Foundations for Wikidata

Peter F. Patel-Schneider
Unaffiliated

Ege Atacan Doğan
Unaffiliated

Rosario Uceda-Sosa
IBM Research

Abstract

In this research we propose to investigate select aspects of Wikidata, with the aim of setting up the foundations for a more consistent and useful knowledge resource that can be leveraged by industry and academia alike. These aspects include the core properties of the Wikidata ontology, class order, the property hierarchy, and the interaction and propagation of property statements between classes and instances. We will determine how these aspects are used in Wikidata and determine the perceived meaning for them based on their usage and documentation. We will go on to produce revised meaning, both informal and formal, for them that can be used as foundations for a more cohesive Wikidata. This will make it easier to extend Wikidata as well as making Wikidata easier to use in other Wikimedia projects, for research, and in industrial applications. This change also has the potential to re-integrate more ”human” structures into AI, ML, and LLM structures, as Wikidata is not used enough in those contexts, mostly due to inconsistencies in it.

Introduction

Wikidata [22] is a large, multi-domain knowledge graph underlying parts of other Wikimedia projects. Not only that, but it has become a ”feeder” knowledge to other graphical information sources, like DBPedia [2] and Conceptnet [18]. Wikidata is often used in research and industrial applications [21, 20] as a referent knowledge graph and is generally considered to be reliable [16]. However, a small fraction of the factual information in Wikidata is incorrect which, in practice, produces quite a few errors when extracting knowledge, due to reasoning that propagates unwanted properties across the graph.

Another challenge in Wikidata is that some domains do not correctly use properties or classes from other domains of Wikidata. For example, it has been the case that information about amphibious aircraft in Wikidata incorrectly uses information relating to ships and boats. This lack of consistency has decidedly negative consequences for the use of Wikidata, particularly in industrial applications.

This internal lack of consistency also happens with the two core ontology properties, ‘instance of’ (P31) and ‘subclass of’ (P279), which are often mixed up, with one being used where the other should [24]. (Items and properties from Wikidata are given as their English label enclosed in single quotes, sometimes followed by their identifier.) For example, ‘paella’ is an ‘instance of’ ‘National dish’ and a ‘subclass of’ ‘Rice Dish’ making it both a subclass and an instance of ‘dish’. It goes against the meanings of these two core ontology properties that something is both a subclass and an instance of a (fixed-order) class like ‘dish’. As these core ontology properties are used throughout Wikidata, their misuse is a large problem for uses of Wikidata.

A related issue is that some aspects of Wikidata do not even have a described intended meaning.

This is sometimes the case with properties on classes. For example, the meaning of ‘Rokin’ (a street in Amsterdam) (physically) ‘connects with’ ‘Dam Square’ is fairly obvious. But it is less obvious, and not documented, what the uses of ‘connects with’ on classes, like ‘hand’ ‘connects with’ ‘forearm’, should mean. Obviously, classes do not physically connect with each other, as they are abstract entities, so this should mean that instances of hand connect with instances of forearm. But this usage is not distinguished from the previous usage, and produces a conflation. This produces problems both when using the information directly and when trying to use the information on a class to determine what kind of information should be put on instances of the class.

This state of affairs has come about largely because the information in Wikidata has been added by many volunteer editors. These editors generally add information according to documentation about Wikidata and the practices of their community, but there is not much that forces them to adhere to a common meaning for the parts of Wikidata that they use or even to provide much of a meaning for novel uses of existing parts of Wikidata. There are weak property constraints, but these can be overridden; there is peer pressure, but that can be ignored; and there is post-addition editing by other editors, but that has not kept up with additions. Plus, enforcing compliance with any of these is also not great, the constraints may be problematic, peer pressure can lead to toxic and unfriendly behaviour, and post-addition editing is not necessarily productive or sustainable.

As a result, use of Wikidata in applications, beyond single domains that are more cohesive and have fewer problems, has been limited and difficult. (Many large uses of Wikidata stick to well-known domains, such as medicine [19, 23].) Sometimes when going beyond single domains the information to be used is projected down into a simpler language, for example, using one ontology property for both ‘instance of’ and ‘subclass of’ [21]. This removes an important distinction and thus the result is less useful than it should be. Sometimes the information to be used is carefully chosen, for example, by picking instances of carefully selected classes lower in the Wikidata ontology instead of all the instances of a general class. (This approach was taken in work done by one of us at Nuance.) This requires considerable human effort and can easily result in missing useful information. Even AI-based industrial applications, where robustness and correctness are both required, can have problems when there are inconsistencies in how information is represented.

What is needed to enhance the utility of Wikidata, particularly for large-scale industrial applications, is a set of Wikidata-wide modelling principles that allow a more cohesive, better unified view of the information in Wikidata and enable easier use of large parts of Wikidata without human intervention. A more cohesive, better unified view of Wikidata is not only beneficial for use in applications. Wikidata editors can struggle when Wikidata has internal inconsistencies in how its information is represented as they often use the information already in Wikidata to guide how they enter new information. Bad information thus begets more bad information. This also applies when the Wikidata ontology is extended to new domains.

Wikidata is becoming a central hub of information, and organization of information, used by other information sources [17]. Internal inconsistencies and lack of modelling principles in Wikidata make integration with external sources more difficult.

What we propose to do in this work is to develop cohesive meanings and modelling principles for certain, important aspects of Wikidata, based on the information that is currently in Wikidata and documentation related to that information, that will point the way towards a full cohesive, integrated view of Wikidata. We do not believe that we can come up with a complete, cohesive, and unified meaning for Wikidata nor do we believe that we can come up with a full set of modelling principles for Wikidata, but we do believe that cohesive meanings and principles for important aspects of Wikidata can make Wikidata easier to edit and extend, make Wikidata a more useful hub for other information sources, and make Wikidata easier to use in and better for applications, particularly large-scale industrial applications.

We also propose to develop methods to effect these meanings throughout Wikidata and embody the principles in Wikidata, moving Wikidata towards a more useful representation resource. Once these principles and methods have been developed we want to promote their use in Wikidata.

Date: We propose to start this work on 1 July 2025 and end before 30 June 2026.

Related work

Wikidata is a knowledge graph and thus this work is related to previous work on formalizing knowledge and property graphs [7, 8] and formalizing Wikidata in property graphs [13]. This in turn depends on work in the general area of formalizing graphical knowledge representation like Description Logics [3]. The work proposed here is more about analyzing the content of Wikidata to find and produce foundations for intended meaning of it than providing content-agnostic foundations for Wikidata.

There has also been much work on formalizing various aspects of graphical knowledge representation, such as metaclasses [11]. The work proposed here will produce variants of this previous work that take into account the data in Wikidata.

Aspects of this proposed work are related to finding problems in Wikidata, because problems are only problems in light of some intended meaning, and fixing problems in Wikidata. There has been previous work in problems related to multi-level modelling in the Wikidata ontology [4], which depends on a good meaning for the core ontology properties. We have done work in this area, not only expanding the finding of problems from the original work, using stronger queries to find more of the same kind of problems and also using different data from Wikidata to find different kinds of problems in this area, but also fixing a few of them [12]. We have also done work on finding and fixing problems with disjointness in Wikidata [5], which depends on having a firm intended meaning for the disjointness information in Wikidata.

We have helped fix up other parts of the Wikidata ontology, including the part of the ontology related to ships (https://www.wikidata.org/wiki/Wikidata:WikiProject_Ships). This work built on previous work establishing class orders for classes related to ships.

There has been work on documenting problems in the Wikidata ontology [14, 1] and proposing some solutions [15]. We created a group to discuss these issues and propose and effect solutions [10], which was the genesis of the work proposed here.

The work proposed here goes beyond previous work by developing theories for aspects of Wikidata and devising modelling principles for these aspects and then proposing and effecting methods for improving Wikidata by incorporating these principles into Wikidata.

Methods

The methods we will use will likely vary somewhat between the different aspects of Wikidata that we investigate but in general we will use a combination of information-gathering and modelling as described below.

The first step will be to pick an aspect of Wikidata to investigate, starting with the ones mentioned above. Here are some aspects that we have identified as (potential) targets.

Ontology Core The core of the Wikidata ontology consists of the notion of classes and the two core ontology properties, 'instance of' and 'subclass of'.

Disjointness Disjointness between classes is an important part of many ontologies. Wikidata has facilities for specifying disjointness by stating that a class is a disjoint union of other classes.

Class Order A class is first-order if all its instances are not classes; second-order if all its instances are first-order; etc. Wikidata has methods for stating the order of a class but these are not supported very well. Class order is a powerful tool for finding problems in an ontology, such as classes that are instances of a first-order class.

Property Hierarchy Properties in Wikidata participate in a generalization hierarchy based on the property `subproperty of'.

Property Conflations Some properties have multiple distinct usages. This causes a sort of genericness that is counter-productive. Distinct phenomena in the real world are best represented by different properties.

Property Characteristics Wikidata can state characteristics of properties by making the property an instance of a class. Some of these characteristics are standard, such as transitivity (Px is transitive iff: A Px B and B Px C implies A Px C)and symmetricity (Px is symmetric iff: A Px B iff B Px A). Other characteristics are unusual, such as various characteristics related to how the property is used in Wikidata.

Class Statements Wikidata permits property statements not only on individuals but also on classes. An issue here is how or even whether a property statement on a class interacts with property statements on its instances.

Constraints Wikidata has a general notion of constraints. Wikidata constraints are weak in practice and only cover part of what constraints can be used for.

Hierarchies involving Other Relations Class and property hierarchies are among the most central, but there are more hierarchies in Wikidata, such as those involving 'has part(s)' and 'has cause(s)', that can be investigated.

Once an aspect has been selected the first step will be to gather information about that aspect. This will involve examining the documentation about that aspect that is present in the Wikidata wiki pages and interacting with Wikidata user groups that use the aspect. We expect to make considerable use here of previous work on problems in Wikidata such as the work on problems in the Wikidata ontology [14, 1].

This information-gathering step will also involve examining the actual use of the aspect in Wikidata, probably using a combination of manual examination of typical uses and querying to find patterns in usage and statistics of usage. We expect to produce usage information such as how much the aspect is used, likely by domain, for example how the aspect is used in the medical domain vs how it is used in the ship domain.

We will have access to our own copy of Wikidata on a powerful desktop computer that will allow us to employ queries that cannot be run on public services providing access to Wikidata.

The next step will be to develop a theory of the aspect. We expect to either build a formal meaning for the aspect, somewhat similar to the formal meaning for RDF [6], or a set of inference rules for the aspect, as in the inference rules for RDF [6], perhaps in combination. Predictions from this theory will be compared with actual usage in Wikidata to find where usage deviates from the theory, again likely by domain. Reasons for these deviations will be produced where possible, For example, class order deviations are concentrated in the biomolecular domain but the reason for this is unclear.

From this theory we will determine modelling principles that can be used in Wikidata to improve use and usability of the aspect. For example, a potential modelling principle is to create as many disjointness statements for high-level classes to detect individuals incorrectly placed in two classes that are disjoint in the real world.

Finally we will develop methods for implementing these principles in Wikidata, aligning actual use with expected use, which will likely involve not only additions to Wikidata but changes to the information already in Wikidata.

We will then publicize this work to the Wikidata community and request feedback.

It is likely that these steps will be somewhat iterative, as doing the modelling may expose the need to go back and gather more information or interaction with the Wikidata community may expose problems with our initial solutions.

Here are some thoughts as to what we might achieve in the target areas described above.

Ontology Core Although there has been much work on core ontology properties similar to 'instance of' and 'subclass of', there are a surprising amount of problems related to these two core ontology properties in Wikidata. The main work we will be doing here will be to come up with ways to develop modelling principles that can be used to improve the Wikidata ontology, probably in conjunction with some of the other aspects we will investigate.

Disjointness We have already investigated disjointness in Wikidata, concentrating on violations of disjointness [5]. We would probably propose new ways of specifying disjointness and principles for how and when to use disjointness.

Class Order There is some work needed to recast existing formalizations of class order in a way suitable for Wikidata. What is also needed here are modelling principles for class order that can be used to improve the Wikidata ontology. For example, a potential principle is that almost all classes should have a fixed order and the ones that don't are explicitly stated to have a variable order. This principle would have to be validated by looking at the (many) classes in Wikidata. This validation would be based on, among other things, logically inferring class order from instance and subclass relations.

Property Hierarchy This aspect of ontologies already has a good formalization that can be adapted to Wikidata. What is missing are principles on when and how to use this hierarchy to create more general and more-specific properties. There have to be decisions made, balancing precision and ease of use, or the introduction of a new approach entirely.

Property Conflations Some properties, such as the aforementioned ‘connects with’, are used in several different ways. The intended meaning (according to the description) is physical connection, but it is widely used as thematic connection as well. There is also the previously detailed issue of this property being used on individuals and classes alike. Our goal would be to either propose to break this property up into multiple properties, or to find a different way of handling this conflation. This effort is thus closely related to the property hierarchy.

Property Characteristics A theory of what property characteristics can mean and principles on when and how to use them might be possible, but depends on a close examination of how property characteristics are used in Wikidata.

Class Statements This issue has been studied in graphical representation for decades, at least since the work of Woods [25]. There is no good model for how this works in general nor is there a good model on how it should work in Wikidata. It is possible that using the techniques above can find some principles to help with its use in Wikidata.

Constraints Developing a stronger notion of constraints, including meaning for constraints, would serve as the start of a better constraint system for Wikidata. These constraints are not limited by the current constraint system, but may include automatic queries, entity schemas, and similar external tools.

Hierarchies of Other Relations The ‘has part(s)’ hierarchy is somewhat similar to the ‘subclass of’ hierarchy and can use some of the same methods. But this hierarchy, often studied as mereology, has pieces that are transitive and some that are not, indicating that ‘has part(s)’ in Wikidata could first be split up into several properties, each with its own formal meaning and modelling principles.

Class Order

To illustrate how our methods might work, here is how part of the proposed work is likely to proceed, building on the work that we have done in investigating class order [12].

We would develop a formal meaning for class order, extending our formal meaning for the core of the Wikidata ontology, and then a set of inference rules for class order that reflect this formal meaning, extending our set of inference rules for the core of the Wikidata ontology. There would be several rules that determine the class order of instances of a class, such the following rule (written informally):

> If C ‘instance of ’ ‘third-order class’ and D ‘instance of ’ C then D ‘instance of ’ ‘second-order class’.

Because we would be building on top of our set of inference rules for the core of the Wikidata ontology all instances of subclasses are already instances of superclasses so we don’t have to consider subclasses in these rules.

Because negation is not a part of Wikidata there would have to be rules for conflicting information, such as:

> C ‘instance of ’ ‘third-order class’ conflicts with C ‘instance of ’ second-order class’.

> D ‘instance of ’ C and E ‘instance of ’ D conflicts with C ‘instance of ’ ‘first-order class’.

This set of inference rules would be used to show the correctness of SPARQL queries that detect class order conflicts in Wikidata and determine how many Wikidata classes have or should have a fixed class order. The conflicts would be used to develop a plan to eliminate problems in the Wikidata ontology related to class order.

If, as expected, these queries show that Wikidata classes generally should have a fixed class order a modelling principle would be developed stating that Wikidata classes should have a fixed class order unless they are instances of the class 'variable-order class'. This principle would provide guidance when editing the Wikidata ontology resulting in more classes with specified class orders, resulting in the detection and potential elimination of more problems in Wikidata and making Wikidata more useful.

Expected output

We will create documents detailing our work in the Wikidata wiki and publicize these documents. These documents will provide information for the Wikidata community on the modelling principles we develop for Wikidata and methods for implementing these principles in Wikidata.

We expect to publish several papers in the International Semantic Web Conference, Extended Semantic Web Conference, or Association for the Advancement of Artificial Intelligence conference or in associated workshops or in the Transactions on Graph Data and Knowledge journal. One of us has been invited to give a keynote talk on our ongoing research related to Wikidata at the 2025 ISWC Wikidata Workshop (https://wikidataworkshop.github.io/2025/). If this proposal is funded the talk will have a section on the work performed up to that point.

These papers and talks will be our way of publicizing the work in the Semantic Web and Knowledge Graph research communities.

Risks

The work has reduced risk because it will concentrate on certain aspects of Wikidata and not aim to produce a complete set of modelling principles for Wikidata. Risk will be further reduced by an initial examination of aspects before they are tackled in detail.

There are aspects, such as the core of the Wikidata ontology, where part of work has already been done and much of the remaining portion appears to be largely straightforward. Even in these aspects there can be challenges, for example determining whether there are automated or semi-automated methods that can eliminate problems with the two core ontology properties.

Once the work has been done for these aspects the more-challenging aspects will be selected, such as how statements on classes interact with statements on their instances. It is entirely possible that formal descriptions will not be obtained for all the aspects selected or that an implementation plan cannot be devised but even partial solutions can be of interest.

There are several risks to the overall goal of this line of work. It may be that some aspects of Wikidata cannot be effectively analyzed, being too underspecified, too fragmented, or just too difficult. It may be that before producing modelling principles work needs to first be done in a community to better document or better coalesce the information they add to Wikidata. If one of these aspects is encountered in the work it will be left for later, after the necessary community work is done.

It is likely that as more aspects are taken up, previous work will have to be modified. This is a challenge, but an expected one, as expanding the scope of representation often requires reworking the previous solution.

It may be that Wikidata is too large to have overall uniform representation principles. If this is the case, the overall work might never be completed, even if Wikidata is changed to be more susceptible to the approach. But partial solutions are useful even if a total solution remains out of reach and partial solutions are often useful even for aspects that are not completely incorporated as domains that use a non-incorporated aspect generally will also use aspects that have representation principles and these principles will provide guidance for adding information from the domain.

The end goal of this work is to improve Wikidata. It may be that some communities in Wikidata are resistant to change and require extra efforts to demonstrate the benefits of incorporating representation principles.

Community impact plan

Our aim is not just to perform the research and publish it but to affect and improve Wikidata. We have already been active in the Wikidata community and have been making edits to Wikidata that remove modelling problems, such as violations of disjointness.

We will discuss our work with the Wikidata community and the Wikidata team at Wikimedia Deutschland as we perform the work. We expect that our ideas will be modified as the result of these discussions.

We will publicize our work both on general Wikidata channels and to specific Wikidata WikiProjects. We will advocate embracing modelling principles and promote ways to implement them in Wikidata. We may create tools that help move towards modelling principles, but tool creation is not a part of the work proposed here.

Evaluation

The ultimate goal of this line of work is changing how Wikidata works so that it fully embodies modelling principles. This is a long-term goal and will not be achieved during a one-year project.

We will consider this one-year project a success if modelling principles can be developed and written up for at least the core Wikidata ontology properties, class order, and the interaction of property values between classes and instances, or three aspects with similar scope, and approaches for at least partially achieving them in Wikidata be formulated and discussed with the affected communities.

We would prefer to as well have buy-in from the affected communities and plans developed to move towards the modelling principles we develop. We would also prefer to expand to several other aspects of Wikidata.

Budget

We are asking for US$39,888 for the following: support for our researcher, Ege Atacan Doğan; partial support for our principal investigator, Peter F. Patel-Schneider; software costs for communication and document preparation; a laptop for our researcher; bank fees; and travel to conferences to present the results of our work. The third member of our team, Rosario Uceda-Sosa, is employed at IBM Research and does not need support.

We have access to a high-end desktop to run queries and other interactions with Wikidata that cannot be run on public services. We also need to control and post-process the results of these queries and interactions. The control can interfere with querying Wikidata and needs to be run on a separate machine. The post-processing will likely be very interactive and is thus better to do on a local machine rather than on a server. The principal investigator already has a good laptop that can be used for control and post-processing but a good laptop for the researcher is desirable to provide him with adequate local computing power.

We expect to publish at conferences and in diamond open access journals so the dissemination costs are only travel costs for conference presentations. We are budgeting for two trips to an international conference like AAAI, IJCAI or ISWC at $3500 per trip ($1500 travel, $1000 living, $1000 registration) and one attendance at RULEML+RR, which is local for our researcher ($500 registration). There will be no overhead involved, aside from bank fees. For more details see the proposal budget at https://docs.google.com/spreadsheets/d/1B-D7o3x8FowGKUpiPlkFrM-RoPvRE0X6i5a8LUQI1iI/edit?usp=sharing.

References

[1] Mohammed Abdulai and Léa Lacroix. Wikidata: Ontology issues prioritization. www.wikidata.org/wiki/Wikidata:Ontology_issues_prioritization, 2023. Accessed 2024-07-17.

[2] Franz Baader, Diego Calvanese, Deborah L. McGuinness, Daniele Nardi, and Peter F. Patel-Schneider, editors. The Description Logic Handbook: Theory, implementation, and applications. Cambridge University Press, 2nd edition, 2007.

[3] Atílio A. Dadalto, João Paulo A. Almeida, Claudenir M. Fonseca, and Giancarlo Guizzardi. Type or individual? Evidence of large-scale conceptual disarray in Wikidata. In 40th International Conference on Conceptual Modeling (ER 2021), pages 367–377, October 2021.

[4] Ege Atacan Doğan and Peter F. Patel-Schneider. Disjointness violations in wikidata. In Fifth International Knowledge Graph and Semantic Web Conference, December 2024. A longer version available as https://arxiv.org/abs/2410.13707.

[5] Patrick Hayes and Peter F. Patel-Schneider. RDF 1.1 semantics. W3C Recommendation, http://www. w3.org/TR/rdf11-mt/, 25 February 2014.

[6] Aidan Hogan, Eva Blomqvist, Michael Cochez, Claudia d'Amato, Gerard de Melo, Claudio Gutierrez, José Emilio Labra Gayo, Sabrina Kirrane, Sebastian Neumaier, Axel Polleres, Roberto Navigli, Axel-Cyrille Ngonga Ngomo, Sabbir M. Rashid, Anisa Rula, Lukas Schmelzeisen, Juan Sequeda, Steffen Staab, and Antoine Zimmermann. Knowledge graphs. ACM Computing Surveys, 54(4):71:1– 71:37, 2021.

[7] Markus Krötzsch, Maximilian Marx, Ana Ozaki, and Veronika Thost. Attributed description logics: Ontologies for knowledge graphs. In ISWC, pages 418–435. Springer, 2017.

[8] Ronald J. Brachman Hector J. Levesque, editor. Readings in Knowledge Representation and Reasoning. Morgan Kaufmann, San Mateo, California, 1985.

[9] Ontology cleaning task force. www.wikidata.org/wiki/Wikidata:

WikiProject_Ontology/Cleaning_Task_Force, 2024. Accessed 2024-07-19.

[10] Jeff Z. Pan and Ian Horrocks. Owl fa: A metamodeling extension of OWL DL. In Proceedings of the 15th International Conference on the World Wide Web (WWW2006), pages 1065–1066, New York, NY, May 2006. ACM Press.

[11] Peter F. Patel-Schneider and Ege Atacan Doğan. Class order disorder in wikidata and first fixes. https://arxiv.org/abs/2411.15550, 2024.

[12] Peter F. Patel-Schneider and David Martin. Wikidata on mars. In 33rd International Workshop on Description Logics, 2020. Extended version available as https://arxiv.org/abs/2008.06599.

[13] Lydia Pintscher. Ontology issues in Wikidata: Everything in neat and tidy boxes? not quite! In WikidataCon 2023, October 2023. Accessed 2024-07-23.

[14] Lydia Pintscher. Wikidata survey on ontology issues, potential solutions. www.wikidata.org/wiki/Wikidata_talk:Ontology_issues_prioritization#Overview_of_potential_solutions, 2023. Accessed 2024-07-17.

[15] K. Shenoy, F. Ilievski, D. Garijo, D. Schwabe, and P. Szekely. A study of the quality of Wikidata. Journal of Web Semantics, 72, April 2022.

[16] Elizabeth Spaulding, Kathryn Conger1, Anatole Gershman, Rosario Uceda-Sosa, Susan Windisch Brown, James Pustejovsky, Peter Anick, and Martha Palmer. The DARPA wikidata overlay: Wikidata as an ontology for natural language processing. In Proceedings of the 19th Joint ACL-ISO Workshop on Interoperable Semantic Annotation, pages 1–10, June 2023.

[17] Houcemeddine Turki, Thomas Shafee, Mohamed Ali Hadj Taieb, Mohamed Ben Aouicha, Denny Vrandečić, Diptanshu Das, and Helmi Hamdi. Wikidata: A large-scale collaborative ontological medical database. Journal of Biomedical Informatics, 99, November 2019.

[18] Rosario Uceda-Sosa, Karthikeyan Natesan Ramamurthy, Maria Chang, and Moninder Singh. Knowledge consistency of llms. https://huggingface.co/datasets/ibm-research/knowledge_consistency_of_LLMs, 2024.

[19] Rosario Uceda-Sosa, Karthikeyan Natesan Ramamurthy, Maria Chang, and Moninder Singh. Reasoning about concepts with LLMs: Inconsistencies abound. https://arxiv.org/abs/2405.20163, 2024.

[20] Denny Vrandečić and Markus Krötzsch. Wikidata: A free collaborative knowledgebase. Communications of the ACM, 57(10):78–85, 2014.

[21] Andra Waagmeester, Gregory Stupp, Sebastian BurgstallerMuehlbacher, Benjamin M Good, Malachi Griffith, Obi L Griffith, Kristina Hanspers, Henning Hermjakob, Toby S Hudson, Kevin Hybiske, Sarah M Keating, Magnus Manske, Michael Mayers, Daniel Mietchen, Elvira Mitraka, Alexander R Pico, Timothy Putman, Anders Riutta, Nuria Queralt-Rosinach, Lynn M Schriml, Thomas Shafee, Denise Slenter, Ralf Stephan, Katherine Thornton, Ginger Tsueng, Roger Tu, Sabah Ul-Hasan, Egon Willighagen, Chunlei Wu, and Andrew I Su. Science forum: Wikidata as a knowledge graph for the life sciences. Elife, March 2020.

[22] Andrea Westerinen and Lydia Pintscher. Wikidata challenges in the semantic web community. In Wikidata Data Modelling Days, 30 November - 2 December 2023. Accessed 2024-07-23.

[23] William A. Woods. What’s in a link: Foundations for semantic networks. In Daniel G. Bobrow and Alan M. Collins, editors, Representation and Understanding: Studies in Cognitive Science, pages 35–82. Academic Press, New York, 1975. Also appears in [8].