

NLP w języku polskim - obecny stan wiedzy

Łukasz Kobyliński, Sages / IPI PAN
@lkobyliniski





Parę słów o mnie

Łukasz (@lkobylinski)

- Research @ IPI PAN (ICS PAS)
- Consulting @ SigDelta
- Management @ Sages
- Founder @ Labellery



SigDelta

sages

Labellery.

Warsaw University
of Technology



O czym chciałbym powiedzieć?

- NLP to nośny temat, ale najczęściej w kontekście **języka angielskiego**.
- Jak wygląda sytuacja przetwarzania języka naturalnego w kontekście **języka polskiego**?
- Czy warto się tym zajmować i jakie są **otwarte problemy**?

Wiemy, że NLP jest istotne

Zastosowania bezpośrednio w biznesie

- speech recognition
- question answering
- sentiment analysis
- coreference resolution
- information extraction
- event identification
- named entity recognition
- sense disambiguation



Jan Tomczyk

Specjalista ds. Marketingu

Jestem innowacyjnym specjalistą ds. marketingu z 6-letnim doświadczeniem w branży ubezpieczeniowej. Samodzielnie koordynowałem projekt rebrandingu oraz wdrożenia nowej identyfikacji wizualnej. Chęć rozwijać moje kluczowe umiejętności jako Koordynator Marketingu w dynamicznym, międzynarodowym środowisku biznesowym.



Dane osobowe

Data ur.
03.04.1975
Adres
Warszawa
Telefon
777 777 777
E-mail
jan.tomczyk@gmail.com
Social media
LinkedIn

Doświadczenie

- 2012.10 - Firma ubezpieczeniowa - Specjalista ds. Marketingu do dziś**
Zakres obowiązków:
- Tworzę i wdrażam strategię maksymalizacji zysków sprzedażowych
 - Nadzoruję publikację treści na stronach internetowych firmy
 - Koordynuję działania merchandisingowe oddziałów
 - Wspieram wdrożenia nowych produktów
- Osiągnięcia:**
Nadzorowałem projekt rebrandingu spółki, który zakończył się wdrożeniem nowej identyfikacji wizualnej oraz zmianami wyglądu **ponad 200 oddziałów** w Polsce.
Tworzyłem skuteczne kampanie marketingowe dla nowych produktów ubezpieczeniowych, które pozwoliły zdobyć **20% rynku** ubezpieczeń majątkowych.
- 2003.03 - Firma ABC - Specjalista ds. Marketingu**
2012.10 Zakres obowiązków:
- Nadzorowałem współpracę z agencjami reklamowymi
 - Akceptowałem druk materiałów marketingowych
 - Prowadziłem firmowe media społecznościowe
- 2003.01 - Firma ABC - Praktykant w dziale marketingu**
2003.03 Zakres obowiązków:
- Tworzyłem teksty marketingowe do ulotek produktowych
 - Analizowałem oferty konkurencji
- Osiągnięcia:**
Dzięki pozytywnej ocenie praktyk otrzymałem **możliwość pracy etatowej** oraz awansu.



A jak to wygląda dla języka polskiego?

Wciąż niewiele praktycznych implementacji bardziej zaawansowanych metod.

Typowe powody:

- język polski jest obiektywnie trudniejszy w przetwarzaniu,
- brak gotowych narzędzi i łatwych w użyciu bibliotek,
- niewystarczające zasoby językowe.

Dlaczego NLP dla polskiego jest trudne? (#1)

Polski jest językiem fleksyjnym

- Liczba możliwych form w korpusie jest dużo większa, niż np. dla języka angielskiego,
- wiele form jest morfologicznie niejednoznacznych
 - człowieka (D. - dopełniacz / B. - biernik)
 - człowieku (Ms. - miejscownik / W. - wołacz)
 - ludzie (M. - mianownik / W. pl. - wołacz l. mn.)
 - ludzi (D. - dopełniacz / B. pl. - biernik l. mn.)

człowiek			
rzeczownik [SJPDor.]			
m1 <input type="radio"/> B3k+człowiek / <input type="radio"/> B3k+człowiek+!			
	l. p.	l. m.	
M.	człowiek	ludzie	ndepr
		ludzie	depr
D.	człowieka	ludzi	
C.	człowiekowi	ludziom	
B.	człowieka	ludzi	
N.	człowiekiem	ludźmi	
Ms.	człowieku	ludziach	
W.	człowieku	ludzie	ndepr
	człowiecze <i>daw.</i>	ludzie	depr

Dlaczego NLP dla polskiego jest trudne? (#2)

Homonimia

admirat

rzeczownik [SJPDor.]

m1 [B4ł+w](#)

	I. p.	I. m.	
M.	admirat	admiratowie	ndepr
		admiraty	depr
D.	admirata	admiratów	
C.	admiratowi	admiratom	
B.	admirata	admiratów	
N.	admiratem	admiratami	
Ms.	admirale	admiratach	
W.	admirale	admiratowie	ndepr
		admiraty	depr

admirat *gatunek motyla*

rzeczownik

m2 [B4ł+w](#)

	I. p.	I. m.
M.	admirat	admiraty
D.	admirata	admiratów
C.	admiratowi	admiratom
B.	admirata	admiraty
N.	admiratem	admiratami
Ms.	admirale	admiratach
W.	admirale	admiraty



Dlaczego NLP dla polskiego jest trudne? (#3)

Swobodny szyk zdania

- Język angielski
 - John loves Mary (SVO)
- Język polski
 - Jan kocha Marię (SVO),
 - Jan Marię kocha,
 - Marię kocha Jan,
 - Marię Jan kocha,
 - Kocha Marię Jan,
 - Kocha Jan Marię.

Dlaczego NLP dla polskiego jest trudne?

(#4)

W konsekwencji: opis języka jest dużo bardziej złożony

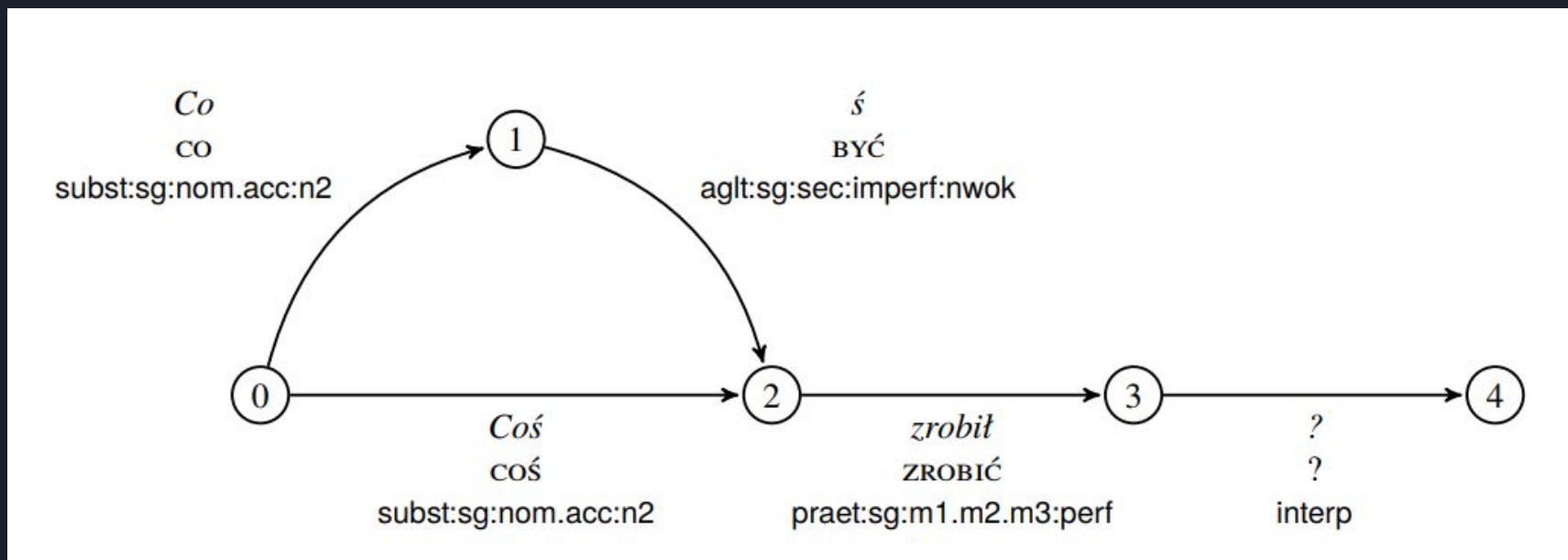
Na przykład: znakowanie morfosyntaktyczne

- j. angielski: 36 - 200 tagów, np. Penn Treebank:
 - NN - noun (singular); NNS - noun (plural);
 - VB - verb, base form; VBD - verb, past tense
- j. polski: 4 000 teoretycznie możliwych tagów! (podobnie dla czeskiego, ok. 2 000 w jęz. słoweńskim).

Number	Tag	Description
1.	CC	Coordinating conjunction
2.	CD	Cardinal number
3.	DT	Determiner
4.	EX	Existential <i>there</i>
5.	FW	Foreign word
6.	IN	Preposition or subordinating conjunction
7.	JJ	Adjective
8.	JJR	Adjective, comparative
9.	JJS	Adjective, superlative
10.	LS	List item marker
11.	MD	Modal
12.	NN	Noun, singular or mass
13.	NNS	Noun, plural
14.	NNP	Proper noun, singular
15.	NNPS	Proper noun, plural
16.	PDT	Predeterminer
17.	POS	Possessive ending
18.	PRP	Personal pronoun
19.	PRPS	Possessive pronoun
20.	RB	Adverb
21.	RBR	Adverb, comparative
22.	RBS	Adverb, superlative
23.	RP	Particle
24.	SYM	Symbol
25.	TO	<i>to</i>
26.	UH	Interjection
27.	VB	Verb, base form
28.	VBD	Verb, past tense
29.	VBG	Verb, gerund or present participle
30.	VCN	Verb, past participle
31.	VBP	Verb, non-3rd person singular present
32.	VBZ	Verb, 3rd person singular present
33.	WDT	Wh-determiner
34.	WP	Wh-pronoun
35.	WPS	Possessive wh-pronoun
36.	WRB	Wh-adverb

Dlaczego NLP dla polskiego jest trudne? (#5)

Bonus: niejednoznaczności segmentacyjne





NLP w języku polskim - czy jest dobrze?

Co umiemy robić z wysoką dokładnością?

Zadanie	Polski	Angielski
Lematyzacja	(Waszczuk '12) ~ 96%	
Znakowanie morfosyntaktyczne	(Krasnowska '17) 94,5%	(Choi '2016) 97,64%
Parsowanie zależnościowe	(Wróblewska '17) 84-93%	(Dozat & Manning '17) 95,75%
Rozpoznawanie nazw własnych	(Marcińczuk et al. '13) 80-85% (F1)	(Strubell et al. '17) 86,81% (F1)



NLP w języku polskim - czy jest dobrze?

Jakie inne główne zadania były podejmowane?

- ujednoznacznianie sensu słów,
- analiza wydźwięku,
- analiza koreferencji,
- tworzenie podsumowań,
- rozpoznawanie sytuacji,

Wyzwania:

- generowanie języka naturalnego,
- transfer learning.



Znakowanie morfosyntaktyczne (tagowanie)

Dotychczasowe podejścia:

- **2007:** znakowanie oparte na regułach (TaKiPI), ok. 88% dokładności,
- **2010:** tager Brilla dostosowany do j. polskiego (PANTERA), ok. 89% dokładności,
- **2012:** uczenie pamięciowe (WMBT), ok. 90% dokładności,
- **2012:** warunkowe pola losowe (WCRFT, Concraft), ok. 91% dokładności,
- **2014:** łączenie klasyfikatorów (PoliTa), ok. 92% dokładności.

Czy istniejące metody można poprawić?

PolEval 2017

- konkurs wzorowany na Kaggle,
- obecnie związany z zadaniami z obszaru NLP dla j. polskiego.

W 2017 roku otrzymaliśmy:

- 16 zgłoszeń od 9 zespołów,
- wszystkie zgłoszenia (!) oparte były na sieciach neuronowych.

POLEVAL 2017

[Home](#) [Tasks](#) [Dates](#) [Submission](#) [Results](#) [Prizes](#) [Publication](#) [Organizers](#)



Task 1: POS Tagging

Introduction

There is an ongoing discussion whether the problem of part of speech tagging is already solved, at least for English (see Manning 2011), by reaching the tagging error rates similar or lower than the human inter-annotator agreement, which is ca. 97%. In the case of languages with rich morphology, such as Polish, there is however no doubt that the accuracies of around 91% delivered by taggers leave much to be desired and more work is needed to proclaim this task as solved.

The aim of this proposed task is therefore to stimulate research in potentially new approaches to the problem of POS tagging of Polish, which will allow to close the gap between the tagging accuracy of systems available for English and languages with rich morphology.

Task definition

Subtask (A): Morphosyntactic disambiguation and guessing

Given a sequence of segments, each with a set of possible morphosyntactic interpretations, the goal of the task is to select the correct interpretation for each of the segments and provide an interpretation for segments for which only 'ign' interpretation has been given (segments unknown to the morphosyntactic dictionary).

Subtask (B): Lemmatisation

Given a sequence of segments, each with a set of possible morphosyntactic interpretations, the goal of the task is to select the correct lemma for each of the segments and provide a lemma for segments for which only 'ign' interpretation has been given (segments unknown to the morphosyntactic dictionary).



Znakowanie morfosyntaktyczne po PolEval 2017


System name	Acc (%)	deep network	hand-crafted features	character-level embeddings	word-level embeddings
<i>Toygger</i>	94.6343	<i>yes</i>	<i>yes</i>	<i>no</i>	<i>yes</i>
<i>KRNNT_AB</i>	93.8083	<i>yes</i>	<i>yes</i>	<i>no</i>	<i>no</i>
<i>NeuroParser</i>	93.6109	<i>yes</i>	<i>no</i>	<i>yes</i>	<i>no</i>
<i>AvgPer_Forced</i>	90.9134	<i>no</i>	<i>yes</i>	<i>no</i>	<i>no</i>
<i>Concraft</i>	91.6115	<i>no</i>	<i>yes</i>	<i>no</i>	<i>no</i>
<i>WCRFT</i>	91.1693	<i>no</i>	<i>yes</i>	<i>no</i>	<i>no</i>
<i>WMBT</i>	90.6722	<i>no</i>	<i>yes</i>	<i>no</i>	<i>no</i>



PolEval 2018

Tegoroczna edycja PolEval

- Task 1: Parsowanie zależnościowe.
- Task 2: Rozpoznawanie jednostek nazewniczych (NER).
- Task 3: Modele językowe.



Czy są biblioteki, z których można skorzystać?

Niestety, w większości bez wsparcia dla języka polskiego:

- **NLTK** - dobra do badań naukowych i edukacji,
- **Spacy.io** - dobra na produkcji - załączek wsparcia dla j. polskiego.

Pozostałe:

- **CoreNLP, OpenNLP** - możliwość trenowania własnych modeli,
- **gensim** - wykorzystanie reprezentacji wektorowych słów (zanurzenia).



Czy są zasoby, z których można skorzystać?

Korpusy tekstowe z dostępnym źródłem

Korpus	Rozmiar (segmentów)	Anotacja
Narodowy Korpus Języka Polskiego	~1,2 mln	Ręczna: POS, NER, WSD, frazy składniowe
KPWr	~ 450 tys.	Automatyczna: POS, NER, WSD, frazy składniowe, sytuacje
Polski Korpus Sejmowy	~ 300 mln	Automatyczna: POS, NER, frazy składniowe
Korpus Polskiej Wikipedii	~ 169 mln	Automatyczna: POS, NER, frazy składniowe
PolEval 2018	> 400 mln	brak

Czy są zasoby, z których można skorzystać?

Słownik morfologiczny (sgjp.pl)

- **SGJP** - tabele odmian gramatycznych słów w j. Polskim
- **Morfeusz** - narzędzie, pozwalające na efektywne przeszukiwanie

hasło	kl. gr.	rdz.
konfekt	rz.	m3
konferansjer	rz.	m1
konferansjerka	rz.	f
konferansjerski	przym.	
konferansjersko	ps.	
konferansjerskość	rz.	f
konferansjerstwo	rz.	n2
konferencja	rz.	f
konferencyjka	rz.	f
konferencyjność	rz.	f
konferencyjny	przym.	
konferować	cz.	ndk
konferować	cz.	ndk/dk
konferowanie	ods.	ndk
konferowanie	ods.	ndk/dk

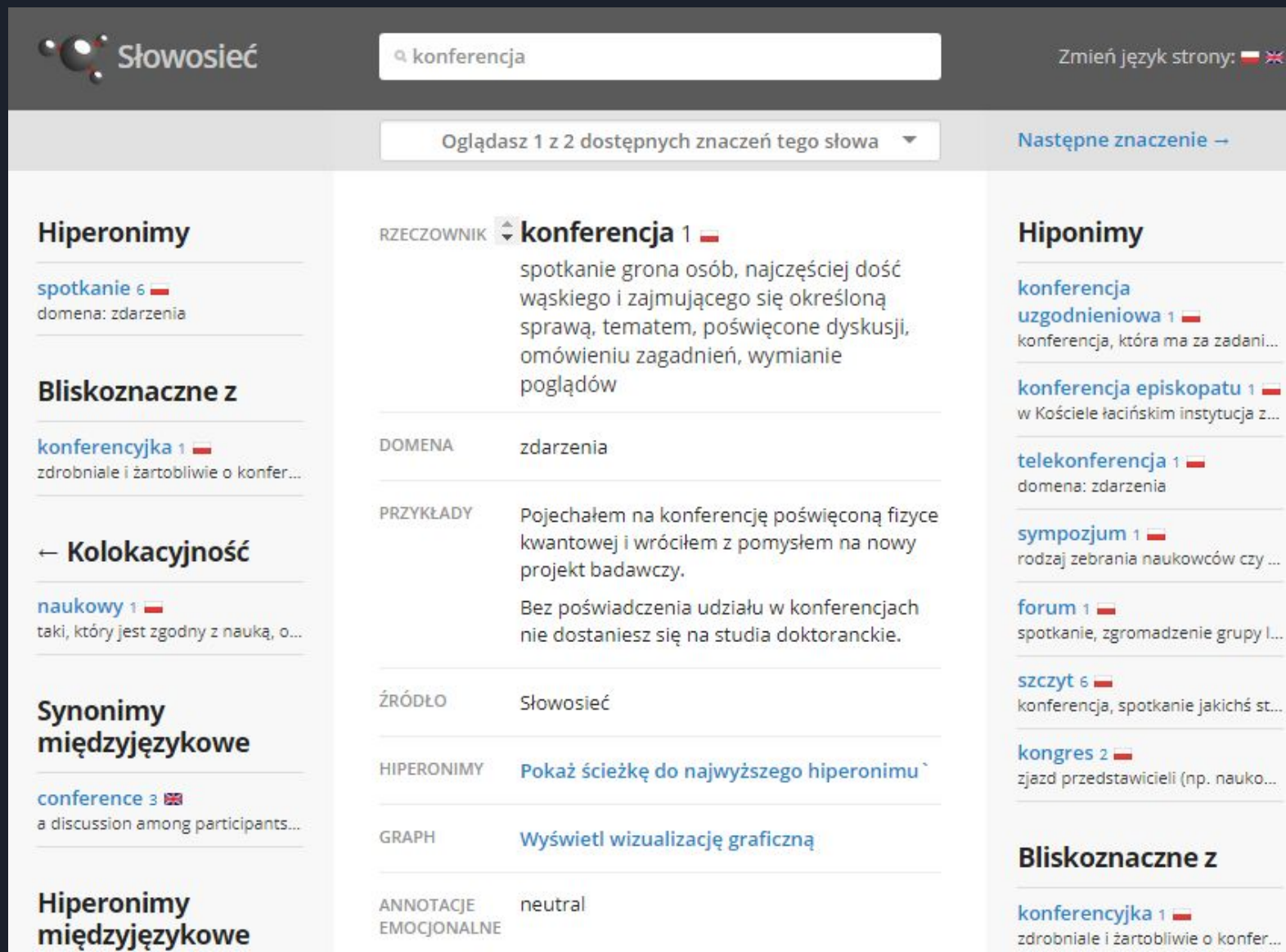
konferencja
rzeczownik [SJPDor.]
f D1j±(yj)

	l. p.	l. m.
M.	konferencja	konferencje
D.	konferencji	konferencji <i>hom.</i> konferencyj <i>char. arch.</i>
C.	konferencji	konferencjom
B.	konferencję	konferencje
N.	konferencją	konferencjami
Ms.	konferencji	konferencjach
W.	konferencjo	konferencje

Czy są zasoby, z których można skorzystać?

SłowoSiec

- Web
- API
- XML



The screenshot displays the SłowoSiec website interface. At the top, there is a search bar with the word 'konferencja' entered. To the right of the search bar, there is a link to 'Zmień język strony: PL EN'. Below the search bar, a navigation bar indicates 'Oglądasz 1 z 2 dostępnych znaczeń tego słowa' and a link to 'Następne znaczenie →'.

The main content area is divided into three columns. The left column contains sections for 'Hiperonimy' (listing 'spotkanie 6' with domain 'zdarzenia'), 'Bliskoznaczne z' (listing 'konferencyjka 1' with domain 'zdrobniale i żartobliwie o konferencjach'), '← Kolokacyjność' (listing 'naukowy 1' with domain 'taki, który jest zgodny z nauką, o...'), 'Synonimy międzyjęzykowe' (listing 'conference 3' with domain 'a discussion among participants...'), and 'Hiperonimy międzyjęzykowe'. The middle column shows the word 'konferencja 1' with its part of speech 'RZECZOWNIK' and a detailed definition: 'spotkanie grona osób, najczęściej dość wąskiego i zajmującego się określoną sprawą, tematem, poświęcone dyskusji, omówieniu zagadnień, wymianie poglądów'. It also lists the domain 'zdarzenia', provides two examples of usage, and identifies the source as 'SłowoSiec'. The right column lists 'Hiponimy' including 'konferencja uzgodnieniowa 1', 'konferencja episkopatu 1', 'telekonferencja 1', 'sympozjum 1', 'forum 1', and 'szczyt 6', each with a brief description and domain.

At the bottom of the middle column, there are links for 'Pokaż ścieżkę do najwyższego hiperonimu`' and 'Wyświetl wizualizację graficzną'. The bottom of the page shows 'ANNOTACJE EMOCJONALNE' with the value 'neutral'.

Czy są zasoby, z których można skorzystać?

Słownik walencyjny

opisuje sposób, w jaki określone jednostki językowe wiążą się ze swoimi podrzędnikami

Hasło	Status
abdykacja	(S) gotowe
abdykować	(S) sprawdzone
abłacja	(S) gotowe
abolicja	(S) gotowe
abominacja	(S) gotowe
abonament	(S) w obróbce
abonować	(S) sprawdzone
aborcja	(S) gotowe
absencja	(S) gotowe
absolutorium	(S) gotowe
absolutyzacja	sprawdzone
absolutyzować	(S) sprawdzone
absorbować	(S) sprawdzone

Schematy [2]Semantyka [2+0]Przykłady [0]

adresować (1M=15,300M=3943)

adresować (_,imperf):

Schemat:	pewny [1]			
Funkcja:	subj	obj		
Typy fraz:	np(str)	np(str)	prepnp(do,gen)	

Schemat:	pewny [70]			
Funkcja:	subj	obj		

Hasło	Status
abdykacja	(S) gotowe
abdykować	(S) sprawdzone
ablacja	(S) gotowe
abolicja	(S) gotowe
abominacja	(S) gotowe
abonament	(S) w obróbce
abonować	(S) sprawdzone
aborcja	(S) gotowe
absencja	(S) gotowe
absolutorium	(S) gotowe
absolutyzacja	sprawdzone
absolutyzować	(S) sprawdzone
absorbować	(S) sprawdzone
absorpcja	(S) gotowe
abstrahować	(S) sprawdzone
abstynencja	(S) gotowe
absurd	(S) sprawdzone
absurdalnie	sprawdzone
absurdalność	(S) sprawdzone
absurdalny	sprawdzone
achać	(S) gotowe
adaptacja	(S) gotowe
adaptować	(S) sprawdzone
adekwatność	(S) sprawdzone
adekwatny	(S) gotowe
adjustacja	sprawdzone


Schematy [2]


Semantyka [2+0]

Przykłady [0]

adresować (1M=15,300M=3943)

adresować (_,imperf):

Schemat:	pewny [1] 		
Funkcja:	subj	obj	
Typy fraz:	np(str)	np(str)	prepn(do,gen)

Schemat:	pewny [70] 		
Funkcja:	subj	obj	
Typy fraz:	np(str)	np(str)	prepn(na,acc)

Przegląd przykładów

Filtrowanie schematów x

Identyfikator:	Przykład:	Źródło:	Ocena:
359006	Nie było dla nikogo tajemnicą, że tym członkiem ONZ do którego sekretarz generalny adresował głównie swe słowa były Stany Zjednoczone.	podkorpus zrównoważony NKJP (300M segmentów)	dobry
	No niech się nie martwi, a nie smuci - ja i tak	podkorpus zrównoważony	

Czy są zasoby, z których można skorzystać?

Zanurzenia słów (embeddings)

<http://dsmodels.nlp.ipipan.waw.pl>

Word2Vec x

dsmodels.nlp.ipipan.waw.pl/w2v.html

Modele Podobieństwo wyrazów Analogie Arytmetyka wyrazów

Modele

[Semantyka dystrybucyjna](#) znalazła w ostatnich latach szerokie zastosowanie w rozwiązywaniu szeregu zadań związanych z językiem naturalnym. U jej podstaw leży hipoteza, że słowa występujące w podobnych kontekstach w dużych zbiorach danych mają podobne znaczenie. Znaczenia słów reprezentowane są przez wektory liczbowe.

Poniższa lista zawiera modele dystrybucyjne (zbiory wektorów reprezentujących znaczenia słów) dla języka polskiego z pakietu [gensim](#) dla dwóch dużych korpusów tekstowych: pełnej wersji Narodowego Korpusu Języka Polskiego (NKJP) z końca 2016 roku.

Źródło:	<input checked="" type="checkbox"/> NKJP	<input checked="" type="checkbox"/> Wiki	<input checked="" type="checkbox"/> NKJP+Wiki
Typ modelu:	<input checked="" type="checkbox"/> Formy	<input checked="" type="checkbox"/> Lematy	
Rozmiar wektora:	<input checked="" type="checkbox"/> 100	<input checked="" type="checkbox"/> 300	

#	Nazwa pliku do pobrania	MB
1	nkjp+wiki-forms-all-100-cbow-hs.txt.gz	802.3
2	nkjp+wiki-forms-all-100-cbow-ns.txt.gz	758.6
3	nkjp+wiki-forms-all-100-skipg-hs.txt.gz	780.0
4	nkjp+wiki-forms-all-100-skipg-ns.txt.gz	771.0
5	nkjp+wiki-forms-all-300-cbow-hs-50.txt.gz	634.3
6	nkjp+wiki-forms-all-300-cbow-hs.txt.gz	2312.1
7	nkjp+wiki-forms-all-300-cbow-ns-50.txt.gz	620.8
8	nkjp+wiki-forms-all-300-cbow-ns.txt.gz	2172.2
9	nkjp+wiki-forms-all-300-skipg-hs-50.txt.gz	590.2
10	nkjp+wiki-forms-all-300-skipg-hs.txt.gz	2242.1
11	nkjp+wiki-forms-all-300-skipg-ns-50.txt.gz	591.4
12	nkjp+wiki-forms-all-300-skipg-ns.txt.gz	2197.3
13	nkjp+wiki-forms-restricted-100-cbow-hs.txt.gz	748.3
14	nkjp+wiki-forms-restricted-100-cbow-ns.txt.gz	707.2
15	nkjp+wiki-forms-restricted-100-skipg-hs.txt.gz	727.3
16	nkjp+wiki-forms-restricted-100-skipg-ns.txt.gz	720.3

Nazwa pliku: corpus-type-stype-dim-arch-alg.txt.gz	
corpus	nazwa korpusu - nkjp , wiki lub nkjp+wiki
type	typ modelu - model oparty na formach (forms) lub lematach (lemmas)
stype	podtyp modelu - wszystkie części mowy (all) lub tylko wybrane części mowy (restricted)
dim	rozmiar wektora - 100 lub 300
arch	architektura sieci neuronowej - CBOW (cbow) lub Skip-Gram (skipg)
alg	algorytm uczący - Hierarchical Softmax (hs) lub Negative Sampling (ns)

Niektóre modele ograniczone zostały tylko do tych słów, które wystąpiły co najmniej 30 lub 50 razy w korpusie. Jest to zaznaczone po nazwie algorytmu uczącego [alg](#). [it100](#) w nazwie pliku oznacza, że dany model został wytrenowany w stu iteracjach.

NLP w języku polskim - nie tylko publikacje naukowe

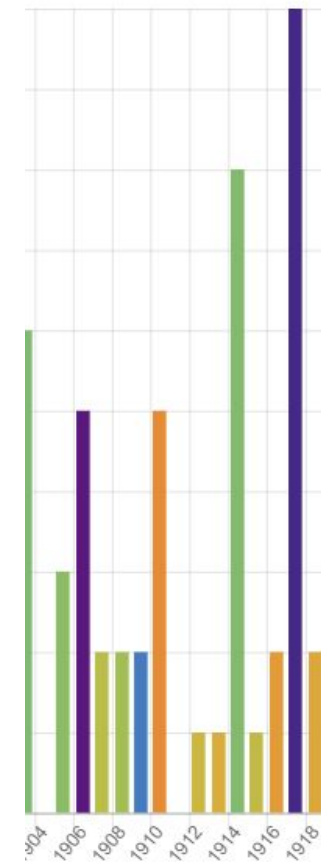
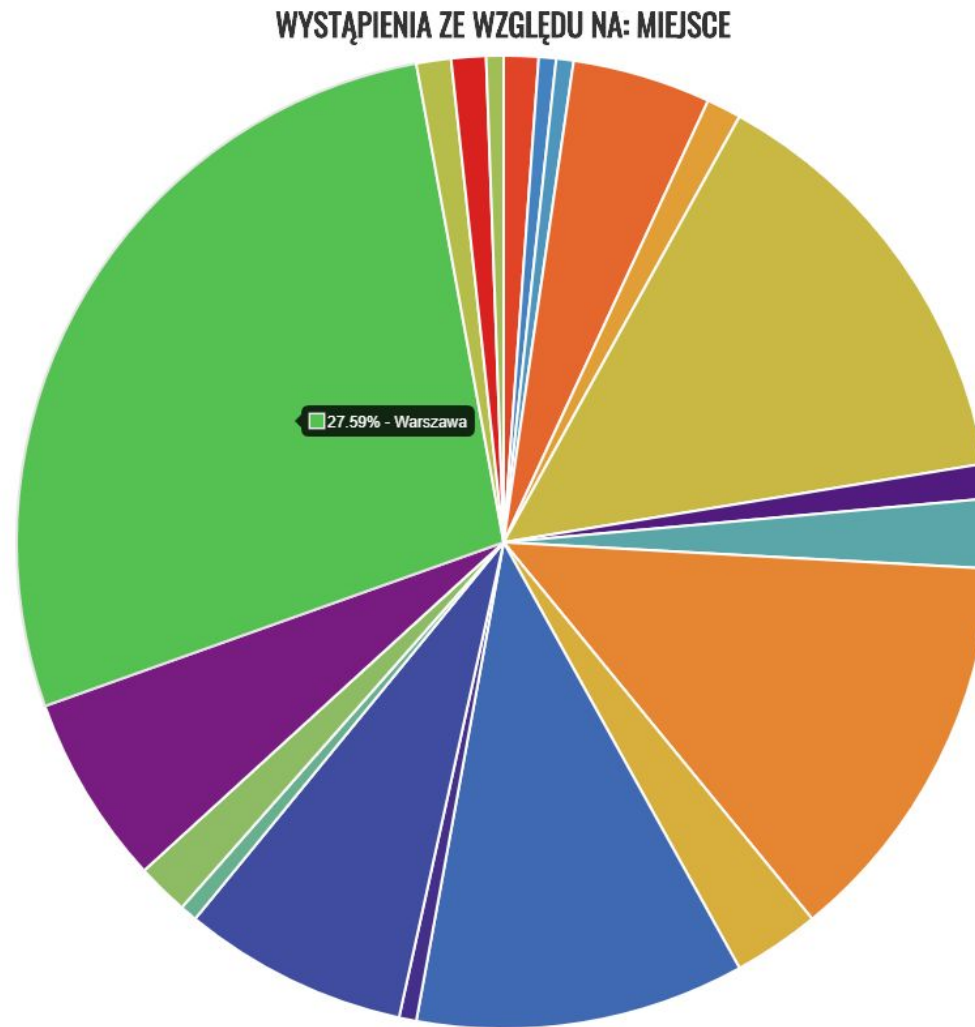
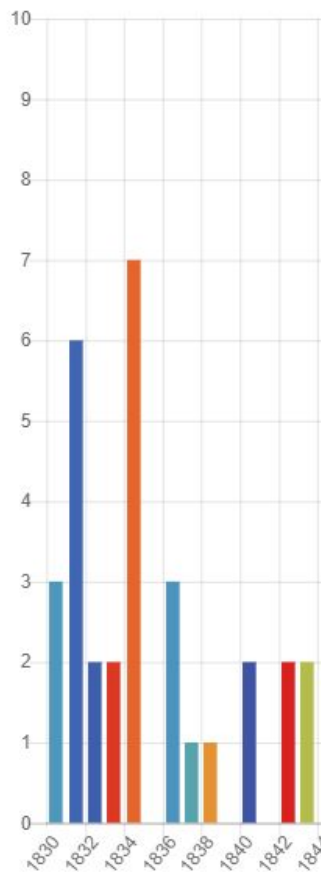
CLARIN-PL - narzędzia

(<http://clarin-pl.eu/en/uslugi/>)

The screenshot displays the 'NARZĘDZIA I APLIKACJE' (Tools and Applications) section of the CLARIN-PL website. The page features a dark header with navigation links: Home page, About project, Services, Media library, Licenses, and flags for Poland and the European Union. A search icon is located in the top right corner. The main content area is a grid of 12 tool cards, each with a purple circular icon, a title, a brief description, and a link to the tool.

Tool Name	Description	Link
Chunker	program do płytkiej analizy składniowej	Zobacz demo ▶
Websty	interfejs webowy do grupowania tekstów w języku polskim	Przejdź do Websty ▶
Nowy Morfeusz	analizator morfologiczny umożliwiający dostosowanie do dziedziny	Przejdź do Nowego Morfeusza ▶
Liner2	rozpoznawanie jednostek identyfikacyjnych i wyrażen temporalnych	Przejdź do Liner2 ▶
Inforex	system do edycji anotowanych korpusów	Zobacz Inforex ▶
klasyfikator tematyczny tekstów WIKNN	pierwsza wersja klasyfikatora tematycznego Wikipedia K-Nearest Neighbours dla tekstów polskich i angielskich	Przejdź do klasyfikatora ▶
Kuźnia	narzędzie do (współ)tworzenia fleksyjnych słowników dziedziny	Przejdź do Kuźni ▶
WNLoom-Viewer	aplikacja do przeglądania Słowosieci	Pobierz WNLoom-Viewer ▶
Mapa Literacka	rozpoznawanie odniesień do nazw geograficznych i miejscowych	Przejdź do Mapy Literackiej ▶
MeWeX	aplikacja do wydobywania z korpusu słowników kolokacji oraz tworzenia słowników jednostek leksykalnych	Przejdź do MeWeX ▶
Mowa	narzędzia i usługi do przetwarzania mowy	Zobacz narzędzia ▶
Transkrypcja fonetyczna	narzędzie do konwersji zapisu ortograficznego na fonetyczny	Przejdź do transkrypcji fonetycznej ▶

NLP w języku polskim - nie tylko publikacje naukowe



NLP w języku polskim - nie tylko publikacje naukowe

Jasnopis

STATYSTYKI

Klasa trudności tekstu:

6 / 7

Tekst trudny w odbiorze dla przeciętnego Polaka

[rozwiń](#)

LEGENDA

Aa Fragment wyraźnie trudniejszy od reszty tekstu

Aa Fragment trudniejszy od reszty tekstu

Aa Bardzo długie zdanie

Aa Trudne słowo wymagające zmiany

DANE

Tekst po analizie

W poniedziałek, 4 czerwca w Urzędzie Komisji Nadzoru Finansowego odbyło się pierwsze spotkanie Grupy roboczej ds. Blockchain. W jego skład weszli członkowie międzyresortowego Zespołu roboczego ds. rozwoju innowacji finansowych (FinTech) oraz dodatkowe podmioty rynkowe reprezentujące branżę kryptowalut i technologii Blockchain. Celem działania Grupy jest rozpoczęcie przez UKNF dialogu z przedstawicielami rynku na temat technologii DLT/Blockchain, w tym walut wirtualnych (tzw. kryptowalut) oraz innych tokenów cyfrowych, w kontekście właściwych uregulowań prawnych w tym zakresie, w tym przygotowanie Raportu końcowego z prac Grupy określającego stanowisko UKNF w przedmiotowym zakresie.

Grupa będzie zajmowała się aspektami operacyjnymi i prawnymi dotyczącymi technologii rozproszonych rejestrów (DLT/Blockchain), ze szczególnym uwzględnieniem tematyki walut wirtualnych oraz innych tokenów cyfrowych, ICO (Initial Coin Offerings), smart kontraktów, funkcjonowania giełd i kantorów wirtualnych, w tym giełd kryptowalut oraz ograniczeń i wyzwań dla rozwoju technologii DLT/Blockchain, w tym dla rozwoju walut wirtualnych.

„Chcielibyśmy rozpocząć dialog dotyczący rozwiązań związanych z technologią blockchain, w tym „walut” wirtualnych, ICO. Chcielibyśmy w ciągu najbliższych miesięcy wspólnie porozmawiać o wyzwaniach, przed którymi stoi branża, o percepcji obecnie funkcjonujących regulacji, o barierach oraz o tym, jak te bariery możemy usuwać lub ograniczać. Chcielibyśmy ustalić wspólny punkt widzenia” – powiedział Marek Chrzanowski, Przewodniczący KNF, który otworzył posiedzenie Grupy.

„W niektórych kwestiach się różnimy, ale w niektórych jesteśmy ze sobą bardzo blisko. Nie jesteśmy do nikogo uprzedzeni. Jesteśmy postrzegani przez pryzmat działań ostrzegawczych, które realizujemy w ramach kampanii dotyczących zagrożeń na rynku finansowym, natomiast chcę podkreślić, że co do zasady nie jesteśmy przeciwnikami technologii blockchain i nie chcemy



Kluczowe odnośniki

- CLIP: <http://clip.ipipan.waw.pl/>
- CLARIN-PL: <https://clarin-pl.eu>
- Słownik Gramatyczny Języka Polskiego: <http://sgjp.pl/>
- Słowosieć: <http://plwordnet.pwr.wroc.pl/wordnet/>
- PolEval: <http://poleval.pl>

Dziękuję za uwagę!

