

Wiktionary Popularity from a Citation Analysis Point of View

Alireza Noruzi 

Editor-in-Chief, Associate Professor, Informology Center, Marseille, France. E-mail: anoruzi@gmail.com

Article Info

Article type:

Editorial note

Keywords:

Wiktionary,
Citation analysis,
Bibliometric

ABSTRACT

Objective: *Wiktionary* is a collaborative web-based project to produce a free-content multilingual dictionary of terms in all natural languages and in a number of artificial languages. This study aims to provide an overview of the citation rate of *Wiktionary*.

Materials and Methods: The primary source of data utilized in this study was the Scopus database. A REFERENCE search was conducted for indexed citations in the Scopus citation index, to find citations to *Wiktionary* in June 2023. Bibliometrix was used to design the keyword co-occurrence network of author-supplied keywords of documents citing *Wiktionary*.

Results: This study determines to what extent the *Wiktionary* is used and cited by papers indexed in Scopus. The total number of citations to *Wiktionary* from 2006 was 1,766 of which the highest number of citations is 161 in the year 2017 and the lowest number of citations is five in the year 2006. *Wiktionary* is highly cited by the subject areas of computer science, social sciences, and arts and humanities. The analysis of the language distribution of citations to *Wiktionary* indicates that the authors of citing papers used *Wiktionary* in different languages. However, the English language was the most dominant language of citing documents with 1,642 citations (i.e., 93%).

Conclusion: *Wiktionary* was cited 1,766 times in Scopus by different languages (especially English, German, and French) in different countries (especially the U.S. with 335 citations, Germany with 295 citations, and France with 122 citations) mainly by the subject areas of computer science, social sciences, and arts and humanities. The significance of *Wiktionary* from a citation analysis point of view goes well beyond open access and enhanced opportunities for citation in linguistics, natural language processing systems, computational linguistics, semantics, and ontology.

Cite this article: Noruzi, A. (2023). Wiktionary popularity from a citation analysis point of view. *Informology*, 2(1), 1-6.



© The Author.

Publisher: Informology Center.

Introduction

Wiktionary is a collaborative web-based project to produce a free-content multilingual dictionary of terms in all natural languages and in a number of artificial languages. *Wiktionary* aims to describe all words of all languages using definitions and descriptions. It is run by the Wikimedia Foundation, a non-profit

foundation registered in Florida, USA. *Wiktionary* was launched on December 12, 2002, using wiki technology to create a collaborative online dictionary.

It is worth noting that since its inception *Wiktionary*, the free international multi-lingual dictionary, has been used for many purposes worldwide, such as natural language processing, information extraction and retrieval, lemmatization, stemming, antithesis detection, context labels extraction, etymology, hypernym extraction, idiom identification, lexical disambiguation in multiple languages, lexical innovations, lexical ontology, lexical semantic knowledge extraction, lexicography, machine-readable structured data, morphological and morphosemantic knowledge, morphological lexicon extraction, morphosyntactic lexicon, multilingual data extraction, multilingual syllabification and stress prediction, named entity disambiguation, ontology construction and matching, part-of-speech tagging, phonology, pronunciation extraction, pronunciation lexicon, quotation corpus, semantic relatedness of concepts or words, synonymy networks detection, terminology extraction, text mining, thesaurus construction, translation data extraction, verb similarity, word sense disambiguation, WordNet building, etc.

This study aims to provide an overview of the citation rate of *Wiktionary*. One way to test the popularity of a dictionary is to use the citation analysis method to determine to what extent it is cited by scientific articles (Noruzi, 2009). This study tries to identify and quantify the number of citations to *Wiktionary* in the articles indexed by the Scopus citation index, studying the citation behavior of authors who have published papers in journals indexed by this database.

Materials and Methods

A REFERENCE search was conducted for indexed citations in the Scopus citation index, to find citations to *Wiktionary* in June 2023. Therefore, the search was performed within the REFERENCES field as follows:

REF (*Wiktionary*)

The keyword co-occurrence network of author-supplied keywords of documents citing *Wiktionary* was designed by Bibliometrix (2023).

Results

The total number of citations to *Wiktionary* in Scopus was 1,766. These document results were analyzed. Table 1 reveals how many *Wiktionary* was cited each year. The pattern shows the increasing or decreasing trends of citations to *Wiktionary* as used per annum. The yearly distribution of citations gives an idea about the scattering of citations. The total number of citations to *Wiktionary* from 2006 was 1,766 of which the highest number of citations is 161 in the year 2017 and the lowest number of citations is five in the year 2006.

Table 1. The total number of citations to Wiktionary per year

Publication Year	No. of citations	Publication Year	No. of citations
2023	41	2014	120
2022	145	2013	89
2021	138	2012	117
2020	158	2011	78
2019	148	2010	73
2018	139	2009	59
2017	161	2008	24
2016	142	2007	8
2015	121	2006	5

The subject areas citing *Wiktionary* are shown (by record count) in Table 2. Table 2 shows that *Wiktionary* is highly cited in the subject areas of computer science, social sciences, and arts and humanities.

Table 2. Subject areas frequently citing Wiktionary

Subject Areas	No. of citations	% out of 1,766
Computer Science	958	54.2
Social Sciences	656	37.1
Arts and Humanities	557	31.5
Engineering	229	13.0
Mathematics	206	11.7
Medicine	113	6.4
Decision Sciences	63	3.6
Business, Management and Accounting	60	3.4
Psychology	43	2.4
Physics and Astronomy	27	1.5
Environmental Science	25	1.4
Materials Science	24	1.4
Biochemistry, Genetics and Molecular Biology	18	1.0
Economics, Econometrics and Finance	18	1.0
Agricultural and Biological Sciences	17	1.0
Energy	17	1.0
Neuroscience	16	0.9
Health Professions	15	0.8
Earth and Planetary Sciences	14	0.8
Multidisciplinary	14	0.8
Chemistry	13	0.7
Chemical Engineering	11	0.6
Nursing	9	0.5
Immunology and Microbiology	4	0.2
Pharmacology, Toxicology and Pharmaceutics	4	0.2
Veterinary	3	0.2
Dentistry	1	0.1

Table 3 shows the distribution of citations to *Wiktionary* by country of publication with at least 30 citations.

Table 3. Countries cited Wiktionary

Country	No. of citations	% out of 1,766
United States	335	19.0
Germany	295	16.7
France	122	6.9
China	99	5.6
United Kingdom	97	5.5
Italy	87	4.9
Russian Federation	80	4.5
Spain	69	3.9
India	69	3.9
Canada	60	3.4
Undefined	59	3.3
Poland	56	3.2
Australia	47	2.7
Ireland	37	2.1
Japan	36	2.0
Netherlands	35	2.0
Denmark	32	1.8
Austria	31	1.8
Switzerland	30	1.7
Sweden	30	1.7

Table 4 shows the language distribution of citations to *Wiktionary*. The authors of citing papers used *Wiktionary* in different languages. The English language was found to be the most dominant language citing *Wiktionary* with 1,642 citations.

Table 4. Languages cited Wiktionary

Language	No. of citations	% out of 1,766
English	1,642	93.0
German	33	1.9
French	33	1.9
Spanish	15	0.8
Russian	13	0.7
Polish	9	0.5
Portuguese	9	0.5
Czech	4	0.2
Chinese	3	0.2
Croatian	3	0.2
Italian	3	0.2

Conclusion

The result showed that *Wiktionary* was cited 1,766 times in Scopus by different languages (especially English, German, and French) in different countries (especially the United States with 335 citations, Germany with 295 citations, and France with 122 citations), especially by the subject areas of computer science, social sciences, and arts and humanities. The authors of citing papers used *Wiktionary* in different languages. However, the English language was the most dominant language of citing documents with 1,642 citations (i.e., 93%). The significance of *Wiktionary* from a citation analysis point of view goes well beyond open access and enhanced opportunities for citation in linguistics, natural language processing systems, computational linguistics, semantics, ontology, and machine learning.

Data Availability Statement

Not applicable.

Ethical considerations

The author avoided from data fabrication and falsification.

Funding

Not applicable.

References

- Bibliometrix (2023). An R-tool for comprehensive science mapping analysis. K-Synth Srl, Academic Spin-Off of the University of Naples Federico II. <https://www.bibliometrix.org/home/>
- Noruzi, A. (2009). Editorial: Wikipedia popularity from a citation analysis point of view. *Webology*, 6(2), editorial 20, 1-3. <http://www.webology.org/2009/v6n2/editorial20.html>
- Perianes-Rodriguez, A., Waltman, L., and van Eck, N. J. (2016). Constructing bibliometric networks: A comparison between full and fractional counting. *Journal of Informetrics*, 10(4), 1178-1195. <https://doi.org/10.1016/j.joi.2016.10.006>
- Wiktionary (2023). Wiktionary, the free dictionary. https://en.wiktionary.org/wiki/Wiktionary:Main_Page