………………….

*(Nr albumu)*

Wydział Elektrotechniki, Elektroniki,

Informatyki i Automatyki

………………………………………………………….

*(Jednostka organizacyjna prowadząca studia)*

Informatyka

………………………………………………………….

*(Kierunek studiów)*

Studia I stopnia inżynierskie stacjonarne

………………………………………………………….

*(Poziom kształcenia i forma studiów)*

Oświadczam, że poinformowano mnie o zasadach dotyczących kontroli oryginalności pracy dyplomowej w Jednolitym Systemie Antyplagiatowym.

.………………………….......

*(Podpis studenta)*

Załącznik nr 2

do Procedury antyplagiatowej prac dyplomowych

w Politechnice Łódzkiej

Łódź, dn. ......……………….

**OŚWIADCZENIE**

**o samodzielności wykonania i oryginalności pracy dyplomowej**

Paweł Wolski

………………………………………………………….

*(Imię i nazwisko studenta)*

Orla 3 m. 10, 90-317 Łódź

………………………………………………………….

*(Adres)*

209509

………………………………………………………….

*(Nr albumu)*

Wydział Elektrotechniki, Elektroniki,

Informatyki i Automatyki

………………………………………………………….

*(Jednostka organizacyjna prowadząca studia)*

Informatyka

………………………………………………………….

*(Kierunek studiów)*

Studia I stopnia inżynierskie stacjonarne

………………………………………………………….

*(Poziom kształcenia i forma studiów)*

Świadomy/a odpowiedzialności karnej za składanie fałszywych zeznań oświadczam, że:  
1. Przedkładana praca dyplomowa **inżynierska** [[1]](#footnote-1)\*) na temat: **Pasjans – implementacja popularnych łamigłówek karcianych z użyciem wtyczek** została wykonana przeze mnie samodzielnie.

2. Wymieniona wyżej praca:

- nie narusza praw autorskich w rozumieniu ustawy z dnia 4 lutego 1994 r. o prawie autorskim i prawach pokrewnych (t.j. Dz. U. z 2018 poz. 1191, z późn. zm.) oraz dóbr osobistych chronionych prawem cywilnym, a także nie zawiera danych i informacji, które uzyskałem/am w sposób niedozwolony,

- nie była wcześniej podstawą żadnej innej urzędowej procedury związanej z nadawaniem dyplomów wyższej uczelni lub tytułów zawodowych.

3. Jednocześnie wyrażam zgodę/nie wyrażam zgody\*\* na wykorzystanie fragmentów mojej pracy dyplomowej w publikacjach naukowych pracowników Politechniki Łódzkiej za zgodą prodziekana właściwego ds. studenckich, na zasadach wynikających z Ustawy z dnia 4 lutego 1994 r. o prawie autorskim i prawach pokrewnych (t.j. Dz. U. z 2018 poz. 1191, z późn. zm.)

.………………………….......

*(Podpis studenta)*

Załącznik nr 3

do Procedury antyplagiatowej prac dyplomowych

w Politechnice Łódzkiej

Łódź, dn. ......……………….

**OŚWIADCZENIE**

**o zgodności wersji elektronicznej pracy dyplomowej  
z przedstawionym wydrukiem komputerowym**

Paweł Wolski

………………………………………………………….

*(Imię i nazwisko studenta)*

Orla 3 m. 10, 90-317 Łódź

………………………………………………………….

*(Adres)*

209509

………………………………………………………….

*(Nr albumu)*

Wydział Elektrotechniki, Elektroniki,

Informatyki i Automatyki

………………………………………………………….

*(Jednostka organizacyjna prowadząca studia)*

Informatyka

………………………………………………………….

*(Kierunek studiów)*

Studia I stopnia inżynierskie stacjonarne

………………………………………………………….

*(Poziom kształcenia i forma studiów)*

Świadomy odpowiedzialności karnej za składanie fałszywych zeznań oświadczam, że przedkładana na nośniku elektronicznym praca dyplomowa **inżynierska** [[2]](#footnote-2)\*) na temat: **Pasjans – implementacja popularnych łamigłówek karcianych z użyciem wtyczek** zawiera te same treści, co oceniany przez promotora i recenzenta wydruk komputerowy.

Jednocześnie oświadczam, że jest mi znany przepis art. 233 § 1 Kodeksu karnego określający odpowiedzialność za składanie fałszywych zeznań.

.………………………….......

*(Podpis studenta)*

1. \*) Wpisać odpowiednio: licencjacka, inżynierska, magisterska.  
   \*\*) niepotrzebne skreślić [↑](#footnote-ref-1)
2. \*) Wpisać odpowiednio: licencjacka, inżynierska, magisterska. [↑](#footnote-ref-2)