

The Coronavirus Disease Ontology (CovidO)*

Sumit Sharma¹[0000-0001-5054-8670] and Sarika Jain²[0000-0002-7432-8506]

Department of Computer Applications,
National Institute of Technology Kurukshetra, Haryana 136119, India
sharma24h@gmail.com, jasarika@nitkkr.ac.in

Abstract. This paper presents the Coronavirus Disease Ontology (CovidO), a superset of the available Coronavirus (Covid-19) ontologies, including all the possible dimensions. CovidO consists of an ontological network of thriving distinct dimensions for storing coronavirus information. CovidO has 175 classes, 169 properties, 4141 triples, 645 individuals with 264 nodes and 308 edges. CovidO is based on standard input of coronavirus disease data sources, activities, and related sources, which collects and validates records for decision-making used to set guidelines and recommend resources. We present CovidO to a growing community of artificial intelligence project developers as pure metadata and illustrate its importance, quality, and impact. The ontology developed in this work addresses grouping the existing ontologies to build a global data model.

Keywords: Coronavirus Ontology · Ontology Learning · Semantic Metadata · Semantic Web · Covid-19.

1 Introduction and Motivation

The World Health Organization declared a Public Health Emergency of International Concern on 30 January 2020 and a pandemic on 11 March 2020 [1]. At the same time, novel coronavirus (Covid-19) pandemic data has been collected by various data sources like the World Health Organization (WHO) and DXY.CN, BNO News, National Health Commission (NHC) of the People's Republic of China, China Centers for Disease Control and Prevention (CCDC), Hong Kong Health Department, Macau Government, Taiwan CDC, US Centers for Disease Control and Prevention (CDC), Government of Canada, Australian Government Department of Health, European Centers for Disease Prevention and Control (ECDC), Ministry of Health Singapore (MOH), and others. The Covid-19 data sources and access links are shown below.

- WHO <https://www.who.in>
- DXY.cn <https://www.dxy.cn>
- BNO News <https://bnonews.com>
- National Health Commission (NHC) <http://en.nhc.gov.cn>

* Supported by NIT Kurukshetra, Haryana, India.

- China CDC (CCDC) <https://www.chinacdc.cn/en/>
- Hong Kong Health Department <https://www.dh.gov.hk>
- Macau Government <https://www.gov.mo/en/>
- Taiwan CDC <https://www.cdc.gov.tw/En>
- US CDC <https://www.cdc.gov/>
- Government of Canada <https://www.canada.ca/en.html>
- Australian Government Department of Health <https://www.health.gov.au/>
- ECDC <https://www.ecdc.europa.eu/en>
- Ministry of Health Singapore (MOH) <https://www.moh.gov.sg/>
- Ministry of Health and Family Welfare <https://www.mohfw.gov.in/>

A significant issue is that the data sources related to Covid-19 are heterogeneous, static, and broad in scope. So many heterogeneous and stationary data sources create situations where data is sometimes under-utilized or, in more extreme cases, not used for the decision-making process [2]. Decision-making impacts all business processes and human life. For example, Covid-19 cases increases are linked to the company's performance as risks to human health and safety. As a result, dealing with multidimensional (e.g., volume, variety, etc.) aspects of information is one of the primary difficulties in today's data management and processing environment. Another vital issue of Covid-19 is to provide semantic (machine understandable) representation of data from various exciting fields such as research, health, resources, drugs, and treatment. Ontology is emerging to solve these issues. Ontology solves the problem of changing user expectations and data integration demands driven by its volatility in a rapidly growing digital market and societal challenges related to resource efficiency [3]. Ontologies have proven practical tools for representing domain knowledge, integrating data from disparate sources, and supporting many semantic applications [4].

This paper presents an OWL-based Coronavirus disease ontology (CovidO) that defines all the possible concepts, features, attributes, and relations to describe Covid-19. To develop CovidO, we have defined seven dimensions that cover all essential aspects of Covid-19: (1) About the Covid-19 infectious disease, symptoms, drugs, and treatment; (2) Information statistics of Covid-19 cases in a geographical region; (3) Covid-19 patients information with the cause of infection and exposure of pandemic; (4) Covid-19 related resources and their availability in a location; (5) Impact of Covid-19 in different verticals like education, finance, business, research and social; (6) Various guidelines and prevention and vaccine mandates by public authority; (7) Global and biomedical research on Covid-19. Several ontologies (e.g., Infectious Disease Ontology (IDO), Virus Infectious Disease Ontology (VIDO), Coronavirus Infectious Disease Ontology (CIDO), etc.) have been created for coronavirus one after another. They comprehensively and thoroughly describe coronavirus disease. These ontologies individually fail to cater to all the dimensions of coronavirus. The ontology developed in this paper aims to group existing ontologies to construct a common global data model with the unified purpose. The main contributions of this paper are:

- to provide a list of dimensions to cover every aspect of coronavirus-related information.

- to develop standard metadata (Providing a schema) called CovidO as a global data model to annotate the Covid-19 information.

This paper presents our work on CovidO version 1.0.1, which consolidates Covid-19 concepts: <https://w3id.org/covido>. To be independent of external ontologies, we defined a new namespace <https://w3id.org/covido> with the prefix `covido` (registered entry at <http://prefix.cc>) for all classes used in the ontology. As a permanent URL service, we use `w3id.org`. The relevant conceptual design, competency question, ontology, data, and code are publicly available to the community through a GitHub project.

The contents of this paper are organized as follows. The related work describes literature on the various existing Covid-19 associated ontologies and their scope boundary. The dimension section describes the incidence of the Covid-19 information dimensions in the primary analysis and provides a summary analysis. The conceptual design and scope of CovidO are described in the CovidO section to build the ontology. The ontology design section defines the competency questions to determine the scope of CovidO with abstract design. The method section outlines the model to develop CovidO. In the results section, we present the main results for the Covid-19 schema, in addition to simple predictions for the future incidence of COVID-19. Some concluding remarks are given in the conclusion.

2 Related Work

Several ontologies represent the Covid-19 pandemic in different contexts. We have found 12 ontologies related to Covid-19, each representing a completely different scope of Covid-19. They are briefed here:

- O1: Infectious Disease Ontology (called IDO):** IDO [5] is an interoperable ontology that comprises domain information on an infectious illness, including elements relevant to the disease’s clinical and biological aspects. The entity ‘disease’ is retrieved from the Ontology for General Medical Science (OGMS) [6] ontology and serves as a foundation of the IDO ontology.
- O2: Virus Infectious Disease Ontology (VIDO):** VIDO [7] adds virus-specific concepts to the IDO ontology and gives factual information on the domain of virus sickness. VIDO, in particular, inherits the entities from IDO by adding the word ‘virus’ to create a subclass, and the logical and textual information about the classes are updated accordingly. VIDO also uses elements from the Open Biomedical Ontologies (OBO) Foundry ontologies [8], such as SARS-CoV-2 proteins from the Protein Ontology.
- O3: Coronavirus Infectious Disease Ontology (CIDO):** CIDO [9] is a coronavirus-specific ontology, broader VIDO. CIDO introduces eight new entities from the coronavirus domain and concepts from other ontologies related to VIDO [7].
- O4: Covid-19 Infectious Disease Ontology (IDO-Covid-19):** IDO-Covid-19 [10] is the most particular version of CIDO, containing information

about Covid-19 and its cause SARS-CoV-2. IDO-Covid-19 adheres to the OBO Foundry design philosophy by extending the CIDO in the same way as the CIDO extends VIDO and the VIDO extends the IDO. IDO-Covid-19 also pulls concepts from other ontologies, such as SARS-CoV-2, imported from NCBITaxon.

- O5: **COviD-19 Ontology for the case and patient information (called CODO):** CODO [12] is an ontology that contains COVID-19 case data in a format that can be used by other ontologies and software systems and is based only on OWL and different W3C standards. CODO tracks specific pandemic cases, including information such as how the patient is considered to have been infected and potential further contacts who may be at risk owing to their association with the infected individual. CODO also tracks clinical tests, travel history, available resources, actual demand (e.g., ICU bed, invasive ventilators), trend analyses, and forecast increases.
- O6: **COVID-19 surveillance Ontology (COVID19):** COVID-19 surveillance Ontology [11] is a Covid-19 application ontology that intends to provide Covid-19 cases and respiratory information by obtaining data from multiple medical records systems. This ontology is constructed as a taxonomy with only 32 classes. COVID-19 verified by a lab test, SARS-CoV-2 identified, Probable Covid-19, Clinical codes, Possible COVID-19, Suspected COVID-19, Under investigation, Exposure, COVID-19 excluded are the ten core ideas of COVID19 ontology. The COVID19 ontology was created using the protégé tool, and its format is based on the OWL language.
- O7: **Covid19-IBO:** The Covid19 Impact on Banking Ontology (Covid19-IBO) [13] is a knowledge graph that gives semantic information regarding Covid-19's influence on India's banking industry. In addition, the authors have provided a schema matching technique with satisfactory results for mapping the Covid-19 ontologies.
- O8: **Kg-COVID-19:** The KG-COVID-19 [14] framework is used to create customized COVID-19 knowledge graphs. The FAIR (Findable, Accessible, Interoperable, and Reusable) approach is followed by KG-COVID-19, which combines various COVID-19-related data.
- O9: **COVID-19 Ontology:** The COVID-19 ontology includes the function of molecular and cellular entities in viral-host interactions throughout the virus life cycle and a wide range of medical and epidemiological concepts associated with COVID-19. A scalable new coronavirus (SARS-CoV-2) entity is represented as an ontology. As a prominent target of ongoing COVID-19 medicinal research, the ontology contains a broad scope on chemical entities ideal for drug repurposing. The ontology's performance was evaluated using Medline and the Allen Institute's COVID-19 corpus.
- O10: **DRUGS4COVID19:** DRUGS4COVID19 [15] identifies drugs and their associations with COVID-19. The ontology's core concepts include drug, effect, disease, symptoms, disorder, chemical substance, and so on.
- O11: **COVIDCRFRAPID:** The World Health Organization's (WHO) COVID-CRFRAPID [16] ontology is a semantic data model for the COVID-19 RAPID case record form. COVIDCRFRAPID ontology provides semantic

references to the form filled by patients during the treatment as questions and responses. It shows a variety of application scenarios, including graph-based machine learning.

- O12: **ROC:** Ontology (Country Responses towards COVID-19) ROC [17] enables data integration from heterogeneous data sources and answers interesting questions. ROC was designed to assist statistical analysis in exploring and analyzing the efficacy and side effects of government responses to COVID-19 in various nations.

We investigate these existing ontologies, focusing on a group of individuals to discuss a specific topic like drug, protein interaction databases, protein function annotations, Covid-19 patient, and cases. So they have a limited scope that does not cover all aspects of COVID-19. We have found that these ontologies refer to the Covid-19 pandemic but represent different aspects and scopes. We fill this gap, and our work follows the same approach to ontology design and has a common motivation. We have conglomerated all these ontologies into a comprehensive design covering all the required distinct dimensions (Covid-19 cases information, patient information, disease-symptom-treatment, resources, Covid-19 impact, research, and event or news related to Covid-19) discussed in the next section. By adopting established models, we aim to facilitate integration, linking, and reuse across the data sources and make data accessible to a wide range of applications. In addition, new entities have been introduced required.

3 Covid-19 Information Dimensions

With a view to allowing stakeholders in the research community and application developers to reach out and benefit, CovidO has been created as a platform through specific dimensions. There are many domains and subdomains; we group these domains into seven significant divisions that cover all the aspects of Covid-19 related knowledge. These seven dimensions are shown in detail in Table 1. The CovidO ontology is developed based on these dimensions representing COVID-19 information in OWL format and other W3C standards utilized by other ontologies and software systems. The last column of Table 1 describes the core CovidO classes associated with a particular dimension.

CovidO allows detailed tracking of specific pandemic dimensions. For example, the diseases and treatment dimension includes clinical test tracking, test report history, illness, symptoms, medication, and clinical measurement and diagnosis. Similarly, CovidO traces other dimensions as well. In brief and with overall dimensions, CovidO monitors the Covid-19 patient's travel history, symptoms, medication, available health care facilities, resources, actual need (e.g., ambulance, invasive ICU bed with ventilators), trend study, impact on business verticals, research publications finding, guidelines for public health Safety, news, and growth projections. To the best of our knowledge, we have not found any ontology that describes all the seven dimensions D1 to D7. Nor was any such ontology found which could provide an interlink between them. All ontologies have different scopes and common goals to provide the schema for Covid-19 data.

Table 1. Seven dimensions covering the Covid-19 information.

S. No	Dimension	Description	Core Classes in CovidO
1	D1: Covid-19 Cases Information	Attribute a Covid-19 case such as active, recovered, deceased, migrated cases daily across the geo-location (district, state, and country).	Statistics
2	D2: Covid-19 Patient Information	Represents the COVID-19 patients. Patient symptoms, suspected covid -19 cases, covid-19 treatment facility, patient travel history, patient nationality, interpersonal relationships between patients, supposed transmission reason, tracking of the patient test, etc.	Patient
3	D3: Covid-19 Disease, Symptom and Treatment	Describe the various diseases, different variants of disease, viruses and discover their symptom, treatment for disease.	Symptom, Disease, Treatment, Vaccine
4	D4: Covid-19 Resources	covid-19 clinical facility (Covid center, hospital, ambulance, available test), Availability of resources (Doctor, Nurse, Medical Equipment, Medicines etc.). Availability of bed, ICU bed with oxygen.	Resources, Covid-19 Clinical Facility
5	D5: Covid-19 Impact on business vertical	Impact of Covid-19 on various sectors like education sector, banking sector, economic etc.	Covid-19 Impact
6	D6: Covid-19 Related Event and Decision	Lockdown, guideline issued by government, requirement of mask, sanitizer spray, awareness program, event hosting, various exposure of covid-19 etc.	Lockdown, Prevention, Exposure
7	D7: Covid-19 Research Domain	Provide the information regarding the research on Coronavirus diseases and findings.	Covid-19 Research , Covid-19 News

4 The Coronavirus Disease Ontology (CovidO)

The CovidO knowledge representation model encodes knowledge in the form of classes, relationships, properties, instances, and axioms [3]. Our work is inspired by COviD-19 Ontology (CODO), and we have taken their work forward with some new features and new dimensions that cover all aspects of COVID-19. We have significantly expanded the capabilities of the codo ontology model made for Covide-19 cases and patients information, i.e., changes to classes, properties, relations, and axioms, emphasizing flexibility and extensibility. We also provide the integration by reusing existing ontology. We are giving annotations for ontology concepts that have already been produced but neither annotated nor labeled.

4.1 Design Methodology

From the survey of the literature one can find that there are several ontology development methodology (ODM) which carried set of activities to create ontologies. In ODM, Knowledge acquisition, integration, and alignment drive the speed of building ontology and these come with the risk of redundancy, consistency, and conflicts.

Invoking the existing ODM, we opted a mixed approach of Diligent and [21] and Methontology [22] development methodology to establish the CovidO. The general procedure, roles, and functions of the CovidO methodology process described as:

- **Ontology Requirement Specification (ORs):** In the ORs process, set the goal for ontology development and study the feasibility of the environment. We examine heterogeneous data sources given in the introduction to building ORs for CovidO and state-of-the-art requirement analysis according to the dimension of CovidO given in Section 3. ORs document help to design competency questions (given in Table 4) to define CovidO's scope. CovidO domain knowledge should be organized as a meaningful ORs model at the knowledge level. After gathering sufficient information and ORs, we create a CovidO conceptual model that describes the problem and its solution (shown in Fig. 4).
- **Adapting Ontology:** In Diligent, the foaf: Agent starts the ontology development process by creating the initial ontology. Thus we have assumed that the initial ontology is already constructed in the form of codo. It saves ontology development time and will help expand ontology in a distributed manner with different stakeholders' objectives. Thus, ontology can have a comprehensive scope and variety of applications and needs.
- **Revision Update:** The analysis attempts to find the similarities in changing requests and users' ontology. The study looks for commonalities between the ontologies of changing requests and users. The current plan is not to combine all user ontologies. Instead, decide what changes should be made to the extended version of the shared ontology.

- **Implementation:** Once have a conceptual design, then implementation can be done with two intermediate way: formalization and ontology design. In formalization, tottransform the conceptual modelinto a formal or semi-compatible model. In the ontology design, to make ontology computable with some ontology editor like protégé with supported formal language.

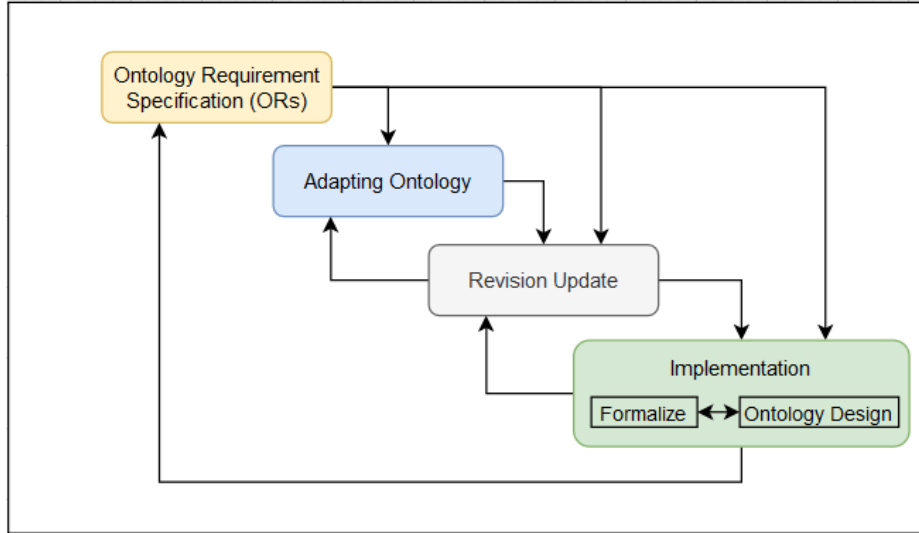


Fig. 1. CovidO development process methodology based on Diligentand [21] and Methontology [22].

4.2 New Features

CovidO is formed by reusing existing ontologies adding new classes and properties to cover all the dimensions listed in section 3. Table 2 represents the core classes of CovidO with referral namespace. We have divided the Statistics class of codo into two parts, codo: Statistics class and CovidO: Resource. Statistics class represents the actual cases information, and Resource class represents the resources utilized for covid-19. Similarly, other classes have been added to be meaningful, easily accessed, and used. Table 3 describes some new relationships between the classes of CovidO that were not present in other Covid ontologies. Currently, CovidO contains 175 classes, 169 relationship types, goes to evaluation. We applied the Pellet Reasoner to verify that CovidO had a good consistency.

The top-level class structure diagram and relationships between core concepts are shown in Fig. 2). There are many types of relationships between concepts.

Table 2. Some latest added CovidO core classes with referral namespaces and their description.

S.No	Core Class	Namespace	Description
1.	dbo:Continent	dbpedia.org	A continent is any of several large landmasses.
2.	dbo:Town	dbpedia.org	A town is a human settlement.
3.	covid0:TownWiseStatistics	w3id.org/covid0	Statistical information of covid case in a town.
4.	codo:Statistics	w3id.org/codo	update, information about the infected cases.
5.	covid0:Status	w3id.org/covid0	provide the status as recovered, hospitalized, deceased.
6.	covid0:TestResult	w3id.org/covid0	Provide test result positive or negative.
7.	covid0:Resources	w3id.org/covid0	Details of Covid-19 related resources.
8.	covid0:ClinicalMeasurement	w3id.org/covid0	Describe clinical measurement required for covid testing i.e. blood test, oxygen saturation, temperature etc.
9.	covid19IBO:Impact	semanticweb.org/archana/ontologies/2021/5/untitled-ontology-6	Covid-19 impact on business verticals.
10.	covid0:Covid-19	w3id.org/covid0	Information about the Covid-19.
11.	covid0:Covid-19Study	w3id.org/covid0	Covid-19 study on application, news, article, global research, origin of covid etc.
12.	covid0:Covid19FungalInfection	w3id.org/covid0	Fungal infections in covid positive patients and different variants.
13.	covid0:Laboratory-Test	w3id.org/covid0	Laboratory Test for covid specific.
14.	covid0:Treatment	w3id.org/covid0	A planned operation based on data for the delivery of health care.
15.	covid0:Vaccine	w3id.org/covid0	Antigenic chemicals are used in preparations to stimulate the immune system and elicit an immunological response.

Table 3. Some newly added core object properties of CovidO with domain and range.

S.No	Object properties	Namespace	Domain	Range
1.	isPartOf	w3id.org/covid0	dbo:Continent	dbo:Continent
2.	hasResources	w3id.org/covid0	codo:Covid-19DedicatedFacility	covid0:Resources
3.	hasClinicalFinding	w3id.org/covid0	schema:MedicalClinic	covid0:ClinicalFinding
4.	isVaccinated	w3id.org/covid0	foaf:Person	dbo:Vaccine
5.	hasImpactOn	w3id.org/covid0	covid0:Covid-19	xmlns:Organization
6.	hasPrevention	w3id.org/covid0	covid0:Mndate	covid0:Prevention
7.	hasDecision	w3id.org/covid0	foaf:Person	covid0:Decision
8.	hasImpactOn	w3id.org/covid0	xmlns:Organization	xmlns:Organization
9.	mandatissued	w3id.org/covid0	covid0:PublicAuthority	covid0:Mndate
10.	hasRiskFactor	w3id.org/covid0	xmlns:Organization	covid0:RiskFactor
11.	hasGovern	w3id.org/covid0	covid0:PublicAuthority	xmlns:Organization

The content ensures the semantic consistency of relationships between concepts and facilitates logical axioms and reasoning definitions.

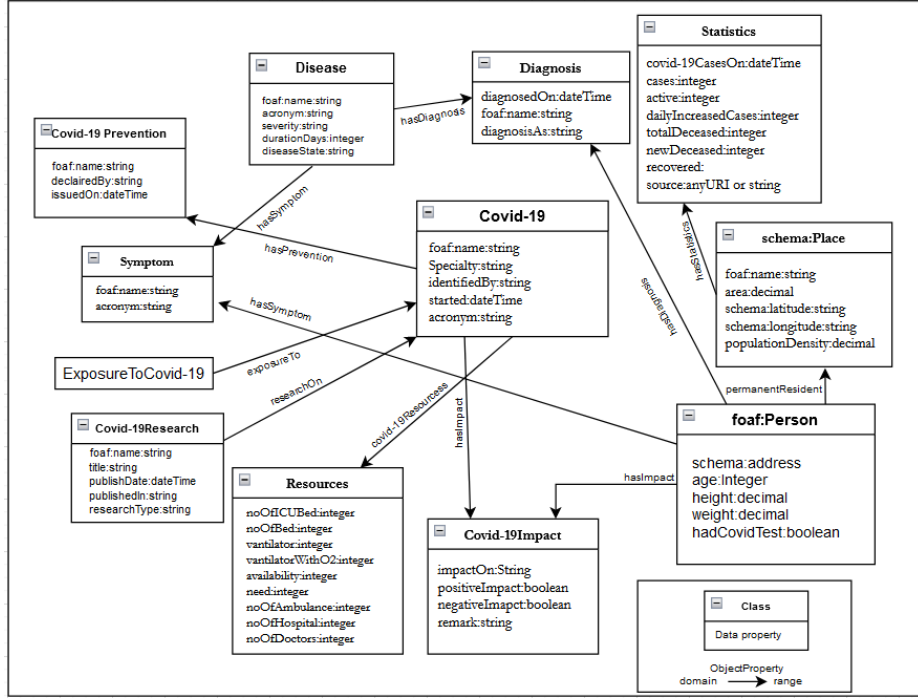


Fig. 2. class structure diagram of the top levels of the core class hierarchy in CovidO.

4.3 CovidO Scope

Competency Questions (CQs) play an essential role in the ontology development lifecycle, representing the ontology requirements. Some recommendations have to formalize requirements through Covid-19 data sources. We could structure and expand the conceptual modeling design of covidO to enhance covidO model capabilities using competency questions (CQs), reasoning requirements, and contextual statements. The competency questions have broad coverage on Covid-19 data, so we are trying to reduce it and map into the seven dimensions of CovidO to cover all the aspects of the pandemic. Table 4 represents some competency questions with the respective dimension that CovidO is expected to answer.

4.4 Reusing the ontology concept

According to [18], there are two perspectives to reuse ontology: (1) assembling, extending, specializing, and adapting other ontologies that are components of

Table 4. CovidO competency questions excerpt.

Dimension	Q.No.	Competency Questions
D1	CQ1.	When did country <i>c</i> have the first Covid-19 case?
	CQ2.	How many Covid-19 cases have been found at any (location) place <i>l</i> on date <i>t</i> ?
	CQ3.	How many patients died in Continent <i>r</i> on date <i>t</i> ?
D2	CQ4.	What is the travel history of a patient <i>p</i> ?
	CQ5.	How many patients travelled from / to Continent <i>r</i> on date <i>t</i> ?
	CQ6.	List all the patients between age 18 and 30.
D3	CQ7.	What are the different variants of Covid-19?
	CQ8.	What are the prevention vaccines for Covid-19?
	CQ9.	What are the drugs for Covid-19 treatment?
D4	CQ10.	What are the clinical measurements of a Covid-19 patient <i>p</i> ?
	CQ11.	What are the health care facilities for Covid-19 in a place <i>l</i> ?
	CQ12.	How many ICU beds are in a hospital <i>h</i> at a place <i>p</i> ?
D5	CQ13.	How many ambulances are available in a hospital <i>h</i> on date <i>t</i> ?
	CQ14.	What are the business verticals on which COVID-19 has a positive impact?
D6	CQ15.	Provide a list of all organizations at high risk?
	CQ16.	When was the first lock down announced in a country <i>c</i> ?
D7	CQ17.	What is the exposure of Covid-19 spread?
	CQ18.	What is prevention advice issued by state public authority for covid-19?
D7	CQ19.	What are the research articles published for Covid-19 in a month?
	CQ20.	What are the news headlines and their sources of covid-19 on date <i>t</i> ?

the final ontology, or (2) integrating other ontologies on the single concept that integrate all concept. We have used second approach, integration of other ontologies. The core concepts for CovidO are determined as prevention, vaccine, hospital facility, disease, infection, disorder, virus, SARS-Cov2, Coronavirus, agent, patient, disease, symptom, drug, treatment, organization, impact, host, diagnosis, statistics, place, etc. Which is already in use in other ontologies O1 to O12. CovidO integrate the concept with existing ontologies to pursue the basic principle of ontology implementation. Fig. 3 represents the terms inherited by CovidO from available schemas. Different colors define each schema. The solid and the dotted line show immediate and remote child-parent relations between classes. Most inherited entities belong to Schema, BFO, and CODO, while Kg4Grug, SYMP, and VO are the least inherited ontologies.

4.5 Conceptual Design of CovidO

We use a permanent URL service, w3id.org, that makes it independent of external ontologies. We have defined a new namespace convention <https://w3id.org/covido> for all classes used in CovidO, with the prefix covido (the entry registered at <http://prefix.cc>). A common concept is integrated into a single concept that unifies them all. The concept selection is based on the best fit for the scope and dimension of CovidO. CovidO relevant file, competency questions, data, code, and owl file have been provided by the GitHub community, advantaging

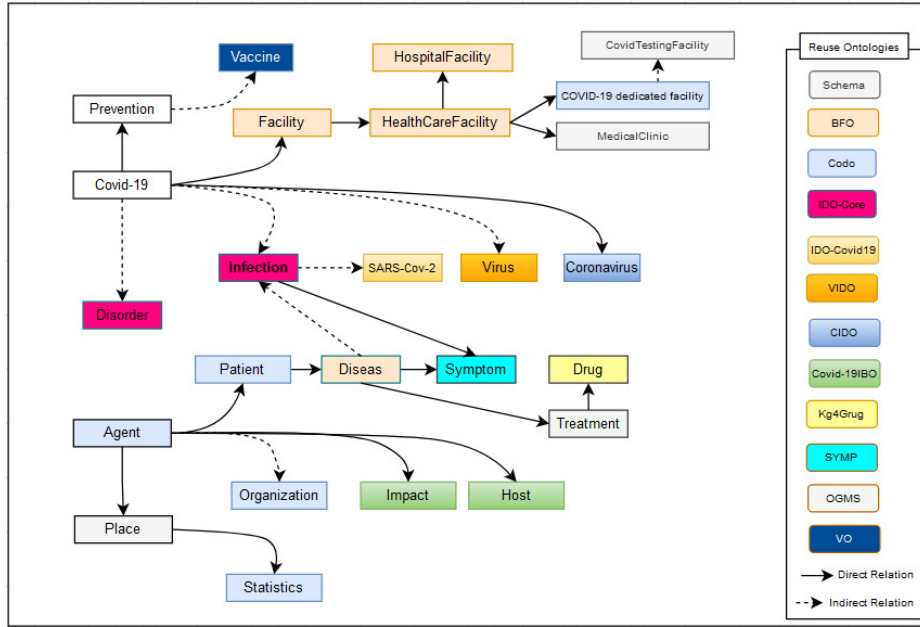


Fig. 3. Reuse core concepts of CovidO ontology extracted from the related ontologies.

to the contributor. An overview of classes and properties of CovidO is represented in Fig. 4. Complete CovidO structure has seven dimensions (Disease and Treatment, Resources, Cases, Patient, Events, Research, and Symptoms), seven dimensions have been represented in different colors. Most of the entities in our ontology belong to the Disease and Treatment dimension. The resources dimension is the smallest in terms of the number of classes. The solid and the dotted arrow represent the property and the subclass relation between two entities respectively. The name over the arrow represents the object property, the arrow tail connects the domain, and the head is the range of property. Vertical containers represent the class with inside data properties. The data properties follow up vocabularynamespace:dataproperty:datatype.

5 Evaluation

We evaluated CovidO in two ways: using SPARQL query and OOPS! Pitfall Scanner. SPARQL query describes the accessibility of elements and OOPS! Pitfall represents the RDF quality of CovidO.

5.1 SPARQL Query Evaluation

SPARQL [19] allows users to query necessary and alternate graph patterns and their conjunctions and disjunctions. CovidO is schema-agnostic, so schema

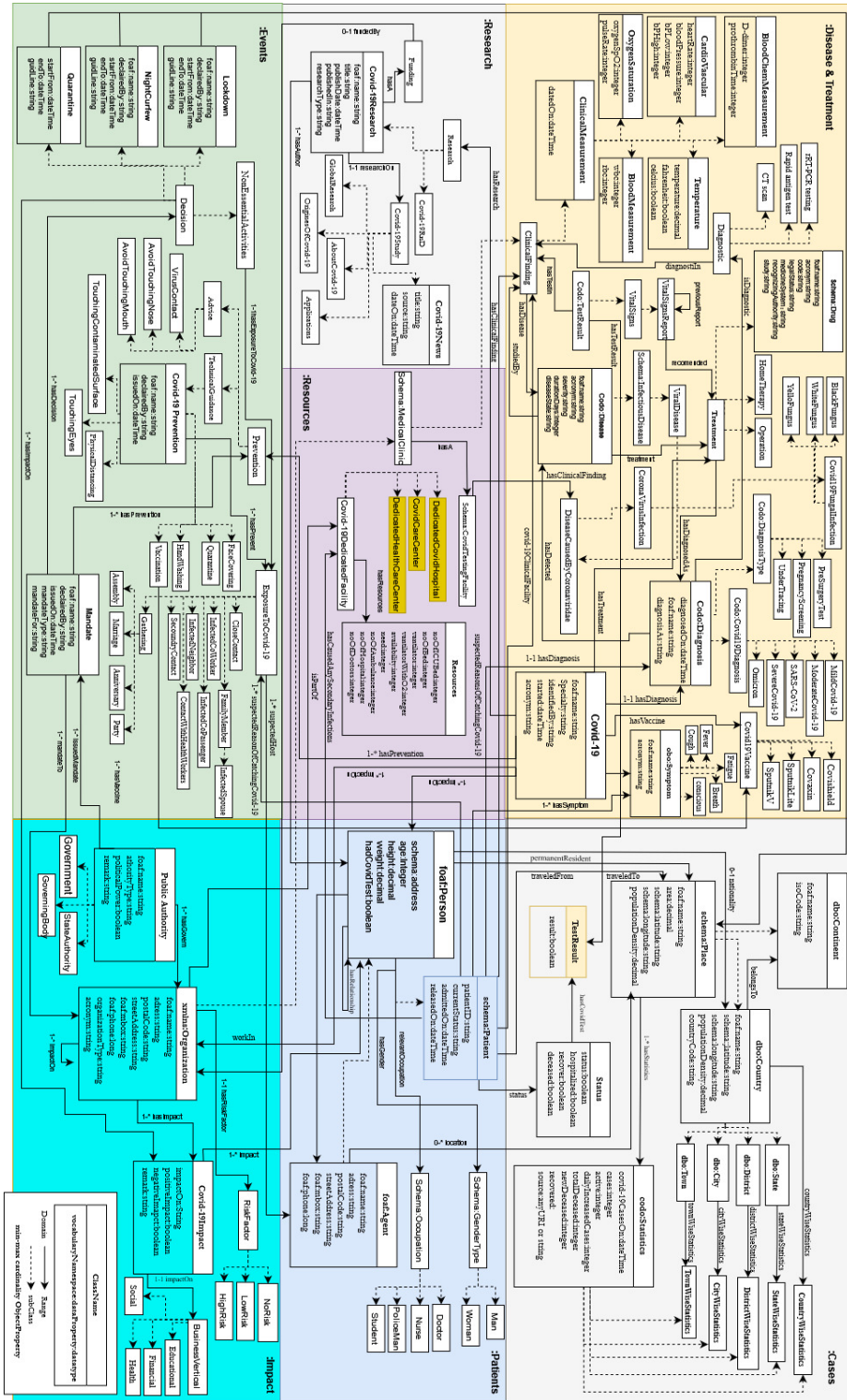


Fig. 4. A figure represent the conceptual design of CovidO.

level SPARQL query is allowed. Once the ordered representation is done as user requirement from the coronavirus distributed and heterogeneous data sources, CovidO instance-level questions are permitted to answer. We evaluate CovidO on schema-level competency questions based on the D3 dimension. The SPARQL queries and their results are shown in Table 5. The prefix and their namespaces used for SPARQL queries are shown in Table 6.

Table 5. CovidO schema level queries and their SPARQL queries and results obtained.

S.No	Question	SPARQL Query	Result
CQ8.	What are the prevention vaccines for Covid-19?	SELECT ?CovidVaccine WHERE { ?Vaccine rdfs:subClassOf covido:Vaccine. ?CovidVaccine rdfs:subClassOf ?Vaccine.}	Novavax, Cov- axin, Moderna, SputnikLife, Zydus- Cadila, Covishield, Corbevax, SputnikV,
CQ10.	What are the clinical measurements of a Covid-19 patient?	SELECT * WHERE ?Vaccine rdfs:subClassOf covid:ClinicalMeasurement.	Temperature, Oxy- genSaturation, BloodMeasurement, CardioVascular, BloodChemMeasure- ment, Diabetes

Table 6. CovidO prefixes table.

Prefix	Namespace
rdf:	http://www.w3.org/1999/02/22-rdf-syntax-ns
owl:	http://www.w3.org/2002/07/owl
rdfs:	http://www.w3.org/2000/01/rdf-schema
xsd:	http://www.w3.org/2001/XMLSchema
foaf:	http://xmlns.com/foaf/0.1/
ndf:	https://www.semintelligence.org/ns/whonto
xmlns:	http://xmlns.com/foaf/0.1/
dbo:	http://dbpedia.org/ontology/
schema:	https://schema.org/
covid:	http://www.w3id.org/covid/

5.2 OOPS! Pitfall Evaluation

We have used OOPS! Pitfall scanner [20] to examine CovidO. OOPS! Pitfall ontology diagnosis online tool detects 40 different types of pitfalls in OWL ontologies, including semantic and structural checks and best practices verification. OOPS! Pitfall scanner has two components, Pitfall Scanner, and Suggestion Scanner. Pitfall Scanner checked the ontology syntax, and Pellet reasoner has

analyzed the logical consistency of ontology. Whereas the Suggestion scanner has thrown some suggestions for possible errors of ontology elements. OOPS! Pitfall based result evaluation on CovidO is shown in the Fig. 5. We resolved the problems reported by OOPS! in the CovidO and push the changes to the target GitHub repository.

[Expand All] | [Collapse All]










Results for P08: Missing annotations.	97 cases Minor 
Results for P13: Inverse relationships not explicitly declared.	52 cases Minor 
Results for P20: Misusing ontology annotations.	4 cases Minor 
Results for P22: Using different naming conventions in the ontology.	ontology* Minor 
Results for P24: Using recursive definitions.	6 cases Important 
Results for P25: Defining a relationship as inverse to itself.	6 cases Important 
Results for P26: Defining inverse relationships for a symmetric one.	6 cases Important 
Results for P30: Equivalent classes not explicitly declared.	3 cases Important 
Results for P34: Untyped class.	7 cases Important 
SUGGESTION: symmetric or transitive object properties.	10 cases

Fig. 5. Screenshot of result summary of CovidO on OOPS! Pitfall Scanner.

6 Conclusion

This work presents Coronavirus Disease Ontology (CovidO): 1) Describes complete knowledge of Covid-19; 2) Maps existing ontologies and create standard metadata to understand and share Covid-19 knowledge that humans perform to interact with the same affordances; 3) Acts as a vocabulary for researchers, engineers, and developers to search/identify the commonly used Covid-19 gesticulation for particular affordances, and to understand the scope and dynamics of a specific gesture. We extracted elements observed from existing coronavirus ontologies defined in previous studies while developing the ontology. The intention and scope of the CovidO can be summarized in seven dimensions as discussed above.

As future work, several possible extensions can be made to the ontology by integrating linguistics and adding more uses case. In addition, we intend to release and deploy the CovidO RESTful service in the Cloud and API clients to the leading coronavirus annotation.

References

1. World Health Organization, 30 January 2020, Geneva, Switzerland, accessed 10 Feb, 2022, [https://www.who.int/news/item/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-\(2005\)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-\(2019-ncov\)](https://www.who.int/news/item/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov)), (2022).

2. Dividino R, Soares A, Matwin S, Isenor AW, Webb S, Brousseau M. Semantic integration of real-time heterogeneous data streams for ocean-related decision making. Defence Research and Development Canada= Recherche et développement pour la défense Canada; (2018).
3. Osborne, Francesco, and Enrico Motta. "Klink-2: integrating multiple web sources to generate semantic topic networks." In International Semantic Web Conference, pp. 408-424. Springer, Cham, (2015).
4. Ding, Li, Pranam Kolari, Zhongli Ding, and Sasikanth Avancha. "Using ontologies in the semantic web: A survey." In Ontologies, pp. 79-113. Springer, Boston, MA, (2007).
5. Babcock, Shane and Beverley, John and Cowell, Lindsay G and Smithd, Barry , The Infectious Disease Ontology in (2020).
6. Scheuermann, R.H., Ceusters, W., Smith, B. (2009). Toward an Ontological Treatment of Disease and Diagnosis. Summit on Translational Bioinformatics, p.116 (2009).
7. Beverley, John, Shane Babcock, Lindsay Cowell, and Barry Smith. "The COVID 19 Infectious Disease Ontology." (2021).
8. Smith, Barry, Michael Ashburner, Cornelius Rosse, Jonathan Bard, William Bug, Werner Ceusters, Louis J. Goldberg et al. "The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration." Nature biotechnology 25, no. 11: 1251-1255, (2007).
9. He, Yongqun, Hong Yu, Edison Ong, Yang Wang, Yingtong Liu, Anthony Huffman, Hsin-hui Huang et al. "Cido: The community-based coronavirus infectious disease ontology." (2021).
10. John Beverley, COVID-19 Infectious Disease Ontology, <https://biportal.bioontology.org/ontologies/IDO-COVID-19>, Accessed 25 Jun (2022).
11. Dylan McGagh, Harshana Liyanage, Simon de Lusignan, John Williams, COVID-19 Surveillance Ontology, <https://biportal.bioontology.org/ontologies/COVID19> Accessed 17 Feb (2022).
12. Dutta, Biswanath, and Michael DeBellis. "CODO: an ontology for collection and analysis of COVID-19 data." arXiv preprint arXiv:2009.01210 (2020).
13. Patel, Archana, Narayan C. Debnath, Ambrish Kumar Mishra, and Sarika Jain. "Covid19-IBO: a Covid-19 impact on Indian banking ontology along with an efficient schema matching approach." New Generation Computing 39, no. 3: 647-676 (2021).
14. Reese, Justin T., Deepak Unni, Tiffany J. Callahan, Luca Cappelletti, Vida Ravanmehr, Seth Carbon, Kent A. Shefchek et al. "KG-COVID-19: a framework to produce customized knowledge graphs for COVID-19 response." Patterns 2, no. 1: 100155 (2021).
15. Badenes-Olmedo, Carlos, David Chaves-Fraga, MarÍA Poveda-Villalón, Ana Iglesias-Molina, Pablo Calleja, Socorro Bernardos, Patricia MartÍN-Chozas et al. "Drugs4Covid: Drug-driven Knowledge Exploitation based on Scientific Publications." arXiv preprint arXiv:2012.01953 (2020).
16. Luiz Bonino, COVIDCRFRAPID, <https://biportal.bioontology.org/ontologies/COVIDCRFRAPID>, Accessed 17 Feb (2022).
17. Qundus, Jamal Al, Ralph Schäfermeier, Naouel Karam, Silvio Peikert, and Adrian Paschke. "ROC: An Ontology for Country Responses towards COVID-19." arXiv preprint arXiv:2104.07345 (2021).
18. Pinto, H. Sofia, and J. P. Martins. "Reusing ontologies." In AAAI 2000 Spring Symposium on Bringing Knowledge to Business Processes, vol. 2, no. 000, p. 7. Karlsruhe, Germany: AAAI, (2000).

19. Eric Prud, Andy Seaborne, Hewlett-Packard Laboratories, Bristol, SPARQL Query Language for RDF, <https://www.w3.org/TR/rdf-sparql-query/>, Accessed 20 Feb (2022).
20. Gómez-Pérez, Asunción. "OOPS!(OntOlogy Pitfall Scanner!): supporting ontology evaluation on-line" (2009).
21. H. S. Pinto, S. Staab, and C. Tempich. Diligent: Towards a fine-grained methodology for distributed, loosely-controlled and evolving. In *Ecai 2004: Proceedings of the 16th European Conference on Artificial Intelligence*, volume 110, page 393. IOS Press, 2004.
22. Fernández-López, Mariano, Asunción Gómez-Pérez, and Natalia Juristo. "Methontology: from ontological art towards ontological engineering." (1997).
23. Author, F.: Article title. *Journal* **2**(5), 99–110 (2016)
24. Author, F., Author, S.: Title of a proceedings paper. In: Editor, F., Editor, S. (eds.) *CONFERENCE 2016, LNCS*, vol. 9999, pp. 1–13. Springer, Heidelberg (2016). <https://doi.org/10.1007/1234567890>
25. Author, F., Author, S., Author, T.: Book title. 2nd edn. Publisher, Location (1999)
26. Author, A.-B.: Contribution title. In: *9th International Proceedings on Proceedings*, pp. 1–2. Publisher, Location (2010)
27. LNCS Homepage, <http://www.springer.com/lncs>. Last accessed 4 Oct 2017