

Wikidane i Wikisłownik



Leszek Manicki
Wikimedia Deutschland

Zlot Zimowy Wikimedian 2018
Rochna, 13.01.2018

Wikidane

- Repozytorium ustrukturalizowanych danych
- Jedna baza danych, dostępna w wielu językach
- Przechowują i udostępniają dane dla projektów Wikimedia (nie tylko dla Wikipedii!)
- Obsługują także inne projekty, serwisy, itp.
- www.wikidata.org

Dlaczego?



Wikisłowniki to skarb

- Mało kto ma mapę prowadzącą do tego skarbu
- Złożony system szablonów, struktur, itd
- Każdy Wikisłownik sobie rzepekę skrobie
- Dane (hasła) nie są zapisane w postaci łatwej do użycia poza Wikisłownikiem

Zróźnicowanie Wikisłowników

```
== koza ({{język polski}}) ==
[[Plik:Goats Go Inspecting..jpg|thumb|[[para]] kóz (1.1)]]
[[Plik:Hofmann uebungshuemmelchen.JPG|thumb|koza (1.3)]]
[[Plik:Knappenrode - Energiefabrik - HEISS GELIEBT 06 ies.jpg|thumb|koza (1.5)]]
{{wymowa}} {{IPA3|'kɔza}}, {{AS3|k''o''za}} {{audio|Pl-koza.ogg}}
{{znaczenia}}
''rzeczownik, rodzaj żeński''
: (1.1) {{zool}} [[ssak]] [[z]] [[rodzina|rodziny]] [[krętorogie|krętorogich]]; {{wikipedia}}
: (1.2) {{pot}} [[areszt]], [[więzienie]]
: (1.3) {{muz}} [[rodzaj]] [[instrument]]u [[muzyczny|muzycznego]] [[dęty|dętego]]; {{wikipedia|dudy}}
: (1.4) {{pot}} [[wydzielina]] [[z]] [[nos]]a
: (1.5) {{pot}} [[mały]], [[żelazny]] [[piecyk]] [[z]] [[rura|rurą]]
: (1.6) {{gwara}} [[wypchany|wypchana]] [[koźlęcy|koźlęca]] [[skóra]] [[nosić|noszona]] [[przez]] [[młodzienc|młodzienc]]
{{odmiana}}
: (1) {{odmiana-rzeczownik-polski}}
|Mianownik lp = koza
|Dopełniacz lp = kozy
|Celownik lp = kozie
|Biernik lp = kozę
|Narzędnik lp = kozą
|Miejscownik lp = kozie
|Wołacz lp = kozo
|Mianownik lm = kozy
|Dopełniacz lm = kóz<ref name="Markowski">{{MarkowskiNSPP2000|strony=364}}</ref>
|Celownik lm = kozom
|Biernik lm = kozv
```

```
==Polish==
===Etymology===
From {{inh|pl|sla-pro|*koza}}.

===Pronunciation===
* {{pl-IPA}}
* {{audio|Pl-koza.ogg|Audio|lang=pl}}
* {{hyphenation|ko|za|lang=pl}}

===Noun===
{{pl-noun|f|dim=kózka}}

# {{l|en|goat}}
# {{l|pl|colloquial}} {{l|en|bogey}}, {{l|en|booger}} (piece of solid or semisolid mucus in the nostril)
# {{l|pl|slang}} {{l|en|jail}}, {{l|en|prison}}

===Declension===
{{pl-decl-noun-f|genp=kóz}}

===Derived terms===
* {{l|pl|kozi}}

===Related terms===
* {{l|pl|koziół}}
```

```
==Puola==
===Substantiivi===
{{pl-subs|ko|za|suku=f}}
# [[vuohi]]

[[Luokka:Puolan kielen eläimet]]
```

Zróżnicowanie Wikisłowników

Hasła

Angielski	5 454 935
Francuski	3 227 897
Niemiecki	663 284
Polski	604 284
Japoński	195 921
Perski	97 800
Czeski	96 989
Arabski	64 668

Aktywni użytkownicy

Angielski	1407
Francuski	452
Niemiecki	202
Polski	120
Czeski	54
Perski	28
Kannada	9
Suahili	5
Kaszubski	2

Stan na 12.01.2018 10:00. Źródło: <https://meta.wikimedia.org/wiki/Wiktionary>.

Search

[Ask Question](#)

315 results

relevance newest votes active

5 votes

Q: Parse Wiktionary

have one that is specifically suited to **wiktionary**, one that could give me what type of word it is and all of the definitions. I would prefer not to write my own parser for this. Any suggestions? ...

2 answers

[c#](#) [.net](#) [wiki](#) [mediawiki-api](#) [wiktionary](#)

asked Dec 5 '11 by [Bananable](#)

33 votes

Q: Has anyone parsed Wiktionary?

Wiktionary is a wiki dictionary that covers many languages. It even has translations. I would be interested in parsing it and playing with the data, has anyone does anything like this before? Is there any library I can use? (Preferably Python.) ...

10 answers

[python](#) [web-services](#) [dictionary](#) [wiktionary](#)

asked Jul 29 '10 by [Rory](#)

6 votes

Q: Wiktionary API for pronunciation?

I was looking for a way to get the pronunciation of any given word by querying an API of some sort. As **Wiktionary** comes in handy to find pronunciations of words I was trying to use their API, but how do I get the pronunciation of a specific word? ...

3 answers

[php](#) [api](#) [mediawiki](#) [mediawiki-api](#) [wiktionary](#)

asked Mar 21 '11 by [baik](#)

0 votes

Q: Wiktionary in Structured Format

How do I acquire a **Wiktionary**, for say English, in structured format, typically RDF? The recommended website <http://downloads.dbpedia.org/wiktionary/> is dead. And I don't understand ...

1 answer

[xml](#) [rdf](#) [semantic-web](#) [wiktionary](#) [wikimedia-dumps](#)

asked Nov 20 '15 by [Nordlow](#)

3 votes

Q: Build Errors When Importing Wiktionary

When importing the **Wiktionary** project into Eclipse, I immediately get error messages related to my strings.xml file. The same errors occur whether I download the project through subversion or from ... the SDK sample files. Console errors: [2010-12-27 15:25:33 - **Wiktionary**] W/ResourceType(2752): Bad XML block: header size 62 or total size 0 is larger than data size 0 [2010-12-27 15:25:33 ...

3 answers

[android](#) [wiktionary](#)

asked Dec 27 '10 by [user555361](#)

[Advanced Search Tips](#)

results found containing

wiktionary

Hot Network Questions

- Forcing Flow Distinct
- Why is the price on the shelves for drinks often different from what I actually pay in supermarkets in Germany?
- How to take keys with me while swimming?
- How to say f*** yeah in Japanese
- How to "burn" through wood in a vacuum?
- How to say "BTW" in German?
- Did Roman unmarried men have a preferred hand and finger for wearing their seal ring?
- Is potential difference the same across each branch in a parallel circuit under ALL circumstances?
- What was a "Qua" in a "Subsidy of Armour" in 1590?
- How was a powerful house so easily defeated?
- Seriously, what volume does a mole of gas occupy?
- Why don't we use low voltage Power Sources for high wattage applications?
- When does one leave turn based action if you are in a hostile area?
- AWK with one decimal place
- Do the symbols that appear in the OP of Sasami-san@Ganbaranai mean anything?
- Is there a safe way to determine how powerful an opponent is?
- Primes other than Optimus

Repositories **364**[Code](#)Commits **34K**Issues **2K**Wikis **404**[Users](#)[Advanced search](#)

364 repository results

Sort: **Best match** ▾

[wikimedia/WiktionaryMobile](#)

Wiktionary on Mobile!

Updated on Dec 16, 2016

● JavaScript

★ 27

[Suyash458/WiktionaryParser](#)

A Python *Wiktionary* Parser

Updated 26 days ago

● Python

★ 34

[benreyнвар/wiktionary-parser](#)

A parser and autocorrection tool for *wiktionary*.

Updated on Dec 4, 2015

● Python

★ 32

[ybourque/Wikparser](#)

Wiktionary Parser

Updated on Feb 10

● PHP

★ 20

[Conradlrwin/wiktionary](#)

A python library for playing with *Wiktionary*

Updated on Jun 27, 2011

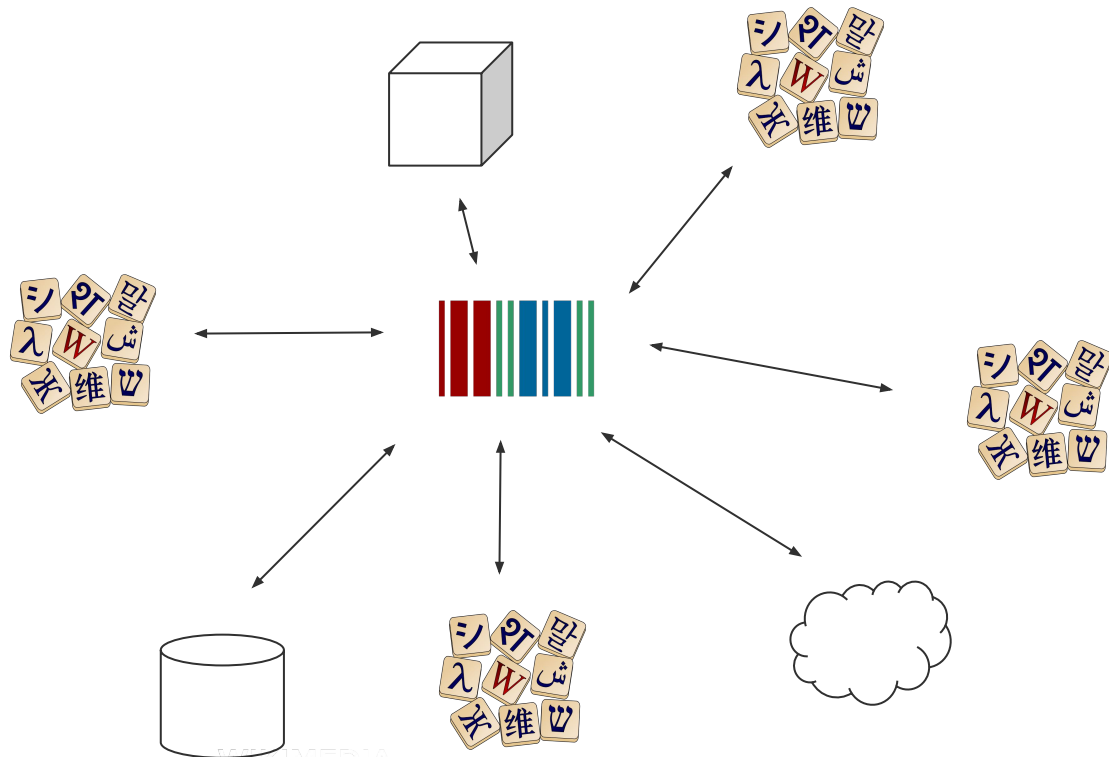
● Python

★ 4

Languages

Java	109
Python	76
JavaScript	47
Ruby	16
Perl	15
C#	14
HTML	11
PHP	11
C++	3
CSS	3

Wikidane mogą pomóc!



Jak Wikidane mogą pomóc?



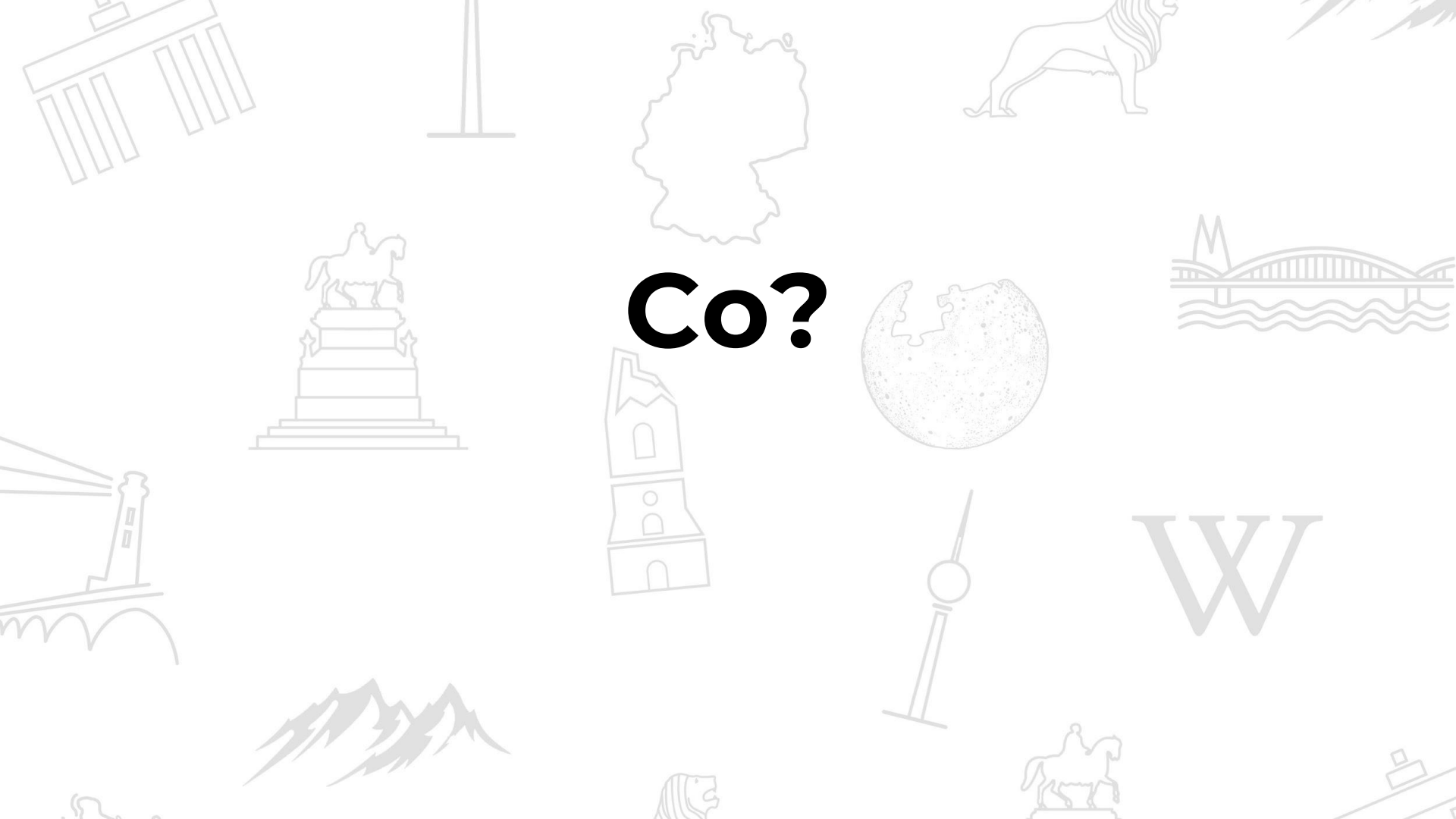
Wikisłownik

- Możliwość wspólnej pracy Wikisłowników (jeśli społeczności zechcą)
- Nowe narzędzia umożliwiające edycję Wikisłownika
- Nowi użytkownicy edytujący Wikisłownik
- Więcej szans dla “małych” społeczności
- **Pomoc a nie zastąpienie Wikisłownika!**

Poza projektami Wikimedia

- Aplikacje słownikowe (jedno- lub wielojęzyczne, tezaury, słowniki synonimów, specjalistyczne)
- Narzędzia wspomagające naukę języków
- Zastosowania badawcze (badania “historyczne”, badanie korpusów)
- Analiza tekstu (anotowanie tekstu częściami mowy, rozpoznawanie jednostek nazwanych, *sentiment analysis*)
- Tłumaczenie automatyczne
- I wiele innych!
- Zapytania SPARQL
- https://www.wikidata.org/wiki/Wikidata:Wiktionary/Ideas_of_tools

Co?



Dane leksykalne w Wikidanych

- Nowe rodzaje encji:
 - Lexeme
 - Form
 - Sense
- Techniczna infrastruktura umożliwiająca modelowanie danych leksykograficznych
 - Minimalna i elastyczna struktura
 - Więcej niż jeden sposób modelowania

(L283)

koza
pl

[edytuj](#)

Język język polski

Kategoria leksykalna rzeczownik

Deklaracje

Rodzaj



żeński

[edytuj](#)

▼ 0 odnośników

+ dodaj przypis

+ dodaj wartość

+ dodaj deklarację

Forms

L283-F1

koza
pl

[edytuj](#)

Grammatical features mianownik, liczba pojedyncza

Statements about L283-F1

Wymowa IPA



['kɔza]

[edytuj](#)

▼ 0 odnośników

+ dodaj przypis

+ dodaj wartość

Plik dźwiękowy z wymową



Pl-koza.ogg

[edytuj](#)

▼ 0 odnośników

+ dodaj przypis

+ dodaj wartość

L283-F2

kozy
pl

[edytuj](#)

Grammatical features liczba pojedyncza, dopełniacz

Statements about L283-F2

Rymuje się z



[płozę](#)

[edytuj](#)

▼ 0 dopełniacz, liczba poj; Leksem: płoza (rzeczownik, polski)

+ dodaj przypis

+ dodaj wartość

+ dodaj deklarację

Senses

283-S1
pl
en

ssak z rodziny kęrotogich

 edytuj

A mammal, *Capra aegagrus hircus*, and similar species of the genus *Capra*.

Statements about L283-S1

oznacza



koza

 edytuj

▼ 0 odnośników

+ dodaj przypis

+ dodaj wartość

tłumaczenie



goat

 edytuj

▼ 0 odnośników

+ dodaj przypis

+ dodaj wartość

+ dodaj deklarację

 edytuj

283-S2
pl
areszt,
więzienie

Statements about L283-S2

rejestr



potoczny

 edytuj

▼ 0 odnośników

+ dodaj przypis

+ dodaj wartość

+ dodaj deklarację

 edytuj

283-S3
pl
mały, żelazny piecyk z
rurą

Więcej informacji Kontakt



<https://www.wikidata.org/wiki/Wikidata:Wiktionary>

https://www.wikidata.org/wiki/Wikidata_talk:Wiktionary

<https://lists.wikimedia.org/mailman/listinfo/wikidata>

https://www.wikidata.org/wiki/Wikidata:Status_updates

System testowy: <http://wikidata-lexeme.wmflabs.org>

lydia.pintscher@wikimedia.de

lea.lacroix@wikimedia.de

leszek.manicki@wikimedia.de