



Politechnika  
Wrocławska

---

# Aplikacja Słownik - Instrukcja obsługi

---

ZADANIE PROJEKTOWE REALIZOWANE W RAMACH  
ZAJĘĆ "ZASTOSOWANIE INFORMATYKI W  
GOSPODARCE"

*Autorzy:*

MICHAŁ SURMACKI  
ADAM SZCZEŚNIAK  
KRZYSZTOF JOPEK  
DAWID TRZEBIŃSKI  
BARTOSZ STAROSTA

*Prowadzący:*

DR INŻ.  
TOMASZ WALKOWIAK

12 czerwiec 2021

## Spis treści

<b>1</b>	<b>Opis i dostępna funkcjonalność narzędzia</b>	<b>2</b>
<b>2</b>	<b>Instrukcja użytkowania</b>	<b>2</b>
2.1	Startowy widok aplikacji . . . . .	2
2.2	Wczytywanie korpusu po numerze ID . . . . .	5
2.3	Tworzenie profili wyrazów . . . . .	6
2.3.1	Wyniki analizy . . . . .	8

## Spis rysunków

1	Widok startowy aplikacji . . . . .	2
2	Widok procesowania wgranego korpusu . . . . .	3
3	Widok po przeprocesowaniu wgranego korpusu wraz z wygenerowanym ID. . . . .	4
4	Przykład niepoprawnego formatu korpusu oraz adresu e-mail. . . . .	4
5	Wynik przesłania korpusu z niedozwolonym formatem plików tekstowych. . . . .	5
6	Widok wczytywania korpusu po danym ID . . . . .	6
7	Widok wpisywania wyrazu do analizy . . . . .	6
8	Widok wpisania wyrazu nie istniejącego w danym korpusie . . . . .	7
9	Wyniki analizy - wystąpienia wyrazu "nie" w korpusie . . . . .	8
10	Wyniki analizy - kolokacje wyrazu "stół" w korpusie . . . . .	9


# 1 Opis i dostępna funkcjonalność narzędzia

Narzędzie Słowik to aplikacja internetowa, która pozwala na tworzenie opisu profili wyrazów w zadanych korpusach. Program wykorzystuje jedno z narzędzi platformy CLARIN - analizator morfo-syntaktyczny. Program rozszerza to narzędzie poprzez przetwarzanie całych korpusów wgrywanych w postaci skompresowanego pliku zawierającego wybrane teksty do analizy. W kolejnych sekcjach zawarto dostępne w narzędziu funkcjonalności wraz z ich opisami tak, aby czytelnik w jasny sposób mógł zrozumieć jak korzystać z aplikacji, aby osiągnąć oczekiwany rezultat.

## 2 Instrukcja użytkowania

W poniższym dokumencie opisano wszystkie dostępne funkcjonalności zrealizowanego serwisu wraz z załączonymi przykładami w postaci zdjęć, gdzie pokazano ich wykorzystanie.

### 2.1 Startowy widok aplikacji



---

3 Load files

4 Get by ID

Select files you want to upload

Select a ZIP file containing a corpus you want to process. The corpus must consist of text files in the following formats: pdf, txt, doc, docx, rtf. Processing your files may take a while. Please submit your e-mail so we can notify you when we're done.

1 Przeglądaj...

Nie wybrano pliku.

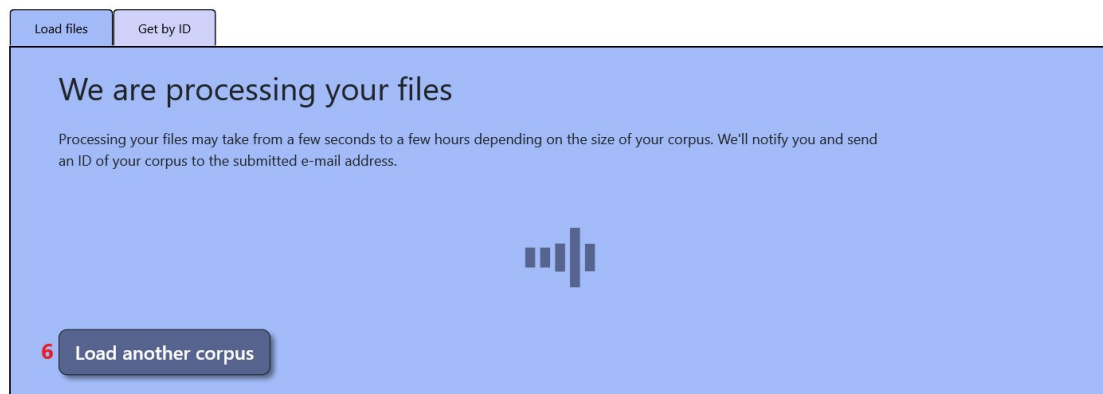
Submit your e-mail:

2 example@example.com

5 Upload

Rysunek 1: Widok startowy aplikacji

- (1) Import korpusu - Po naciśnięciu przycisku "Przeglądaj" wybieramy skompresowany plik ZIP zawierający teksty do analizy. Pliki tekstowe muszą być w jednym z formatów: ".pdf", ".txt", ".doc", ".docx", ".rtf". W przypadku gdy użytkownik wybierze korpus w złym formacie pliku, zostanie wyświetlony komunikat, aby wprowadzić plik w dozwolonym formacie. Podobnie w przypadku, gdy w korpusie znajdują się pliki tekstowe, które posiadają niedozwolony format.
- (2) Adres email - Wpisujemy adres e-mail, na który ma zostać przesłane ID korpusu. Jest to potrzebne do identyfikacji korpusu w bazie, aby była możliwość wykonania na nim dalszych czynności. W przypadku podania nieprawidłowego adresu e-mail, zostanie wyświetlony komunikat.
- (3) Zakładka "Load files" - zakładka przenosi użytkownika do widoku wgrywania korpusu, który jest domyślnie widokiem startowym.
- (4) Zakładka "Get by ID" - zakładka przenosi użytkownika do widoku pobierania przetworzonego korpusu, aby móc wykonać na nim profile wyrazów.
- (5) Przycisk "Upload" - Przycisk zatwierdzający wysłanie korpusu do przetworzenia wraz z adresem e-mail, na który zostanie wysłane ID korpusu.



Rysunek 2: Widok procesowania wgranego korpusu

- (6) Przycisk "Load another corpus" - Przycisk umożliwiający powrót do widoku wgrania nowego korpusu.

Load files


Get by ID

### Your corpus has been processed

We've managed to successfully process your files. You can proceed further or save your corpus ID for future use

7

036abf1c-5e8e-43cb-8e2e-74d88fe9143f



8

Results

Rysunek 3: Widok po przeprocesowaniu wgranego korpusu wraz z wygenerowanym ID.

- (7) Wygenerowane ID korpusu wraz z ikoną umożliwiającą skopiowanie tekstu do schowka.
- (8) Przycisk "Results" - Przycisk umożliwiający przejście do widoku tworzenia profili wyrazów.

Load files

Get by ID

### Select files you want to upload

Select a ZIP file containing a corpus you want to process. The corpus must consist of text files in the following formats: pdf, txt, doc, docx, rtf. Processing your files may take a while. Please submit your e-mail so we can notify you when we're done.

Przeglądaj...

wrongfile.jpg

Please upload a .zip file

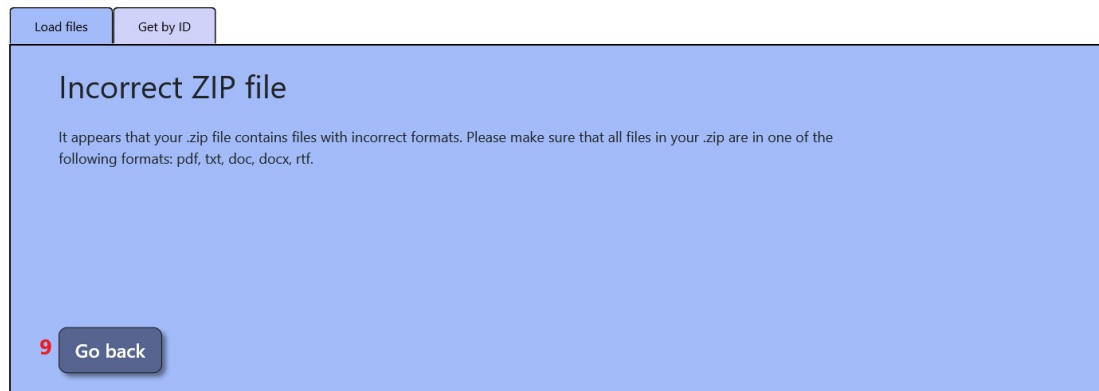
Submit your e-mail:

wrongemail@example

Please enter a valid email

Upload

Rysunek 4: Przykład niepoprawnego formatu korpusu oraz adresu e-mail.



Rysunek 5: Wynik przesłania korpusu z niedozwolonym formatem plików tekstowych.

- (9) Przycisk "Go back" - Przycisk umożliwiający powrót do widoku wgrania pliku korpusu.

## 2.2 Wczytywanie korpusu po numerze ID

W zakładce "Get by Id" możemy wczytać wybrany korpus potrzebny do wykonania profilów wyrazów.

- (11) W polu tekstowym możemy wpisać numer ID wybranego korpusu, który otrzymaliśmy na podany adres mailowy lub w komunikacie po prawidłowym wgraniu nowego.
- (12) Przycisk "Next" - pozwala przejść do widoku tworzenia profilów wyrazów.

Load files

Get by ID

Enter your corpus ID

Enter an ID of a previously uploaded corpus

11

12 Next

Rysunek 6: Widok wczytywania korpusu po danym ID

## 2.3 Tworzenie profili wyrazów

W tym widoku możemy wpisać w pole tekstowe który wyraz interesuje użytkownika i ma zostać poddany analizie we wczytanym korpusie.

Enter a word

Enter a word you're interested in so we can provide statistics regarding colocations and occurrences

13

14 Analyse

Rysunek 7: Widok wpisywania wyrazu do analizy

- (13) Pole tekstowe, gdzie należy podać wyraz do analizy.
- (14) Przycisk "Analyse" - uruchamia proces tworzenia profilu zadanego wyrazu we wskazanym korpusie.

### Enter a word

Enter a word you're interested in so we can provide statistics regarding colocations and occurrences

Analyse

### Word not found!

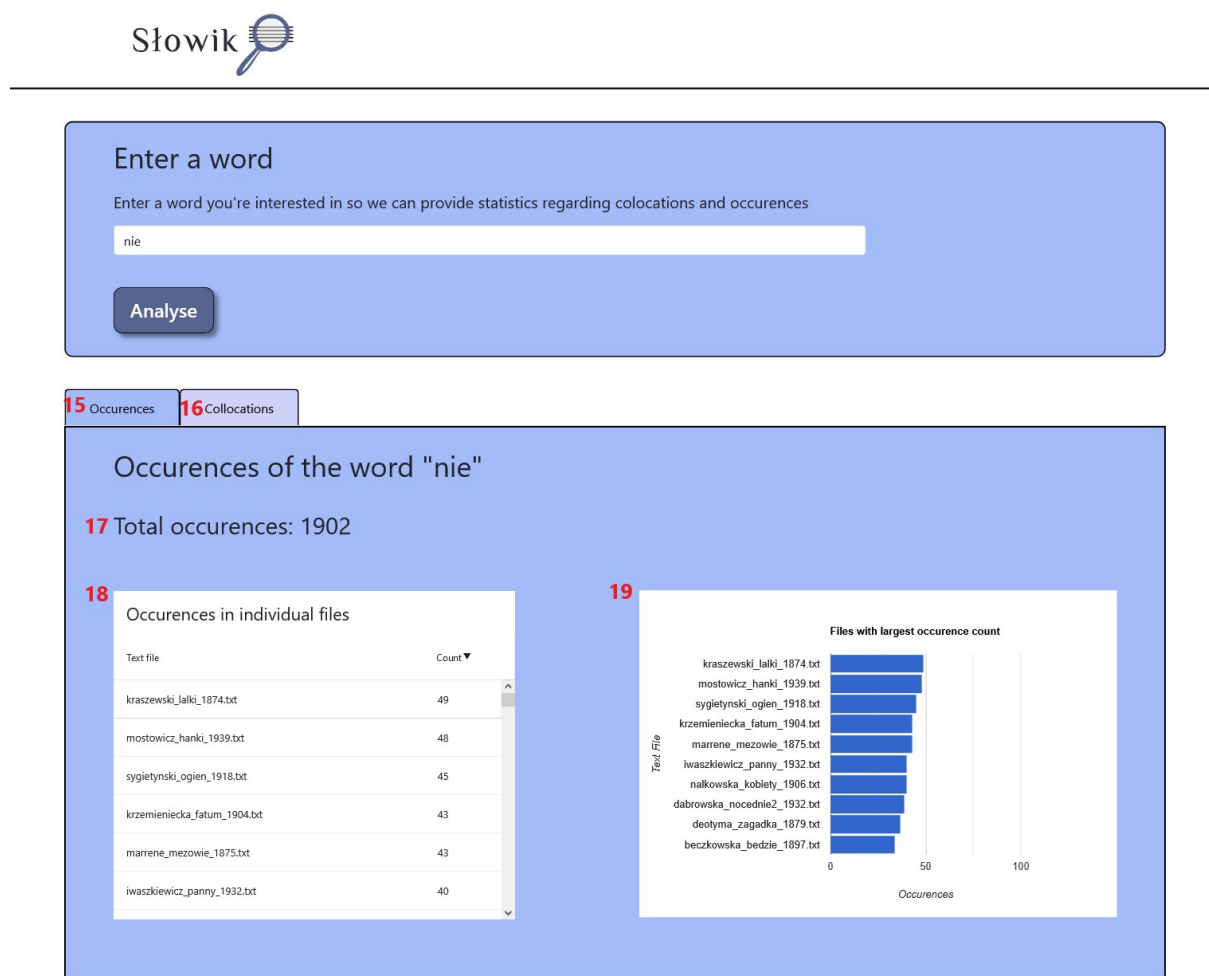
It appears that the word you are looking for is not present in your corpus. Try a different word.

Rysunek 8: Widok wpisania wyrazu nie istniejącego w danym korpusie



### 2.3.1 Wyniki analizy

Po wpisaniu wyrazu istniejącego w korpusie poniżej zostanie wygenerowana sekcja z wynikami analizy.



Rysunek 9: Wyniki analizy - wystąpienia wyrazu "nie" w korpusie

- (15) Zakładka "Occurrences" - pokazuje informacje o ilości wystąpień wybranego wyrazu w korpusie,
- (16) Zakładka "Collocations" - pokazuje informacje o kolokacjach wybranego wyrazu w korpusie,
- (17) Informacja o łącznej liczbie wystąpień wyrazu w korpusie,
- (18) Tabela z ilością wystąpień wyrazu w poszczególnych plikach tekstowych,
- (19) Wykres 10 plików tekstowych z największą liczbą wystąpień wybranego wyrazu.

**Enter a word**

Enter a word you're interested in so we can provide statistics regarding collocations and occurrences

stół

Analyze

Occurrences

Collocations

**Collocations of the word "stół"**

**21** Collocation range

**22** ☐ Do not cross sentences

**Left collocations 23**

Word	Lexem	Part of speech
cicho	cicho	przysłówek
.	.	interpunkcja
ostre	ostry	przymiotnik
wieku	wiek	rzeczownik
nie	nie	kubik
oknem	okno	rzeczownik

**24** Collocation range

**25** ☒ Do not cross sentences

**Right collocations 26**

Word	Lexem	Part of speech
nakryty	nakryć	czasownik - imiesłów przymiotniko...
utopił	utopić	czasownik - pseudoimiesłów
Ja	ja	zaimek nietrzeciosobowy
serwetą	serweta	rzeczownik
ciebie	ty	zaimek nietrzeciosobowy
.	.	interpunkcja

**20**

Refresh

Rysunek 10: Wyniki analizy - kolokacje wyrazu "stół" w korpusie

- (20) Przycisk "Refresh" - służy do odświeżania wyników analizy dla ustawionej konfiguracji,
- (21) Pasek przesuwany to ustalania zakresu dla lewej kolokacji,
- (22) Znacznik ustawiający, aby kolokacje nie przekraczały zdań,
- (23) Tabela z lewymi kolokacjami wyrazu "stół",
- (24) Pasek przesuwany to ustalania zakresu dla prawej kolokacji,
- (25) Znacznik ustawiający, aby kolokacje nie przekraczały zdań,
- (26) Tabela z prawymi kolokacjami wyrazu "stół",

9