

Projekt PJN — Analiza językowa

Projekt demonstracyjny do analizy korpusowej i wizualizacji metryk językowych.

Krótki opis

- Backend: narzędzia przetwarzania korpusu, ekstrakcja rzeczowników, analiza częstości i metryk Zipfa.
- Frontend: interaktywne UI zbudowane w React + Vite, prezentujące tabele, wykresy i graf rdzenia językowego.

Co robi projekt

Analizuje korpus językowy, wylicza częstotliwości słów, liczy metryki Zipfa oraz buduje graf „language core”, a następnie udostępnia wyniki w interfejsie webowym z tabelami i wykresami.

RANGA ^	SŁOWO	LICZBA WYSTĄPIEŃ	CZĘSTOŚĆ WZGLĘDNA
#1	v	24 339	2,742%
#2	a	20 881	2,353%
#3	byt	18 915	2,131%
#4	na	13 322	1,501%
#5	se	10 442	1,176%
#6	rok	8452	0,952%

[Analiza statystyczna języka](#)[Tabela częstości wyrazów](#)[Prawo Zipfa](#)[Rdzeń języka](#)[Top 50 rzeczowników](#)[Analiza semantyczna](#)PUNKTY WYKRESU
2000

Każdy punkt odpowiada rangi i częstości wyrazu.

WSPÓŁCZYNNIK KIERUNKU
-0.880

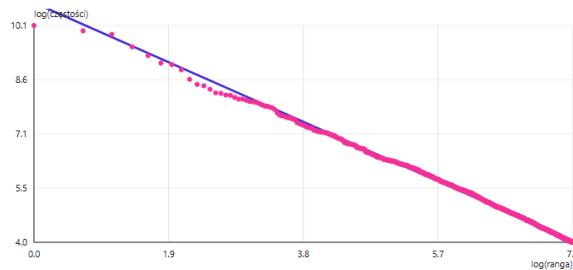
Im bardziej ujemny, tym silniejsza zależność Zipfa.

PRZECIĘCIE
10.738

Logarytm oczekiwanej częstości dla rzędu 1.

Wykres log-log

Punkty pokazują zależność logarytmicznego rzędu od logarytmicznej częstości, a czerwona linia to dopasowanie regresji z metadanych.



WSPÓŁCZYNNIK DETERMINACJI: 0.998

Porównanie częstotliwości

Porównuje rzeczywistą częstość z przewidywaną wartością z modelu Zipfa oraz pokazuje stosunek modelu do obserwacji.

Ranga

Wyraz

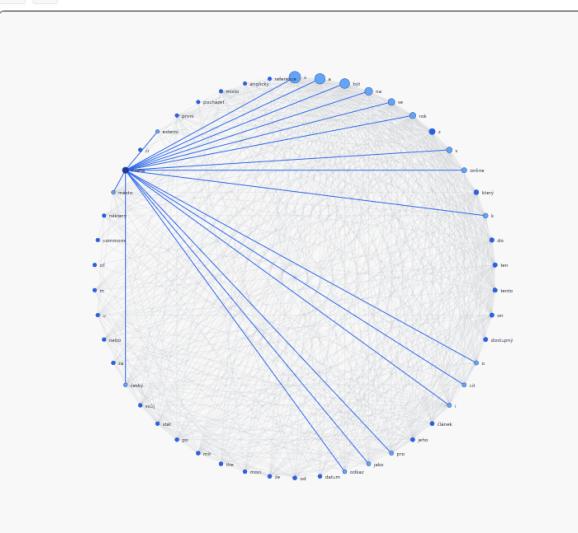
Częstość

Model

Stosunek

[Analiza statystyczna języka](#)[Tabela częstości wyrazów](#)[Prawo Zipfa](#)[Rdzeń języka](#)[Top 50 rzeczowników](#)[Analiza semantyczna](#)

Zoom: 1.00x

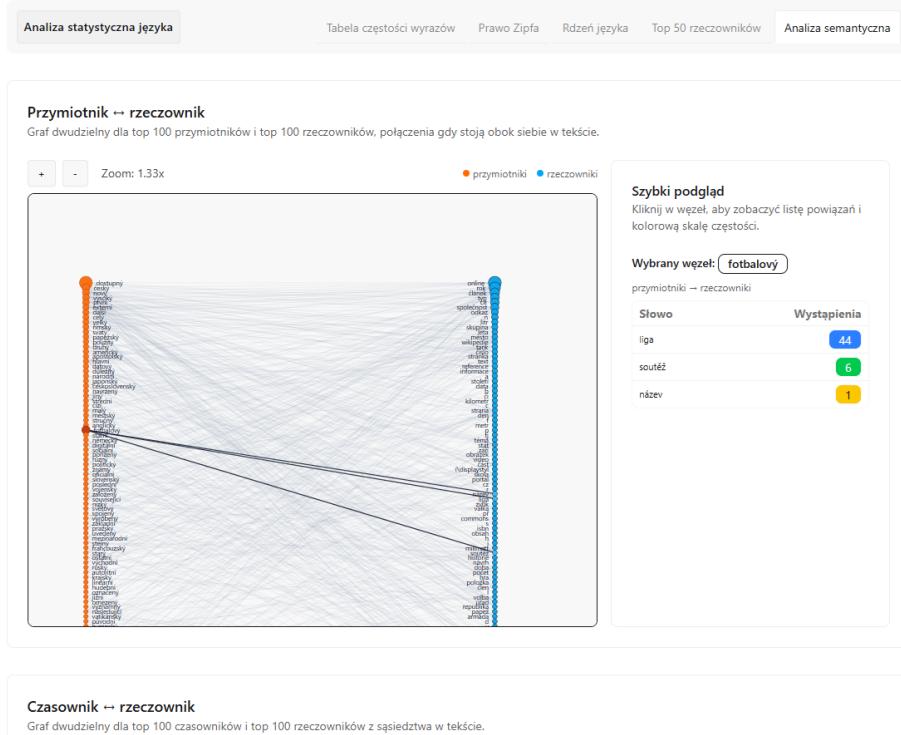
**Praha**Frequency: 1529
Unique neighbors: 904
Connection weight: 3041

Top połączenia

- v — 275
- český — 64
- š — 47
- online — 44
- úroveň — 28
- a — 28
- být — 27
- na — 26
- rok — 26
- díl — 25
- externí — 19
- portály — 18

Top 50 rzeczowników

Ranga	Słowo	Liczba	Tłumaczenie (PL)
#6	rok	8452	rok
#20	článek	2975	artykuł
#28	moci	2442	Móc
#34	český	1905	czeski
#42	město	1557	miasto
#43	Praha	1529	Praga
#47	místo	1398	miejscie
#51	Wikipédie	1351	Wikipedia
#53	obrázek	1276	zdjęcie
#56	téma	1233	temat
#57	datový	1225	dane
#58	položka	1216	przedmiot
#59	video	1213	wideo
#61	portály	1203	portale



Główne foldery

- apps/backend: kod przetwarzania danych i analizy (zobacz README tamże).
- apps/frontend: aplikacja kliencka (React + Vite + Tailwind/DaisyUI).
- img: grafiki używane w README i w interfejsie.

Pliki danych w frontendzie

- Statyczne dane używane przez UI znajdują się w apps/frontend/src/data.

Uruchomienie frontendu (lokalnie)

1. Wejdź do katalogu frontendu:

```
cd apps/frontend
```

2. Zainstaluj zależności:

```
npm install
```

3. Uruchom środowisko deweloperskie:

```
npm run dev
```

Domyślnie Vite udostępnia aplikację pod adresem `http://localhost:5173` — otwórz tę stronę w przeglądarce.

4. (Opcjonalnie) Zbuduj i podglądaj produkcyjną wersję:

```
npm run build  
npm run preview
```

Jeżeli używasz `pnpm` lub `yarn`, zastąp odpowiednio krok instalacji poleceniem `pnpm install` lub `yarn`.

Jeżeli potrzebujesz, mogę dodać skrypt uruchamiający backend lub przykładowe dane do szybkiego startu.