

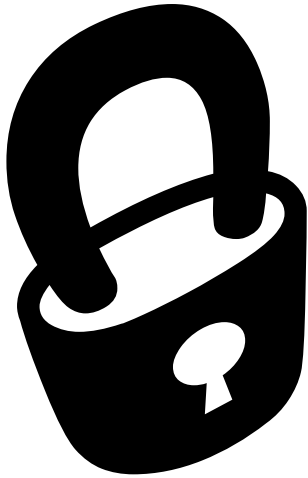
Wikidata + Wiktionary

Lexicographical data for everyone

Lydia Pintscher
@nightrose

What's the problem?





Wiktionary is a locked treasure chest!

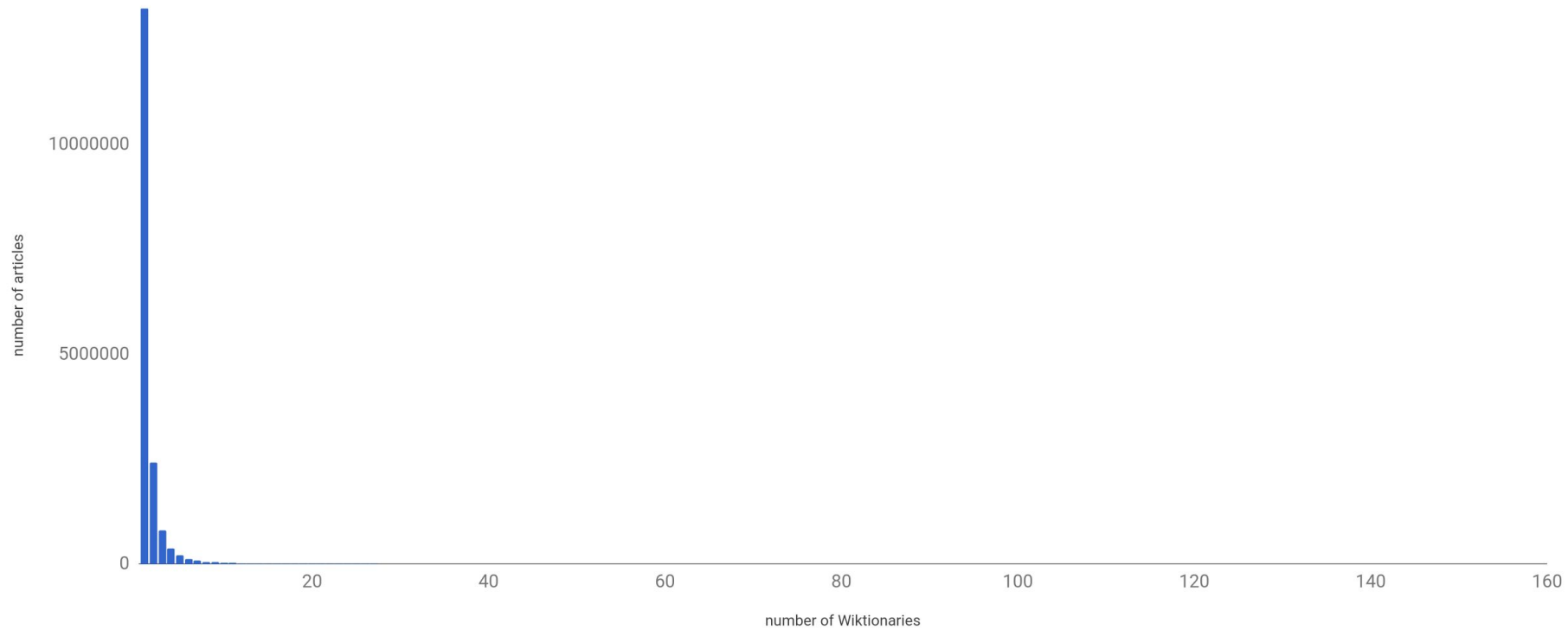
- Hard to understand the structures and templates
- Sharing workload between Wiktionaries is hard to non-existent
- The data is not machine-readable and therefore not re-usable

Page views per day:

Total:	2.6 M
English:	1.2 M
French:	336 000
Russian:	268 000
German:	163 000

Articles:

Total:	27.8 M
English:	5.2 M
French:	3.1 M
Russian:	838 000
German:	623 000



Active Wiktionarians:

Total: 1028
English: 403
French: 120
Armenian: 84
German: 59

Very active Wiktionarians:

Total: 302
English: 104
French: 35
German: 21
Armenian: 19

66 of 149

Wiktionaries have not been
active in the last 3 months

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

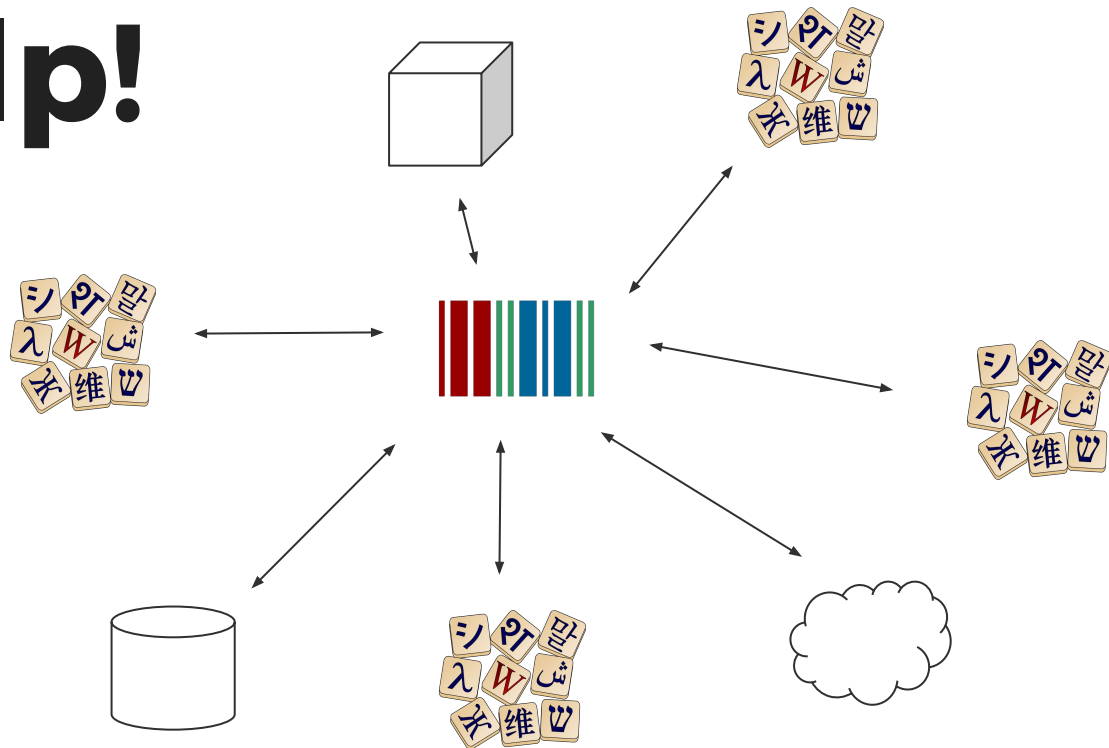
● Python

★ 4

Languages

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

Wikidata can help!



**What can we do once
we have it?**



Workload sharing and new ways to contribute to Wiktionary

- Working together on the same data (if wanted!)
- New tools to make contributing easier and open it up to new contributor groups

Potential users: Wiktionary,
Wikidata Game

Dictionary applications and more

- Looking up definitions and translations
- Special purpose dictionaries (rhyme, specific topics)
- Thesauri and synonym dictionaries

Potential users: dict.leo.org

Language learning tools

- Creating word lists and lessons
- Illustrating words

Potential users: Parley, Duolingo

Research

- How do languages evolve over time, social class and more?
- Do classes of words change their meaning over time?

Potential users: The Rosetta Project

Text analysis

- Sentiment analysis
- Part of speech tagging
- Named entity
recognition

Potential users: TextRazor

Automated translation

- Build translation tools (especially for underserved languages that don't have any yet)

Potential users: Apertium

Demo

