COVIWD: COVID-19 Wikidata Dashboard

Fariz Darari[1,2]

[1] Faculty of Computer Science, Universitas Indonesia, Depok, Indonesia
[2] Tokopedia-UI AI Center of Excellence, Depok, Indonesia

Email: fariz@ui.ac.id

Abstract

COVID-19 (short for coronavirus disease 2019) is an emerging infectious disease that has had a tremendous impact on our daily lives. Globally, there have been over 95 million cases of COVID-19 and 2 million deaths across 191 countries and regions. The rapid spread and severity of COVID-19 call for a monitoring dashboard that can be developed quickly in an adaptable manner. Wikidata is a free, collaborative knowledge graph, collecting structured data about various themes, including that of COVID-19. We present COVIWD, a COVID-19 Wikidata dashboard, which provides a one-stop information/visualization service for topics related to COVID-19, ranging from symptoms and risk factors to comparison of cases and deaths among countries. The dashboard is one of the first that leverages open knowledge graph technologies, namely, RDF (for data modeling) and SPARQL (for querying), to give a live, concise snapshot of the COVID-19 pandemic. The use of both RDF and SPARQL enables rapid and flexible application development. COVIWD is available at http://coviwd.org.

Keywords: *COVID-19, Wikidata, Dashboard, Data Modeling, Data Analysis, Data Visualization*

1. Introduction

The global pandemic of coronavirus disease 2019 (abbreviated as COVID-19) is now impacting nearly every aspect of human lives. COVID-19 is caused by the severe acute respiratory syndrome coronavirus 2 (or for short SARS-CoV-2) [1], which was first discovered in December 2019 in Wuhan, China. At the time of writing (January 2020), there have been more than 95 million cases of COVID-19 and 2 million deaths across 191 countries and regions. On the other hand, the number of recoveries has reached around 68 million worldwide [2]. Nevertheless, it is suggested that COVID-19 survivors might experience some after-effects of the virus, both physically and psychologically [3]. This shows how the pandemic has devastating consequences, not only at the moment, but also in the future.

The quick emergence and harmful effects of COVID-19 have called for a multitude of actions in combating the spread of COVID-19. From a medical perspective, this could mean developing vaccines for the coronavirus, at pandemic speed [4]. From a public health perspective, novel mitigation strategies are deemed necessary, particularly in the form of enhanced communication regarding COVID-19-related topics to general as well as vulnerable populations [5]. In this regard, developing a dashboard that serves as a one-stop information/visualization service might come as a viable method to raise awaraness as well as inform the public about COVID-19 and its related topics.

A dashboard provides a visual display, consolidating a wide range of data about the topics of interest. The information presented in a dashboard focuses on what is important so that the information can be viewed at a glance [6]. The use of dashboards has span across different domains, such as government [7], private sector [8], and healthcare [9]. As for COVID-19, several dashboards have been developed, such as those by WHO [10], Johns Hopkins University [11], and Worldometers [12]. Indeed, these dashboards give essential information on COVID-19 like the number of cases & deaths worldwide as well as the growth over time and comparison among countries. However, none of these dashboards relies on open knowledge graph technologies.

Contributions. This paper presents COVIWD, a COVID-19 Wikidata dashboard, which complements the joint[1] initiative in communicating essential information and latest updates about COVID-19. Our dashboard is unique in the sense that it relies on an open knowledge graph (i.e., Wikidata) and open knowledge graph technologies (i.e., RDF and SPARQL). We argue that such combination enables rapid and flexible development of the dashboard, which is especially crucial for taming the fast-emerging and volatile nature of COVID-19-related information. Our dashboard is available at http://coviwd.org.

Methodology. Our research methodology is as follows. First, we start by shaping research directions based on gaps from existing work. Second, we study necessary background knowledge. Next, we analyze the modeling of the data of interest. Finally, we develop a proof-of-concept of our solution based on open knowledge graph technologies and demonstrate its potential usage.

Outline. The rest of the paper is structured as follows. Section 2 gives preliminaries on knowledge graphs and their technologies. Section 3 reports on how COVID-19 and its related topics are modeled in Wikidata. Section 4 presents COVIWD, our Wikidata-based COVID-19 dashboard. In Section 5, we discuss lessons learned, and in Section 6, we conclude our paper.

2. Preliminaries

2.1. Knowledge Graphs & Wikidata

A knowledge graph is a graph describing entities and their relationships [13]. Knowledge graphs are closely associated with the Semantic Web, introduced in 2001 [14], which is an extended form of the Web where information is made more structured and meaningful. Knowledge graphs are recently gaining more attention from the industry [15], e.g., Google, Facebook, and Microsoft, thanks to their compact representation for factual (product) knowledge. Aside from industry, knowledge graphs can be leveraged for other domains, such as geography, government, life sciences, and many more [16]. Wikidata [17] is an openly available, cross-domain knowledge graph that is backed by the Wikimedia Foundation, whose sister projects include the more well-known Wikipedia. To put it simply,

Wikidata is just like, Wikipedia for structured data. Wikidata can be edited in a collaborative manner, both by humans and machines/bots. Its data can be exported into JSON and RDF, and accessed via SPARQL queries. Additionally, data in Wikidata is well-linked to other datasets, such as the Internet Movie Database (IMDb) or the International Classification of Diseases (ICD).

2.2. RDF & SPARQL

Resource Description Framework (RDF) is a framework to publish and link (structured) data on the Web [18]. RDF can be used to describe real-world items and their interrelations. Data in RDF is represented using Subject-Predicate-Object (SPO) triples, and that a collection of triples is often referred to as a graph. An RDF triple is composed of literals (e.g., strings and numbers), Internationalized Resource Identifiers (IRIs), or blank nodes. An RDF graph can be serialized into concrete syntaxes, such as Turtle and JSON-LD.

SPARQL is a query language for RDF data [19]. The building blocks of SPARQL queries are triple patterns, which are like triples but with the addition of variables. Triple patterns can be joined together forming so-called Basic Graph Patterns (BGPs). SPARQL includes a rich set of query constructs ranging from UNION and OPTIONAL to FILTER and GROUP BY. When evaluated over RDF data, SPARQL queries return variable bindings as their results.[2] For more details regarding the syntax and semantics of SPARQL, we refer the reader to the W3C specification of SPARQL [19].

3. COVID-19 on Wikidata

Our COVID-19 dashboard relies on Wikidata as its main data source. This section describes how Wikidata models COVID-19 and its surrounding topics. Our understanding of the data model would be crucial in the development of the dashboard.

3.1. COVID-19

Wikidata provides identifiers for its items (Qid) and properties (Pid). COVID-19, as a disease, is identified by Q84263196. The very first types of information noticed from the Wikidata page of COVID-19[3] are its label, description, and aliases, as shown in Fig. 1.

While the above information concerns more on naming, the next types of information describe what

---

[1] By joint, we mean we share the same goal, albeit with different approaches to achieve the goal.

[2] In this work, we focus on the SELECT query form.
[3] https://www.wikidata.org/wiki/Q84263196

Figure 1. COVID-19 Label, Description, and Aliases on Wikidata

COVID-19 is and how COVID-19 relates to other items on Wikidata. Such types of information are called statements in the Wikidata jargon, each of which consists of a property and a value, and can be augmented with qualifiers (e.g., point in time) and references. We identify more than 270 statements in 60 distinct properties characterizing COVID-19.[4] These properties can be further categorized into generic ones and medical-specific ones. Examples of the generic properties are instance of (P31), image (P18), significant event (P793), and time of discovery or invention (P575). The medical-specific properties include health specialty (P1995), symptoms (P780), possible treatment (P924), drug used for treatment (P2176), and number of recoveries (P8010). Fig. 2 illustrates how the property health specialty is used to characterize COVID-19 on Wikidata. Additionally, there are also external ID properties, linking the COVID-19 item on Wikidata to that of external sources, such as Library of Congress (P244), Disease Ontology (P699), eMedicine (P673), and ICD-11 (P7807).

Figure 2. COVID-19 Health Specialties on Wikidata

3.2. SARS-CoV-2

The value of the has cause (P828) property of COVID-19 is SARS-CoV-2 (Q82069695), which is the virus responsible for COVID-19. The SARS-CoV-2 item, as observed from its Wikidata page,[5] has aliases like 2019-nCoV and Coronavirus. The generic properties of SARS-CoV-2 include instance of (P31) and image (P18), as also used for the COVID-19 item, as well as video (P10) and country of origin (P495). The virus-specific properties of the SARS-CoV-2 item encompass host (P2975), genome size (P2143), and parent taxon (P171). The parent taxon property is especially interesting, as we can trace not only the parent (i.e., severe acute respiratory syndrome-related coronavirus), but also the grandparent (i.e., Sarbecovirus), and so on, of SARS-CoV-2.

3.3. COVID-19 Pandemic by Country

Aside from COVID-19 (as a disease) and SARS-CoV-2, information about the COVID-19 pandemic in numbers by country is particularly important to monitor how the pandemic has spread across countries. Wikidata covers such information in the following way: there is a specific Wikidata item for the COVID-19 pandemic in every country. Such a distinction brings a better data organization, as otherwise it would be too cumbersome to store the COVID-19 pandemic information of all countries in just one Wikidata item.

As an illustration, consider the Wikidata item of COVID-19 pandemic in India (Q84055514). The statement with the instance of (P31) property for that item is shown in Fig. 3. We can observe that in addition to the property (i.e., instance of) and the value (i.e., disease outbreak), there are qualifiers providing context to the statement, e.g., the of (P642) qualifier with the qualifier value of COVID-19. Note that the pandemic information in other countries also follows the same pattern, for example, that in Indonesia (Q86913546) and Germany (Q83889294).

---

[4]Note that a property can be used in several statements.

[5]https://www.wikidata.org/wiki/Q82069695
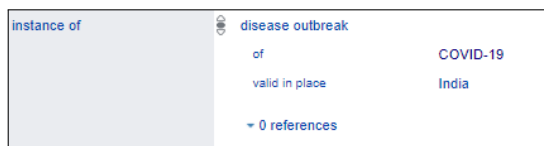
Figure 3. Statement with the instance of (P31) Property for COVID-19 Pandemic in India
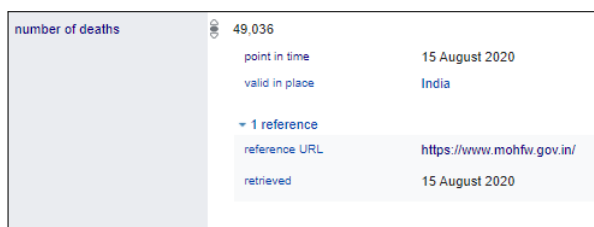
Figure 4. COVID-19 Deaths in India

Besides the typing information, there is also information concerning the number of cases, recoveries, and deaths. Fig. 4 depicts a statement about the number of deaths (as of August 15, 2020) regarding the COVID-19 pandemic in India. As seen, the statement not only makes use of property (i.e., number of deaths) and value (i.e., 49036), but also the point in time (P585) qualifier information and the reference URL (P854) information. The referred URL actually points to the website of the Ministry of Health and Family Welfare of India.

Other notable properties describing the item of COVID-19 pandemic in India include coordinate location (P625), start time (P580), distribution map (P1846), and has part (P527).

3.4. Other Related Topics

In addition to what is mentioned above, Wikidata provides information about individual COVID-19 victims as well as publications related to COVID-19. COVID-19 victims in Wikidata can be recognized by the value of COVID-19 given to the property cause of death (P509) for Wikidata items, e.g., Herman Cain (Q491019) and Imam Suroso (Q17410598). On the other hand, COVID-19-related publications can be identified by an item of the type scholarly article (Q13442814) or preprint (Q580922) with its main subject (P921) of COVID-19, for example, the publication of "Covid-19: should the public wear face masks?" (Q91785260).[6]

3.5. RDF Representation

Up until now, one might wonder how to represent the information above in RDF. The Turtle snippet in Fig. 5 shows how the information of COVID-19 deaths in India (as in Fig. 4) is captured in RDF.[7]

Let us explain the RDF snippet in Fig. 5. Line 1–11 of the snippet denote prefix declarations. The item of COVID-19 pandemic in India is described in Line 13–25. In Wikidata, there are two different cases in representing statements as RDF triples [20]: the `wdt:` case for direct, simple statements where qualifiers and references are omitted (see Line 15 for a triple representing a simple statement of the number of deaths); and the `p:` case for full statements featuring qualifiers and references (see Line 16–25 for triples representing a full statement of the number of deaths). Note that the point in time (P585) qualifier is given in Line 20, whereas the reference URL (P854) information is provided in Line 24.

4. COVIWD: COVID-19 Wikidata Dashboard

This section reports on how we develop COVIWD, a dashboard for COVID-19, based on the data model we have described in the previous section. We first give a user story in order to sketch the information/visualization requirements for the dashboard. Then, we describe our COVIWD system architecture, and finally, we present the SPARQL queries we use to retrieve COVID-19-related information/visualization for the dashboard.

4.1. User Story

Meet Bob, who is keen in understanding the state of the COVID-19 pandemic. He would like to know how the COVID-19 pandemic is spread around the world (Req-01),[8] and how the number of cases, deaths, and recoveries compares among countries (Req-02). He is also curious as to who are (some of) the COVID-19 victims (Req-03). Finally, to learn more about COVID-19, he feels the need to look for COVID-19-related publications (Req-04).

4.2. System Architecture

Before listing the SPARQL queries to cater the user story above, we describe the system architecture of COVIWD, as displayed in Fig. 6. At its core, COVIWD relies on the RDF and SPARQL

---

[6]The authors of the publication suggest a yes answer, in case the reader of our paper wonders.

[7]https://www.wikidata.org/wiki/Special:EntityData/Q84055514.ttl
[8]Req stands for request.

```
1   @prefix p: <http://www.wikidata.org/prop/> .
2   @prefix pq: <http://www.wikidata.org/prop/qualifier> .
3   @prefix pr: <http://www.wikidata.org/prop/reference/> .
4   @prefix prov: <http://www.w3.org/ns/prov#> .
5   @prefix ps: <http://www.wikidata.org/prop/statement/> .
6   @prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
7   @prefix ref: <http://www.wikidata.org/reference/> .
8   @prefix s: <http://www.wikidata.org/entity/statement/> .
9   @prefix wd: <http://www.wikidata.org/entity/> .
10  @prefix wdt: <http://www.wikidata.org/prop/direct/> .
11  @prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
12
13  wd:Q84055514
14    rdfs:label "COVID-19 pandemic in India"@en ;
15    wdt:P1120 "+49036"^^xsd:decimal ;
16    p:P1120 s:Q84055514-9926D495-31D7-4B4A-ABCA-128DB79B2793 .
17
18  s:Q84055514-9926D495-31D7-4B4A-ABCA-128DB79B2793
19    ps:P1120 "+49036"^^xsd:decimal ;
20    pq:P585 "2020-08-15T00:00:00Z"^^xsd:dateTime ;
21    prov:wasDerivedFrom ref:b46375caf93b45a322376c7796620539d7a6f05e .
22
23  ref:b46375caf93b45a322376c7796620539d7a6f05e
24    pr:P854 <https://www.mohfw.gov.in/> ;
25    pr:P813 "2020-08-15T00:00:00Z"^^xsd:dateTime .
```

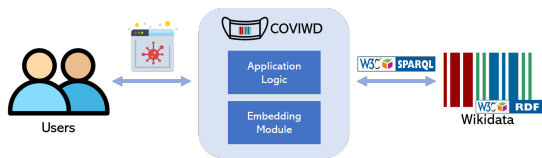**Figure 5.** RDF Turtle Snippet of COVID-19 Deaths in India

Figure 6. COVID-19 System Architecture

technologies. Wikidata provides a SPARQL endpoint for its RDF data.[9] Query results from Wikidata can be presented in various view modes[10] and embedded in HTML pages via the `<iframe>` element. We use Google Sites[11] as the platform underlying our COVIWD site mainly since it supports adding content from other sources through embedding. As a bonus, editing Web content through Google Sites is relatively easy and intuitive. The application logic for COVIWD can be interpreted as how to decide which information/visualization type is placed where. When accessed by end-users, COVIWD performs live queries to Wikidata, ensuring the data presented is always up-to-date, and shows a single webpage consisting of essential COVID-19 information to the users.

4.3. Queries in COVIWD

Now that we have described the system architecture, we are ready to fulfill Bob's requests by providing relevant SPARQL queries. The `Req-01` request about the spread of COVID-19 around the world can be realized by querying disease outbreak (Q3241045) items with the of property (P642) having the COVID-19 value (Q84263196). The information can be visualized into a map based on the coordinate location (P625) of the disease outbreak, as displayed in Fig. 7. The SPARQL query to satisfy `Req-01` is provided as follows.[12]

```
1   #defaultView:Map
2   SELECT ?outbreak ?outbreakLabel ?coord
3   WHERE
4   {
5     ?outbreak p:P31 [ ps:P31 wd:Q3241045 ;
6                       pq:P642 wd:Q84263196 ] .
7     ?outbreak wdt:P625 ?coord .
8     ?outbreak rdfs:label ?outbreakLabel .
9     FILTER(LANG(?outbreakLabel)="en")
10  }
```
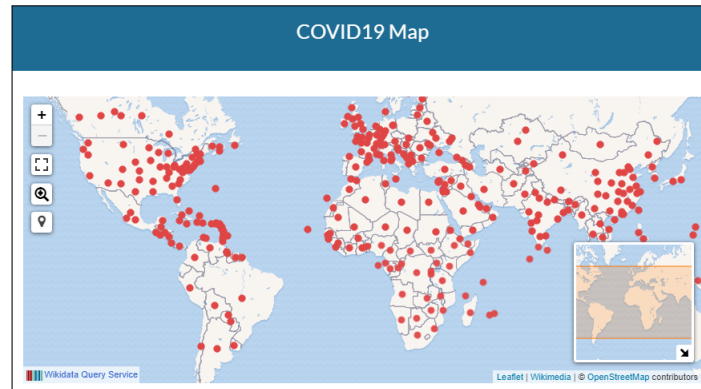
Bob's second request (`Req-02`) is the comparison on the number of cases, deaths, and recoveries among countries. We show the SPARQL query for comparing the number of deaths (as seen below), and that the other comparison can be done analogously.

```
1   #defaultView:BubbleChart
2   SELECT ?countryLabel
3     (MAX(?numDeaths) AS ?numOfDeaths)
4   WHERE {
5     ?outbreak p:P31 [ ps:P31 wd:Q3241045 ;
6                       pq:P642 wd:Q84263196 ] .
7     ?outbreak wdt:P276 ?country .
```

---

[9] https://query.wikidata.org/

[10] https://www.wikidata.org/wiki/Wikidata:SPARQL_query_service/Wikidata_Query_Help/Result_Views

[11] https://sites.google.com/

[12] https://w.wiki/v3U

Figure 7. COVID-19 Outbreak Map on COVIWD

```
8     ?country wdt:P31 wd:Q3624078 .
9     ?outbreak p:P1120 [ ps:P1120 ?numDeaths ] .
10    ?country rdfs:label ?countryLabel .
11    FILTER(LANG(?countryLabel)="en")
12   } GROUP BY ?countryLabel
```

Again, the query looks for disease outbreaks of COVID-19 (Line 5–6), and the outbreaks should be the ones for countries (Line 7–8),[13] as opposed to, say, cities. Then, the query retrieves the number of deaths (P1120) information, and takes the largest number for each country (so we get the latest one). The addition of a special command in Line 1 enables a bubble chart visualization of the query results. Besides the absolute number, we also provide a comparison of the number of deaths per 100k population by country. The two visualizations, as appearing on COVIWD, are shown in Fig. 8. From the figure, we can see that while the USA has the highest absolute number of COVID-19 deaths, it is San Marino which has the highest number of COVID-19 deaths per 100k population.

The third information request of Bob (Req-03) is about (a subset of) the individual victims of COVID-19. This can be done by executing a query for Wikidata items whose cause of death (P509) is COVID-19. The query can be augmented with information about the nationality, occupation, and date of death of the victim.[14]

The final request of Bob (Req-04) is about COVID-19-related publications. The query for such a request should then retrieve all preprints (Q580922) or scholarly articles (Q13442814) having the main subject (P921) of COVID-19. Information such as the publication date (P577), venue (P1433), title (P1476), and DOI (P356) for the publication can enrich the query even more. The query of Req-04

is given below, and that its result is illustrated in Fig. 9.

```
1    SELECT ?date ?venue ?title ?url
2    WHERE {
3      ?pub wdt:P921 wd:Q84263196 .
4      { { ?pub wdt:P31 wd:Q580922 }
5        UNION
6        { ?pub wdt:P31 wd:Q13442814 } }
7      ?pub wdt:P1476 ?title .
8      ?pub wdt:P356 ?doi .
9      BIND(IRI(CONCAT("https://doi.org/",
10       STR(?doi))) AS ?url)
11     ?pub wdt:P577 ?date .
12     ?pub wdt:P1433 ?venueRes .
13     ?venueRes wdt:P1476 ?venue .
14   }
```

In fact, COVIWD serves more information needs than only those requested by Bob, listed as follows: factbox, growths over time (for the number of cases, deaths, and recoveries, as shown in Fig. 10), symptoms, risk factors, possible treatments, health specialties, taxonomy, images, and external links. Whenever possible, COVIWD visualizes information as a graph (e.g., symptoms, risk factors), a line chart (e.g., growths), and a tree (e.g., taxonomy).

5. Lessons Learned

In this section, we highlight lessons learned from developing COVIWD: data quality, joint efforts by the community, and replicability to other diseases.

5.1. Data Quality

Data quality is defined as fitness for use by data users [21]. We identify three important aspects of quality for data about COVID-19: accuracy (whether the data is correct and reliable), completeness (whether all essential information is contained in the data), and timeliness (whether the data is timely enough for the task at hand). It is also of

---

[13]Or to be more precise, sovereign states (Q3624078).
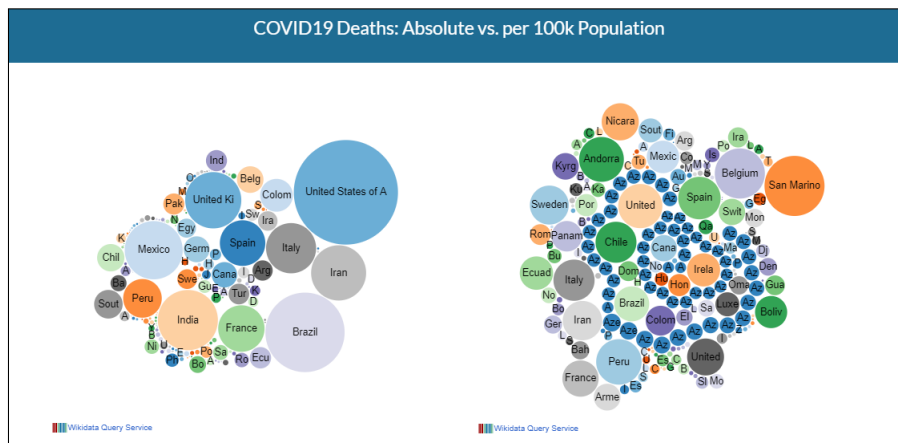
[14]The query can be accessed at https://bit.ly/covid19victims.

Figure 8. Comparison of Number of COVID-19 Deaths on COVIWD
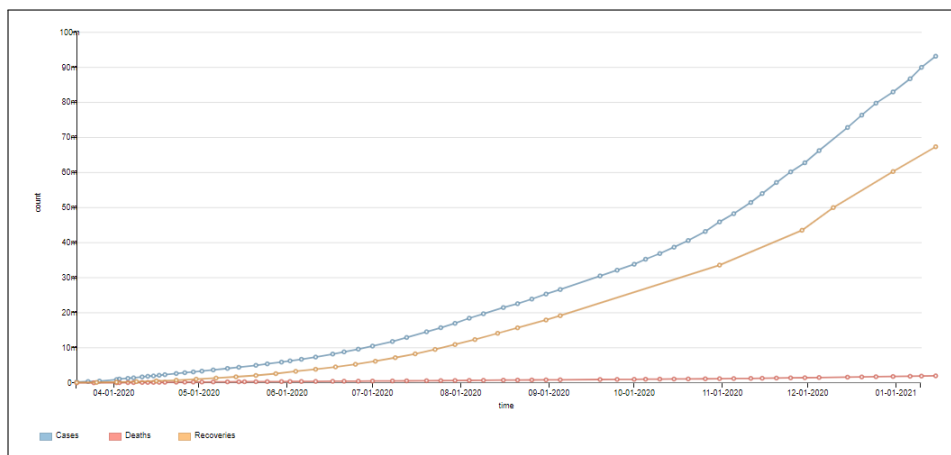

Figure 9. COVID-19 Publications on COVIWD

Figure 10. Growths of COVID-19 Cases, Deaths, and Recoveries on COVIWD

importance to assess whether the COVID-19 data in Wikidata is balanced or not [22], [23], to have a fair, unbiased view of the COVID-19 pandemic. The quality of information presented in COVIWD is as good as that in Wikidata. Hence, maintaining the quality of Wikidata is especially crucial.

5.2. Joint Efforts by the Community

The realization of COVIWD would be impossible without the joint efforts by the community in adding data to Wikidata. A WikiProject has been launched to collect Wikidata resources about COVID-19 as well as its epidemiological events [24]. The WikiProject may also improve the coordination in contributing COVID-19 data, and also serve as a discussion place among Wikidata collaborators interested in COVID-19 data. Additionally, outreach activities can be instantiated to involve more and more people in editing Wikidata, particularly those with medical expertise, such as medical researchers, doctors, nurses, and pharmacists.

5.3. Replicability to Other Diseases

The development of COVIWD is motivated by the ongoing COVID-19 pandemic. In the past, there were other pandemics, such as the 1918 H1N1 pandemic (also called the Spanish flu) and the 2009 H1N1 pandemic (also known as the swine flu). In the future, no one knows what lies ahead, so we could not exclude the possibility of a new pandemic. The approach underlying COVIWD can be replicated to other cases as long as the modeling of diseases, viruses, and other epidemiological-related information, remains consistent in Wikidata. In regard to the queries as presented in Sec. 4, nothing much has to be changed to accommodate other types of diseases: the only thing to do is replace the Wikidata item of COVID-19 (Q84263196) in the queries to other disease of interest.

6. Conclusions

The COVID-19 pandemic has affected almost every part of our lives due to its serious risks and quick emergence globally. The availability of a dashboard of COVID-19 could be useful in understanding and monitoring the latest state of COVID-19. We have presented COVIWD (http://coviwd.org), a COVID-19 Wikidata dashboard, providing a one-stop information/visualization service for topics relevant to COVID-19. The distinctive characteristic of the dashboard is that it is built based on an open

knowledge graph (i.e., Wikidata) and open knowledge graph technologies (i.e., RDF and SPARQL). We have analyzed how data about COVID-19 and its related topics are modeled in Wikidata. We have also reported on how COVIWD is developed: its system architecture and SPARQL queries that enable rich presentation of COVID-19 information. Furthermore, we have touched the aspects of data quality, joint efforts, and replicability as lessons learned. As future work, we would like to keep improving COVIWD, particularly with respect to its informativeness and usability aspect. Optimizing the SPARQL queries used in COVIWD is also in our todo list.

References

[1] World Health Organization, "Naming the coronavirus disease (COVID-19) and the virus that causes it," Available online: https://bit.ly/who-covid-naming (accessed on 13 August 2020).

[2] Statista, "Number of coronavirus (COVID-19) cases, recoveries, and deaths worldwide as of January 2021," Available online: https://www.statista.com/statistics/1087466/covid19-cases-recoveries-deaths-worldwide/ (accessed in January 2021).

[3] V. Balachandar, I. Mahalaxmi, M. Subramaniam, J. Kaavya, N. Senthil Kumar, G. Laldinmawii, A. Narayanasamy, P. Janardhana Kumar Reddy, P. Sivaprakash, S. Kanchana, G. Vivekanandhan, and S.-G. Cho, "Follow-up studies in COVID-19 recovered patients - is it mandatory?" *Science of The Total Environment*, vol. 729, p. 139021, 2020.

[4] N. Lurie, M. Saville, R. Hatchett, and J. Halton, "Developing COVID-19 Vaccines at Pandemic Speed," *New England Journal of Medicine*, vol. 382, pp. 1969–1973, 2020.

[5] D. L. Heymann and N. Shindo, "COVID-19: what is next for public health?" *The Lancet*, vol. 395, no. 10224, pp. 542–545, Feb 2020.

[6] S. Few, *Information Dashboard Design: The Effective Visual Communication of Data.* O'Reilly Media, Inc., 2006.

[7] D. Edwards and J. C. Thomas, "Developing a Municipal Performance-Measurement System: Reflections on the Atlanta Dashboard," *Public Administration Review*, vol. 65, no. 3, pp. 369–376, 2005.

[8] H. Tokola, C. Gröger, E. Järvenpää, and E. Niemi, "Designing Manufacturing Dashboards on the Basis of a Key Performance Indicator Survey," *Procedia CIRP*, vol. 57, pp. 619–624, 2016.

[9] K. Lee, S. Y. Jung, H. Hwang, S. Yoo, H. Y. Baek, R.-M. Baek, and S. Kim, "A novel concept for integrating and delivering health information using a comprehensive digital dashboard: An analysis of healthcare professionals' intention to adopt a new system and the trend of its real usage," *International Journal of Medical Informatics*, vol. 97, pp. 98–108, 2017.

[10] World Health Organization, "WHO COVID-19 Dashboard," Available online: https://covid19.who.int/ (accessed on 13 August 2020).

[11] Johns Hopkins University, "COVID-19 Dashboard," Available online: https://coronavirus.jhu.edu/map.html (accessed on 13 August 2020).

[12] Worldometer, "COVID-19 Coronavirus Pandemic Dashboard," Available online: https://www.worldometers.info/coronavirus/ (accessed on 13 August 2020).

[13] H. Paulheim, "Knowledge graph refinement: A survey of approaches and evaluation methods," *Semantic Web*, vol. 8, no. 3, pp. 489–508, 2017.

[14] T. Berners-Lee, J. Hendler, and O. Lassila, "The Semantic Web," *Scientific American*, vol. 284, no. 5, pp. 34–43, 2001.

[15] N. Noy, Y. Gao, A. Jain, A. Narayanan, A. Patterson, and J. Taylor, "Industry-Scale Knowledge Graphs: Lessons and Challenges," *Queue*, vol. 17, no. 2, pp. 48—-75, 2019.

[16] J. P. McCrae, "The Linked Open Data Cloud," Available online: https://lod-cloud.net/ (accessed on 13 August 2020).

[17] D. Vrandecic and M. Krötzsch, "Wikidata: a free collaborative knowledgebase," *Commun. ACM*, vol. 57, no. 10, pp. 78–85, 2014.

[18] R. Cyganiak, D. Wood, and M. Lanthaler, Eds., *RDF 1.1 Concepts and Abstract Syntax*. W3C Recommendation, 25 February 2014. [Online]. Available: http://www.w3.org/TR/rdf11-concepts/

[19] S. Harris and A. Seaborne, Eds., *SPARQL 1.1 Query Language*. W3C Recommendation, 21 March 2013. [Online]. Available: https://www.w3.org/TR/sparql11-query/

[20] B. DuCharme, "The Wikidata data model and your SPARQL queries," Available online: http://www.snee.com/bobdc.blog/2017/04/the-wikidata-data-model-and-yo.html (accessed on 13 August 2020).

[21] R. Y. Wang and D. M. Strong, "Beyond Accuracy: What Data Quality Means to Data Consumers," *J. Manag. Inf. Syst.*, vol. 12, no. 4, pp. 5–33, 1996.

[22] A. Wisesa, F. Darari, A. Krisnadhi, W. Nutt, and S. Razniewski, "Wikidata Completeness Profiling Using ProWD," in *K-CAP*, 2019.

[23] N. H. Ramadhana, F. Darari, P. O. H. Putra, W. Nutt, S. Razniewski, and R. I. Akbar, "User-Centered Design for Knowledge Imbalance Analysis: A Case Study of ProWD," in *VOILA Workshop co-located with ISWC*, 2020.

[24] Wikidata, "WikiProject COVID-19," Available online: https://www.wikidata.org/wiki/Wikidata:WikiProject_COVID-19 (accessed on 15 August 2020).