

POLIMORF

The ultimate inflectional dictionary of Polish

PoliMorf

is the new morphological dictionary for Polish resulting from the merger of two large existing dictionaries: **SGJP** and **Morfologik**.

SGJP

Słownik gramatyczny języka polskiego
(Grammatical Dictionary of Polish)
<http://sgjp.pl>

Authors Zygmunt Saloni, Marcin Woliński,
Robert Wołosz, Włodzimierz Gruszczyński,
Danuta Skowrońska

Lexemes 244,341

Forms 4,223,981

Aim scientific – development of a sound description
of Polish inflection

Development closed-source

Morfologik

<http://morfologik.blogspot.com/>

Author Marcin Miłkowski

Based on collaborative spell-checking dictionary
SJP.pl enriched with morphological information
using heuristic methods

Lexemes 216,992

Forms 3,475,809

Aim practical – word games, spell-checking,
grammar-checking

Development open-source

Other sources

- additions based on the IPI PAN Corpus
- proper names described for other IPI PAN projects
- lexemes added within the scope of the Cesar project

PoliMorf in numbers (by source dictionary)

| | common | SGJP only | Morfologik only | total |
|----------------|---------------|---------------|-----------------|---------------|
| nouns | 74342 | 94587 | 50922 | 219851 |
| adjectives | 66434 | 32271 | 33224 | 131929 |
| verbs | 17442 | 12362 | 4023 | 33827 |
| non-inflecting | 9111 | 17281 | 2304 | 28696 |
| total | 167329 | 156501 | 90473 | 414303 |

Internal representation

Patterns

Inflectional pattern

| | sg | pl |
|------|-------|-------|
| nom | cko | ci |
| gen | cka | ci |
| dat | cku | ciom |
| acc | cko | ci |
| inst | ckiem | ćmi |
| loc | cku | ciach |
| voc | cko | ci |

dzie +

stem + ending

Inflection

| | prefix | stem | ending | suffix | bafoitag | tag |
|-----|--------|-------|--------|--------|----------|--------------------|
| naj | | bia | tego | | 2 | adj:sg:gen:m1 |
| | | biels | zego | | 2 | adjsup:sg:gen:m1 |
| | | mag | a | | sg:gen | sg:gen:m1 |
| | | mag | a | | sg:gen | sg:acc:m1 |
| | | sag | a | | sg:gen | sg:gen:m3 |
| | | czyta | ł | o | 8 | praet:sg:n1:ter |
| | | czyta | ł | y | 8 | praet:pl:f:ter |
| nie | | czyta | n | y | 10 | ppas:sg:nom:m1:neg |

Export format

```
...
słowników słownik subst:pl:gen:m3
słownikiem słownik subst:sg:inst:m3
słownika słownik subst:sg:gen:m3
słowniki słownik subst:pl:voc:m3
słownikom słownik subst:pl:dat:m3
słownikami słownik subst:pl:inst:m3
słownikach słownik subst:pl:loc:m3
słowniki słownik subst:pl:acc:m3
słownik słownik subst:sg:acc:m3
słownik słownik subst:sg:nom:m3
słowniki słownik subst:pl:nom:m3
słowniku słownik subst:sg:loc:m3
słowniku słownik subst:sg:voc:m3
słownikowi słownik subst:sg:dat:m3
...
```

**liberal
BSD
license!**



<http://zil.ipipan.waw.pl/PoliMorf>

