Ten years of Wikidata: A bibliometric study

Houcemeddine Turki[1,*], Mohamed Ali Hadj Taieb[1], Mohamed Ben Aouicha[1], Lane Rasberry[2] and Daniel Mietchen[3,4]

[1]Data Engineering and Semantics Research Unit, Faculty of Sciences of Sfax, University of Sfax, Sfax, Tunisia
[2]School of Data Science, University of Virginia, Charlottesville, VA, United States of America
[3]Ronin Institute for Independent Scholarship, Montclair, New Jersey, United States of America
[4]Leibniz Institute of Freshwater Ecology and Inland Fisheries, Berlin, Germany

Abstract
In this research paper, we analyzed the scientific research dealing with Wikidata from its creation in 2012 until late 2022, as revealed by Scopus. We identified 945 relevant scholarly publications, mostly at conferences. This landscape is characterized by small groups of experts and Wikidata contributors from the Global North. The same applies to the funders of Wikidata research, which are mainly governmental institutions from the Global North. Further networking and outreach should be done for better diversity and inclusion inside the Wikidata research community. The analysis also finds an emphasis on research around computer science perspectives on the development of Wikidata. Most outputs are mainly focused on developing methods for the creation, enrichment, reuse, and evaluation of open knowledge graphs, particularly Wikidata. However, there is also a significant but narrower interest in application-oriented research about the use of Wikidata in digital humanities, biology, and healthcare.

Keywords
Wikidata, Open Knowledge Graphs, Bibliometrics, Wikimedia Research

1. Introduction

Wikidata's influence in media and research ecosystems prompts questions and curiosity about the nature and extent of its impact, reach, and user community. Many commentators have presented their reactions to Wikipedia's 2001 establishment and ongoing development [1]. Wikidata's 2012 establishment included sharing website infrastructure and exchanging information with Wikipedia [2, 3]. Soon after starting the flow of data between Wikipedia and Wikidata, the next endeavor was connecting Wikidata with other structured data collections. The intent was to promote knowledge transfer back and forth among Wikipedia, Wikidata, and the Linked Open Data ecosystem, thus positioning the Wikimedia platform as a global hub for importing and exporting general reference information [4, 5].

Wikidata'23: Wikidata Workshop at ISWC 2023
*Corresponding author.
turkiabdelwaheb@hotmail.fr (H. Turki); mohamedali.hajtaieb@fss.usf.tn (M. A. Hadj Taieb); mohamed.benaouicha@fss.usf.tn (M. Ben Aouicha); lr2ua@virginia.edu (L. Rasberry); daniel.mietchen@ronininstitute.org (D. Mietchen)
0000-0003-3492-2014 (H. Turki); 0000-0002-2786-8913 (M. A. Hadj Taieb); 0000-0002-2277-5814 (M. Ben Aouicha); 0000-0002-9485-6146 (L. Rasberry); 0000-0001-9488-1870 (D. Mietchen)

Commentaries review Wikidata from various perspectives, including general review of the literature [6], topical coverage [7], quality of content [8], use in research libraries [9], and use in digital humanities [10]. Although a 2014 paper took for granted that ”The relevance of Wikidata for researchers in semantic technologies, linked open data, and Web science thus hardly needs to be argued for” [4], these reports establish at a high level that researchers are using Wikidata. However, understanding the extent and nature of researchers' use of Wikidata remains an open question. To address this gap, our study aims to explore the patterns and trends in scientific research related to Wikidata for ten years (2012-2022) through an analysis based on *Scopus*, one of the largest controlled bibliographic databases [11]. To achieve our research objectives, we are focused on answering key questions. This includes analyzing usage patterns to understand how frequently researchers integrate Wikidata into their work and whether there is a noticeable upward trend over time. We are also investigating the primary publication outlets for Wikidata-related research, identifying if specific academic journals or conferences dominate in this regard. Furthermore, our research involves examining researcher demographics, aiming to uncover who engages with Wikidata, including their nationalities and affiliations. Lastly, we are exploring acknowledgments to determine the sponsors or funding sources that researchers credit for supporting their Wikidata-related research endeavors.

Through addressing these questions, our study aims to offer a comprehensive overview of the scholarly landscape surrounding Wikidata. We acknowledge the importance of understanding Wikidata's utilization and impact in research for assessing its role in advancing knowledge and for identifying potential biases or areas for improvement in the Knowledge Graph. Our research strives to provide valuable insights into how Wikidata is employed in academia, serving as a foundation for further discussions and potential enhancements of this collaborative platform.

2. Method

We extract bibliographic metadata of research publications related to Wikidata between 2012 and 2022 as indexed by *Scopus*[1]. This is mainly done despite the better coverage of several other bibliographic databases like *OpenAlex*, particularly because *OpenAlex* and other automatically generated databases include several problems in author disambiguation and data formatting by contrast to controlled bibliographic databases like *Scopus* that have consistent data modeling [12]. Also, controlled bibliographic databases verify that included scholarly venues meet research integrity standards by contrast to automatically generated databases that include all kinds of scholarly publications without proper validation [13]. It is true that controlled bibliographic databases involve significant biases and can provide a distorted image of the scholarly production on Wikidata [14]. However, this is the opportunity cost to pay for having a quality bibliometric study with minor efforts and a limited allocated time.

The query extracts all the publications issued before 2023 that mention Wikidata in their title, abstract, or keywords. These publications are verified by hand to eliminate irrelevant papers. Extracted data involves *author names*, *author and index keywords*, *source titles*, *years of publication*, *affiliations*, *funding information*, *document types*, *conference names*, and *subject areas.* We use this information to generate statistical data about the characteristics of the considered

[1]Query: `(TITLE-ABS-KEY(Wikidata) AND PUBYEAR > 2011 AND PUBYEAR < 2023)`.

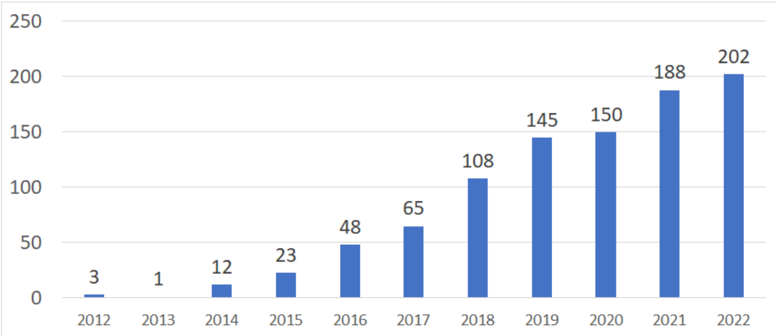

Figure 1: Distribution of Wikidata-related research publications per year, according to Scopus.

publications using the Scopus user interface. Then, we generate the co-occurrence network of the most common keywords and the co-authorship network of the most productive countries in the dataset using *VOSViewer*, software for generating bibliographic networks [15]. We assign weights to nodes based on their total link strength and we choose the colors of the nodes based on their recency in the networks (so-called *Overlay representation*) where recency ranges between -1 (oldest) and 1 (newest).

3. Results and Discussion

As of July 1, 2023, we identified 945 research publications related to Wikidata and indexed by *Scopus*. These papers have been issued since the early days of Wikidata in 2012, and their yearly productivity has linearly grown since 2014, reaching 188 publications in 2021 and 202 publications in 2022, as shown in Figure 1. This goes in line with the linear growth of the yearly production of open knowledge graph research between 2013 and 2022 [16]. This growth is less dynamic than that for knowledge graph research more generally, which is growing exponentially every year [17]. More emphasis on open knowledge graph research should be done to achieve yearly exponential growth, a characteristic of active and trendy topics in research [16].

When considering the types of Wikidata-related scholarly publications, we found that 78.1% (738 out of 945) of them are proceedings papers, mostly showing original research and a limited number of conference reviews, as shown in Figure 2. This finding seems to be applicable to computer science research, in general, [18], particularly for knowledge graph-related topics such as *graph neural networks* [19]. However, it does not seem to be valid for computer science research in developing regions like Africa [20]. The dependency of the community on conference papers rather than journal articles is motivated by the possibility of sharing works under development and implementation during conferences, allowing the community to coordinate their efforts regarding the ongoing development of Wikidata [18, 19]. That being said, the Wikidata research community did not try shorter and easier types of journal publications like *letters to the editor* that can stimulate discussion around the issues related to the development of Wikidata as an open knowledge graph [21].

When seeing the target conferences for Wikidata-related research, we found that most of

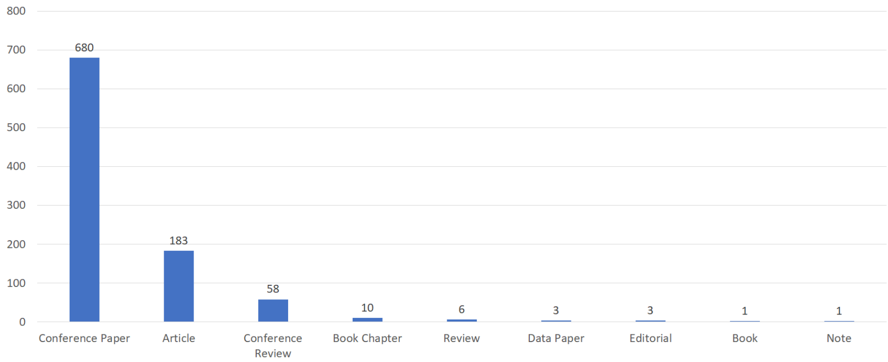

Figure 2: Distribution of Wikidata-related research publications per document type.

Table 1
Top conferences publishing Wikidata-related scholarly research

Conference	Abbreviation	CORE	Publications
International Semantic Web Conference	ISWC	A	68
International World Wide Web Conference	WWW	A*	44
Extended Semantic Web Conference	ESWC	A	42
Wikidata Workshop @ ISWC	Wikidata	N/A	41
Empirical Methods in Natural Language Processing	EMNLP	A	22
Semantic Web Challenge on Tabular Data to Knowledge Graph Matching @ ISWC	SemTab	N/A	17
International Symposium on Open Collaboration	OpenSym	C	17
Metadata and Semantics Research Conference	MTSR	N/A	14
Language Resources and Evaluation Conference	LREC	C	14
ACM International Conference on Information and Knowledge Management	CIKM	A	11

them are specific ones dealing with semantic web and knowledge engineering as shown in Table 1. According to the CORE 2021 Rankings, most of these research venues are classified as top-tier ones, having a rating of A (e.g., *ISWC* and *ESWC*) or A* (e.g., *WWW*) [22]. Despite this, Wikidata researchers have also published significant contributions to CORE C conferences, particularly *OpenSym* and *LREC*, as well as to several workshops at CORE A and A* conferences (e.g., *Wikidata* and *SemTab*), where their topics of interest are closely related to the development of the open knowledge graph. The ratings of these conferences have been the same as in previous editions of CORE since 2013. These events serve as a forum for Wikidata researchers to share and discuss their preliminary findings and source code [23].

As for the main journals that publish Wikidata-related research, we found that they are mainly

Table 2
Top journals publishing Wikimedia-related scholarly research

Journal	Publications	Publisher	SJR
Semantic Web	22	IOS Press BV, Netherlands	0.828
Jlis.it	13	Università di Firenze, Italy	0.208
Database	7	Oxford University Press, United Kingdom	1.786
Journal Of Web Semantics	7	Elsevier, Netherlands	0.955
Cataloging And Classification Quarterly	4	Routledge, United States of America	0.199
IEEE Access	4	IEEE, United States of America	0.926
AIB Studi	3	Associazione Italiana Biblioteche, Italy	0.229
Information Switzerland	3	MDPI, Switzerland	0.662
Information Systems	3	Elsevier Ltd., United Kingdom	0.976
International Journal Of Metadata Semantics And Ontologies	3	Inderscience Enterprises Ltd, United Kingdom	0.138
Journal Of Medical Internet Research	3	JMIR Publications Inc., Canada	1.992
Komp'juternaja Lingvistika i Intellektual'nye Tehnologii	3	Komp'juternaja Lingvistika i Intellektual'nye Tehnologii, Russian Federation	0.203
Nucleic Acids Research	3	Oxford University Press, United Kingdom	8.234
Peerj Computer Science	3	PeerJ Inc., United States of America	0.638

scholarly journals having a high citation impact (SJR > 0.8) [24] and dealing with semantic technologies, database management, and information processing as shown in Table 2. However, there are open-access mega-journals covering a wide range of research topics that also publish Wikidata-related research like *IEEE Access* and *PeerJ Computer Science*. Furthermore, several high-impact journals not related to computer science like *Nucleic Acids Research* can also be a target for describing large-scale multidisciplinary applications of Wikidata. Moreover, scientists also aim for national-level research journals that are closely related to database management like *Jlis.it*, *AIB Studi*, and *Komp'juternaja Lingvistika i Intellektual'nye Tehnologii* to show their preliminary results and to communicate the recent advances of Wikidata to local research communities in their mother tongues [25]. This situation seems to be similar to the one of research about knowledge graphs [17], specifically open knowledge graphs [16], except for the use of nationwide specialized journals as targets for Wikidata-related research.

Most conferences and journals publishing Wikidata-related research are predominantly in English, with 97.2% (919 out of 945) of Wikidata research publications being in English, reflecting the prevailing language bias in scholarly research [26]. Open-access publications account for only 32.3% (305 out of 945) of Wikidata-related scholarly research, primarily in the form of green open access (235 publications) and a limited number of gold open-access papers (64 publications), possibly due to challenges related to open-access publication fees. Figure 3

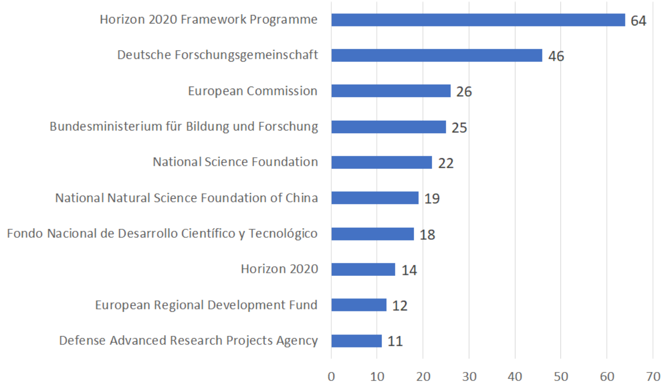

Figure 3: Distribution of Wikidata-related research publications per funding sponsor.

illustrates limited support from the Wikimedia Foundation for open-access scholarly publishing within its research community, with only 7 publications receiving such support. Major funders of Wikidata-related research are government-led funding agencies located in developed countries, including Germany (e.g., *Deutsche Forschungsgemeinschaft* and *Bundesministerium für Bildung und Forschung*), the United States of America (e.g., *National Science Foundation* and *Defense Advanced Research Projects Agency*), and countries with established traditions of semantic web research like China (e.g., *National Natural Science Foundation of China*) and Chile (e.g., *Fondo Nacional de Desarrollo Científico y Tecnológico*). Significant support also comes from the European Commission and its Horizon Europe program (formerly Horizon 2020), a continent-level research funding initiative. While these sponsors contribute significantly to Wikidata-related research, their coverage is limited to specific countries. In areas beyond their reach, restrictions on publishing updated research findings under permissive licenses often occur, especially when research funding is scarce [27]. Funding agencies in developing regions tend to focus more on capacity building and the latest computer science technology advancements rather than large-scale knowledge engineering projects, particularly those related to open knowledge [28].

This bias in funding did not only affect open-access publishing but also research productivity in underserved countries [29]. This is clearly revealed by the distribution of Wikidata-related research publications per country as shown in Figure 4. Wikidata-related research is mainly dominated by developed countries, mostly from Europe and North America, led by *Germany* (239 publications) and the *United States of America* (175 publications). BRICS nations (i.e., *Brazil*, *Russian Federation*, *India*, and *China*) are also among the most productive nations of Wikidata-related scholarly research, ranked between 6th and 16th. A surprising fact is the standings of *Chile* among the top countries in Wikidata-related research with 29 publications (Ranked 11th). Despite its good standings in the Human Development Index[2], Chile is ranked 52nd in computer science research productivity and 43rd in computer science citations [24]. Although knowledge graph research is mainly dominated by developed countries, the standings of the top countries in publishing knowledge graph research are quite different from the ones of the same countries in publishing Wikidata-related scholarly research [17, 16]. *China* and the *United*

[2]https://hdr.undp.org/data-center/country-insights#/ranks.

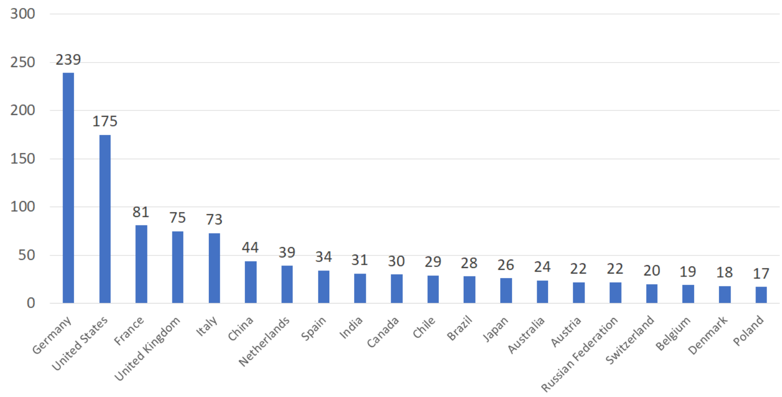

Figure 4: Distribution of Wikidata-related research publications per country.

States of America are the two main countries that dominate knowledge graph research [17], particularly in the context of open knowledge graphs [16].

Germany's prominence in Wikidata-related research is partly attributed to its close association with Wikidata maintenance, overseen by Wikimedia Deutschland, the German chapter of the Wikimedia Foundation [30]. The collaborative landscape of Wikidata research (Figure 5) reflects its origins in partnerships between *Germany* and several European nations, including *Italy*, *Denmark*, and *Belgium*, depicted in Purple. Non-European developed countries such as the *United States*, *Canada*, and *Australia* entered the field later through extensive collaborations with *Germany* and the *United Kingdom*, as indicated in Green. In recent years, BRICS countries like *China*, *Brazil*, and *India*, Eastern European nations like *Poland*, East Asian countries like *South Korea*, and developing nations such as *Tunisia* and *Malaysia* have also integrated into Wikidata-related research, as shown in Yellow. This hierarchical collaboration network differs from the general knowledge graph research landscape, where the collaboration between the United States and China holds significant influence [17].

Germany's dominant position is further evident in the list of top contributing institutions, with seven of the 18 most productive institutions hailing from Germany (Figure 6). These institutions, all universities and research institutes, showcase Germany's robust presence in Wikidata-related scholarly research. Other countries on the list typically have one representative institution, except for France, having significant publications from *Université de Lyon* and the *Centre National de Recherche Scientifique*, and the United States of America, featuring the *University of Southern California*, its *Information Sciences Institute*, and two prominent AI corporations, *Google* and *IBM*. The involvement of tech giants in Wikidata research may be attributed to the United States' orientation towards private sector-driven AI research [31]. The substantial university involvement in European Wikidata-related research aligns with the research policies of European countries, particularly Germany, which prioritize universities as key players in AI research within the Triple Helix model framework [31].

When seeing the list of the main authors of Wikidata-related research, we found that most of the universities that are among the most productive ones in Wikidata-related research are included thanks to the efforts of individual scientists as shown in Table 3. Based on an examina-

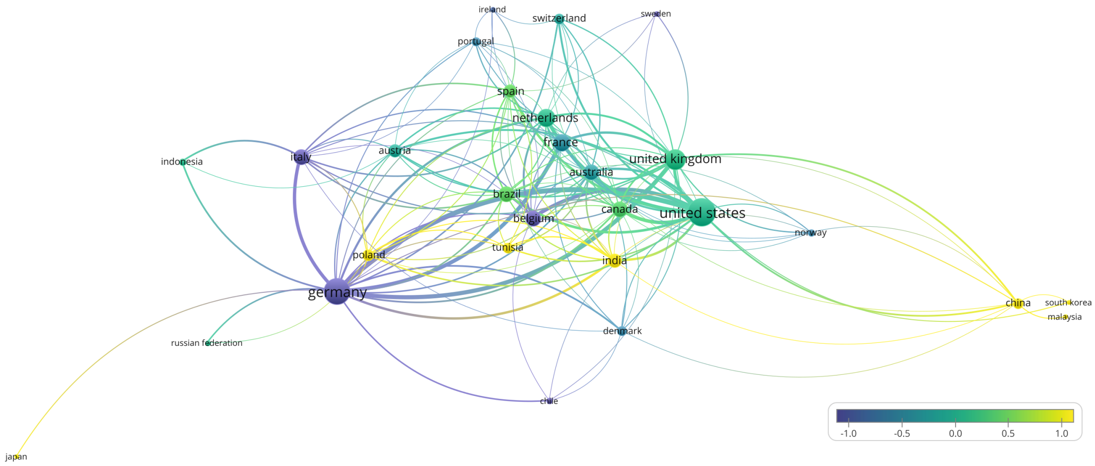

Figure 5: Overlay representation of the co-authorship network for the top countries by number of publications of Wikidata-related research. The oldest nodes are in purple. The newest ones are in yellow.

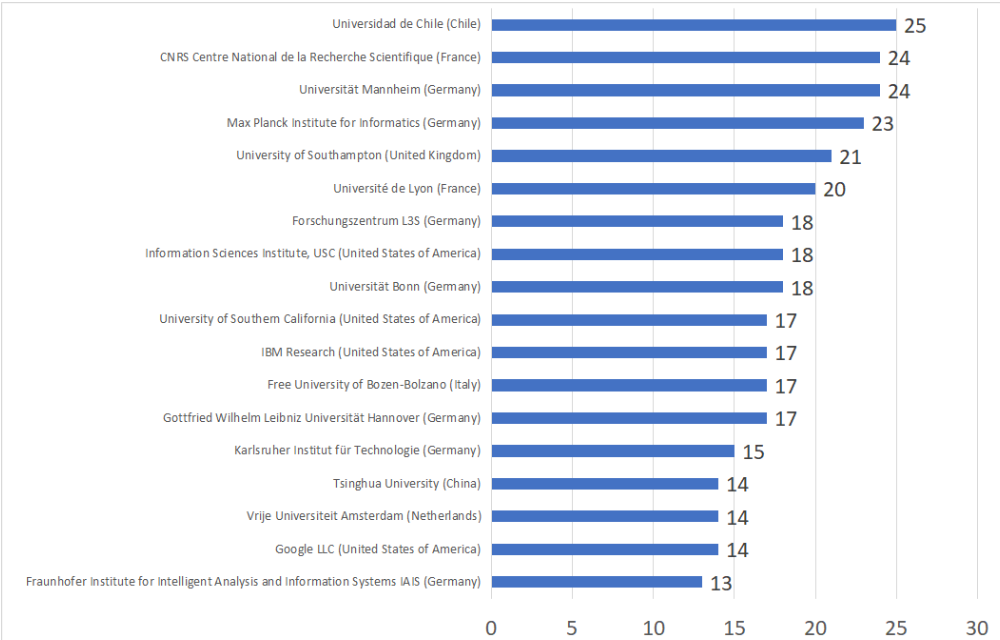

Figure 6: Distribution of Wikidata-related research publications per institution.

tion of Google Scholar profiles as of July 1, 2023, many of these scientists have a background in knowledge engineering and an h-index > 30, as imported from Google Scholar profiles [32]. For example, the status of *Universidad de Chile* and *Universität Mannheim* respectively as the first and third most productive institutions mainly occurred thanks to the contributions of *Aidan Hogan* and *Heiko Paulheim*. This information could be insightful for efforts to establish long-term Wikidata research communities and traditions inside research institutions. Sometimes, younger

Table 3
Top authors publishing Wikidata-related scholarly research

Author and Institution	Publications	H-Index
Hogan, Aidan (Universidad de Chile, Chile)	26	40
Razniewski, Simon (Bosch Center for AI, Germany)	21	18
Simperl, Elena (King's College London, United Kingdom)	20	42
Paulheim, Heiko (Universität Mannheim, Germany)	14	48
Szekely, Pedro (Information Science Institute, USC, United States of America)	14	40
Darari, Fariz (Universitas Indonesia, Indonesia)	13	9
Diefenbach, Dennis (Université de Lyon, France)	13	15
Kaffee, Lucie-Aimée (University of Copenhagen, Denmark)	13	11
Waagmeester, Andra (Micelio, Belgium)	13	26
Ilievski, Filip (Information Science Institute, USC, United States of America)	12	15
Schubotz, Moritz (University of Wuppertal, Germany)	12	21
Lehmann, Jens (Technische Universität Dresden, Germany)	11	63
Nutt, Werner (Free University of Bozen-Bolzano, Italy)	11	39
Gipp, Bela (University of Göttingen, Germany)	9	41
Mihindukulasooriya, Nandana (IBM Research, United States of America)	9	17
Su, Andrew I. (Scripps Research Institute, United States of America)	9	67

scientists having an h-index < 30 could be behind the establishment of research traditions related to Wikidata development [32]. Successful examples are *Simon Razniewski*, *Lucie-Aimée Kaffee*, *Fariz Darari*, and *Dennis Diefenbach*. This proves that Ph.D. or post-doc works can catalyze the involvement of individuals and institutions in Wikidata-related research, and is supporting evidence that early-career scientists can establish Wikidata research communities and traditions in their scholarly institutions. Rarely, we can find several prolific authors of Wikidata research that are also involved in the Wikimedia Community as active members of the Wikimedia movement and as Wikidata, editors such as *Andra Waagmeester* and *Andrew I. Su*. Such individuals are key for the development of research at the intersection of Wikidata community priorities and semantic web challenges.

When dealing with the main scientific disciplines that shape Wikidata-related research, it is clear that most of the publications deal with Wikidata research from the perspective of *computer science* (812 publications), *mathematics* (229 publications), and *engineering* (54 publications), as shown in Figure 7. This is a common characteristic of knowledge graph research [17]. The development of generic algorithms for the enrichment and validation of knowledge graphs is a priority to ensure the accuracy and consistency of such resources [17]. Beyond this, a significant number of works have studied the practical applications of Wikidata in multiple research areas. Most of these applications are related to *Social Sciences* (148 publications) and *Arts and Humanities* (89 publications). Such works mainly aim for turning Wikidata into a large-scale cultural heritage and archival database and using it to enhance free access to

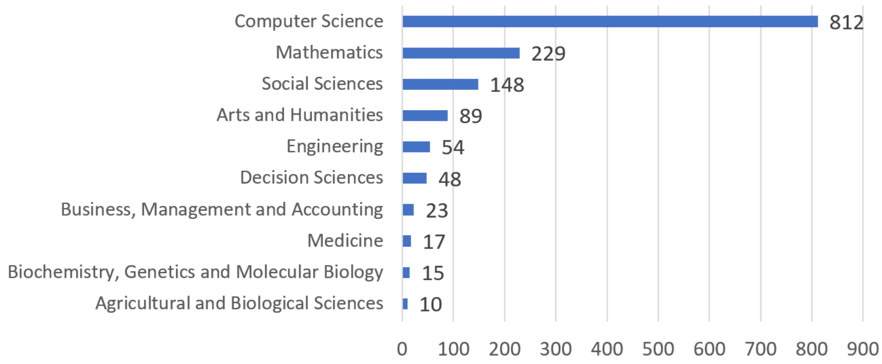

Figure 7: Distribution of Wikidata-related research publications per subject area.

digital humanities data for research and development purposes [33, 34]. Considerable effort has been provided to develop Wikidata applications for *Decision Sciences* (48 publications) and *Business, Management, and Accounting* (23 publications). This is mostly linked to the study and adjustment of user behavior and data governance when contributors are collaboratively editing Wikidata [7]. Finally, there is a good amount of research work about using Wikidata in *Medicine* (17 publications), *Biochemistry, Genetics, and Molecular Biology* (15 publications), and *Agricultural and Biological Sciences* (10 publications). These works integrate biological and medical knowledge into Wikidata and use them to develop systems to inform decisions in medicine and biology such as dashboards and user interfaces [35].

The top featured keywords for the considered publications highlight the role of computer science in Wikidata. Although the first four keywords represent a general terminology of the topic as shown in Figure 8 (i.e., *Knowledge Graphs*, *Wikidata*, *Semantic Web*, and *Knowledge Graph*), most of the top keywords show the main applications of Wikidata in computer science: *Knowledge-Based Systems* (144 publications), *Natural Language Processing Systems* (123 publications), *Knowledge Representation* (81 publications), *Computational Linguistics* (71 publications), *Data Mining* (69 publications), *Question Answering* (69 publications), *Information Retrieval* (47 publications), and *Data Handling* (46 publications). These applications are confirmed by previous studies on knowledge graph research [17], particularly the ones related to open knowledge graphs [16]. Subsequently, the top keywords also reveal the main secondary resources that are used to enrich and maintain Wikidata and comparatively evaluate its evolution: *Wikipedia* (129 publications), *Linked Data* (96 publications), *Ontology* (78 publications), *DBpedia* (70 publications), and *Open Data* (64 publications). Research trends on open knowledge graphs identify *Wikidata* and *DBpedia* as the most studied open knowledge graphs that are comparatively evaluated [16]. The same trends reveal the development of methods to use *Wikipedia*, *Ontologies*, and *Linked Open Data* for the construction of open knowledge graphs, particularly between 2013 and 2016 [16].

When examining the topics covered in Wikidata-related scholarly research through the co-occurrence network of common keywords shown in Figure 9, we identified three stages in the evolution of Wikidata research: The *Creation stage* (2012-2015, Purple) mainly focused on demonstrating how Wikidata can serve as a collaborative knowledge graph to handle large-scale

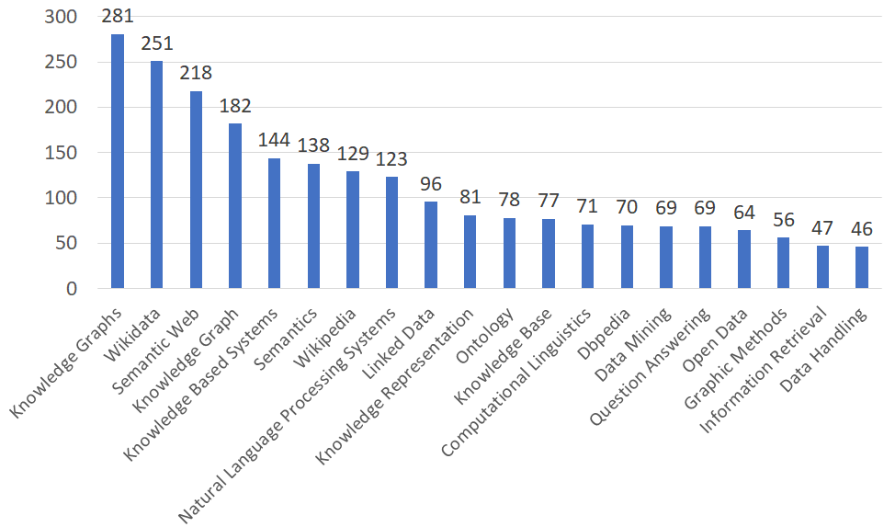

Figure 8: Top keywords of the Wikidata-related scholarly research publications.

data. The *Enrichment and Application stage* (2016-2018, Blue-Green) aimed to enrich and assess Wikidata using online sources like Linked Open Data, ontologies, and DBpedia. It also explored applications of Wikidata in natural language processing, data classification, search engines, knowledge management, and named entity recognition. Additionally, it analyzed the feasibility of developing techniques for extracting semantic relations from text and integrating these retrieved statements into Wikidata through entity linking. The *Machine Learning and AI stage* (since 2019, Yellow) built upon previous research by focusing on the use of machine learning, embeddings, and language models to enhance Wikidata. Researchers worked on advanced applications of Wikidata, such as question answering, and aimed to predict missing statements within Wikidata. These phases of Wikidata's development coincided with similar stages in the evolution of knowledge graph research over the past decade [17, 16]. Early knowledge graph research primarily emphasized the development of semantic web standards and information retrieval techniques, particularly in the context of applications in computational biology. Over time, knowledge graph research expanded to support various disciplines, including natural language processing and cultural heritage. This expansion also involved the adoption of query-based methods like SPARQL and the incorporation of machine learning techniques, particularly *embeddings* and *graph learning* [17, 16].

4. Conclusion

In this research paper, we identified 945 scholarly publications about Wikidata from 2012 to 2022, as indexed by *Scopus*. These works are mainly developed thanks to the personal initiatives of several highly-cited scientists, Wikidata contributors, and early-career researchers working in developed countries receiving funds from governmental institutions in Europe and North America, rather than through the Wikimedia Foundation and other non-profit

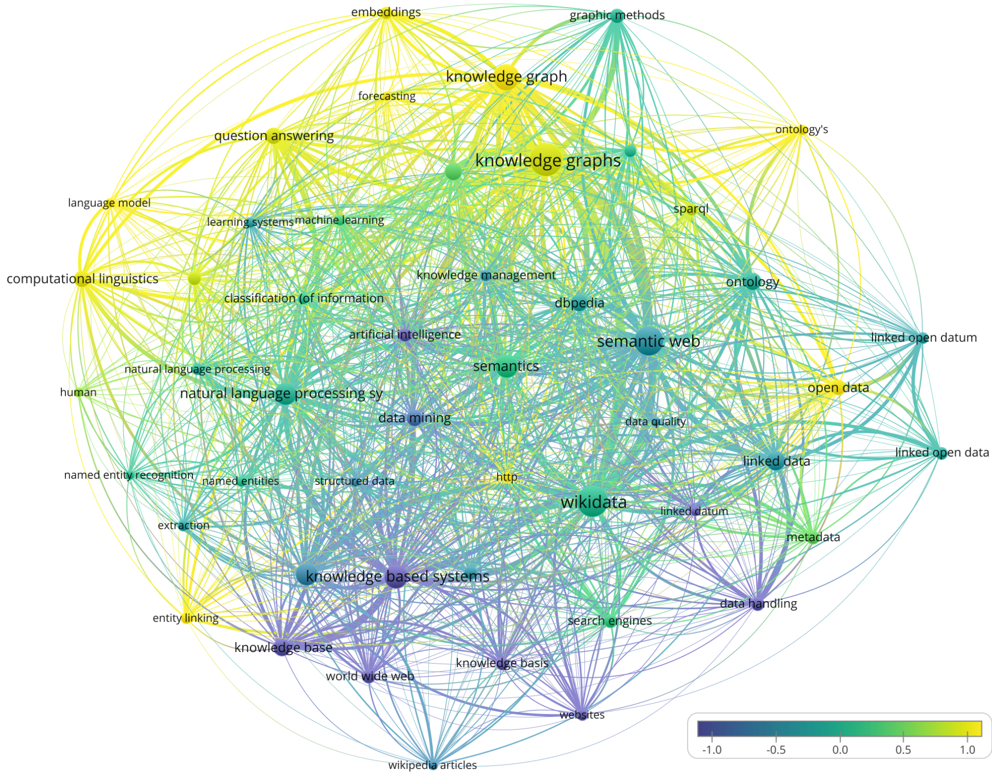

Figure 9: Overlay representation of the co-occurrence network for the most common keywords of Wikidata-related research. The oldest nodes are in purple. The newest ones are in yellow.

organizations. Many of these works emphasize the computer science perspective of the creation and sustainability of Wikidata and to a lesser extent the application-oriented research about the use of Wikidata in several fields like Molecular Biology, Digital Humanities, and Medicine. The computer science-related works about Wikidata are keeping pace with the latest advances in knowledge graph research (e.g., *entity linking*) and artificial intelligence (e.g., *embeddings* and *language models*). In terms of limitations, this paper relied on a single database (Scopus) to identify Wikidata-related research and did not explore other such databases, including Wikidata itself. The paper did not analyze the dynamics behind research collaborations around Wikidata either. As future directions of this research work, we intend to expand it by analyzing other databases like *Dimensions* and by studying the correlation between the different characteristics of the respective publications, including ones not studied in the present work, such as *affiliations*.

Acknowledgments

This research is funded by the Wikimedia Research Fund of the Wikimedia Foundation (San Francisco, California, United States of America) through the *Adapting Wikidata to support clinical practice using Data Science, Semantic Web and Machine Learning* Project. Source data is made available upon request.

References

[1] M. M. Mostafa, Twenty years of wikipedia in scholarly publications: a bibliometric network analysis of the thematic and citation landscape, Quality & Quantity (2023). URL: https://doi.org/10.1007/s11135-023-01626-7. doi:10.1007/s11135-023-01626-7.

[2] D. Vrandečić, Wikidata: a new platform for collaborative data collection, in: Proceedings of the 21st International Conference on World Wide Web, ACM, Lyon France, 2012, pp. 1063–1064. URL: https://dl.acm.org/doi/10.1145/2187980.2188242. doi:10.1145/2187980.2188242.

[3] D. Vrandečić, M. Krötzsch, Wikidata: a free collaborative knowledgebase, Communications of the ACM 57 (2014) 78–85. URL: https://dl.acm.org/doi/10.1145/2629489. doi:10.1145/2629489.

[4] F. Erxleben, M. Günther, M. Krötzsch, J. Mendez, D. Vrandečić, Introducing Wikidata to the Linked Data Web, in: P. Mika, T. Tudorache, A. Bernstein, C. Welty, C. Knoblock, D. Vrandečić, P. Groth, N. Noy, K. Janowicz, C. Goble (Eds.), The Semantic Web – ISWC 2014, volume 8796, Springer International Publishing, Cham, 2014, pp. 50–65. URL: http://link.springer.com/10.1007/978-3-319-11964-9_4. doi:10.1007/978-3-319-11964-9_4, series Title: Lecture Notes in Computer Science.

[5] D. Kinzler, L. Pintscher, Wikidata: How We Brought Structured Data to Wikipedia, in: Proceedings of The International Symposium on Open Collaboration, ACM, Berlin Germany, 2014, pp. 1–1. URL: https://dl.acm.org/doi/10.1145/2641580.2641583. doi:10.1145/2641580.2641583.

[6] M. Mora-Cantallops, S. Sánchez-Alonso, E. García-Barriocanal, A systematic literature review on Wikidata, Data Technologies and Applications 53 (2019) 250–268. URL: https://www.emerald.com/insight/content/doi/10.1108/DTA-12-2018-0110/full/html. doi:10.1108/DTA-12-2018-0110.

[7] M. Farda-Sarbas, C. Müller-Birn, Wikidata from a Research Perspective – A Systematic Mapping Study of Wikidata, 2019. arXiv:1908.11153.

[8] A. Piscopo, E. Simperl, What we talk about when we talk about wikidata quality: a literature survey, in: Proceedings of the 15th International Symposium on Open Collaboration, ACM, Skövde Sweden, 2019, pp. 1–11. URL: https://dl.acm.org/doi/10.1145/3306446.3340822. doi:10.1145/3306446.3340822.

[9] K. Tharani, Much more than a mere technology: A systematic review of Wikidata in libraries, The Journal of Academic Librarianship 47 (2021) 102326. URL: https://linkinghub.elsevier.com/retrieve/pii/S0099133321000173. doi:10.1016/j.acalib.2021.102326.

[10] F. Zhao, A systematic review of Wikidata in Digital Humanities projects, Digital Scholarship in the Humanities 38 (2023) 852–874. URL: https://academic.oup.com/dsh/article/38/2/852/6964525. doi:10.1093/llc/fqac083.

[11] J. Baas, M. Schotten, A. Plume, G. Côté, R. Karimi, Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies, Quantitative Science Studies 1 (2020) 377–386. doi:10.1162/qss_a_00019.

[12] L. S. Adriaanse, C. Rensleigh, Web of Science, Scopus and Google Scholar, The Electronic Library 31 (2013) 727–744. doi:10.1108/el-12-2011-0174.

[13] V. K. Singh, P. Singh, M. Karmakar, J. Leta, P. Mayr, The journal coverage of Web of Science,

Scopus and Dimensions: A comparative analysis, Scientometrics 126 (2021) 5113–5142. doi:10.1007/s11192-021-03948-5.

[14] S. Khanna, J. Ball, J. P. Alperin, J. Willinsky, Recalibrating the scope of scholarly publishing: A modest step in a vast decolonization process, Quantitative Science Studies 3 (2022) 912–930. doi:10.1162/qss_a_00228.

[15] N. J. van Eck, L. Waltman, Software survey: VOSviewer, a computer program for bibliometric mapping, Scientometrics 84 (2009) 523–538. doi:10.1007/s11192-009-0146-3.

[16] H. Turki, A. T. Owodunni, M. A. Hadj Taieb, R. F. Bile, M. Ben Aouicha, A Decade of Scholarly Research on Open Knowledge Graphs, 2023. arXiv:2306.13186.

[17] X. Chen, H. Xie, Z. Li, G. Cheng, Topic analysis and development in knowledge graph research: A bibliometric review on three decades, Neurocomputing 461 (2021) 497–515. doi:10.1016/j.neucom.2021.02.098.

[18] D. Fiala, G. Tutoky, Computer Science Papers in Web of Science: A Bibliometric Analysis, Publications 5 (2017) 23. doi:10.3390/publications5040023.

[19] A. Keramatfar, M. Rafiee, H. Amirkhani, Graph Neural Networks: A bibliometrics overview, Machine Learning with Applications 10 (2022) 100401. doi:10.1016/j.mlwa.2022.100401.

[20] M. Harsh, R. Bal, A. Weryha, J. Whatley, C. C. Onu, L. M. Negro, Mapping computer science research in Africa: using academic networking sites for assessing research activity, Scientometrics 126 (2020) 305–334. doi:10.1007/s11192-020-03727-8.

[21] H. Turki, M. A. Hadj Taieb, M. Ben Aouicha, The value of letters to the editor, Scientometrics 117 (2018) 1285–1287. URL: https://doi.org/10.1007/s11192-018-2906-4. doi:10.1007/s11192-018-2906-4.

[22] Computing Research and Education, CORE Rankings Portal, 2021. URL: https://www.core.edu.au/conference-portal.

[23] L. Kaffee, O. Tifrea-Marciuska, E. Simperl, D. Vrandecic, Preface: Wikidata workshop 2020, CEUR Workshop Proceedings 2773 (2020). URL: https://ceur-ws.org/Vol-2773/Preface_Wikidata_Workshop.pdf, 1st Wikidata Workshop, Wikidata 2020 ; Conference date: 02-11-2020 Through 06-11-2020.

[24] SCImago, Scimago Journal and Country Rank, 2023. URL: https://scimagojr.com/.

[25] H. Lrhoul, H. Turki, B. Hammouti, O. Benammar, Internationalization of the Moroccan Journal of Chemistry: A bibliometric study, Heliyon 9 (2023) e15857. doi:10.1016/j.heliyon.2023.e15857.

[26] P. Mongeon, A. Paul-Hus, The journal coverage of Web of Science and Scopus: a comparative analysis, Scientometrics 106 (2015) 213–228. doi:10.1007/s11192-015-1765-5.

[27] D. J. Solomon, B.-C. Björk, Publication fees in open access publishing: Sources of funding and factors influencing choice of journal, Journal of the American Society for Information Science and Technology 63 (2011) 98–107. URL: https://doi.org/10.1002/asi.21660. doi:10.1002/asi.21660.

[28] H. Turki, H. Sekkal, A. Pouris, F.-A. M. Ifeanyichukwu, C. Namayega, H. Lrhoul, M. A. Hadj Taieb, S. A. Adedayo, C. Fourie, C. B. Currin, M. N. Asiedu, A. L. Tonja, A. T. Owodunni, A. Dere, C. C. Emezue, S. H. Muhammad, M. M. Isa, M. Banat, M. Ben Aouicha, Machine learning for healthcare: A bibliometric study of contributions from africa, Preprints.org (2023). doi:10.20944/preprints202302.0010.v2.

[29] Y.-H. Lee, Determinants of research productivity in Korean Universities: the role of research funding, The Journal of Technology Transfer 46 (2020) 1462–1486. doi:10.1007/s10961-020-09817-2.

[30] D. Vrandečić, L. Pintscher, M. Krötzsch, Wikidata: The Making Of, in: Companion Proceedings of the ACM Web Conference 2023, ACM, 2023, pp. 615–624. doi:10.1145/3543873.3585579.

[31] M. G. Jacobides, S. Brusoni, F. Candelon, The Evolutionary Dynamics of the Artificial Intelligence Ecosystem, Strategy Science 6 (2021) 412–435. doi:10.1287/stsc.2021.0148.

[32] Google, Google Scholar, 2023. URL: https://scholar.google.ca/.

[33] E. Kapsalis, Wikidata: Recruiting the Crowd to Power Access to Digital Archives, Journal of Radio & Audio Media 26 (2019) 134–142. doi:10.1080/19376529.2019.1559520.

[34] F. Zhao, A systematic review of Wikidata in Digital Humanities projects, Digital Scholarship in the Humanities 38 (2022) 852–874. doi:10.1093/llc/fqac083.

[35] H. Turki, M. A. Hadj Taieb, T. Shafee, T. Lubiana, D. Jemielniak, M. Ben Aouicha, J. E. Labra Gayo, E. A. Youngstrom, M. Banat, D. Das, D. Mietchen, WikiProject COVID-19, Representing COVID-19 information in collaborative knowledge graphs: The case of Wikidata, Semantic Web 13 (2022) 233–264. doi:10.3233/sw-210444.