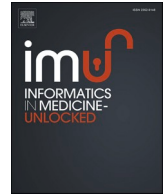




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



## Ontological engineering for the definition of a COVID-19 pandemic ontology

Alexandra González-Eras<sup>a,b,c</sup>, Ricardo Dos Santos<sup>a,c</sup>, Jose Aguilar<sup>a,c,d,e,\*</sup>, Alberto Lopez<sup>a,c</sup>

<sup>a</sup> CEMISID, Facultad de Ingeniería- Universidad de Los Andes, 5101, Mérida, Venezuela

<sup>b</sup> Departamento de Ciencias de la Computación y Electrónica - Universidad Técnica Particular de Loja, Cda. Universitaria San Cayetano Alto, 1101608, Loja, Ecuador

<sup>c</sup> Tepy R+D Group. Artificial Intelligence Software Development. Mérida, Venezuela

<sup>d</sup> GIDITIC, Escuela de Ingeniería, Universidad EAFIT, Medellín, Colombia

<sup>e</sup> Universidad de Alcalá, Departamento de Automática, Spain

### ARTICLE INFO

#### Keywords:

COVID ontology

COVID-19

Ontology integration

Ontological engineering

### ABSTRACT

COVID-19 has generated a lot of information in different formats, and one of them is in the ontology format. Also, there are previous ontologies from other disciplines that can help to analyze the COVID-19 pandemic. Thus, due to the large quantity of COVID-19 information in the form of ontologies, approaches to ontology integration and interoperability could be beneficial. In this context, this research proposes a new ontology, called COVID-19 Pandemic ontology, which is the product of an ontological engineering process proposed in this research that allows the integration of several ontologies to cover all the aspects of this infectious disease. The ontological engineering process defines tasks of fusion, alignment, and linking for integrating the ontologies. The resulting pandemic ontology provides a simple repository for storing information about the COVID-19, reusing existing ontologies, to offer multiple views about the disease, including the social context. This ontology has been tested in different case studies to prove its capabilities to infer useful information about the COVID-19 pandemic.

### 1. Introduction

The year 2020 started with a pandemic due to the outbreak of coronavirus (COVID-19 henceforth). This virus has been quickly extended to the world, with significant effects on society: deaths, economic catastrophes, etc. Also, the COVID-19 pandemic has generated many researches in different domains to help health institutions to fight against this virus. Currently, a large amount of information about this disease is produced at an impressive speed, which grows exponentially every day, with an absence of clear criteria to order, comprehend, use and connect it.

In the literature, there are several works about ontologies related to the COVID-19 pandemic, for example: the Coronavirus Infectious Disease Ontology (CIDO) is a community-driven open-source biomedical ontology in the area of coronavirus infectious disease [40]. The CIDO provides a standardized computer representation of coronavirus disease, including its etiology, transmission, epidemiology, pathogenesis, diagnosis, prevention, and treatment. Another one is the COVID-19 Ontology that is for cases and patient information (CODO) [39], which provides a standard-based reusable vocabulary to annotate and describe COVID-19

information. It focuses on the daily data for COVID-19 cases (e.g., active, recovered, deceased, migrated), their geo-location (district, state (province), and country); and patient data like nationality, symptoms, suspected level of COVID -19, patient's travel history, inter-personal relationships between patients, supposed transmission reason, and others. Finally, the authors of [38] describe the Infectious Disease Ontology (IDO) in the context of COVID-19, which is a suite of interoperable ontology modules that aims to provide coverage of all aspects of the infectious disease domain, including biomedical research, clinical care, and public health. The center of this suite was IDO Core, a disease- and pathogen-neutral ontology covering just those types of entities and relations that are generally relevant to infectious diseases. IDO Core represents disease and pathogen-specific ontology modules. They discuss how this ontology might assist with the COVID-19 pandemic and accelerate data discovery in the early stages of future pandemics.

Due to the amount of information and ontologies from the COVID-19 pandemic and related domains, their management and integration are complex and must be accessible to both humans and machines (or systems). In this sense, approaches about ontological engineering processes are required to integrate different ontologies according to the

\* Corresponding author. CEMISID, Facultad de Ingeniería, Unviersidad de Los Andes, Merida, Venezuela.

E-mail addresses: [alexisge@gmail.com](mailto:alexisge@gmail.com) (A. González-Eras), [aguilar@ula.ve](mailto:aguilar@ula.ve) (J. Aguilar).

context, which may involve tasks of merging/mixing ontologies or linking between them. In each case, this will previously imply the alignment between them. Remarkably, an ontological engineering process enables the integrating of several ontologies that have modeled COVID-19 characteristics from different perspectives.

Thus, given the number of ontologies of the COVID-19 pandemic and other related domains, their integration and interoperability are a necessity to be able to use together. In this sense, approaches about ontological engineering processes are required to integrate them, which may involve tasks of merging/mixing ontologies or linking between them. In each case, this will previously imply the alignment between them. Thus, an ontological engineering process allows the integrating of these ontologies that have modeled the COVID-19 pandemic from different perspectives.

In this research, we propose a COVID-19 Pandemic ontology based on the integration of different COVID-19 ontologies, even with ontologies from other domains, but necessary for treating the pandemic. Our ontology allows each ontology to provide information on the pandemic from different perspectives, providing a complete multi-dimensional view. For this purpose, this research defines an ontological engineering process for the integration of the ontologies, and for the definition of the COVID-19 Pandemic ontology.

According to this ontological engineering process, this research analyzes each one of the ontologies to determine its relevance in the context of the COVID-19 pandemic. The basis ontologies are extracted from the literature and are highly specialized to one domain (e.g., IDO). Also, the relationships between them are specified, following the integration strategy defined by the ontological engineering process [16,33,34,36,37]. Thus, the first step is the alignment between them, and according to their domains, the integration can be a merge/mixture or a linking between them. When the ontologies are in the same knowledge domain, then they must be merged/fused; and when they are complementary, then they must be linked [16,36,37].

In this way, we define a new ontology with the use of different ontologies, where each one answers to a particularly useful aspect for the treatment of COVID-19. In this sense, this ontology is related to the pandemic outbreak and it can cover the most remarkable aspects to understand how and why this virus is propagating. This research also provides information about the critical aspects related to the disease. It considers several aspects about basic classes associated with COVID-19 like "Treatment", "Causes", "Symptoms", "Transmission Mechanism", and "Epidemiology", but also, it considers different classes from other domains such as "Sociocultural", "Socioeconomics", and "Demography". Thus, our ontology includes ontologies about Infectious Disease (e.g., IDO) integrated with information of the contexts (e.g., "Sociocultural" class), carrying out a more complete and deep reasoning process to exploit all the knowledge gathered through these ontologies. We have conducted several experiments with our ontology to answer crucial aspects of the pandemic. In general, the main contributions of this research are:

- An ontological engineering process to build ontologies about COVID-19
- An Ontology, called Tepuy-COVID, which mixes ontologies from the COVID-19 domain.
- An Ontology, called Covid-19 Pandemic, which models the knowledge about COVID-19 from different dimensions (symptoms, treatments, socio-cultural aspects).
- Case studies that analyze the behavior of the pandemic from the developed ontology.

This research is divided into the following sections: Section 2 presents the related researches, with the main ontologies linked to our work. Section 3 defines the main concepts in the domain of Ontological Engineering, section 4 describes the procedure to build our COVID-19 Pandemic ontology based on our ontological engineering process.

Section 5 presents some experiments related to our ontology through several case studies, and finally, the conclusions arising from this research are presented.

## 2. Related works

In this section are described several researches relying on ontological engineering and integration techniques of ontologies directed to different purposes and COVID-19 related ontologies. This section analyzes the related works from the following perspectives: 1. COVID-19 ontologies and ontologies of other domains used to represent the information of the pandemic, 2. Techniques and methods of ontological engineering used to model a knowledge domain and, 3. Metrics and validation schemes used in the ontological domain.

According to FAIR principles, several COVID-19 ontologies offer robustly supported data integration, sharing, reproducibility, and computer-assisted data analysis [22], indicating that all research data should be findable, accessible, interoperability and reusable [24]. These include CIDO ontology that brings together various models to represent aspects such as similarity to other viruses, common symptoms and drugs that have been attempted to treat the virus, etc. COVID-19 Surveillance Ontology supports surveillance in primary care. DRUGS4COVID195 defines the relationship between medications and COVID-19 symptoms. The COVIDCRFRAPID7 ontology provides semantic references of quiz questions and answers [25]. The CODO ontology defines patient, clinical tests, travel history, available resources, current need, trend study, and growth projections. The latter describes real cases of the pandemic [22]. However, no studies of ontologies above have been integrated with ontologies from other domains to represent the connotations of the disease in different contexts [23].

Regarding the techniques and methods, in the research [1], the authors reported a strategy for reusing existing ontologies, many of them related to our approach and applied in this publication. They defined a High-Level collaborative Architecture (HLA) to specify the semantics of objects and their interaction since existing ontologies. Thus, they constructed the ontologies interoperation based on an automatic transforming method embedded in HLA. The output was verified through a consistency verification method, guaranteeing the feasibility of the ontology management strategy proposed. In Refs. [2] were presented multi-dimensional collaborative ontology models that integrate a series of sub-ontologies through processes, such as mapping and merging of the concepts, properties, and instances between them. The authors of [2] proposed two ideas: the core ontology and the stage ontology. The stage ontology describe the different ontologies to be integrated; the core ontology was constructed based on the integration of the ontologies in the stage ontology using different techniques, such as mapping and merging. In the research [4] was defined an ontology-based on the federated collaboration mechanism, which involves a fusion ontological strategy and a weighted approach to leverage the integration of the ontologies.

In [3,6], the authors defined an integration process of ontologies by using different engineering methods in the following order: first, an ontology mapping to establish relationships between terms of other ontologies; second, an ontology alignment that looks for connections between different ontologies, and finally, an ontology merging that gets new ontological models based on the previous steps. Notably, in research [3], HLA was extended to integrate ontologies using different engineering methods, such as mapping, merging, and aligning. On the other hand, the authors of research [6] defined an approach to find out the semantic relations in a set of ontologies. Then, they propose an automatic semantic retrieval process to visualize the model that describes the ontological integration. This produced integration is an automatic semantic expansion that the final users in their queries could use. The research [5] proposed an approach to facilitate semantic integration by using different ontological techniques to solve two problems, the reutilization of ontologies and their heterogeneities. They defined an



Fig. 1. Ontology linking of two ontologies.

architecture with four layers: the *presentation layer* that describes the meta-information of the ontologies (language, domain, etc.); the terminology layer defines synonyms, polysemy, etc.; the concepts layer shows the structure of the ontologies (classes). Finally, the Semantic layer defines the semantic relationships, properties, etc.

### 3. Ontological engineering

Ontological Engineering refers to the activities linked to the ontology development process, including the methodologies, tools, and languages required for building ontologies [16,33,34,36,37]. One of the current main domains is Ontological Mining, which consists of extracting behavior patterns, knowledge, and other characteristics, using mining techniques to build or enrich ontologies [16]. Thus, ontological mining is the discovery of new ontology knowledge, from its concepts and relationships, including their structures, to the instances related to each concept. In a context with a high number of ontologies, ontological mining is necessary to extract global knowledge from a set of ontologies. According to Ref. [16], the main ontological mining techniques are Ontology alignment, Ontology linking, Ontology fusing/mixing, which are described below.

#### 3.1. Ontology alignment

The ontology alignment consists of making the comparison (matching) between the concepts of the ontologies analyzed, which is the process of finding relationships or correspondences between entities of different ontologies [16]. For that reason, ontology alignment performs semantic correspondence analysis between two or more ontologies. The technique compares ontology concepts, obtaining the relationships between entities of different ontologies. This relationship can be between other classes, individuals, properties, or formulas.

The objective of performing ontology alignments is to find relationships between the entities expressed in different ontologies to discover equivalences determined through similarity measures between

these entities. The process starts with mapping between the ontology classes, applying measure similarities between them. The comparison is based on the calculation of similarity measures, which can be: linguistic (names of entities), between properties (classes), graphs (taxonomic structure), among others.

#### 3.2. Ontology linking

It is also called mapping and its objective is to establish identity relationships between entities of different ontologies through their common characteristics (ex: superclass\_of, subclass\_of) properties. The mapping result comprises a link ontology that contains the equivalent entities and properties in the ontologies, through which the two ontologies are connected.

Thus, it allows the creation of a general ontology through the integration of different linked ontologies. The identification and definition of the concepts that link the ontologies require a certain consensus. Alignment techniques allow finding the set of relationships and properties potentially used for the link between the ontologies. It may also need an expert to create new concepts that are not considered in any ontology to be linked (see Fig. 1).

#### 3.3. Ontology merging

It is also called fusing or mixing. It is a process where several ontologies within the same knowledge domain come together to standardize knowledge, make knowledge grow, or have locally complete knowledge [36,37]. Mixing is required when ontologies handle the same domain but with different or partial representations, such that the ontologies can coincide in certain concepts and not in others. In that sense, it is necessary to integrate them (see Fig. 2).

Classically, the mixture of ontologies implies obtaining a new ontology, considering aspects such as inconsistencies (that a relationship contradicts another relationship within the same ontology), synonyms, contradictions, and discrepancies between the ontologies. There are two

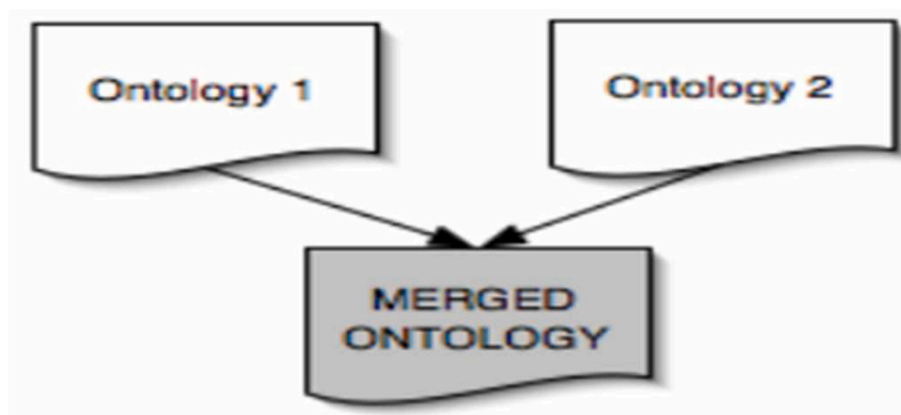


Fig. 2. Ontology merging of two ontologies.

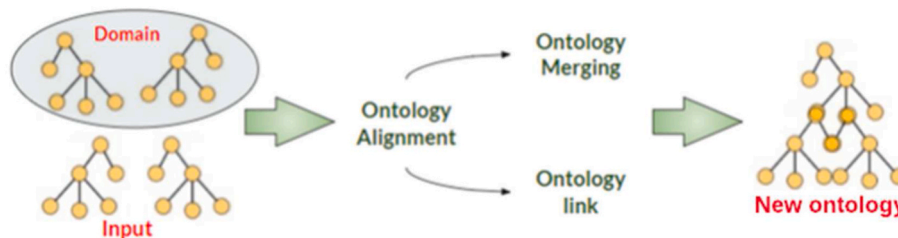


Fig. 3. Methodology of creation of COVID-19 Pandemic ontology.

**Table 1**  
Ontological Engineering Process for the building of the COVID-19 Pandemic ontology.

Input: Ontologies of the COVID domain and ontologies of other domains
<b>Procedure:</b>
1. Create domain ontology (Tepuy-COVID ontology)
1.1. Create pairs without repetition of ontologies depending on the number of existing domain ontologies (see Section 4.2).
1.2. Repeat for each pair on the list.
1.2.1. Calculate Ontology Alignments (see Section 4.1)
1.3. Merge the domain ontologies using the ontology alignments in the Tepuy-COVID ontology
1.4. Validate Tepuy-COVID ontology using competency questions (see Section 4.4)
2. Create Linking Ontologies between Tepuy-COVID and other domains (see Section 4.3).
2.1. Create pairs of ontologies between Tepuy-COVID and each ontology of another domain depending on the number of existing ontologies of other domains (see Section 4.2).
2.2. Repeat for each pair on the list.
2.2.1. Calculate Ontology Alignments (see Section 4.1)
2.2.2. Transform Alignments to Linking Ontology.
2.2.3. Validate Linking Ontology using competency questions (see Section 4.4)
<b>Output:</b> COVID-19 Pandemic ontology

types of merges: a weak merge of ontologies, where it is possible to leave ontology concepts without being mixed, or a strong merge done in two parts, a first part where the weak merge is carried out, and a second part where concepts and relationships left out are added. Some of the principles used during the merging process are described in the following [16,36,37],

- When a concept of one ontology matches with one of the concepts in the other ontology, this concept is picked up to integrate the new ontology, increasing size and enriching itself.
- When a concept in an ontology is the same as in the other, but the name is uniquely different, the concepts are synonyms in the integration process.
- When both names and content are different, a new concept involving the concept name and content is created in the new ontology.

**4. Ontological engineering process to build the COVID-19 pandemic ontology**

Our study contemplates the use of an ontological engineering process to ensemble different COVID-19 ontologies that have been reported in the literature. Also, this merged ontology must be extended with ontologies from other domains to analyze the pandemic from different points of view, not only in the health domain, such as economic, social, etc. Fig. 3 shows the general method for the creating of the Pandemic ontology. There is a first step where the classes and properties of the different ontologies must be aligned according to a similarity measure, no matter if they are in the same domain or not. Once similar classes and properties are obtained between the ontologies of the same domain, a fusion (merger) process of ontologies is carried out. The result is a domain ontology that can be linked with ontologies from other domains. Then, based on the alignment between ontologies of other domains, it is

**Table 2**  
Tools for Ontology Engineering tasks.

Ontology Engineering task	Alignment Process	Fusion	Ontology Validation
<b>Ontology Merging</b>	Align	Protégé	Pellet, Align, and Protégé
<b>Ontology Linking</b>	Align	n/a	Pellet, Align, and Protégé

carried out the linking of all related classes and properties. The final model contains all classes and properties similar to the domain of COVID-19 and its extension to other domains.

Table 1 presents the description of the ontology construction process. The Ontological Engineering Process initially defines the Tepuy-COVID ontology, which brings together ontologies of the COVID domain. This ontology is created by mixing pairs of domain ontologies, based on the alignment between them (see step 1). With this COVID domain ontology (Tepuy-COVID ontology) are linked ontologies of other domains. For that, the next procedure is repeated between Tepuy-COVID ontology and each one of other domain: first, the alignment between them, and then, the link between the ontologies (see step 2). Finally, in each step competency questions are used to validate the quality of the resulting ontology (Tepuy-COVID ontology and COVID-19 Pandemic ontology).

Additionally, an essential part of the process is selecting the tools that are carrying out the ontological engineering process. Although various applications make the processes of alignment [7–9], linking [10, 11], and mixing [10,11], no one efficiently supports the three processes [12]. So, they must be manually integrated to support an entire ontological engineering process. In this research, we have chosen the Alignment API application (Align) for the alignment and mapping tasks [7] and Protégé for the mixing of the domain ontologies [10]. We develop an ontological engineering process integrating the partial results of these tools using ontological languages like OWL and RDF (Table 2).

Below, we present each ontological engineering process according to the chosen tools and the ontological engineering process defined in Table 1. For this, we will use two ontologies of the COVID-19 domain: COVID-19 Ontology for Cases and Patient information (CODO)<sup>1</sup> and WHO COVID-19 Rapid Version CRF semantic data model (COVIDCRFRAPID).<sup>2</sup> In addition, we have used an ontology of the hospital management domain: Hospital Management (Presence Ontology (PREO)).<sup>3</sup> With the first two domain ontologies, we will carry out alignment and fusion processes. The resulting ontology will be linked with the ontology of the hospital management domain, thus generating

<sup>1</sup> CODO, available on the Internet <https://bioportal.bioontology.org/ontology/es/CODO>.

<sup>2</sup> COVIDCRFRAPID, available on the Internet <https://bioportal.bioontology.org/ontologies/COVIDCRFRAPID>.

<sup>3</sup> Presence Ontology, available in <http://data.bioontology.org/ontologies/PREO/submissions/1/download?apikey=8b5b7825-538d-40e0-9e9e-5ab9274a9aeb>.

**Table 3**  
Alignment of the CODO and COVIDCRFRAPID ontologies.

CODO	RELATION STRENGTH	COVIDCRFRAPID	Type
codo#Negative	1.0	whocovid19crfsemdatamodel#instances/Negative	Instance
codo#Positive	1.0	whocovid19crfsemdatamodel#instances/Positive	Instance
codo#Hospitalized	1.0	whocovid19crfsemdatamodel#instances/Hospitalized	Instance
codo#VitalSigns	0.90909	whocovid19crfsemdatamodel#Vital_signs	Class
codo#ViralDisease	0.61538	whocovid19crfsemdatamodel#Liver_disease	Class
codo#Tired	0.6	whocovid19crfsemdatamodel#Piped	Instance
codo#LaboratoryTestFinding	0.57142	whocovid19crfsemdatamodel#Laboratory_question	Class
codo#hasUncle	0.55555	whocovid19crfsemdatamodel#has_value	Property
codo#CoronavirusInfection	0.55	whocovid19crfsemdatamodel#corona_virus_list	Class
codo#Disease	0.53846	whocovid19crfsemdatamodel#Liver_disease	Class
codo#Fever	0.5	whocovid19crfsemdatamodel#instances/Better	Instance
codo#UP	0.5	whocovid19crfsemdatamodel#instances/avpu-list/P	Instance
codo#Recovered	0.5	whocovid19crfsemdatamodel#Transferred	Instance

our Pandemic ontology. To validate the results, we will use competency questions and some metrics such as Precision, Recall, and F1 measure [7], which will give us a vision of the quality of the obtained ontology.

4.1. Aligning

The process begins with the entry of the CODO and COVIDCRFRAPID ontologies in the Alignment API tool [19]. For this, we selected the similarity measure to find the similar names (terms) of classes, properties, and instances. The tool offers a set of lexical and semantic similarity methods, typical of text disambiguation context, which analyze terms according to their linguistic structure (comparing terms according to the characters that make them up) [16] and their semantic domain (comparing terms against a dictionary or thesaurus) [15,16]. In this way, the result represents the relationship strength between pairs of ontology elements assigning a value between zero and one (where one means a maximum similarity and zero that there is no similarity).

According to Ref. [17], several similarity methods validate the alignments between different language units, as is the case of pairs of terms, which are also applicable to the context of ontology alignment. We get similar pairs of elements from ontologies using the Alignment API tool and two approaches of similarity measures: terminological matching and Linguistic-based similarity. Terminological matching considers that “same concepts are likely to be modeled using quite similar names” [27], and it uses string-based techniques [29] to concepts

comparison, similar as the case of Levenshtein and SMOA (String Metric for Ontology Alignment) measures. Linguistic-based similarity [15,29] focuses on semantic domain terms, using distance methods and lexical databases, such as WordNet, obtaining a “lexical semantic relatedness measure that represents the strength of the semantic relationship between terms according to the shortest path between nodes in the semantic network” [18].

The used measures were: 1. **EditDistNameAlignment**: it uses the Levenshtein distance between entity names to look for a similarity between pairs of terms [7,14,33]; 2. **SMOANameAlignment**: it considers two features: the commonalities and differences between terms [26,27]; 3. **JWNLAlignment**: it computes a substring distance between the entity names of the first ontology and the entity names of the second ontology expanded with WordNet 3.0 synset [30]. Additionally, it uses the WordNet thesaurus.

Table 3 presents the alignments of the CODO and COVIDCRFRAPID ontologies using the Levenshtein Distance method [13] with a threshold of 0.8 [19,20]. Thus, the pairs of elements that have a relative strength equal to 1 correspond to exact matches in the name of terms in the two ontologies. However, as we can see in the “Type” column, they do not necessarily belong to the same ontological element (class, property, and instance). In the case of the pair [codo#VitalSigns, whocovid19crfsemdatamodel#Vital\_signs], we observe that they have a relation force of 0.9 that places this pair above the established threshold (0.8). On the other hand, for the remaining pairs, which do not exceed

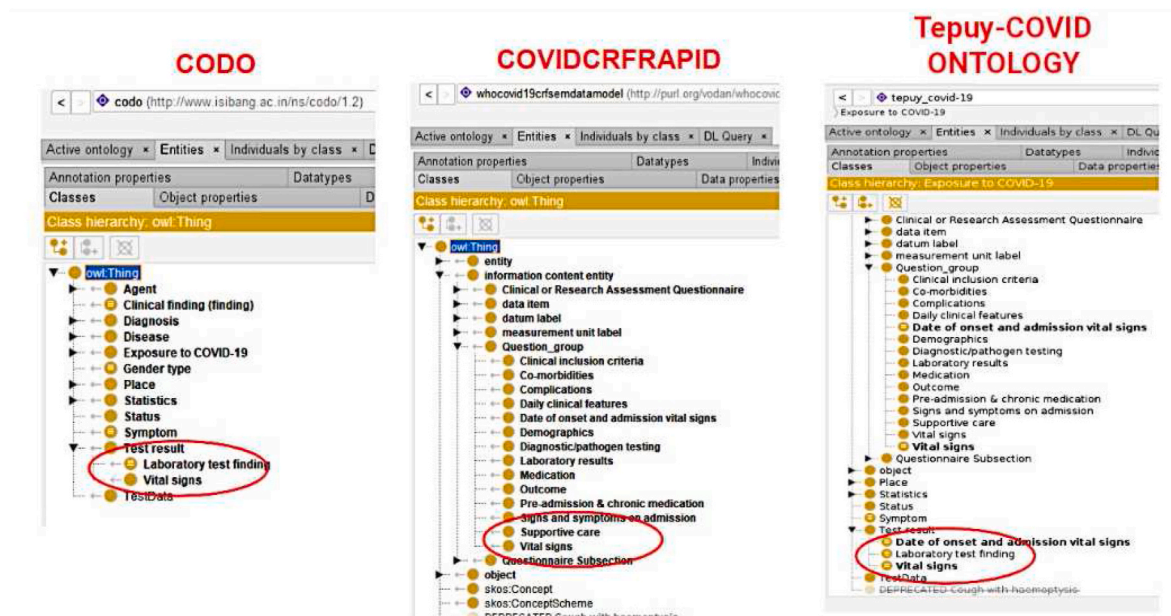


Fig. 4. Ontologies for the merging process and its result (Tepuy-COVID Ontology).

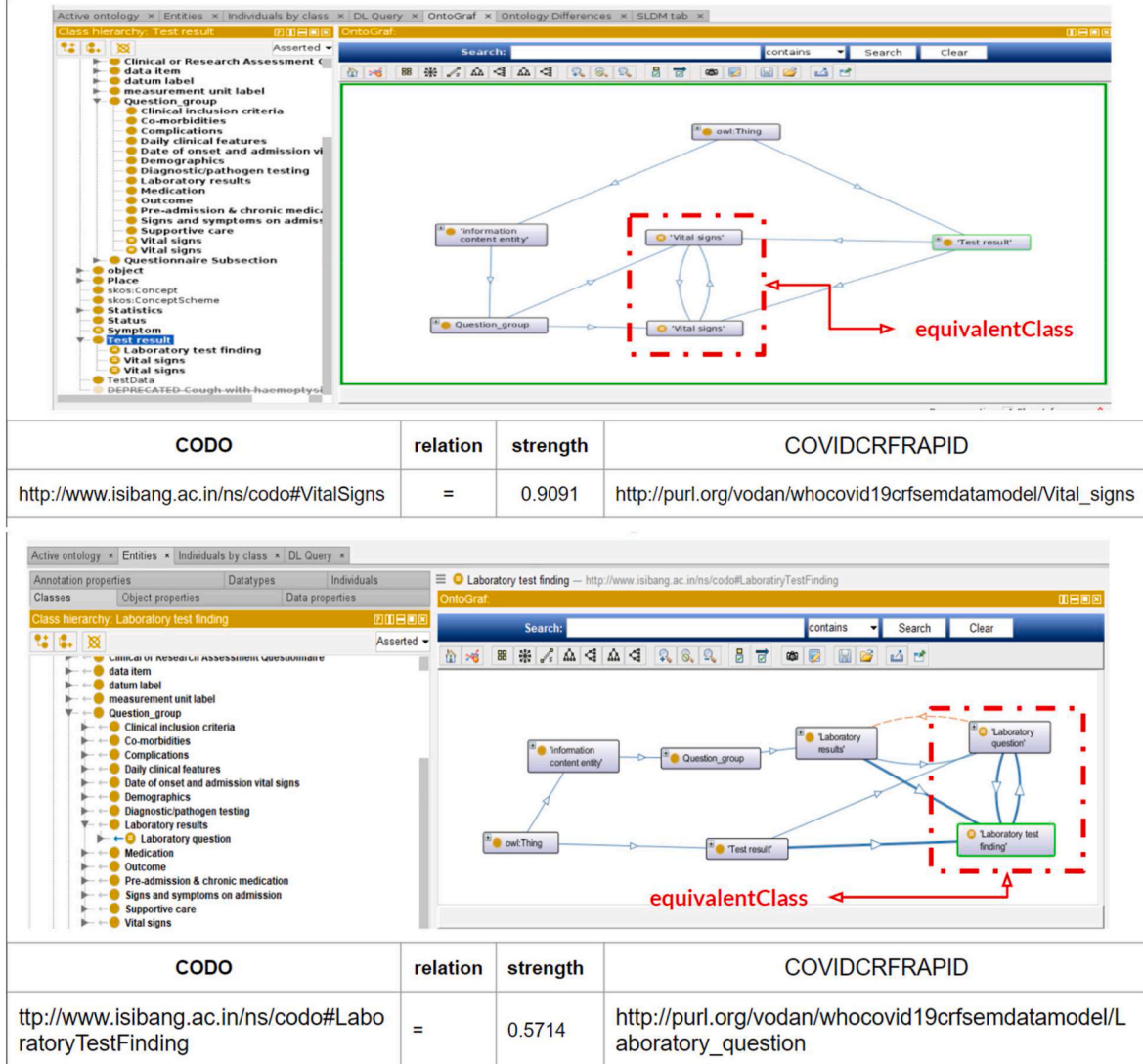


Fig. 5. Some results of the merging process.

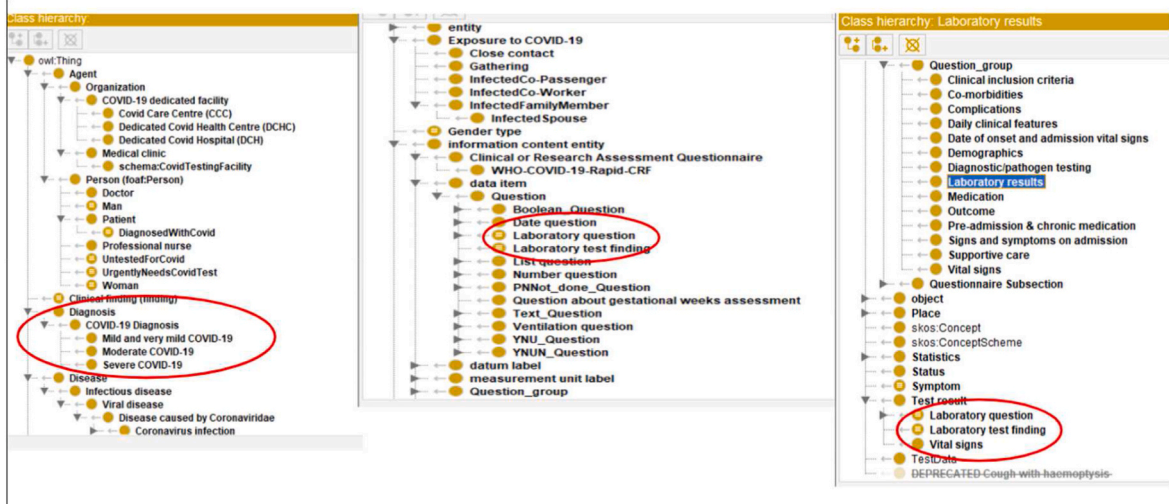


Fig. 6. Tepuy-COVID ontology.

**Table 4**  
Alignment for the Linking process.

Tepuy-COVID Ontology	RELATION STRENGTH	Presence Ontology (PREO)	Type
<a href="#">whocovid19crfsemdatamodel#Person</a>	1.0	presence-ontology.org#ontology/Person	Class
<a href="#">whocovid19crfsemdatamodel#Hypertension</a>	1.0	presence-ontology.org#ontology/Hypertension	Class
<a href="#">codo#Nurse</a>	1.0	presence-ontology.org#ontology/Nurse	Class
<a href="#">whocovid19crfsemdatamodel#Outcome</a>	1.0	presence-ontology.org#ontology/Outcome	Class
<a href="#">whocovid19crfsemdatamodel#Diabetes</a>	1.0	presence-ontology.org#ontology/Diabetes	Class
<a href="#">whocovid19crfsemdatamodel#Age</a>	1.0	presence-ontology.org#ontology/Age	Class
<a href="#">schema.org/Patient</a>	1.0	presence-ontology.org#ontology/Patient	Class
<a href="#">foaf/0.1/Person</a>	1.0	presence-ontology.org#ontology/Person	Class
<a href="#">codo#Diagnosis</a>	1.0	presence-ontology.org#ontology/Diagnosis	Class
<a href="#">whocovid19crfsemdatamodel#Kidney_disease</a>	0.92857	presence-ontology.org#ontology/KidneyDisease	Class
<a href="#">whocovid19crfsemdatamodel#has_value</a>	0.88888	presence-ontology.org#ontology/hasValue	Property

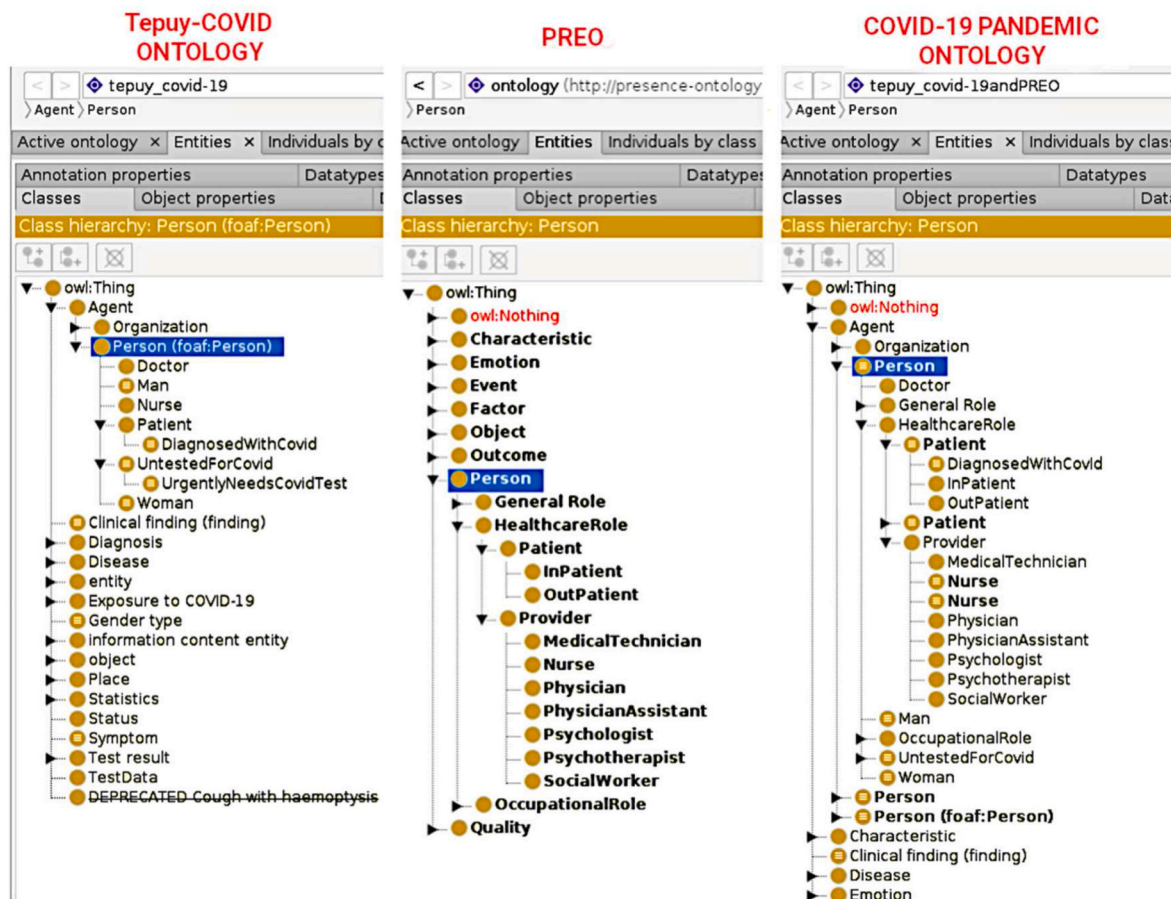


Fig. 7. Ontologies for the Linking process.

the threshold, we observe that although the terms share a specific group of characters, semantically speaking, the domain of each one does not seem to be related (example:[[codo#LaboratoryTestFinding](#), [whocovid19crfsemdatamodel#Laboratory\\_question](#)]). The final result is an ontological model formed by pairs alignments used as a link ontology between the ontology CODO and COVIDCRFRAPID.

#### 4.2. Merging

The merging process follows the steps indicated in Ref. [10]. Thus, the CODO and COVIDCRFRAPID ontologies are analyzed from the alignment obtained with the Alignment API tool. For that, a transformation of the alignments to OWL axioms is carried out using the Alignment API Rendering method (specifically, the OWLAXiomsRendererVisitor) [19], which generates a document that only

includes the alignments found (which we will call the Alignment Ontology). Then, we create a fusion with the Protégé’s merge option between the two COVID-19 ontologies and the Alignment Ontology to generate the Tepuy-COVID Ontology (see Fig. 4). The result is a new ontology that contains the elements of the two ontologies and includes the alignments between these ontologies.

Fig. 5 shows some alignments detected that have been included in the mix, such as the classes “VitalSigns” (CODO) and “Vital\_Signs” (COVIDCRFRAPID), which have a relationship of “equivalent classes” because their similarity exceeds the threshold of 0.8. Also, we see that the new ontology includes equivalence relations with similarities less than the threshold, as is in the case of “LaboratoryTestFinding” (CODO) and “Laboratory\_question” (COVIDCRFRAPID) with a similarity measure of 0.57, which is included due to the expert opinion that considers that both classes belong to the same knowledge domain. As a result, we



obtain the Tepuy-COVID Ontology (see Fig. 6) using Protégé. Furthermore, we validate the model using the Pellet reasoner (which comes with Protégé by default), analyzing the inconsistencies that can occur during the merging process. Pellet does the following tests: class hierarchy, object property hierarchy, data property hierarchy, class assertions, object property assertions, and same individuals [28]. If no errors are found in each test, then it is considered that the ontological model is consistent, and it is possible to infer knowledge, as it can be seen in Fig. 6 (see red circles) for the case of Fig. 5.

#### 4.3. Linking

Once we have obtained the Tepuy-COVID Ontology, we can extend the model to other domains through the Presence Ontology (PREO) ontology.<sup>4</sup> The process begins with the alignment of the Tepuy-COVID Ontology with the PREO ontology. Table 4 shows the alignment result using the **Levenshtein Distance method (EditDistNameAlignment)**, but again, we use the same metric for the alignment phase than for the merging case. The pair's similarity must exceed the threshold (0.8). For example, pairs of properties such as "has\_value" (Tepuy-COVID Ontology) and "hasValue" (PREO) represent nearby domains. In summary, the "Type" column shows the type of alignment (classes, property) found.

Then, using the Alignment API tool, we transform the alignment result into a format that links the ontologies involved with their respective alignments. For this process, the tool offers some methods.<sup>5</sup> We use OWL Axioms Renderer Visitor, since this method delivers a linking ontology that invokes the ontologies involved and the alignments between them. In addition, it is a compatible format with Protégé, which allows displaying and using the generated ontology [19].

Fig. 7 shows the alignment between Tepuy-COVID Ontology and PREO ontology and the generated linked ontology (COVID-19 Pandemic ontology). In the alignment ontology, we can observe how certain classes shared by the two origin ontologies complement their subclasses and properties in the resulting ontology. For example, the Person class of the Tepuy-COVID ontology aligns with the PREO ontology's Person class. Likewise, patient found in the two origin ontologies, it is enriched in the alignment ontology with two new subclasses, InPatient and OutPatient, which allow identifying whether the patient is hospitalized or not. Another case is the Nurse class, which became a subclass of Provider where Provider is a subclass of Health care Role, which confirms a better classification of the classes related to medical personnel according to their role. It is good to clarify that particularly in this process; the linking ontology of the two source ontologies is embedded in the alignment ontology that contains the links to the sub-trees of the source ontologies that complement the information between them.

#### 4.4. Validation

For the validation of our ontological models, we adapted the method described in Ref. [33], considering three types of validations:

- (i) Application of competence questions: to establish the coherence of the ontology information, resulting from the ontological engineering process.
- (ii) Quality validation through metrics: to determine the alignment precision of concepts, relationships and individuals, that is carried out during the ontological alignment process.
- (iii) Component validation using Protégé: to verify the ontology consistency resulting from the ontological mixing processes, in terms of their hierarchical and axiomatic structure.

<sup>4</sup> PREO, available online in <http://data.bioontology.org/ontologies/PREO/submissions/1/download?apikey=8b5b7825-538d-40e0-9e9e-5ab9274a9aeb>.

<sup>5</sup> <http://alignapi.gforge.inria.fr/builtin.html>.

##### 4.4.1. Application of competence questions

Competency questions are user-oriented questions to evaluate an ontology [42]. In other words, they are questions that users would want to have answers to by querying the ontology. Particularly, in this research, the competence questions verify the integrity of the data at the end of the mixing and linking processes. In the case of Tepuy-COVID ontology, Q1, Q2, and Q3 demonstrate that the result of the information provided by the Tepuy-COVID ontology is the same as the sum of the results of each ontology used in mixing. In the same way of COVID-19 Pandemic ontology, Q4 and Q5 demonstrate that the result of the information of the COVID-19 Pandemic ontology is the product of the linking of concepts present in the Tepuy-COVID ontology and the ontologies of other domains. For that, queries were executed in Protégé to determine the quality of the inferred information [33]:

##### 4.4.1.1. Quality of the Tepuy-COVID ontology.

- Q1: Which individuals belong to the Person class?  
R1: Persons
- Q2: Which individuals are diagnosed with COVID-19?  
R2: Person and ('has diagnosis' some 'COVID-19 Diagnosis')
- Q3: Which individuals have not tested for COVID-19?  
R3: Person and (hadCovidTest value false)

##### 4.4.1.2. Quality of COVID-19 pandemic ontology. - Competency Question: PREO.

- Q4: Which individuals are diagnosed with COVID-19 and who are on a stretcher?  
R4: Person and ('has diagnosis' some 'COVID-19 Diagnosis') and (hasFactor value Stretcher)

##### - Competency Question: PersonaOnto.

- Q5: Which individuals have not taken the COVID-19 test and are happy?  
R5: Person and (hadCovidTest value false) and (hasAffectiveState value Happy)

These competence questions are used in the experimentation of section 6.

##### 4.4.2. Quality validation through metrics

The quality validation through metrics of the new ontology begins from the phase of alignment of ontological models. The first step evaluates the alignments of the classes, properties, and instances obtained in the alignment phase. For this, we use the Precision, Recall, and F-Measure metrics defined in Refs. [7,21]. In order to use these measures, it is necessary to compare against a manual alignment (expert). For this purpose, two experts made the annotations of the alignments between the ontologies considered in the experimentation and the results were consolidated in a single description, using the Kappa index to solve the inconsistencies between the annotations of the experts according to the procedure indicated in Ref. [41]. Additionally, we use a threshold (0.8) to define the correspondences. In this way, we evaluate the alignments that we use both in the linking and merging processes.

Particularly, we define TP, FN, FP and TN to calculate Precision, Recall, and F-Measure, according to the comparison between the automatic alignments with the manual alignments (see an example in Table 5), which are explained below:

- **True Positives (TP):** true positives are matches that are recognized by both the manual and automatic approaches. In the case shown in

**Table 5**  
Examples of Automatic VS Manual Alignment using Threshold of 0.8

Example	Ontology A	RELATION STRENGTH	Ontology B	Automatic	Manual
TP	<a href="#">codo#Hospitalized</a>	1.0	<a href="#">whocovid19crfsemdatamodel#instances/Hospitalized</a>	True	True
FN	<a href="#">codo#LaboratoryTestFinding</a>	0.57142	<a href="#">whocovid19crfsemdatamodel#Laboratory_question</a>	False	True
FP	<a href="#">codo#SecondaryContact</a>	0.81	<a href="#">personasonto.owl#SecondaryType</a>	True	False
TN	<a href="#">codo#Tired</a>	0.6	<a href="#">whocovid19crfsemdatamodel#Piped</a>	False	False

**Table 6**  
Results of the alignment process.

Test	Precision	Recall	F-Measure
COVID-19 Pandemic Ontology	1.00	0.80	0.89

the example of [Table 5](#), the concept Hospitalized is the same in both ontologies and, it is true in the manual approach. Now, in the automatic approach, it generates a relation strength of 1.0, which is greater than the threshold of 0.8; therefore, it is also true;

- **False Negatives (FN):** false negatives are alignments in the manual approach but not automatically identified. In the example of [Table 5](#), the automatic approach generates a relation strength of 0.57; which is a false alignment because it is below the threshold of 0.8. In the manual approach, the expert considers that there is a match between the Laboratory Test Finding and Laboratory\_question concepts; therefore, it determines that it is true;
- **False Positives (FP):** the false positives are alignments falsely proposed by our automatic approach. In the example of [Table 5](#), the concepts Secondary Contact and Secondary Type are considered aligned (true) by the automatic approach since it indicates a relation strength of 0.81, which exceeds the threshold of 0.8. Now, in the manual approach, the expert determines that they are not aligned (false);
- **True negatives (TN):** the true negatives are false alignments, which the automatic approach has correctly discarded. In this case, Tired and Piped’s concepts are not aligned, so in manual approach is false. In the automatic approach is generated a relation strength of 0.6, lower than the threshold of 0.8; therefore, it is also false.

The Precision is the ratio between TP and the sum of FP and TP; Recall as the relation between TP and the sum of FN and TP and, finally, F-Measure as the measure of the connection between the Precision and the Recall. [Table 6](#) presents the results of the alignment validation obtained from the CODO, COVIDCRFRAPID, and PREO ontologies, using the Precision, Recall, and f-measure metrics provided by the Alignment API tool. In this case, the automatic alignment is when the Alignment API tool uses the Levenshtein measure with a threshold of 0.8. As we can see, the Precision in our automatic approach is outstanding, and the

Recall and F-measure values are affected by the number of FN and FP recognized by the automatic approach. Consequently, the automatic alignment presents a recall of 0.8. That is, the measure of similarity recognizes 80% of all correct alignments (completeness) [30].

4.4.3. *Component validation using Protégé*

In addition to the metrics mentioned above, we evaluate the new ontology to determine its consistency. In this case, we are using the reasoner Pellet from Protégé. As we see in [Fig. 8](#), the ontology does not show problems of inconsistencies. The classes, properties, and individuals of the three ontologies were correctly integrated into the ontology since there are no errors in the inferences made by the reasoner at the hierarchy level of classes, properties, objects, and individuals.

5. Experimentation

This section is divided into three parts. The first part shows the merging of ontologies from the same domain (section 5.1), obtaining the Tepuy-COVID Ontology. The second part shows the linking cases with ontologies from other fields (section 5.2), getting the COVID-19 Pandemic ontology. Finally, the third part presents the validations of the COVID-19 Pandemic ontology in different cases, and analyzes the results obtained from the entire process (section 5.3). The resulting ontologies from our ontological engineering process can be found in <http://bit.do/fSFy4>.

5.1. MERGING CASE: Tepuy-COVID ontology

For this stage, we chose five ontologies representative of the COVID-19 domain, which were all the ontologies that were available in the BioPortal repository about COVID (<https://bioportal.bioontology.org/>). Following our ontological engineering process, we obtained the **Tepuy-COVID Ontology** from merging selected ontologies with their respective classes and sources ([Table 7](#)). It is worth mentioning that these ontologies describe different sub-domains of knowledge, aiming to create a new ontology that represents the COVID-19, including the vast majority of the characteristics or concepts related to this disease.

The first step for the merging is to determine the possible alignments between the ontologies to be merged. However, since the Alignment API tool only obtains the alignments between two ontologies, and in this



Fig. 8. Results of validation with reasoners.

**Table 7**  
List of ontologies about COVID-19.

N	Ontology	Class	URL
1	COviD-19 Ontology for Cases and Patient information (CODO)	52	<a href="https://bioportal.bioontology.org/ontologies/CODO">https://bioportal.bioontology.org/ontologies/CODO</a>
2	COVID-19 Surveillance Ontology (COVID19)	52	<a href="https://bioportal.bioontology.org/ontologies/COVID19">https://bioportal.bioontology.org/ontologies/COVID19</a>
3	The COVID-19 Infectious Disease Ontology (IDO-COVID-19)	486	<a href="https://bioportal.bioontology.org/ontologies/IDO-COVID-19">https://bioportal.bioontology.org/ontologies/IDO-COVID-19</a>
4	COVID-19 Ontology (COVID-19)	2286	<a href="https://bioportal.bioontology.org/ontologies/COVID-19">https://bioportal.bioontology.org/ontologies/COVID-19</a>
5	WHO COVID-19 Rapid Version CRF semantic data model (COVIDCRFRAPID)	398	<a href="https://bioportal.bioontology.org/ontologies/COVIDCRFRAPID">https://bioportal.bioontology.org/ontologies/COVIDCRFRAPID</a>

case, there are more than two ontologies to be merged, it is necessary to combine in pairs between all the ontologies to find all the possible alignments. Table 8 shows the alignments according to each similarity technique with a threshold higher than 0.8 for the CODO and COVIDCRFRAPID ontologies. There is an important semantic similarity between classes and properties of ontologies; in such a way, the value of 1 in the four similarity measures indicates an exact match in the name of

**Table 8**  
Alignments detected by Align API with COVID-19 ontologies.

N	Ontology A	Ontology B	Type	Similarity Technique (strength)				
				1	2	3	4	Expert
1	<a href="#">codo#Negative</a>	<a href="#">whocovid19crfsemdatamodel#instances/Negative</a>	Instance	1.0	1.0	1.0	1.0	Yes
2	<a href="#">codo#Positive</a>	<a href="#">whocovid19crfsemdatamodel#instances/Positive</a>	Instance	1.0	1.0	1.0	1.0	Yes
3	<a href="#">codo#Hospitalized</a>	<a href="#">whocovid19crfsemdatamodel#instances/Hospitalized</a>	Instance	1.0	1.0	1.0	1.0	Yes
4	<a href="#">codo#VitalSigns</a>	<a href="#">whocovid19crfsemdatamodel#Vital_signs</a>	Class	0.91	1.0	-	-	Yes
5	<a href="#">codo#Disease</a>	<a href="#">whocovid19crfsemdatamodel#Liver_disease</a>	Class	-	0.87	-	-	Yes
6	<a href="#">codo#date</a>	<a href="#">whocovid19crfsemdatamodel#date_of_birth</a>	Property	-	0.86	-	-	Yes
7	<a href="#">codo#CountryWiseStatistics</a>	<a href="#">whocovid19crfsemdatamodel#Country</a>	Class	-	0.85	-	-	Yes
8	<a href="#">codo#NasalCongestion</a>	<a href="#">whocovid19crfsemdatamodel#Nasal_prongs</a>	Instance	-	0.81	-	-	Yes
9	<a href="#">codo#CoronavirusInfection</a>	<a href="#">whocovid19crfsemdatamodel#Coronavirus</a>	Class	-	0.91	-	-	Yes
10	<a href="#">codo#COVID-19</a>	<a href="#">whocovid19crfsemdatamodel#Coronavirus</a>	Class	-	-	-	1.0	Yes
<b>Total</b>				<b>4</b>	<b>9</b>	<b>3</b>	<b>4</b>	<b>10</b>

1. Levenshtein similarity; 2. SMOA similarity 3. Linguistic similarity using WordNet thesaurus 4. Linguistic similarity using the English Wordnet version.

classes and properties (for example, cases 1 to 3 of Table 8). Furthermore, in cases 4 to 9, the concepts have a lexical similarity (measure 2). Finally, in case 10, the semantic relationship between the concepts occurs due to the context to which they belong. The Wordnet thesaurus relates the two concepts as synonyms. In summary, alignment was achieved between classes (cases 4, 5, 7, 9 and 10), properties (case 6) and instances (cases 1, 2, 3 and 8).

Fig. 9 shows the process of merging the five ontologies to generate the Tepuy-COVID Ontology. As we can see, compelling cases of this process appear, for example, the addition of the *Coronavirus* and *Coronavirus infection* classes (obtained from the alignment process (see Table 5)) to the *disease caused by Coronavirus* class. Also, we can see how the *Nasal prongs* entity was aligned with *Nasal congestion*. In this way, we have achieved the Coronavirus domain model representing the union of the five ontologies that address different aspects of the disease.

5.2. Linking cases: different domains

In this case, the aim is to extend the domain of the Tepuy-COVID Ontology to complementary domains, exploiting the structure (classes, properties, individuals, etc.) of the ontologies to allow the exchange of information between them. Moreover, this domain expansion allows establishing generalizations and specializations for the Tepuy-COVID Ontology. Table 9 presents a list of ontologies from other domains

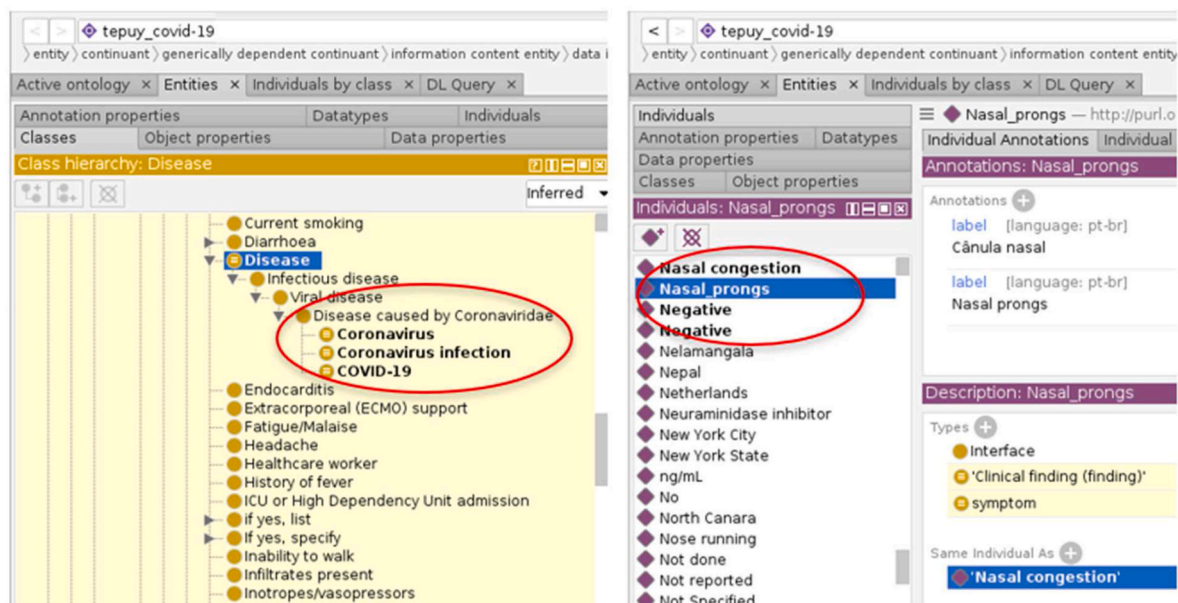


Fig. 9. Tepuy-COVID Ontology merging the five ontologies.

**Table 9**  
List of ontologies from different domains.

N	Ontology	Domain	Class	URL
1	Clinical MetaData Ontology (CMDO)	It models the Clinical MetaData on the different studies and clinical documents	188	<a href="https://bioportal.bioontology.org/ontologies/CMDO">https://bioportal.bioontology.org/ontologies/CMDO</a>
2	Provenance Ontology (PROVO)	It models a set of classes, properties, and constraints that can represent and exchange source information generated in different systems and contexts.	31	<a href="https://bioportal.bioontology.org/ontologies/PROVO">https://bioportal.bioontology.org/ontologies/PROVO</a>
3	Ontology of Medically Related Social Entities (OMRSE)	It models the scope of social entities related to health care, such as demographic information (social entities to record gender (but not sex) and marital status, for example) and the roles of various individuals and organizations (patient, hospital, etc.).	582	<a href="https://bioportal.bioontology.org/ontologies/OMRSE">https://bioportal.bioontology.org/ontologies/OMRSE</a>
4	Components Ontology (DoCO)	It models all the characteristics of a document, allowing the description of different aspects related to the content of scientific and other academic texts.	122	<a href="http://www.sparontologies.net/ontologies/doco">http://www.sparontologies.net/ontologies/doco</a>
5	Medical Technology Innovation in healthcare centers (ITEMAS)	It models the management of medical innovation technology in public health institutions.	134	<a href="https://bioportal.bioontology.org/ontologies/ITEMAS">https://bioportal.bioontology.org/ontologies/ITEMAS</a>
6	PersonasOnto	It models people, incorporating concepts and properties such as context, personality, emotions, affectivity, education, and experiences.	53	<a href="http://blankdots.com/open/personasonto.owl">http://blankdots.com/open/personasonto.owl</a>
7	Presence Ontology (PREO)	It models the encounters that take place each day between providers, patients, and family members or friends in settings such as hospitals and clinics.	276	<a href="https://bioportal.bioontology.org/ontologies/PREO">https://bioportal.bioontology.org/ontologies/PREO</a>
8	Publishing Roles Ontology (PRO)	It models the roles of agents (individuals, corporate entities, and computational agents) in the process of publishing scientific articles.	37	<a href="https://sparontologies.github.io/pro/current/pro.html">https://sparontologies.github.io/pro/current/pro.html</a>
9	DBpedia	Global and Unified Access to Knowledge.	828	<a href="https://wiki.dbpedia.org/Downloads2014">https://wiki.dbpedia.org/Downloads2014</a>

that are of interest for the growth of the Tepuy-COVID Ontology.

At the following are instantiated two case studies in which different ontological engineering processes will be applied to show the generation of the linking between Tepuy-COVID Ontology with other domains.

### 5.2.1. Case 1: Extension with personal information

- **Objective:** Tepuy-COVID Ontology enrichment with information related to people (Doctor, Nurse, Patient, Family members, etc.), incorporating concepts and properties such as context, personality, emotions, affectivity, education, experiences, among others. The enrichment result is reflected in the COVID-19 Pandemic Ontology.

- **Intervening Ontologies:** Tepuy-COVID Ontology (see section 5.1) and PersonasOnto (see Fig. 10).
- **Ontological Engineering Process:** aligning both ontologies and thus determining relationships between their classes and properties.

Table 10 shows the alignments found using the Alignment API tool with different similarity techniques and a threshold of 0.8, achieving alignment between classes (cases 1, 2, 5, 7, 8, 13, 15, 18, 19, 26, and 27), and properties (cases 3, 4, 6, 9–12, 14, 16, 17, 23–25, 28, and 29). Notably, there are an important number of alignments with an exact match of the names of the concepts and properties (cases 1 to 9). Cases 10 to 12 present a similarity according to three of the four measures used, where the relationships between of the concepts occur at the level of character strings (SMOA similarity measure) and linguistic similarity (3 and 4). We highlight that the expert discards a relationship between the concepts for cases 15, 17, 20, 23, 24, and 25 (where only a high similarity value is obtained in measure 2). However, the measure obtained exceeds the threshold. Regarding cases 26 to 29, a linguistic similarity of the concepts is obtained based on synonymy, detected by the two thesauri of WordNet used in measures 3 and 4. It should be mentioned that the absence of a value in the table represents that the obtained similarity value does not exceed the threshold, for which it is not considered. The second ontological engineering process links the ontologies since the alignments are transformed into a linking ontology describing their relationships.

In Fig. 10 is presented in Protégé the ontologies involved in this linking and the final result. In addition, it shows three specific cases where the Tepuy-COVID Ontology is enriched from the domain of the PersonasOnto ontology: **Organization Concept:** Tepuy-COVID Ontology is increased with the subclasses *Corporation*, *EducationalOrganization*, *GovernmentOrganization*, *SportsTeam*, among others. **Person Concept:** COVID-19 Pandemic Ontology is related to the *Participant* and *PersonaType* subclasses. **AffectiveState Concept:** Tepuy-COVID Ontology is augmented with the new *AffectiveState* class and its subclasses.

### 5.2.2. Case 2: extension of the Tepuy-COVID ontology with information for the hospital management

- **Objective:** Tepuy-COVID Ontology enriched with information related to the providers, patients' family members, and other aspects. The enrichment result can see it in the COVID-19 Pandemic Ontology.
- **Intervening ontologies:** Tepuy-COVID Ontology (see section 5.1.) and PREO (see Table 8).
- **Ontological Engineering Process:** As in the previous case, the first step is to apply an alignment between both ontologies to determine the relationship between their classes and properties (see section 4.1).

Table 11 indicates alignments found by the Alignment API using different similarity techniques and a threshold of 0.8. As we can see, most of the alignments between concepts are defined based on the measure of similarity 2 (SMOA similarity), according to the character strings that make them up. On the other hand, for cases 1 and 2, we see that the similarity measures 3 and 4 determine that the concepts are synonymous based on each thesaurus in the alignment.

Finally, we proceed to generate the linking ontology (see section 4.3). Fig. 11 shows the final ontology together with the two ontologies used in the linking process. In general, PREO enriches the Tepuy-COVID Ontology with many classes. Specifically, we can see in Fig. 11 the following: *Gender\_type* of Tepuy-COVID Ontology is increased with *Gender* of PREO ontology, acquiring the entire tree PREO's *Gender* subclasses, which includes classes such as *Female*, *Male* and *OtherGender* (*Gender-NonConforming* and *Transgender*). Tepuy-COVID Ontology receives new PREO concepts, such as *MaritalStatus*, *Race*, *ReligiousAffiliation*, *SexualOrientation*, *SocioEconomicStatus*, among many others.

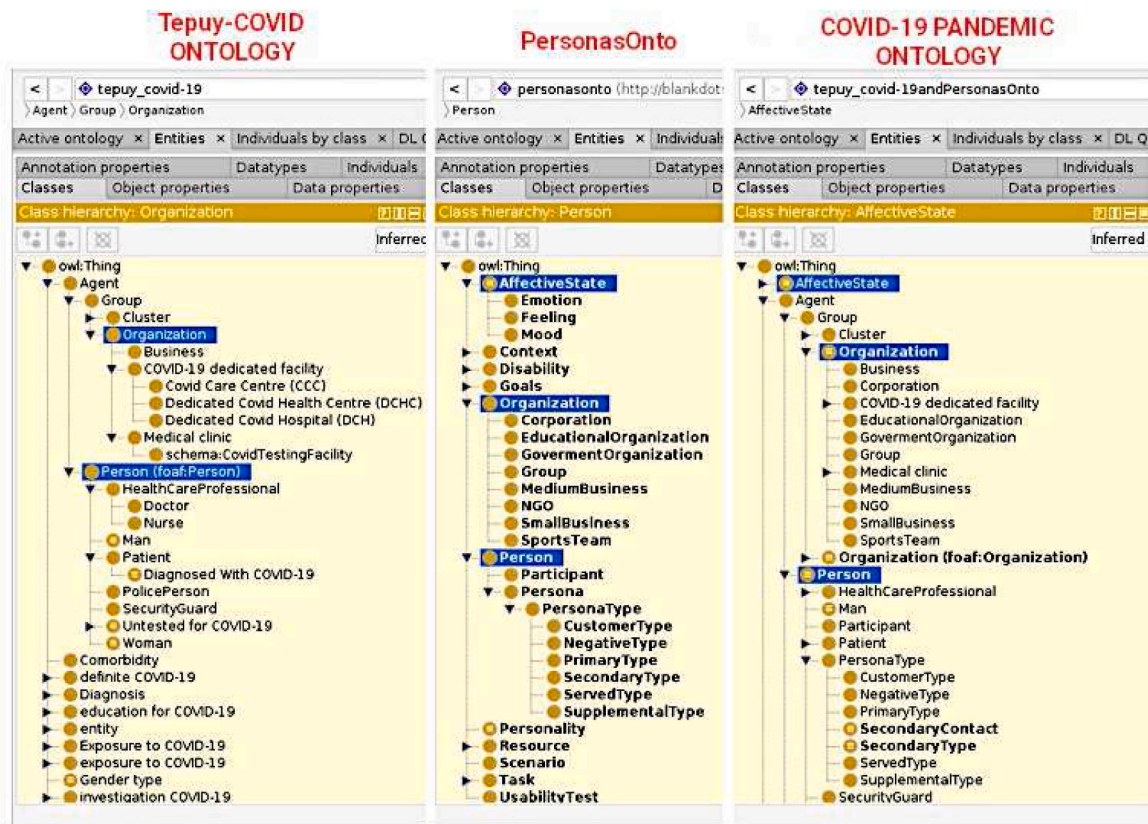


Fig. 10. Ontologies of the linking process in case 1.

Table 10  
List of alignments between Tepuy-COVID Ontology and PersonasOnto.

N	Tepuy-COVID Ontology	PersonasOnto	Type	Similarity Technique (strength)				
				1	2	3	4	Expert
1	codo#Group	personasonto.owl#Group	Class	1.0	1.0	1.0	1.0	Yes
2	foaf#rganization	personasonto.owl#Organization	Class	1.0	1.0	1.0	1.0	Yes
3	codo#age	personasonto.owl#age	Property	1.0	1.0	1.0	1.0	Yes
4	codo#status	personasonto.owl#status	Property	1.0	1.0	1.0	1.0	Yes
5	foaf#Person	personasonto.owl#Person	Class	1.0	1.0	1.0	1.0	Yes
6	schema#gender	personasonto.owl#gender	Property	1.0	1.0	1.0	1.0	Yes
7	whocovid19crfsemdatamodel#Person	personasonto.owl#Person	Class	1.0	1.0	1.0	1.0	Yes
8	schema#Place	personasonto.owl#Place	Class	1.0	1.0	1.0	1.0	Yes
9	codo#isMemberOf	personasonto.owl#isMemberOf	Property	1.0	1.0	1.0	1.0	Yes
10	codo#hasLocation	personasonto.owl#location	Property	-	0.92	0.84	0.84	Yes
11	codo#height	personasonto.owl#maxHeight	Property	-	0.9	0.8	0.8	Yes
12	codo#weight	personasonto.owl#maxWeight	Property	-	0.90	0.8	0.8	Yes
13	codo#Business	personasonto.owl#BusinessGoals	Class	-	0.93	-	-	Yes
14	codo#isLocationFor	personasonto.owl#location	Property	-	0.88	-	-	Yes
15	whocovid19crfsemdatamodel#Age	personasonto.owl#Image	Class	-	0.88	-	-	No
16	codo#hasState	personasonto.owl#hasAffectiveState	Property	-	0.88	-	-	No
17	codo#hasSon	personasonto.owl#hasPersonality	Property	-	0.86	-	-	No
18	codo#TimeCluster	personasonto.owl#Time	Class	-	0.86	-	-	Yes
19	codo#PolicePerson	personasonto.owl#Person	Class	-	0.85	-	-	Yes
20	codo#hasWife	personasonto.owl#hasLifeGoal	Property	-	0.84	-	-	No
21	foaf#name	personasonto.owl#taskName	Property	-	0.83	-	-	Yes
22	codo#hasMember	personasonto.owl#isMemberOf	Property	-	0.82	-	-	Yes
23	codo#hasCity	personasonto.owl#hasDisability	Property	-	0.81	-	-	No
24	codo#SecondaryContact	personasonto.owl#SecondaryType	Property	-	0.81	-	-	No
25	whocovid19crfsemdatamodel#has_specification	personasonto.owl#hasPublications	Property	-	0.8	-	-	No
26	foaf#Person	personasonto.owl#Persona	Class	-	-	0.92	0.92	Yes
27	whocovid19crfsemdatamodel#Person	personasonto.owl#Persona	Class	-	-	0.92	0.92	Yes
28	codo#weight	personasonto.owl#minWeight	Property	-	-	0.8	0.8	Yes
29	codo#height	personasonto.owl#minHeight	Property	-	-	0.8	0.8	Yes
<b>Linkings Total</b>				<b>9</b>	<b>25</b>	<b>16</b>	<b>16</b>	<b>29</b>

1. Levenshtein similarity; 2. SMOA similarity 3. Linguistic similarity using WordNet thesaurus 4. Linguistic similarity using the English Wordnet version.

**Table 11**  
A partial list of alignments between Tepuy-COVID Ontology and PREO.

N	Tepuy-COVID Ontology	PREO	Type	Similarity Technique (strength)				
				1	2	3	4	Expert
1	<a href="#">codo#Doctor</a>	presence-ontology.org#ontology/Physician	Class	-	-	1.0	1.0	Yes
2	<a href="#">whocovid19crfsemadatamodel#Sex</a>	presence-ontology.org#ontology/Gender	Class	-	-	1.0	1.0	Yes
3	<a href="#">codo#DiagnosisType</a>	presence-ontology.org#ontology/Diagnosis	Class	-	0.95	0.82	0.82	Yes
4	<a href="#">Schema#GenderType</a>	presence-ontology.org#ontology/Gender	Class	-	0.93	-	-	Yes
5	<a href="#">whocovid19crfsemadatamodel#Temperature</a>	presence-ontology.org#ontology/RoomTemperature	Class	-	0.92	0.85	0.85	Yes
6	<a href="#">codo#InfectedFamilyMember</a>	presence-ontology.org#ontology/FamilyMember	Class	-	0.88	-	-	Yes
7	<a href="#">codo#PolicePerson</a>	presence-ontology.org#ontology/Person	Class	-	0.85	-	-	Yes
8	<a href="#">codo#COVID-19Diagnosis</a>	presence-ontology.org#ontology/Diagnosis	Class	-	0.85	-	-	Yes
9	<a href="#">codo#InfectedDaughter</a>	presence-ontology.org#ontology/Daughter	Class	-	0.83	-	-	Yes
10	<a href="#">codo#InfectedNeighbor</a>	presence-ontology.org#ontology/Neighbor	Class	-	0.83	-	-	Yes
11	<a href="#">codo#InfectedBrother</a>	presence-ontology.org#ontology/Brother	Class	-	0.82	-	-	Yes
12	<a href="#">codo#InfectedMother</a>	presence-ontology.org#ontology/Mother	Class	-	0.8	-	-	Yes
13	<a href="#">codo#InfectedSister</a>	presence-ontology.org#ontology/Sister	Class	-	0.8	-	-	Yes
14	<a href="#">codo#InfectedFather</a>	presence-ontology.org#ontology/Father	Class	-	0.8	-	-	Yes
15	<a href="#">codo#InfectedSpouse</a>	presence-ontology.org#ontology/Spouse	Class	-	0.8	-	-	Yes
16	<a href="#">codo#contains</a>	presence-ontology.org#ontology/contains	Property	1.0	1.0	1.0	1.0	Yes
17	<a href="#">codo#containedIn</a>	presence-ontology.org#ontology/contains	Property	-	0.89	-	-	Yes
<b>Linkings Total</b>				<b>13</b>	<b>54</b>	<b>24</b>	<b>24</b>	<b>65</b>

1. Levenshtein similarity; 2. SMOA similarity 3. Linguistic similarity using WordNet thesaurus 4. Linguistic similarity using the English Wordnet version.

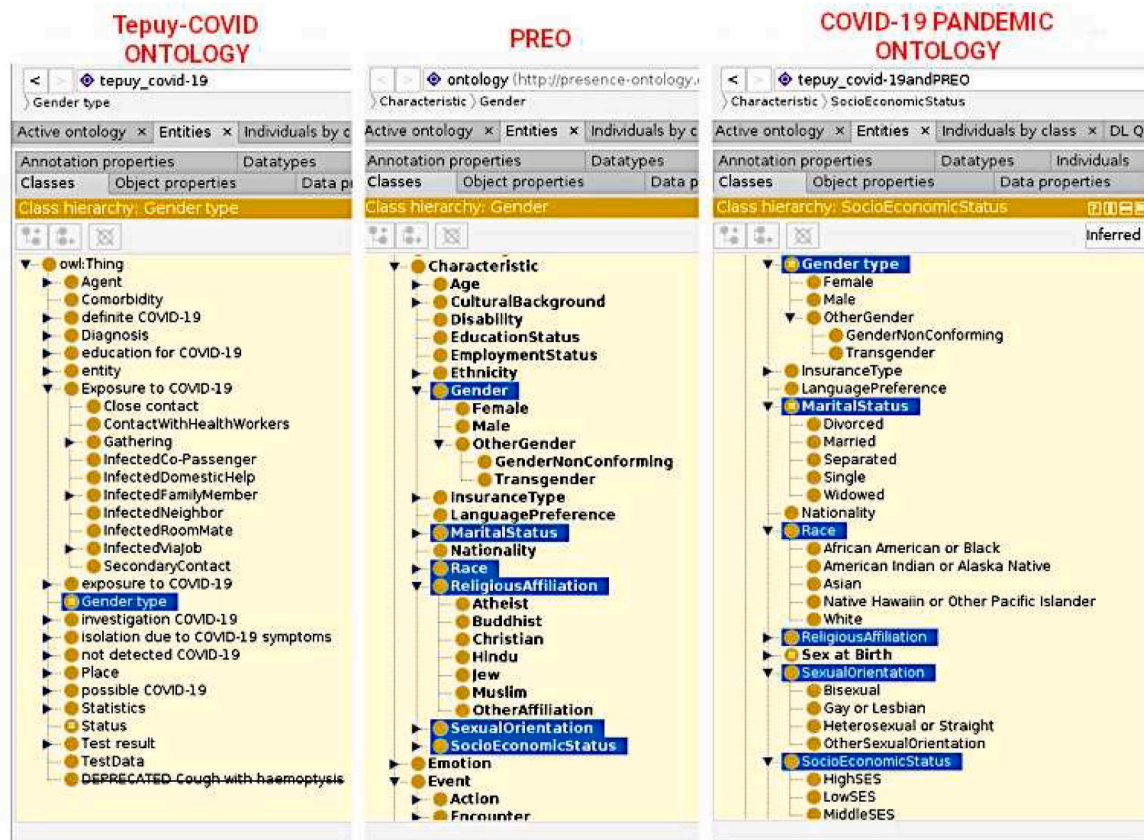


Fig. 11. Ontologies of the linking process in case 2.

## 6. Discussion

This section presents the discussion of the results of our research organized as follows: first, we will talk about the findings made during the merging process of the five domain ontologies of COVID-19 to obtain the Tepuy-COVID Ontology (section 5.1). Next, we will describe the results obtained in linking Tepuy-COVID Ontology with ontologies of other domains (see Table 9) and their performance results.

### 6.1. Merging process to obtain the Tepuy-COVID ontology

Table 12 presents the application of competence questions Q1, Q2, and Q3 to the Tepuy-COVID Ontology. As can be seen for each question, the individuals are maintained by mixing the base ontologies. The new ontology has the individuals with the CODO and COVIDCRFRAPID ontologies since the rest of the ontologies do not provide more individuals.

In Fig. 12, the result of the Q2 competence question in DL Query shows three individuals who coincide with a diagnosis associated with

**Table 12**  
Individuals present in Tepuy-COVID Ontology from the base ontologies.

Ontology	Question 1	Question 2	Question 3
CODO	P000002Diagnosis, p000001, p000002, p000003, p000004, p000005, p000006, p000007, p000008, p000009, p000010, p000011, p000012, p000013, p000014, p000015, p000016, p000017, p000018, p000019, p000020, p99992, p99993	p000001, p000002, p000003	P000002Diagnosis, p000004, p000005, p000006, p000007, p000008, p000010, p000012
COVIDCRFRAPID	'Bruna dos Santos Vieira', 'Jeroen A.M. Belien', 'L.O. Bonino da Silva Santos', 'Lana Meiqari', 'Marc Nieuwland', 'Rajaram Kaliyaperumal'	n/a	n/a
Tepuy-COVID Ontology	P000002Diagnosis, p000001, p000002, p000003, p000004, p000005, p000006, p000007, p000008, p000009, p000010, p000011, p000012, p000013, p000014, p000015, p000016, p000017, p000018, p000019, p000020, p99992, p99993, 'Bruna dos Santos Vieira', 'Jeroen A.M. Belien', 'L.O. Bonino da Silva Santos', 'Lana Meiqari', 'Marc Nieuwland', 'Rajaram Kaliyaperumal'	p000001, p000002, p000003	P000002Diagnosis, p000004, p000005, p000006, p000007, p000008, p000010, p000012

COVID-19. In addition, on the right, it is shown all the characteristics that the individual p000001 with COVID-19 has, for example, the symptoms that it has (Fever and Upper Respiratory Tract Infection), gender (Male), among others.

Table 13 presents the quality validation through metrics using Precision, Recall, and F-measure measures obtained in the alignments

between the five domain ontologies of COVID-19 in Table 7. The precision values for the four similarity measures indicate that they perform the alignments correctly since “a perfect precision score of 1.0 means that every correspondence computed by the algorithm was correct (correctness)” [30]. However, the same does not happen with the recall values of the four similarity measures. SMOA similarity technique presents a value of 0.9, that is, 90% of all correct correspondence (F-Measure of 0.95), which gives greater completeness to the alignments obtained with this measure than those identified by the Levenshtein measure. Note that SMOA analyzes commonalities and differences between concepts [26], while Levenshtein (recall 0.40) only establishes the calculation based on the sets of similar strings between concepts [7]. On the other hand, the similarity measures based on the WordNet thesauri present a recall of 0.30 and 0.40, which shows that only 30% of the ontology concepts were found in the thesauri.

On the other hand, the consistency component validation of the ontological model carried out by the Pellet reasoner indicates the presence of an inconsistency in the alignment [codo#date whocovid19crfsemdata#date\_of\_birth], caused by the difference between the type of ontological element to be aligned (the first is an individual and the second is a subclass). Thus, this alignment is removed from the model to ensure the ontology’s integrity.

6.2. Linking process between the Tepuy-COVID ontology and ontologies of other domains

Table 14 presents the application of competence questions Q4 and Q5 to the COVID-19 Pandemic Ontology. It is observed how in all linked ontologies; the individuals of the Tepuy-COVID Ontology are maintained (see Table 12). Thus, the ontologies resulting from the linking of Tepuy-COVID Ontology with each of the domain ontologies do not affect the integrity of this ontology despite having acquired knowledge of these domains.

6.2.1. Case Tepuy-COVID and PersonasOnto ontologies

To demonstrate the enrichment of the Tepuy-COVID Ontology with knowledge of the external PersonaOnto ontology, the following is required:

- Fill the class AffectiveState- > Emotion of PersonaOnto with individuals that represent the emotions, for this specific case, the individuals Happy and Sad were added.

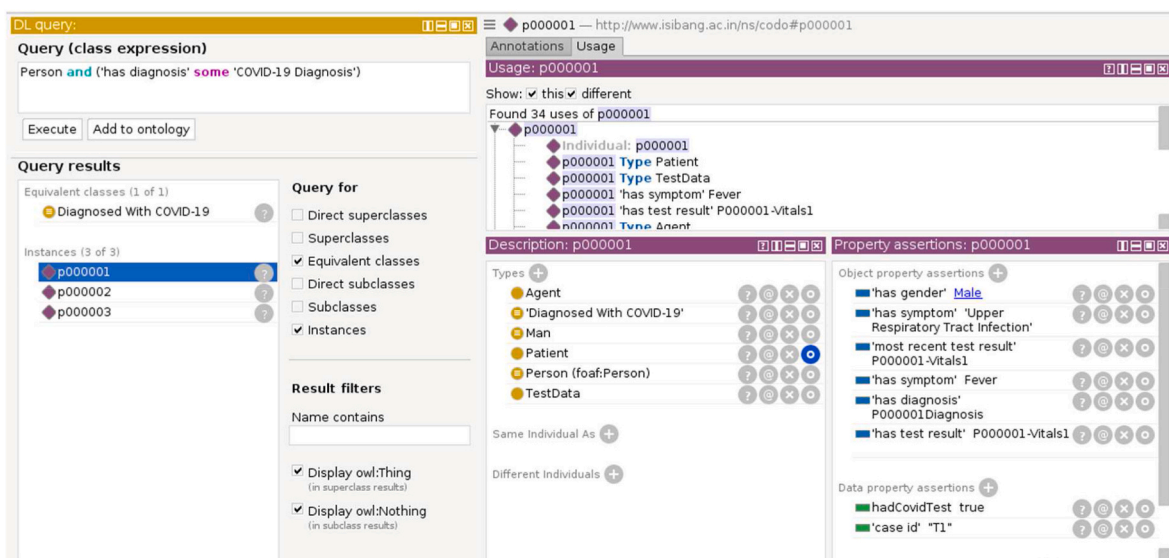


Fig. 12. Result of the Q2 competence question in DL Query.

**Table 13**  
Evaluating alignments for merging process using Alignment API.

Ontology	Levenshtein (1)			SMOA (2)			WordNet (3)			English-WordNet (4)		
	Prec	Rec	Fmeas	Prec	Rec	FMeas	Pre.	Rec	FMeas	Prec	Rec	Fmeas
<b>Tepuy-COVID Ontology</b>	1.00	0.40	0.57	1.00	0.90	0.95	1.00	0.30	0.46	1.00	0.40	0.57

**Table 14**  
Individuals obtained with the competence questions Q4 and Q5.

Ontology	Question 4	Question 5
Tepuy-COVID Ontology with CMDO	p000001, p000003, p000004, p000005	p000004, p000005, p000007, p000008, p000010,
Tepuy-COVID Ontology with PROVO		
Tepuy-COVID Ontology with OMRSE		
Tepuy-COVID Ontology with DoCO		
Tepuy-COVID Ontology with ITEMAS		
Tepuy-COVID Ontology with PersonasOnto		
Tepuy-COVID Ontology with PREO		
Tepuy-COVID Ontology with PRO		
Tepuy-COVID Ontology with Dbpedia		

- Relate the individuals of the Person class of Tepuy-COVID with the emotional states (hasAffectiveState) present in PersonaOnto, for this specific case, the Happy state was added to the individual p000004 and the Sad state to p000005.

In Fig. 13, two aspects are observed. On the right, the Affective State property belonging to the PersonaOnto ontology is presented, with a domain in Person and Range in Affective State. And on the left, the result of the execution of the competence question (Q5) is presented using the knowledge of both linked ontologies. On the one hand, it uses the Covid Test property of Tepuy-COVID, and on the other hand, it uses the Affective State property by PersonaOnto. The result of the query indicates that person p000004 has not undergone the COVID test and has a Happy emotional state.

### 6.2.2. Case Tepuy-COVID and PREO ontologies

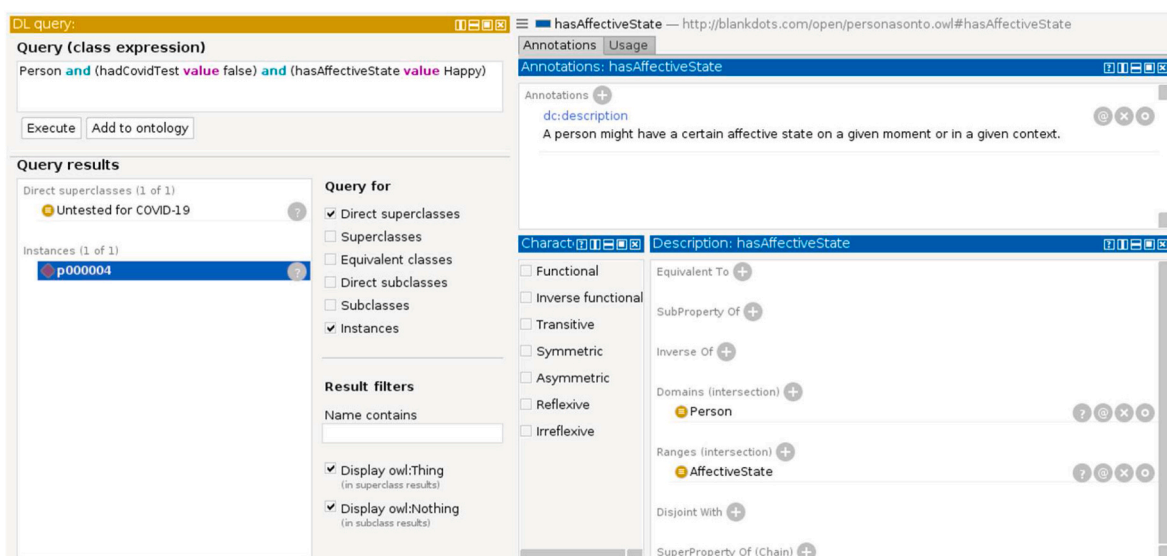
To demonstrate the enrichment of the Tepuy-COVID ontology with knowledge of the PREO external ontology, the following is required:

- Fill PREO’s Environment Factor: Health care System Factor class with individuals that represent factors in the health system, for this specific case, the individuals Wheelchair and Stretcher were added.
- Relate the individuals of the Person class of Tepuy-COVID with the environmental factors (hasEnvironmentFactor) present in PREO. For this specific case, the Wheelchair state was added to the individual p000003 and the Stretcher state to p000001.

In Fig. 14, two aspects are observed. On the right, there is the hasEnvironmentFactor property belonging to the PREO ontology, where the hasEnvironmentFactor is a subproperty of hasFactorByType. And on the left, the result of the execution of the competence question (Q4) is presented, which makes use of the knowledge of both linked ontologies. On the one hand, it uses the ‘has diagnosis’ property of Tepuy-COVID, and on the other hand, it uses the hasFactor property of PREO. The result of the query indicates that person p000001 has been diagnosed with COVID and is on Stretcher.

Table 15 shows the quality validation through metrics, starting with the results obtained from the linking processes between the Tepuy-COVID Ontology and the ontologies of other domains (Table 9 of section 5.2.) As we can see, a set of alignments is obtained (which comply with the threshold) over the total possible linking. Also, there are cases in which none of the similarity measures found possible alignments (for example, cases 1 and 3); in others, such as case 2, the alignments found are not enough because they did not exceed the threshold of 0.8.

For the other cases, it is observed that a large number of alignments were found (some repeated between methods), but few exceeded the defined threshold. For example, alignments between Tepuy-COVID Ontology and PersonasOnto exceeding a 0.8 threshold are 9 of 2.943, 25 of 2.925, 16 of 132.089, and 16 of 132.091. Thus, only ontologies 9, 25 and 16, respectively, met the threshold. The same happens with the



**Fig. 13.** Competency question specific to PersonaOnto.



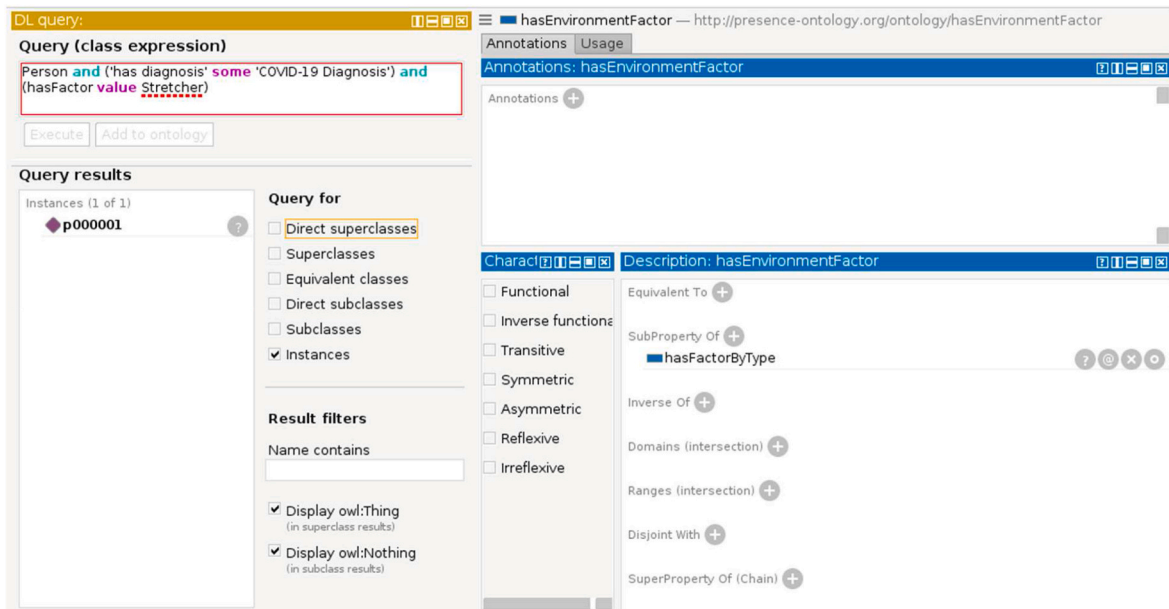


Fig. 14. Competency question specific to PREO.

**Table 15**  
Linking results of Tepuy-COVID Ontology with ontologies from other domains.

N	Ontologies	Linking Number by Similarity Technique				Linking accepted
		1	2	3	4	
1	Tepuy-COVID Ontology with CMDO	0 of 0	0 of 0	0 of 0	0 of 0	0
2	Tepuy-COVID Ontology with PROVO	0 of 345	0 of 54	0 of 481	0 of 481	0
3	Tepuy-COVID Ontology with OMRSE	0 of 0	0 of 0	0 of 0	0 of 0	0
4	Tepuy-COVID Ontology with DoCO	0 of 2.793	9 of 1.678	1 of 116.354	1 of 116.360	10
5	Tepuy-COVID Ontology with ITEMAS	1 of 3.329	6 of 3.045	3 of 368.336	3 of 368.343	9
6	Tepuy-COVID Ontology with PersonasOnto	9 of 2.943	25 of 2.925	16 of 132.089	16 of 132.091	29
7	Tepuy-COVID Ontology with PREO	13 of 2.946	54 of 2.930	24 of 638.459	24 of 638.482	65
8	Tepuy-COVID Ontology with PRO	0 of 2.886	0 of 554	0 of 20.523	0 of 20.523	0
9	Tepuy-COVID Ontology with Dbpedia	40 of 2.949	129 of 2.941	158 of 2.024.274	158 of 2.024.353	193

65 alignments used to link Tepuy-COVID Ontology and PREO ontologies.

Table 16 shows the Precision, Recall, and F-measure metrics for the nine cases of linking ontologies from other domains with the Tepuy-COVID Ontology. We observe that in cases where the precision value is 1, the alignments have been correctly recognized; even when there are cases in which none of the similarity measures found possible alignments (for example, cases 1 and 2). Similarly, the Recall measure for these cases is 1, which indicates that all possible alignments between the two ontologies have been recognized. In Case 4 for measure corresponding to Levenshtein, the Recall and F-measure values are zero.

Thus, no alignments exceed the threshold of 0.8 since, as observed in Table 15, 2,793 alignments have been achieved with this measure, but they are not relevant.

Recall and Metrics of cases 5, 6, 7, and 9 show that each method detects different concepts, which indicates that using the alignments of the other methods allows us to obtain a more significant number of links between Tepuy-COVID Ontology and the ontologies of other domains. Besides, similarity measures 3 and 4 are low (Wordnet and English-Wordnet). That is, in this case, the most significant value comes from the SMOA method.

Concerning results obtained in case 1 and case 2, we see that the SMOA similarity measure reaches a high recall value (0.86 for PersonasOnto and 0.83 for PREO, respectively), while similarity measures 3 and 4 reach recall values of 0.55 for PersonasOnto and 0.38 for PREO, respectively. The 55% and 38% of all ontology alignments have been found according to the similarities with the thesauri terms.

On the other hand, for the consistency component validation of the ontological model, case 4 presents inconsistencies found with the Pellet Reasoner, by linking the Tepuy-COVID Ontology with DoCO ontology (codo#Business with doco#Line), due to the difference between the two concepts. In case 6, connecting Tepuy-COVID Ontology with PersonasOnto ontology in the Place and Group concepts is a different type of ontological element (see Table 10, cases 1 and 8). Also, between Businesses with BusinessGoals (see Table 10, case 13), it generates a conflict with the alignment of the concept Organization. In case 7, it causes inconsistencies in the alignment of the concept Contains, although the concepts are the same in the two ontologies (see Table 11, cases 16 and 17). Finally, in the rest of the cases where Tepuy-COVID Ontology is linked to other domains, no inconsistencies were presented (cases 1, 2, 3, 5, 8, and 9), which means that consistency is maintained in the model at the level of its classes, objects, properties, and individuals.

## 7. Conclusions

This research presents an ontological engineering process for integrating ontologies related to the COVID-19 disease and other contexts for the treatment and representation of the large amounts of information generated by the pandemic. First, the set of disease domain ontologies and other domains are selected, performing alignment and merging processes between the COVID ontologies and mapping processes with ontologies from different contexts. Precision, Recall, and F-Measure

**Table 16**  
Evaluating Alignment of the Linking Process using Alignment API.

N	Ontology	Levenshtein (1)			SMOA (2)			WordNet (3)			English-WordNet (4)		
		Prec	Rec	FMeas	Prec	Rec	FMeas	Prec	Rec	FMeas	Prec	Rec	FMeas
1	Tepuy-COVID Ontology with CMDO	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2	Tepuy-COVID Ontology with PROVO	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
3	Tepuy-COVID Ontology with OMRSE	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
4	Tepuy-COVID Ontology with DOCO	1.00	0.00	0.00	1.00	0.90	0.95	1.00	0.10	0.18	1.00	0.10	0.18
5	Tepuy-COVID Ontology with ITEMAS	1.00	0.11	0.20	1.00	0.67	0.80	1.00	0.33	0.50	1.00	0.33	0.50
6	Tepuy-COVID Ontology with PersonasOnto	1.00	0.31	0.47	1.00	0.86	0.93	1.00	0.55	0.71	1.00	0.55	0.71
7	Tepuy-COVID Ontology with PREO	1.00	0.29	0.34	1.00	0.83	0.91	1.00	0.38	0.55	1.00	0.38	0.55
8	Tepuy-COVID Ontology with PRO	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
9	Tepuy-COVID Ontology with DBpedia	1.00	0.19	0.32	1.00	0.61	0.76	1.00	0.62	0.77	1.00	0.62	0.77

metrics evaluated each partial result, determining the quality and consistency of the resulting COVID-19 Pandemic Ontology. The experimentation results confirm that the proposed method guarantees an adequate ontological generation. The correctness and completeness of alignment sets by the four similarity measures reach over 60%, which are finally used in the ontological model construction process.

Concerning the Tepuy-COVID Ontology, Precision reaches 100% in recognizing all correct alignments between the five ontologies of the COVID-19 domain. Similarly, recall gets 0.90 with the SMOA similarity measure, which implies that 90% of all alignments between the ontologies have been recognized. Furthermore, the resulting model has been validated using the Pellet reasoner, obtaining that the new model is consistent in terms of classes, properties, and integrated instances. In the literature, these ontologies of the COVID-19 domain have been used to represent the information of the pandemic. But these ontologies have not been combined or related to ontologies of other fields to represent complex contexts of the pandemic. Our proposal offers the COVID-19 Pandemic Ontology that mixes five ontologies from the COVID domain with nine ontologies from different disciplines, generating a model capable of representing various types of information about the pandemic. Thus, it will allow the utilization of a large amount of information generated by the pandemic from a comprehensive perspective.

Future researches are framed in improving the experiments with the use of other semantic sources such as domain thesauri in such a way as to increase the recognition of synonymous concepts. In addition, it is necessary to test other similarity techniques based on neural networks and deep learning to increase the number of recognized alignments. Similarly, other reasoners should validate the ontological model consistency.

Another important aspect is that many of these ontological models do not contain individuals in their structure. Thus, one of the possible challenges is to enrich the ontologies with information from the different contexts involved, with the help of Linked Data and Natural Language Processing techniques for the extraction and population of the ontological model. Some formal schemes have been developed [35], which allow automating this enrichment process of ontologies, which could be considered. Also, using approaches based on context ontologies [31,32] is another alternative approach. In addition, we will establish new metrics that allow us to validate the resulting model according to its completeness and robustness. With this, we will evaluate the consistency of the ontology and its capacity to represent the contexts linked in the new model.

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### References

- [1] Sun H, Fan W, Shen W, Xiao T. Ontology-based interoperation model of collaborative product development. *J Netw Comput Appl* 2012;35:132–44.
- [2] Li L-S, Dang Y-Z, J S, Ping J-Y. A multi-dimensional ontology model for product lifecycle knowledge management. *International Conference on E-Product E-Service and E-Entertainment*; 2010. p. 1–4.
- [3] Sun H, Fan W, Shen W, Xiao T. Ontology fusion in high-level-architecture-based collaborative engineering environments. *IEEE Trans. Syst. Man Cybern. B Cybern.: Systems* 2013;43:1–12.
- [4] Sun H, Fan W, Shen W, Xiao T, Chen X. Ontology maintenance in a hierarchical federated collaborative product development environment. *15th International Conference on Computer Supported Cooperative Work in Design* 2011;1:1–9.
- [5] Rijiang B, Xiaoyue W, Xiaofan Y. Study of the methods and tools for ontology integration. *Adv. Mater. Res.* 2011;187:716–27.
- [6] Wang X, Hu Z, Bai R, Mou Y. Automatic semantic retrieval and visualization model based on the integrated ontology library. *J Comput Inf Syst* 2010;6:139–45.
- [7] David J, Euzenat J, Scharffe F, Trojahn dos Santos C. The alignment API 4.0. *Semantic Web* 2011;2(1):3–10.
- [8] Erdmann M, Waterfeld W. Overview of the neon toolkit. In: *Ontology engineering in a networked world*. Berlin, Heidelberg: Springer; 2012. p. 281–301.
- [9] October Seganti A, Kaplański P, Zarzycki P. Collaborative editing of ontologies using fluent editor and orion. In: *International experiences and directions workshop on OWL*. Cham: Springer; 2015. p. 45–55.
- [10] Ameen A, Khan KUR, Rani BP. Semi-automatic merging of ontologies using protégé. *Int J Comput Appl* 2014;85(12).
- [11] October Massmann S, Raunich S, Aumüller D, Arnold P, Rahm E. Evolution of the COMA match system. vol. 814. In: *Proceedings of the 6th international conference on ontology matching*. CEUR-WS. org; 2011. p. 49–60.
- [12] September Stellato A, Turbati A, Fiorelli M, Lorenzetti T, Costetchi E, Laaboudi C, Keizer J. Towards VocBench 3: pushing collaborative development of thesauri and ontologies further beyond. 1937. In: *17th European networked knowledge organization systems workshop, NKOS 2017*. CEUR-WS; 2017. p. 39–52.
- [13] February Levenshtein VI. Binary codes capable of correcting deletions, insertions, and reversals. 10. In: *Soviet physics doklady*; 1966. p. 707–10. 8.
- [14] González-Eras A, Dos Santos R, Aguilar J. Análisis de las contradicciones en las competencias profesionales en los textos digitales usando lógica dialéctica. *Revista Iberoica de Sistemas e Tecnologías de Información*; 2020. p. 150–63. E27.
- [15] González-Eras A, Aguilar J. Determination of professional competencies using an alignment algorithm of academic profiles and job advertisements, based on competence thesauri and similarity measures. *Int J Artif Intell Educ* 2019;29(4): 536–67.
- [16] Aguilar J, editor. *Introducción a la Minería semántica*. Universidad Nacional Experimental del Táchira; 2018.
- [17] Harispe S, Ranwez S, Janaqi S, Montmain J. Semantic similarity from natural language and ontology analysis. *Synthesis Lectures on Human Language Technologies* 2015;8(1):1–254.
- [18] December Cross V. Semantic similarity: a key to ontology alignment. In: *OM@ ISWC*; 2018. p. 61–5.
- [19] November Euzenat J. An API for ontology alignment. In: *International semantic web conference*. Berlin, Heidelberg: Springer; 2004. p. 698–712.
- [20] September Dijkman R, Dumas M, Garcia-Banuelos L, Kaarik R. Aligning business process models. In: *2009 IEEE international enterprise distributed object computing conference*. IEEE; 2009. p. 45–53.
- [21] October Do HH, Melnik S, Rahm E. Comparison of schema matching evaluations. In: *Net. ObjectDays: international conference on object-oriented and internet-based technologies, concepts, and applications for a networked world*. Berlin, Heidelberg: Springer; 2002. p. 221–37.
- [22] Dutta B, DeBellis M. CODO: an ontology for collection and analysis of COVID-19 data. *arXiv preprint arXiv:2009.01210*. 2020.
- [23] He Y, Yu H, Ong E, Wang Y, Liu Y, Huffman A, Chen L. CIDO, a community-based ontology for coronavirus disease knowledge and data integration, sharing, and analysis. *Sci Data* 2020;7(1):1–5.
- [24] Yousefianzadeh O. COVID-19 ontologies and their applications in medical sciences: reviewing BioPortal. *Inf Technol* 2020;1(1):30–5.
- [25] Babcock S, Beverley J, Cowell L, Smith B. "The infectious disease ontology in the age of COVID-19", technical report. University at Buffalo; 2020.

- [26] Stoilos Giorgos, Stamou Giorgos, Kollias Stefanos. A string metric for ontology alignment. In: *Proceedings of the 4th international semantic web conference*. Berlin, Heidelberg: SpringerVerlag; 2005. p. 624–37.
- [27] May Dos Santos CT, Quaresma P, Vieira R. An API for multilingual ontology matching. In: *Proc. 7th conference on language resources and evaluation conference (LREC)*; 2010. p. 3830–5. No commercial.
- [28] Sirin E, Parsia B, Grau BC, Kalyanpur A, Katz Y. Pellet: a practical owl-dl reasoner. *J. Web Semantics* 2007;5(2):51–3.
- [29] December Khan MH, Jan S, Khan I, Shah IA. Evaluation of linguistic similarity measurement techniques for ontology alignment. In: *2015 international conference on emerging technologies (ICET)*; 2015.
- [30] Mascardi V, Locoro A, Rosso P. Automatic ontology matching via upper ontologies: a systematic evaluation. *IEEE Trans Knowl Data Eng* 2009;22(5):609–23.
- [31] Aguilar J, Jerez J, Exposito E, Villemur T. CARMiCLOC: context awareness middleware in cloud computing. In: *2015 Latin American computing conference (CLEI)*; 2015.
- [32] Aguilar J, Rodriguez J, Flores B J, Sola J, Bas A, Soria T, Ramirez P, Parrilla P. Value of repeated fine-needle aspiration cytology and cytologic experience on the management of thyroid nodules. *Otolaryngology-Head Neck Surg (Tokyo)* 1998; 119(1):121–4.
- [33] Puerto E, Aguilar J, López C, Chávez D. Using multilayer fuzzy cognitive maps to diagnose autism spectrum disorder. *Appl Soft Comput* 2019;75:58–71.
- [34] Mendonça M, Perozo N, Aguilar J. An approach for multiple combination of ontologies based on the ants colony optimization algorithm. In: *2015 Asia-Pacific conference on computer-Aided system engineering*; 2015. p. 140–5.
- [35] Mendonça M, Aguilar J, Perozo N. Application of category theory. *Ingénierie Des Systèmes Inf* 2018;23(2):11–38.
- [36] Rangel C, Aguilar J, Cerrada M, Altamiranda J. Int'l Conf. In: *An approach for the emerging ontology alignment based on the bees colonies*. Artificial Intelligence; 2015. p. 536–41.
- [37] Rangel CR, Altamiranda J, Cerrada M, Aguilar J. Procedure based on semantic similarity for merging ontologies by non-redundant knowledge enrichment. *Int J Knowl Manag* 2018;14(2):16–36.
- [38] Babcock S, Beverley J, Cowell L, Smith B. “The infectious disease ontology in the age of COVID-19”, technical report. University at Buffalo; 2020.
- [39] Dutta B, DeBellis M. CODO: an ontology for collection and analysis of COVID-19 data. *Proc. of 12th Int. Conf. on Knowledge Engineering and Ontology Development*. 2020. 2020.
- [40] He Y, Yu H, Ong E, Wang Y, Liu Y, Huffman A, Huang H, Beverley J, Hur J, Yang X, Chen L, Omenn G, Athey B, Smith B. CIDO, a community-based ontology for coronavirus disease knowledge and data integration, sharing, and analysis. *Sci Data* 2020;7(181).
- [41] January Gómez P, González-Eras A, Torres-Carrión P. Method for emotion corpus validation from the consensual identification of patterns in alzheimer’s patients. In: *International conference on information technology & systems*. Cham: Springer; 2018. p. 148–59.
- [42] Bandeira J, Bittencourt I, Espinheira P, Isotani SFOCA. A methodology for ontology evaluation. *Cornell university, Tech. Rep.*; 2016 [Online]. Available, <http://arxiv.org/abs/1612.03353>.