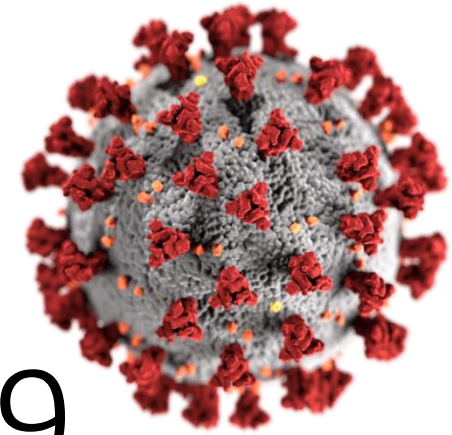




# Knowledge Graph and the current pandemic COVID-19



Dr. Biswanath Dutta

Associate Professor

Documentation Research and Training Centre

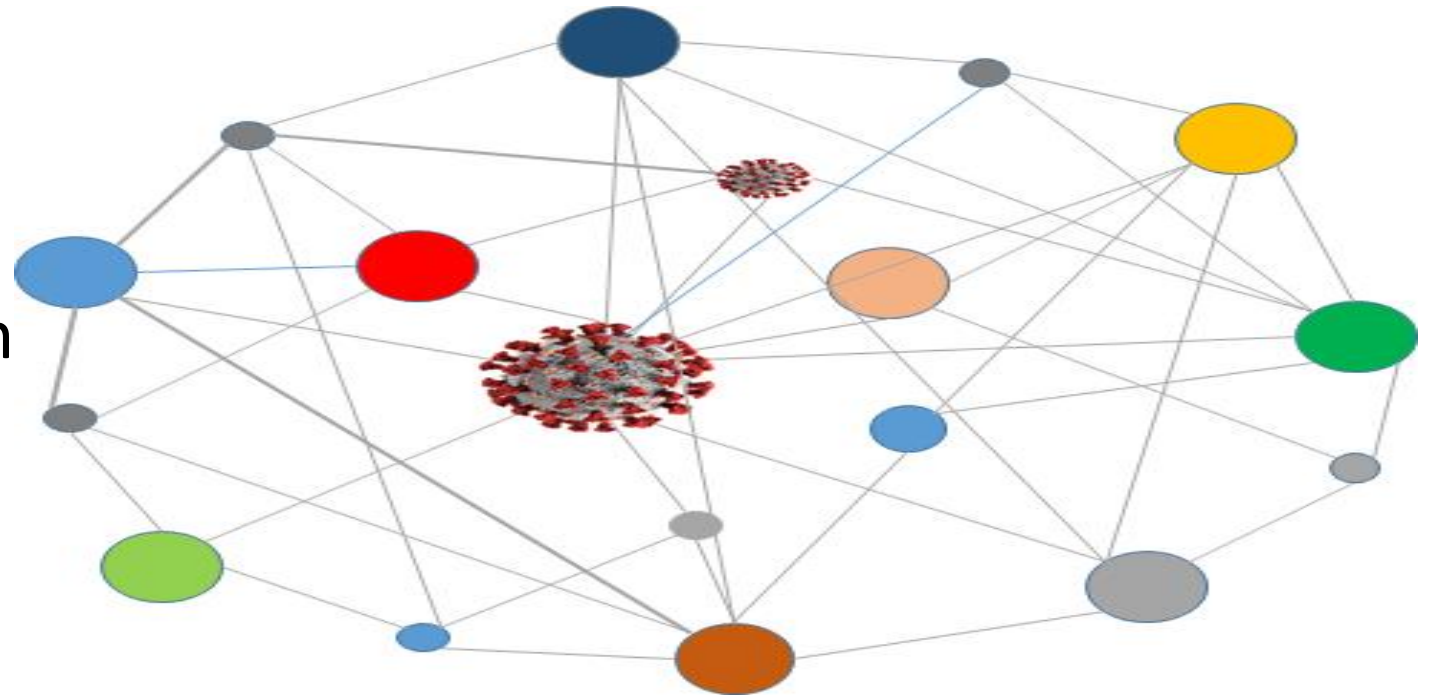
Indian Statistical Institute – Bangalore Centre

Bangalore 560059, INDIA

Email: [dutta2005@gmail.com](mailto:dutta2005@gmail.com), [bisu@drtc.isibang.ac.in](mailto:bisu@drtc.isibang.ac.in)

# Outline

- Knowledge Graph
- KG Core Technologies
- CODO Ontology
- CODO Knowledge Graph



# Background

A	B	C	D	E	F	G	H
Case	Date	Age	Sex	City	State	Cluster	Reason
1	2020-03-09T00:00:00	41	Male	Bangalore	Karnataka	From USA	Texas US
2	2020-03-10T00:00:00	0	Female	Bangalore	Karnataka	From USA	Spouse
3	2020-03-10T00:00:00	13	Female	Bangalore	Karnataka	From USA	Daughter
4		0		Bangalore	Karnataka	From Unit	London
5	2020-03-13T00:00:00	26	Male	Bangalore	Karnataka	From the	Greece
6	2020-03-12T00:00:00	76	Male	Kalburgi	Karnataka	From Mid	Saudi Ara
7		0			Karnataka	Unknown	No detail:
8	2020-03-17T00:00:00	32	Male	Bangalore	Karnataka	From Unit	Co passer
9	2020-03-17T00:00:00	63	Male	Kalburgi	Karnataka	From Unit	Co passer
10	2020-03-17T00:00:00	20	Female	Bangalore	Karnataka	From Unit	UK

## Situation by Country, Territory & Area

Name	Cases - cumulative total	Cases - newly reported in last 24 hours	Deaths - cumulative total
Global	23,518,343	206,382	810,492
United States o...	5,649,928	37,765	175,813
Brazil	3,605,783	23,421	114,744
India	3,167,323	60,975	58,390
Russian Feder...	966,189	4,696	16,568

## In-depth: Excess mortality

Excess mortality refers to the number of deaths *from all causes* above and beyond what we would have expected to see under 'normal' conditions.<sup>1</sup> In this case, we're interested in how deaths during the COVID-19 pandemic compare to the average number of deaths over the same period in previous years.

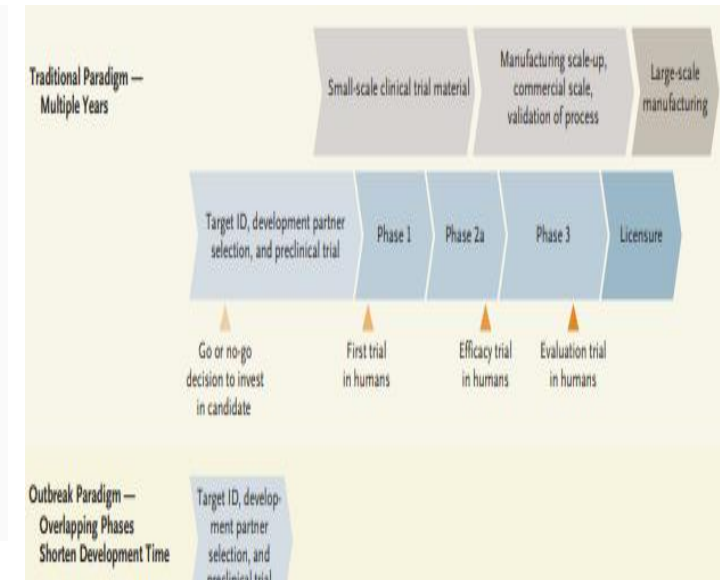
Looking at excess mortality is helpful for understanding the total impact of the pandemic on deaths – both direct and indirect. It helps us understand the direct impact by capturing deaths caused by COVID-19 that were not correctly diagnosed and reported, for example because no test for the virus was conducted. It helps us understand the indirect

# Data

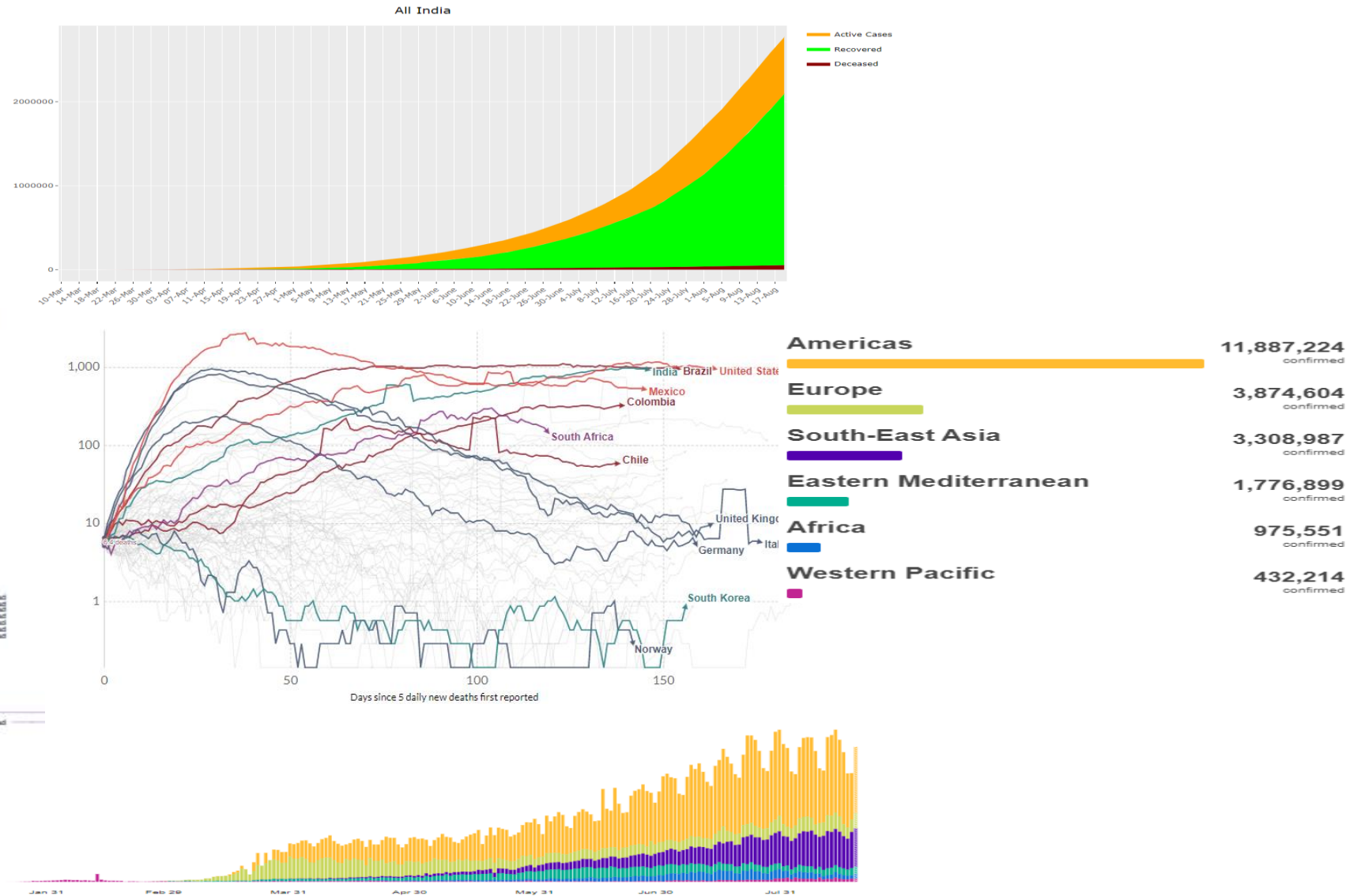
## P59595 · NCAP\_SARS

<b>Name</b>	Nucleoprotein
<b>Organism</b>	Severe acute respiratory syndrome coronavirus 2
<b>Gene</b>	Name N <span>1 Automatic Annotation</span> ORF names 9a
<b>Evidence</b>	1: Evidence at protein level
<b>Annotation score</b>	5/5

Dept. of Comp.Sc. & Eng., UVCE, Bangalore University  
(27Aug2020)



# Background



# Where are the problems?

# What we advocate for

User empowerment

Machine empowerment

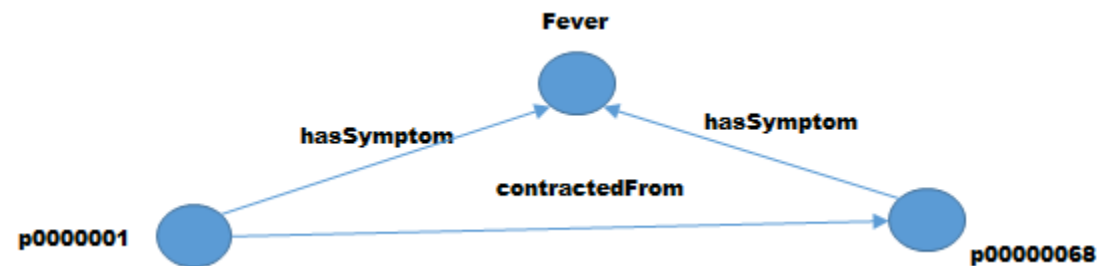
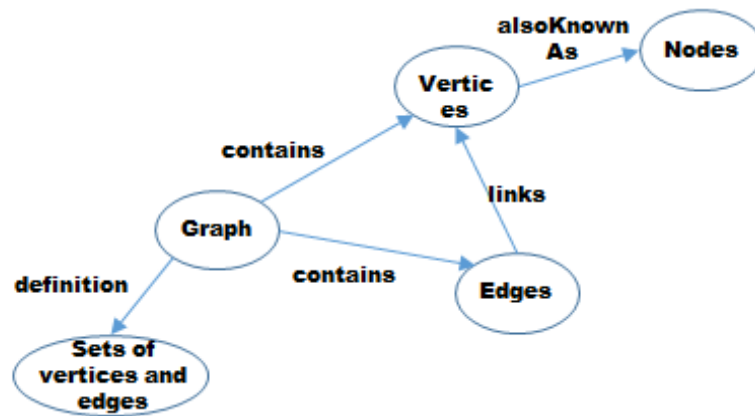
**Knowledge Graph Approach**

# What is Knowledge Graph?

# What is a Graph?

$$\text{Graph } G = (V, E)$$

where  $V$  is a set whose elements are called vertices (or, nodes), and  $E$  is a set of two-sets of vertices, whose elements are called edges (or, links)



A typical example of composition and decomposition technique.

  
[Bender et al., 2010; Graph, 2002]



# What is Knowledge Graph?

It is a manifestation of an intelligent **Web of Data** informed by an **ontology**.

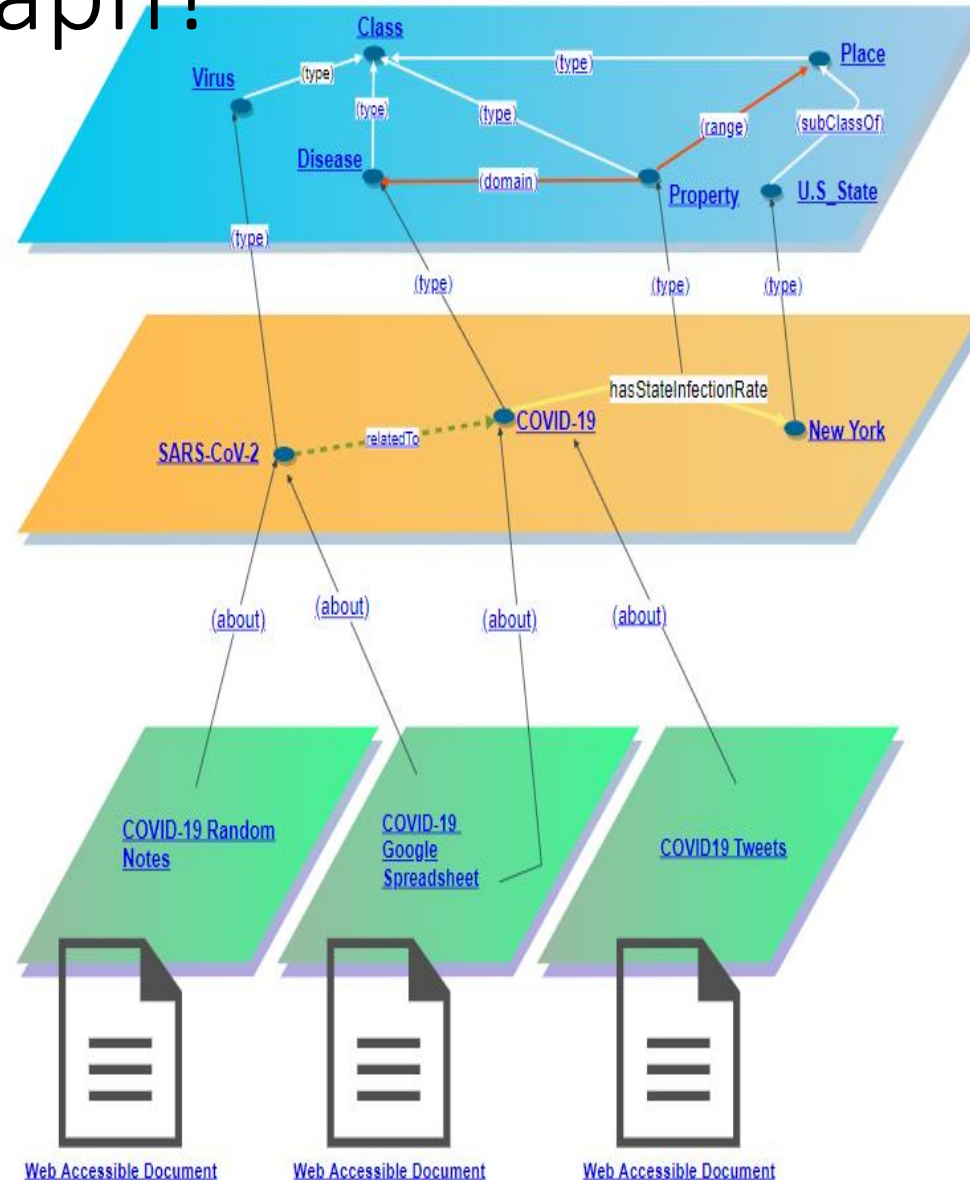
[Idehen, Kingsley U., 2020]



HyperLink Artifacts

HyperData

HyperText



Entity Definition Knowledge Graph or Ontology (Classes & Properties)

Entity Description Knowledge Graph (Instances)

Document Text (Literal Sentences)

Named Graphs

Document

# What is Knowledge Graph?

**KG** can be seen as a

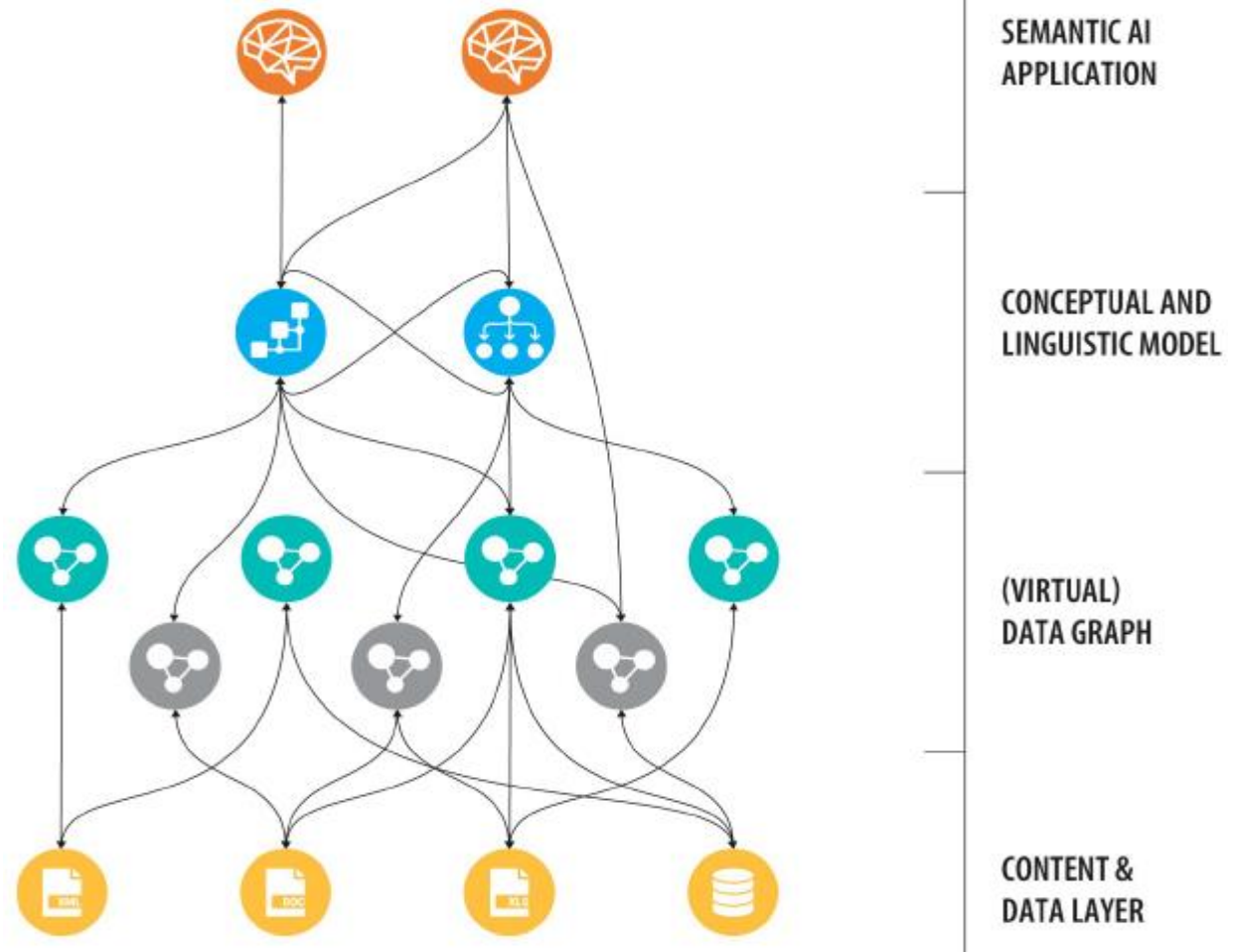
**Database:** that can be queried

**Graph:** that can be analyzed as  
a network of data

**Knowledge base:** new facts can  
be inferred

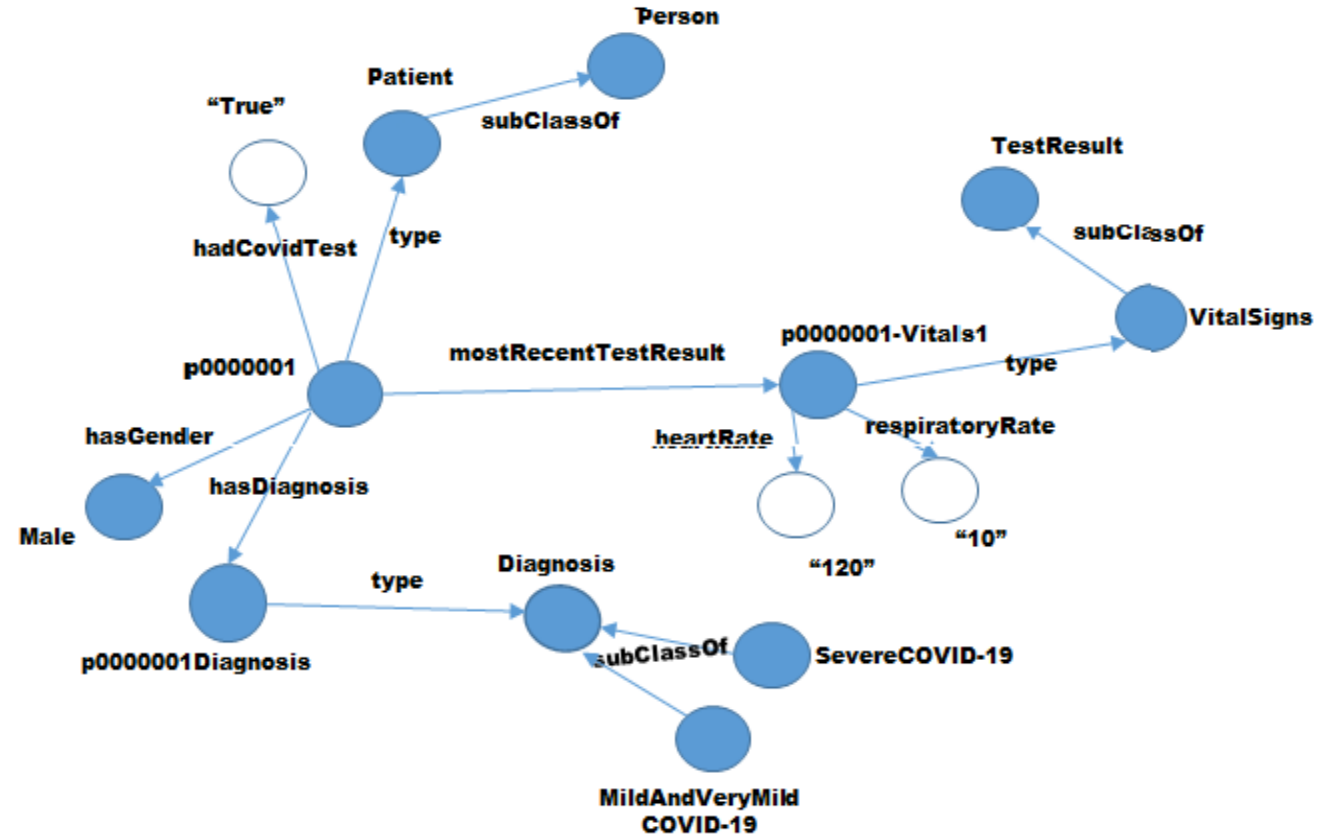


[Blumauer and Kiryakov, 2020]



# What is Knowledge Graph?

KG is a graph-structured representation of the world of human knowledge consisting of definitions and inter-relationships of the concepts and entities.



# Why Knowledge Graph?

# Why Knowledge Graph?

**KG Enables the search of Things (e.g., people, organization, place, event, artifacts).**

**Enables the retrieval of related information relevant to a query**

## Introducing the Knowledge Graph: things, not strings

Amit Singhal  
SVP, Engineering

Published May 16, 2012

Search is a lot about discovery—the basic human need to learn and broaden your horizons. But searching still requires a lot of hard work by you, the user. So today I'm really excited to launch the Knowledge Graph, which will help you discover new information quickly and easily.

Take a query like [taj mahal]. For more than four decades, search has essentially been about matching keywords to queries. To a search engine the words [taj mahal] have been just that—two words.

But we all know that [taj mahal] has a much richer meaning. You might think of one of the world's most beautiful monuments, or a Grammy Award-winning musician, or possibly even a casino in Atlantic City, NJ. Or, depending on when you last ate, the nearest Indian restaurant. It's why we've been working on an intelligent model—in geek-speak, a "graph"—that understands real-world entities and their relationships to one another: things, not strings.

visvesvaraya



All Images Maps News Books More Settings Tools

About 56,10,000 results (0.53 seconds)

en.wikipedia.org › wiki › M.\_Visvesvaraya

## M. Visvesvaraya - Wikipedia

Sir Mokshagundam **Visvesvaraya** KCIE FASc, more commonly known as Sir MV (15 September 1860 – 14 April 1962), was an Indian civil engineer and statesman and the 19th Diwan of Mysore, serving from 1912 to 1919. He received India's highest honour, the Bharat Ratna, in 1955.

**Nationality:** Indian **Born:** 15 September 1860; **Muddenahalli**, ...  
**Awards:** **Bharat Ratna** (1955) **Died:** 14 April 1962 (aged 101); **Bangalore**, ...

[Biography](#) · [Career timeline](#) · [Diwan of Mysore](#) · [Awards and honours](#)

### People also ask

- Who is the father of Engineering in India? ▾
- Why we celebrate Engineer's Day in India? ▾
- What did visvesvaraya invent? ▾
- Why Engineering Day is celebrated on 15th September? ▾

[Feedback](#)

### Videos

## M. Visvesvaraya

Indian civil engineer



Sir Mokshagundam Visvesvaraya KCIE FASc, more commonly known as Sir MV, was an Indian civil engineer and statesman and the 19th Diwan of Mysore, serving from 1912 to 1919. He received India's highest honour, the Bharat Ratna, in 1955. [Wikipedia](#)

**Born:** 15 September 1860, [Muddenahalli](#)

**Died:** 12 April 1962, [Bengaluru](#)

**Education:** [College Of Engineering Pune \(1883\)](#), [Central College Bangalore \(1881\)](#), [United Mission High School](#)

**Awards:** [Bharat Ratna](#)

**Parents:** [Mokshagundam Srinivasa Shastry](#), [Venkatalakshamma](#)

### Books



About 41,80,00,000 results (0.83 seconds)

en.wikipedia.org › wiki › Bangalore

### Bangalore - Wikipedia

Bangalore /bəŋgəˈlɔːr/, officially known as Bengaluru is the capital of the Indian state of Karnataka. It has a population of more than 8 million and a ...

Area code(s): +91-(0)80

Official language: Kannada

State: Karnataka

Country: India

[History of Bangalore](#) · [Bangalore Palace](#) · [Economy of Bangalore](#) · [Bangalore Fort](#)

### Top stories

[Karnataka Bengaluru Coronavirus Live Updates: BBMP allows only one Ganesha idol per ward this year](#)

The Indian Express · 6 hours ago

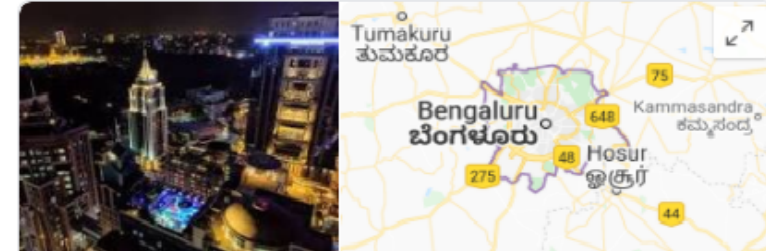


[Faulty results from private labs in Bengaluru on the rise](#)

Times of India · 19 hours ago



[Bengaluru violence: NIA to join hands with city police for further investigation](#)



## Bengaluru

City in Karnataka

Bengaluru (also called Bangalore) is the capital of India's southern Karnataka state. The center of India's high-tech industry, the city is also known for its parks and nightlife. By Cubbon Park, Vidhana Soudha is a Neo-Dravidian legislative building. Former royal residences include 19th-century Bangalore Palace, modeled after England's Windsor Castle, and Tipu Sultan's Summer Palace, an 18th-century teak structure.

**Area:** 709 km<sup>2</sup>

**Metropolitan area:** 8,005 km<sup>2</sup>

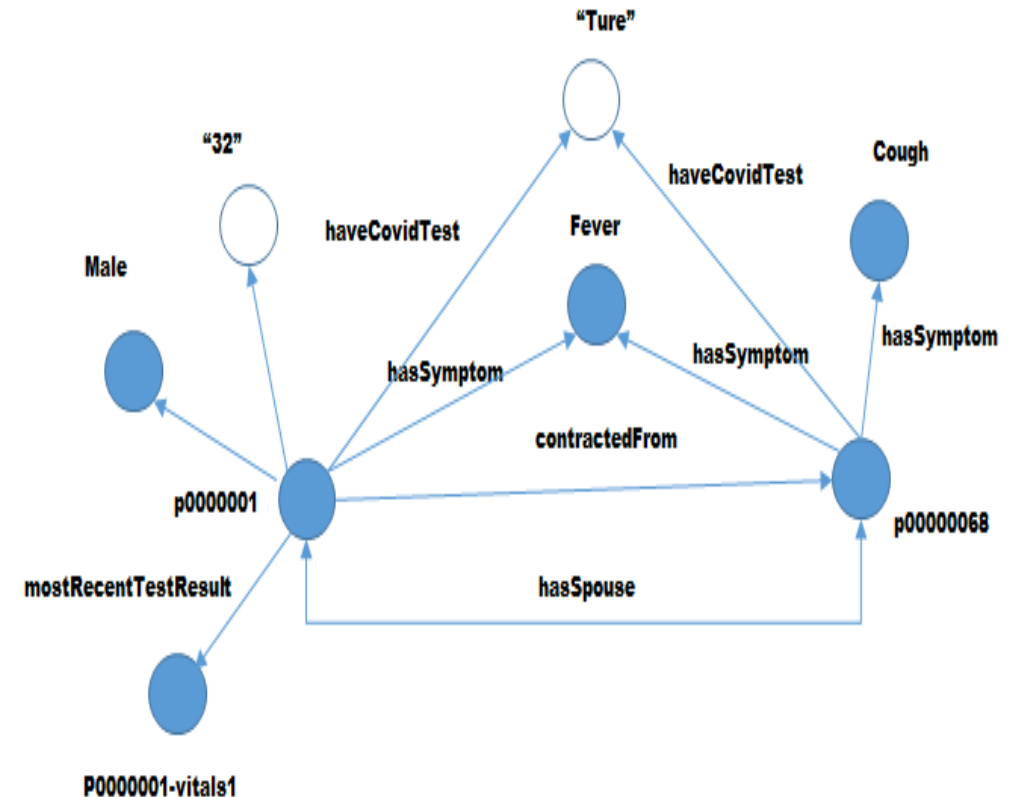


# Why Knowledge Graph?

Enables the capture and explicit expression of human knowledge by **connecting(linking)** the objects and their relationships.

A tool for connecting **various pieces of data** scattered across the silos of databases, text documents, etc.

Enables easy fusion and development of **context**





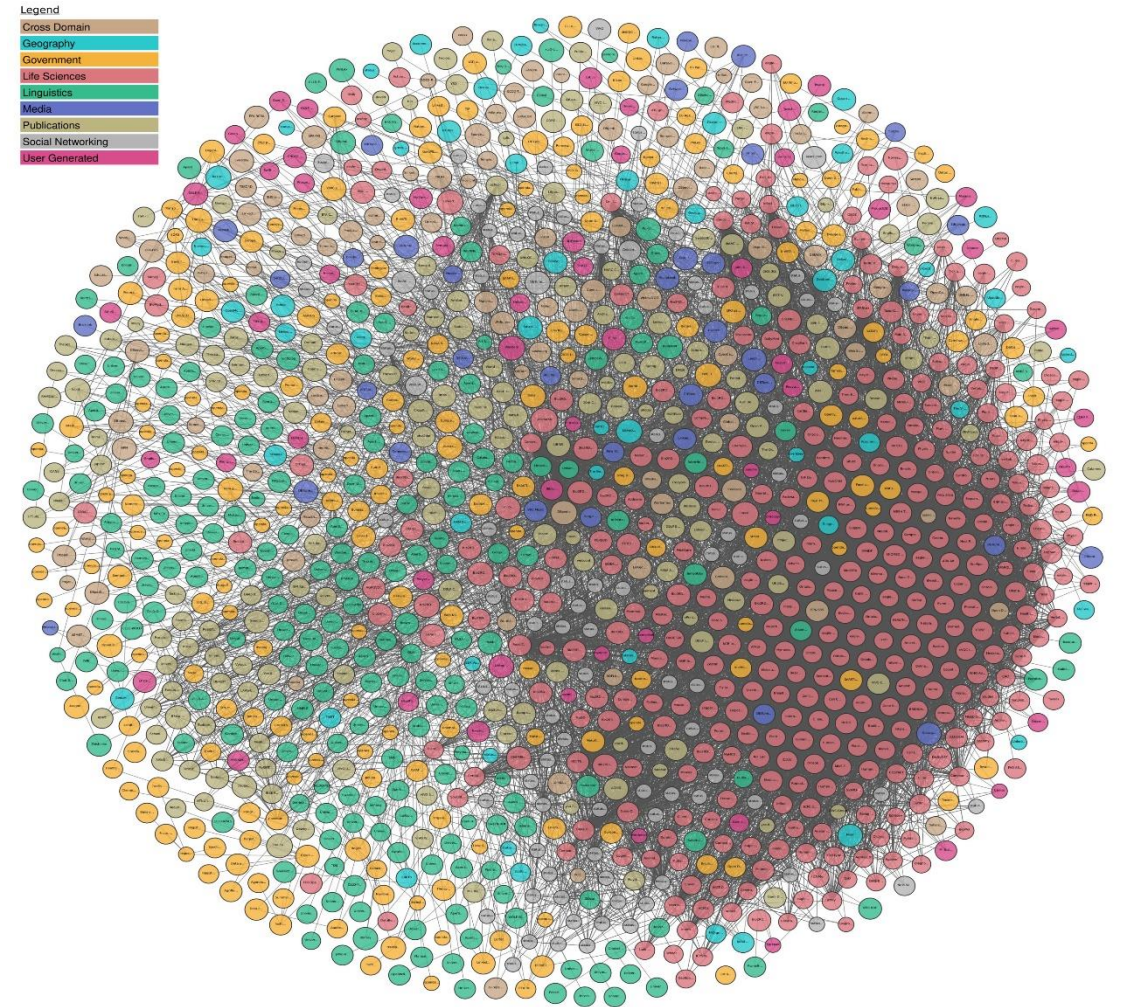
# Why Knowledge Graph?

Facilitate easy linking with the other external resources

**Implications:  
Enriched knowledge**

Creates a collaborative space towards building a **comprehensive** knowledge base (*graphs are by nature composable*)

Linking of all relevant information about the objects (e.g., enterprise knowledge space, education, health)

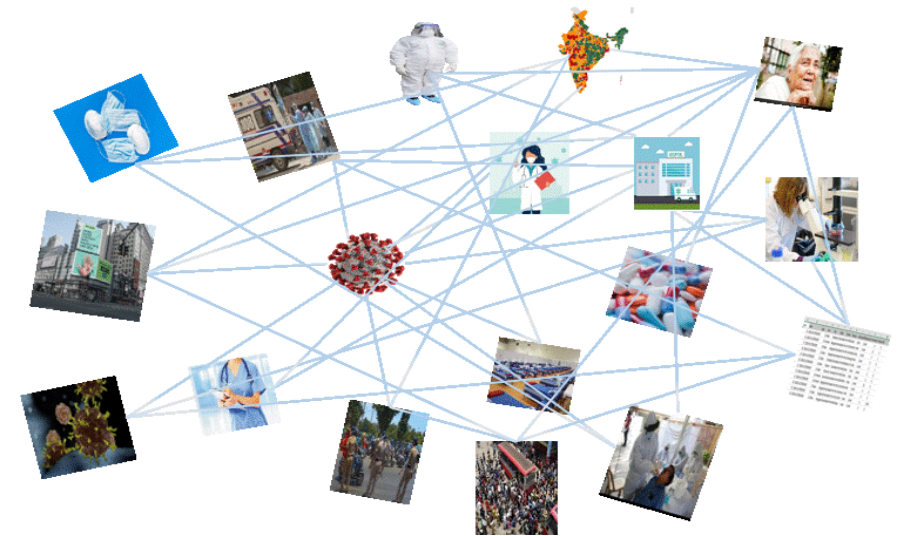


The Linked Open Data Cloud from [lod-cloud.net](http://lod-cloud.net)



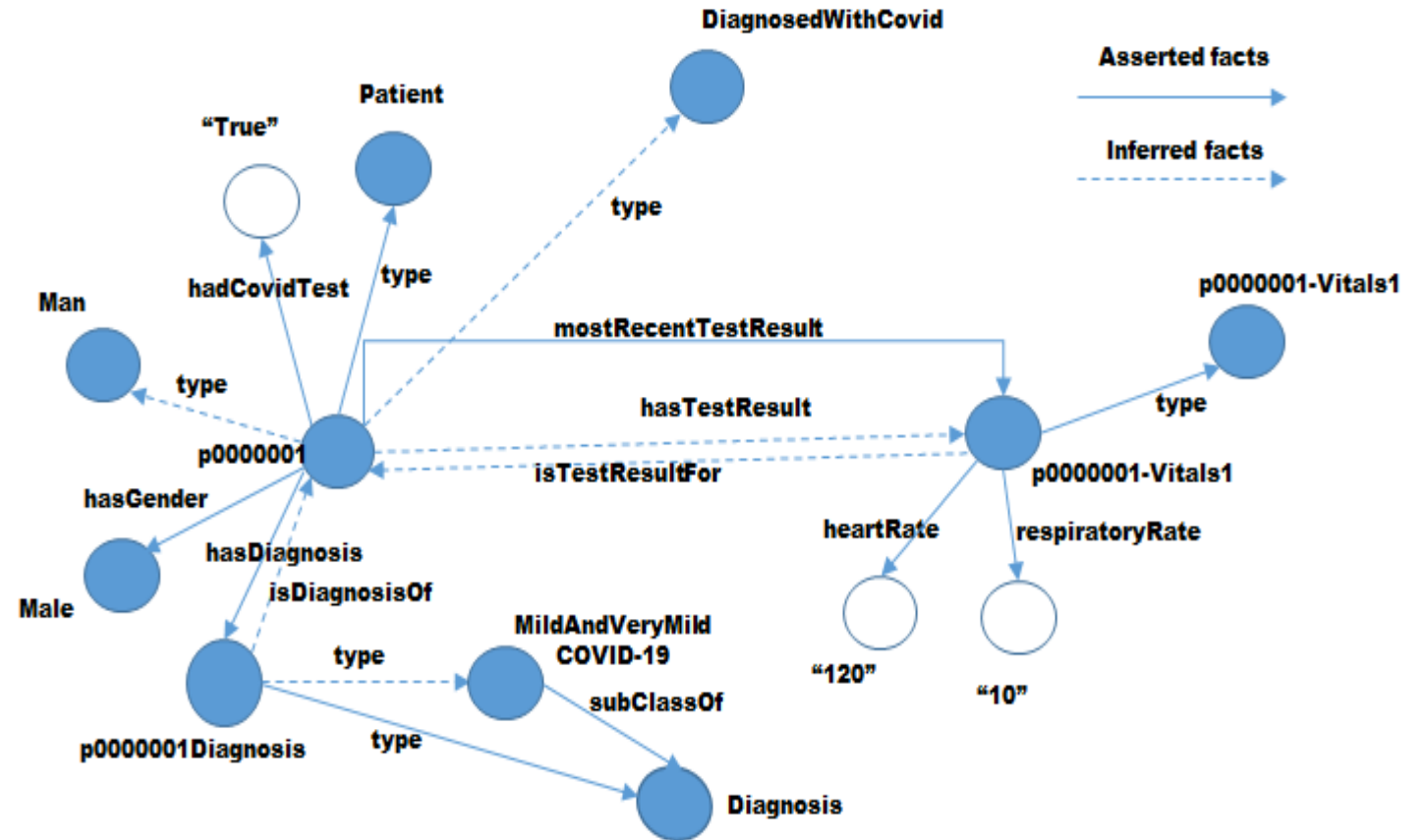
# Why Knowledge Graph?

A data model for learning heterogeneous knowledge



# Why Knowledge Graph?

A tool to **extract insight** from data by interlinking and analyzing







# Why Knowledge Graph?

How many patients had 'gallbladder calculus' in 2010 and later

## Simplified queries

[Aasman (2020)]



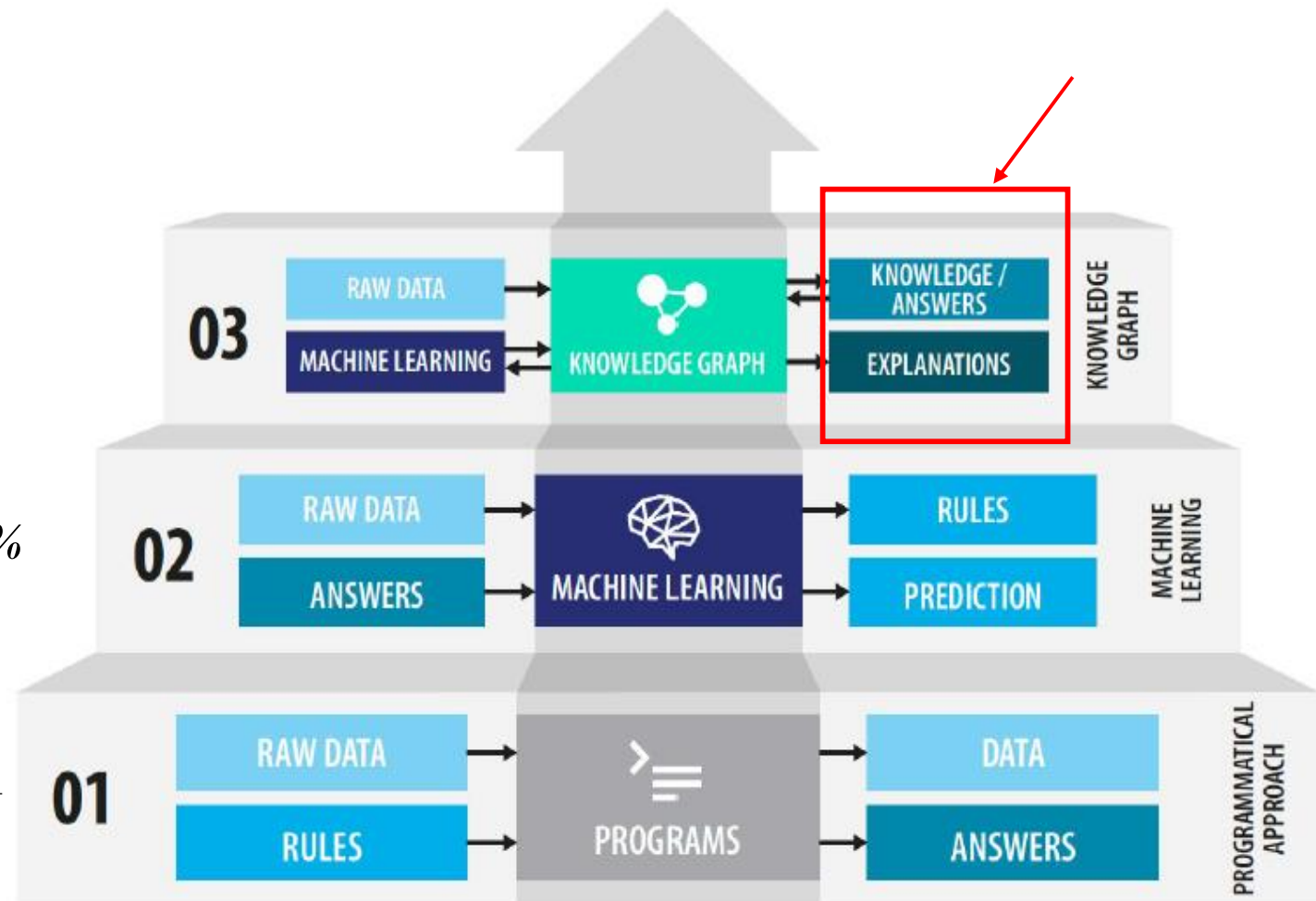
SPARQL	SQL
<pre>Select (count(distinct ?pt) as ?count) where {   ?pt cdm:encounter ?enc .   ?enc timew3c:begin/upper:value ?date ;   cdm:diagnosis/upper:correspondsTo/skos:exactMatch ?cui.   filter (?date &gt; '2010-01-01T00:00:00'^^xsd:dateTime)   ?cui mth:chd* mth:C0497327.}</pre>	<p><b>(Part 1 of 3)</b></p> <pre>Select dx.dxCode, dx.dxCount, cl.dxClassName, cl.ClassCount From (select distinct c.person_source_value as person, c.condition_source_concept_id as dxCode, f.concept_name as dxName, a.concept_id_2 as dxClass, e.concept_name as dxClassName from omopv5.concept_relationship a join cdrn.condition_occurrence c on a.concept_id_1=c.condition_source_concept_id join omopv5.concept_ancestor d on d.descendant_concept_id=a.concept_id_2 join omopv5.concept e on e.concept_id=d.ancestor_concept_id join omopv5.concept f on f.concept_id=d.descendant_concept_id where a.relationship_id='Maps to' and d.min_levels_of_separation=2 and e.standard_concept='S') dx JOIN (select distinct c.person_source_value as person, c.condition_source_concept_id as dxCode, f.concept_name as dxClassName from omopv5.concept_relationship a join cdrn.condition_occurrence c on a.concept_id_1=c.condition_source_concept_id join omopv5.concept_ancestor d on d.descendant_concept_id=a.concept_id_2 join omopv5.concept e on e.concept_id=d.ancestor_concept_id join omopv5.concept f on f.concept_id=d.descendant_concept_id where a.relationship_id='Maps to' and d.min_levels_of_separation=2 and e.standard_concept='S') cl on dx.dxClass=cl.dxClass order by cl.dxClass desc</pre>

# Why Knowledge Graph?

Forms a backbone for AI and analytics platforms

A ML algorithm can say "*person X has a Y% chance of their tumor being cancer*" but most ML algorithms can't explain why.

Integrating ML and KG is a way forward in addressing this issue.



[Blumauer and Kiryakov (2020)]

# KG usage: a quick review

- Web search
- Question answering
- Data integration
- Data collection and analysis
- Data visualization
- Machine learning and advanced analytics

# KG Technology



# Semantic Web

“*A web of data that can be processed directly and indirectly by machines*” – Tim Berners-Lee

An **extension**, not a replacement of the current web

A metadata based infrastructure for reasoning on the Web

Goal: provide a common framework to share data on the Web across application boundaries

Scientific American: Feature Article: The Semantic Web: May 2001



## The Semantic Web

A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities

by [TIM BERNERS-LEE](#), [JAMES HENDLER](#) and [ORA LASSILA](#)

**SUBTOPICS:**  
[Expressing Meaning](#)  
[Knowledge Representation](#)  
[Ontologies](#)  
[Agents](#)  
[Evolution of Knowledge](#)

**SIDEBARS:**  
[Overview / Semantic Web](#)

The entertainment system was belting out the Beatles' "We Can Work It Out" when the phone rang. With his phone turned the sound down by sending a message to all the other *local* devices that had a *volume* Lucy, was on the line from the doctor's office: "Mom needs to see a specialist and then has to have a series of sessions. Biweekly or something. I'm going to have my agent set up the appointments." Pete immediately agreed to share the doctor's office, Lucy instructed her Semantic Web agent through her handheld Web browser. The agent retrieved information about Mom's *prescribed treatment* from the doctor's agent, looked up services and checked for the ones *in-plan* for Mom's insurance within a *20-mile radius* of her home and *excellent* or *very good* on trusted rating services. It then began trying to find a match between a *times* (supplied by the agents of individual providers through their Web sites) and Pete's and Lucy's (The emphasized keywords indicate terms whose semantics, or meaning, were defined for the Semantic Web.)

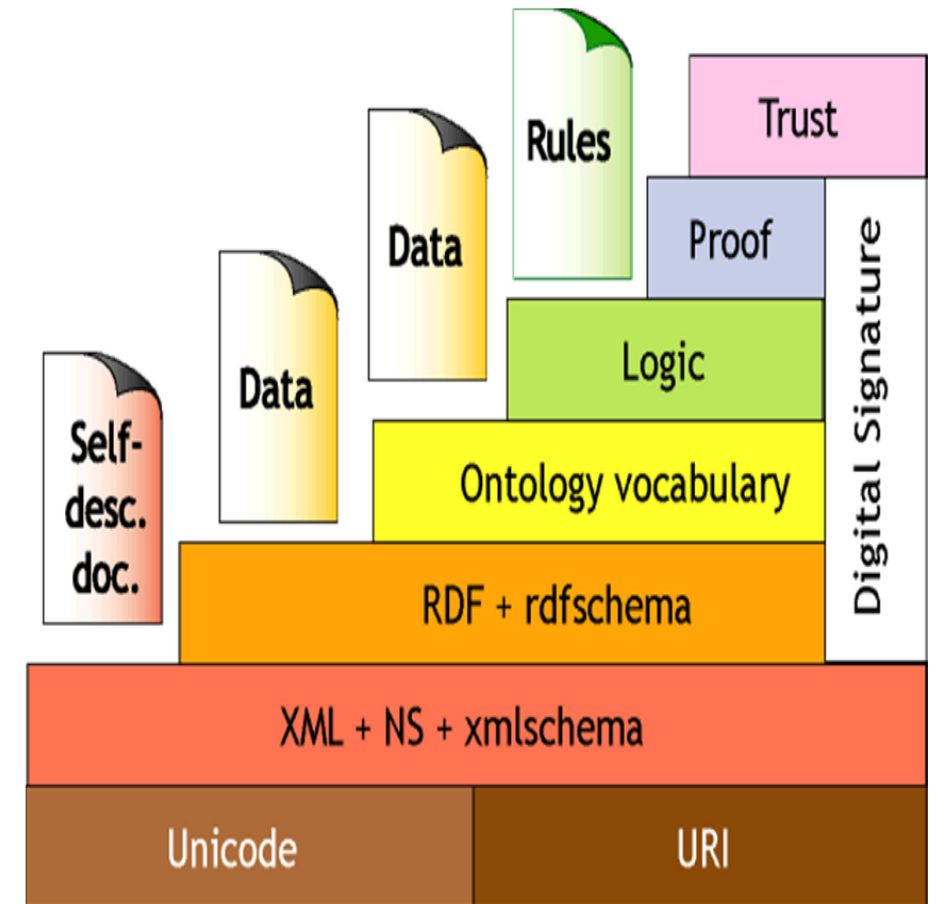


In a few minutes the agent presented them with a plan. Pete didn't like it—University Hospital was a 10-minute drive from Mom's place, and he'd be driving back in the middle of rush hour. He set his own agent with stricter preferences about *location* and *time*. Lucy's agent, having *complete trust* in Pete's agent in the context of the problem, automatically assisted by supplying access certificates and shortcuts to the data it had already sorted through.

# Technologies

## [W3C Standards]

- International Resource Identifier (IRI)
- Resource Description Framework (RDF/RDF Schema)
- Web Ontology Language (OWL)
- SPARQL Protocol and RDF Query Language (SPARQL)
- Semantic Web Rule Language (SWRL)
- Reasoner



# IRI

An IRI looks very much like a URL.

IRI's are more general than URLs and can describe resources to a finer level of granularity than an HTML page.

An IRI can be any resource such as a class, a property, an individual, etc.



[DuCharme, 2011]

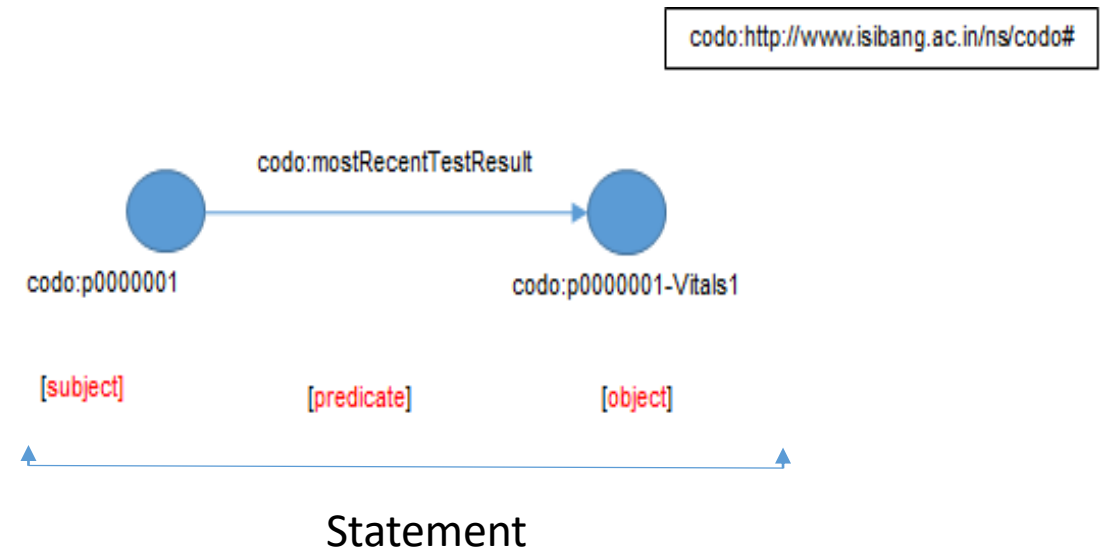
Statements with a subject of [d8eedc167d63eda1e272642595088335](http://www.isibang.ac.in/ns/codo#d8eedc167d63eda1e272642595088335)

Predicate	Object	
status	Hospitalized	<a href="http://www.isibang.ac.in/ns/codo#d8eedc167d63eda1e272642595088335">x</a>
nationality	India	<a href="http://www.isibang.ac.in/ns/codo#d8eedc167d63eda1e272642595088335">x</a>
rdf:type	owl:NamedIndividual	<a href="http://www.isibang.ac.in/ns/codo#d8eedc167d63eda1e272642595088335">x</a>
rdf:type	Patient	<a href="http://www.isibang.ac.in/ns/codo#d8eedc167d63eda1e272642595088335">x</a>
hasCity	Bangalore-Urban	<a href="http://www.isibang.ac.in/ns/codo#d8eedc167d63eda1e272642595088335">x</a>
gender	Male	<a href="http://www.isibang.ac.in/ns/codo#d8eedc167d63eda1e272642595088335">x</a>
rdfs:label	"patient11910"	<a href="http://www.isibang.ac.in/ns/codo#d8eedc167d63eda1e272642595088335">x</a>
clusterString	"27-June Trace History Absent"	<a href="http://www.isibang.ac.in/ns/codo#d8eedc167d63eda1e272642595088335">x</a>
hasCausedAnySecondaryInfections	"false"	<a href="http://www.isibang.ac.in/ns/codo#d8eedc167d63eda1e272642595088335">x</a>
pString	"0"	<a href="http://www.isibang.ac.in/ns/codo#d8eedc167d63eda1e272642595088335">x</a>
hasID	"11910"	<a href="http://www.isibang.ac.in/ns/codo#d8eedc167d63eda1e272642595088335">x</a>
hasState	Karnataka	<a href="http://www.isibang.ac.in/ns/codo#d8eedc167d63eda1e272642595088335">x</a>
age	"36"	<a href="http://www.isibang.ac.in/ns/codo#d8eedc167d63eda1e272642595088335">x</a>
diagnosedOn	"2020-06-27T00:00:00"	<a href="http://www.isibang.ac.in/ns/codo#d8eedc167d63eda1e272642595088335">x</a>
diagnosedOn	"2020-06-27T00:00:00"	<a href="http://www.isibang.ac.in/ns/codo#d8eedc167d63eda1e272642595088335">x</a>
age	"36"	<a href="http://www.isibang.ac.in/ns/codo#d8eedc167d63eda1e272642595088335">x</a>
hasState	Karnataka	<a href="http://www.isibang.ac.in/ns/codo#d8eedc167d63eda1e272642595088335">x</a>
hasID	"11910"	<a href="http://www.isibang.ac.in/ns/codo#d8eedc167d63eda1e272642595088335">x</a>
pString	"0"	<a href="http://www.isibang.ac.in/ns/codo#d8eedc167d63eda1e272642595088335">x</a>
hasCausedAnySecondaryInfections	"false"	<a href="http://www.isibang.ac.in/ns/codo#d8eedc167d63eda1e272642595088335">x</a>
rdfs:label	"patient11910"	<a href="http://www.isibang.ac.in/ns/codo#d8eedc167d63eda1e272642595088335">x</a>
gender	Male	<a href="http://www.isibang.ac.in/ns/codo#d8eedc167d63eda1e272642595088335">x</a>

# RDF

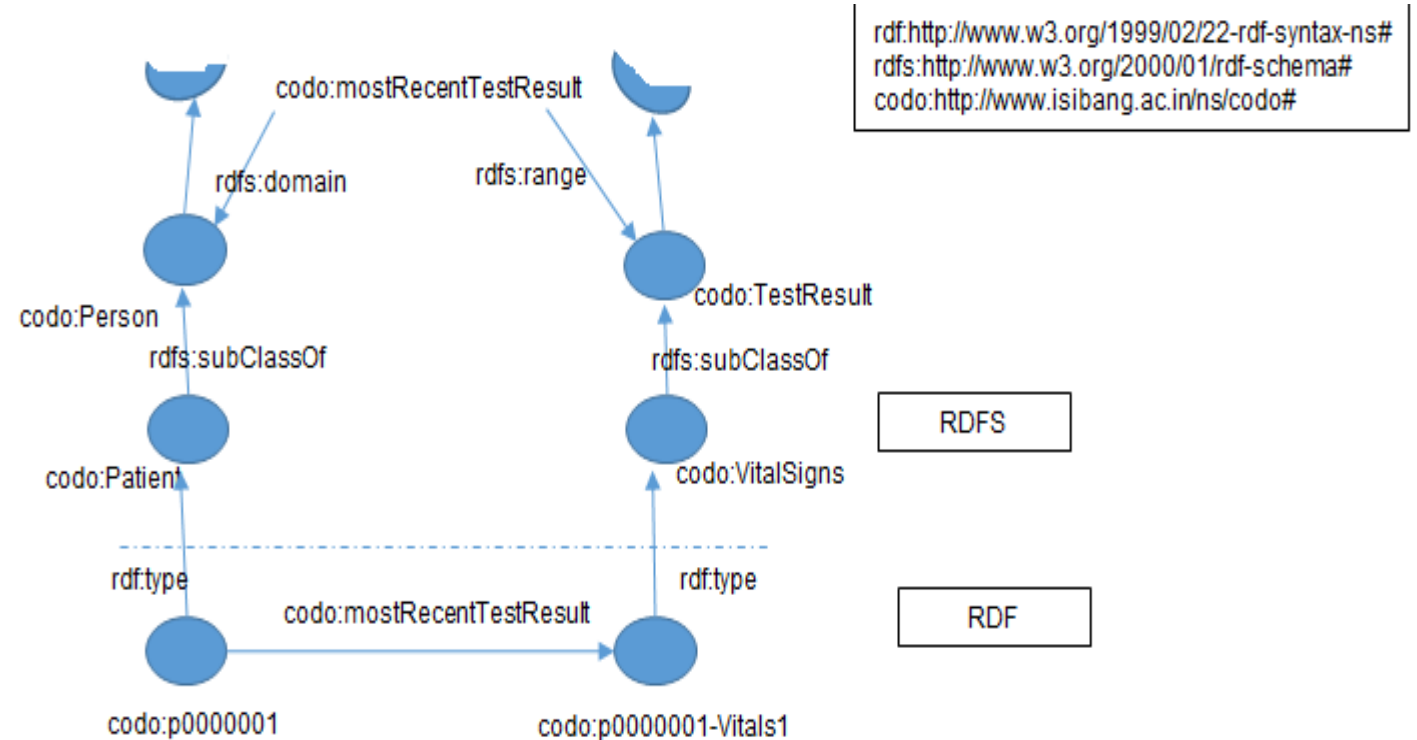
An abstract metadata data model.

It is the foundation language for describing IRI data as a graph.



# RDF Schema

RDFS is layered on top of RDF and provides basic concepts such as classes, properties, collections, etc.



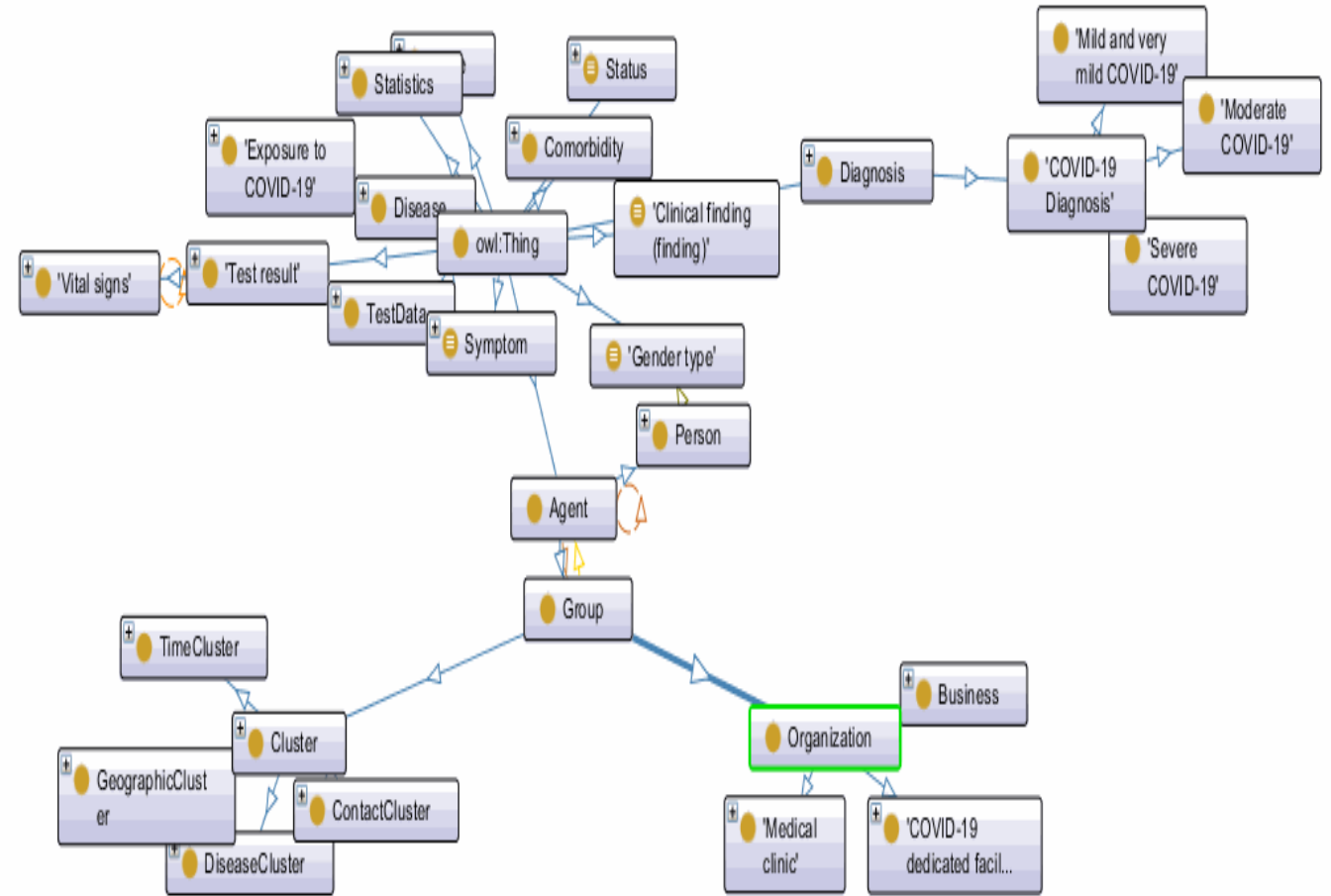
# OWL

- OWL is layered on top of RDFS and provides the **semantics** for knowledge graphs.
- An implementation of Description Logics, a decidable subset of First Order Logic (W3C 2012).
- OWL enables the definition of reasoners which are automated theorem provers.

# Ontology

A formal model that represents knowledge as a set of concepts within a domain and the relationship between these concepts

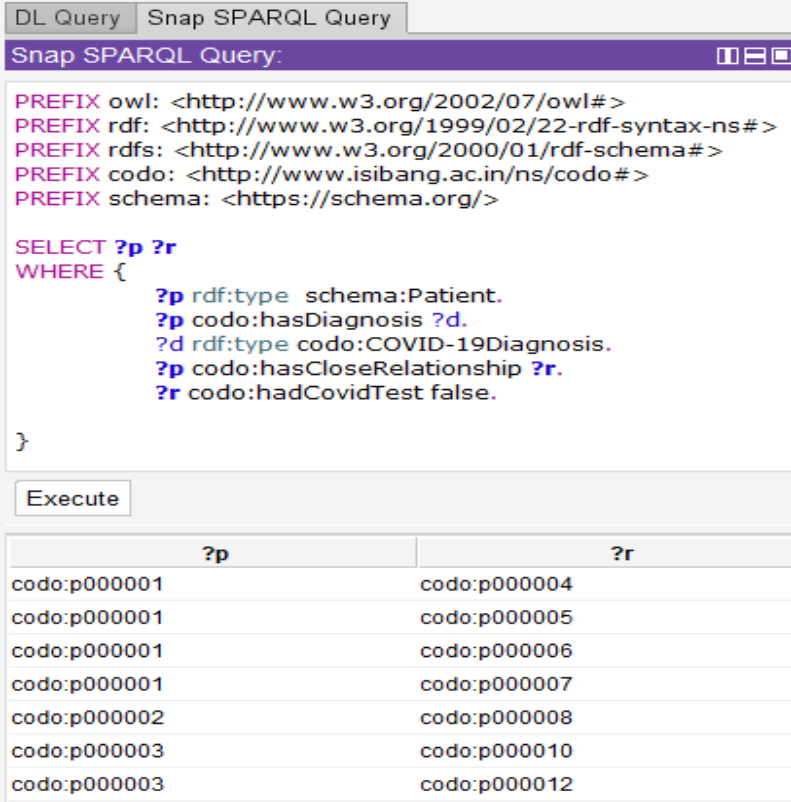
“A formal explicit specification of a shared conceptualization”  
[Gruber, 1993]



# SPARQL

A SPARQL query defines a graph pattern that is matched against the available data sources and returns the data that matches the pattern.

Allows federated queries across heterogeneous sources of data.



The screenshot shows a web interface for executing a SPARQL query. The title bar indicates 'DL Query' and 'Snap SPARQL Query'. The main area contains the following query:

```
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX codo: <http://www.isibang.ac.in/ns/codo#>
PREFIX schema: <https://schema.org/>

SELECT ?p ?r
WHERE {
  ?p rdf:type schema:Patient.
  ?p codo:hasDiagnosis ?d.
  ?d rdf:type codo:COVID-19Diagnosis.
  ?p codo:hasCloseRelationship ?r.
  ?r codo:hadCovidTest false.
}
```

Below the query is an 'Execute' button. The results are displayed in a table with two columns: '?p' and '?r'.

?p	?r
codo:p000001	codo:p000004
codo:p000001	codo:p000005
codo:p000001	codo:p000006
codo:p000001	codo:p000007
codo:p000002	codo:p000008
codo:p000003	codo:p000010
codo:p000003	codo:p000012



[DuCharme, 2011]



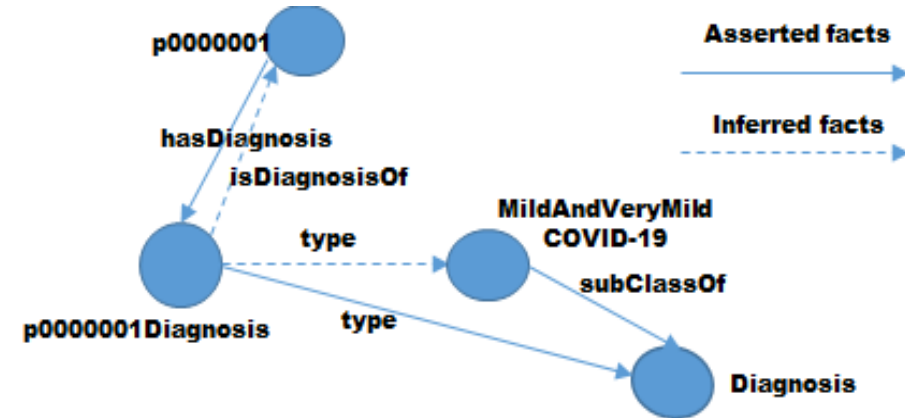
# SWRL

**A rule-based language that extends  
OWL reasoners with additional  
constructs beyond what can be  
described with OWL's Description  
Logic language.**

```
codo:hasDiagnosis(?p, ?d) ^ codo:hasSymptom(?p, codo:URTI) ^ codo:hasSymptom(?p, codo:Fever) ->  
codo:MildAndVeryMildCOVID-19(?d)
```

# Reasoner

- Reasoners are automated theorem provers.
- Reasoners first ensure that an **ontology model** is consistent.
  - If the model is not consistent the reasoner will highlight the **probable source of the inconsistency**.
  - If the model is consistent reasoners can then **deduce additional information** based on concepts described, such as transitivity, inverses, value restrictions, etc.



Explanation for P000001Diagnosis Type 'Mild and very mild COVID-19'

- Show regular justifications
- Show laconic justifications
- All justifications
- Limit justifications to

Explanation 1  Display laconic explanation

Explanation for: P000001Diagnosis Type 'Mild and very mild COVID-19'

- p0000001 'has symptom' 'Upper Respiratory Tract Infection' ?
- 'has diagnosis'(?p, ?d), 'has symptom'(?p, 'Upper Respiratory Tract Infection'), 'has symptom'(?p, Fever) -> 'Mild and very mild COVID-19'(?d) ?
- p0000001 'has symptom' Fever ?
- p0000001 'has diagnosis' P000001Diagnosis ?

# **CODO**

- (1) CODO Ontology
- (2) CODO Knowledge Graph

# CODO: An Ontology for Collection and Analysis of Covid-19 Data

CODO v1.3 consists of  
# of classes: 84  
# of object property: 73  
# of data property: 52

CODO: COVID-19 Ontology for Cases and Patient information

Release April 27, 2020

This version:

<http://www.isibang.ac.in/ns/codo/1.0>

Latest version:

<http://www.isibang.ac.in/ns/codo/1.0>

Authors:

[Biswanath Dutta](#), ([Indian Statistical Institute](#))

[Michael DeBellis](#), ([Semantic Consultant](#))

Publisher:

[Indian Statistical Institute](#), ([null](#))

Download serialization:

[Format JSON LD](#) [Format RDF/XML](#) [Format N Triples](#) [Format TTL](#)

License:

[License Creative Commons Attribution 4.0](#)  [License](#)

Cite as:

Dutta, B. and DeBellis, M.(2020). CODO: an ontology for collection and analysis of COVID-19 data. In Proc. of 12th Int. Conf. on Knowledge Engineering and Ontology Development (KEOD), 2-4 November 2020 (accepted).

Abstract

Available from <https://isibang.ac.in/ns/codo/index.html>

<https://github.com/biswanathdutta/CODO>



Dutta, B. and DeBellis, M.(2020). CODO: an ontology for collection and analysis of COVID-19 data. In Proc. of 12th Int. Conf. on Knowledge Engineering and Ontology Development (KEOD), 2-4 November 2020 (accepted)

# CODO Ontology Goals

- To serve as an explicit ontology for use by data and service providers to publish COVID-19 data using FAIR principles
- To develop and offer distributed, heterogeneous, semantic services and applications
  - E.g., decision support system, advanced analytics, such as behavior analysis of the disease, factors of disease transmission, etc.
- To provide a standards-based reusable vocabulary for the use of various organizations (e.g., government agencies, hospitals) to annotate and describe COVID-19 information

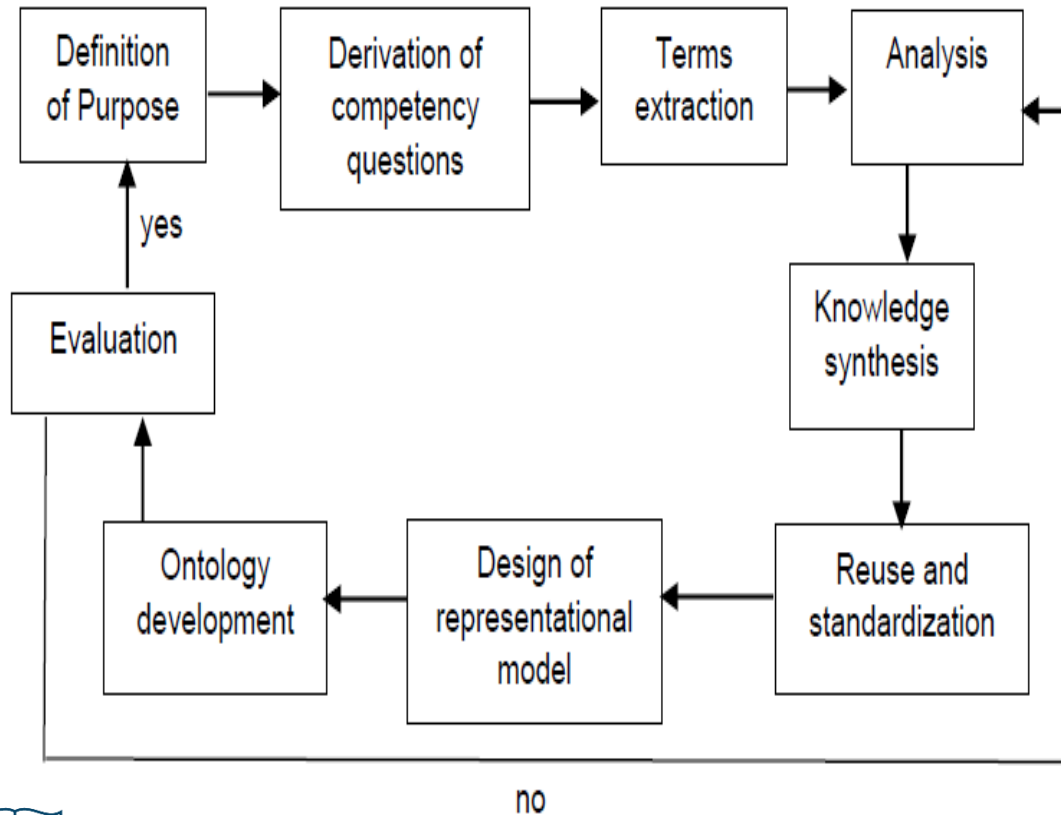
# CODO Ontology Use cases

- **Knowledge Graph creation**
- **Annotation of COVID-19 literature**
- **Application design (e.g., COVID-19 risk detection system)**

# CODO Ontology success stories, so far

- COVID-19 risk detection system for older people in residential aged care using ontology and machine learning technology (<http://biportal.bioontology.org/projects/Ping>)
- Research community/ academicians expressed their interest to use CODO for building KGs

# CODO Ontology Design Approach



## S2: Competency questions

- i. Find all People  $p$  who are related to someone  $r$  who has been diagnosed with COVID-19 and who has not yet been tested.
- ii. Give me the primary reasons  $i$  for the maximum number of COVID-19 patients  $p$ .
- iii. Give me the most prevalent symptoms  $s$  of Severe COVID-19  $d$ .

Description: UrgentlyNeedsCovidTest

Equivalent To +

- foaf:Person and (hasCloseRelationship some DiagnosedWithCovid) and (hadCovidTest value false)

SubClass Of +

- UntestedForCovid and (hasCloseRelationship some DiagnosedWithCovid)
- UntestedForCovid

General class axioms +

SubClass Of (Anonymous Ancestor)

- nationality some Place
- address some xsd:string
- foaf:Person and (hadCovidTest value false)

Instances +

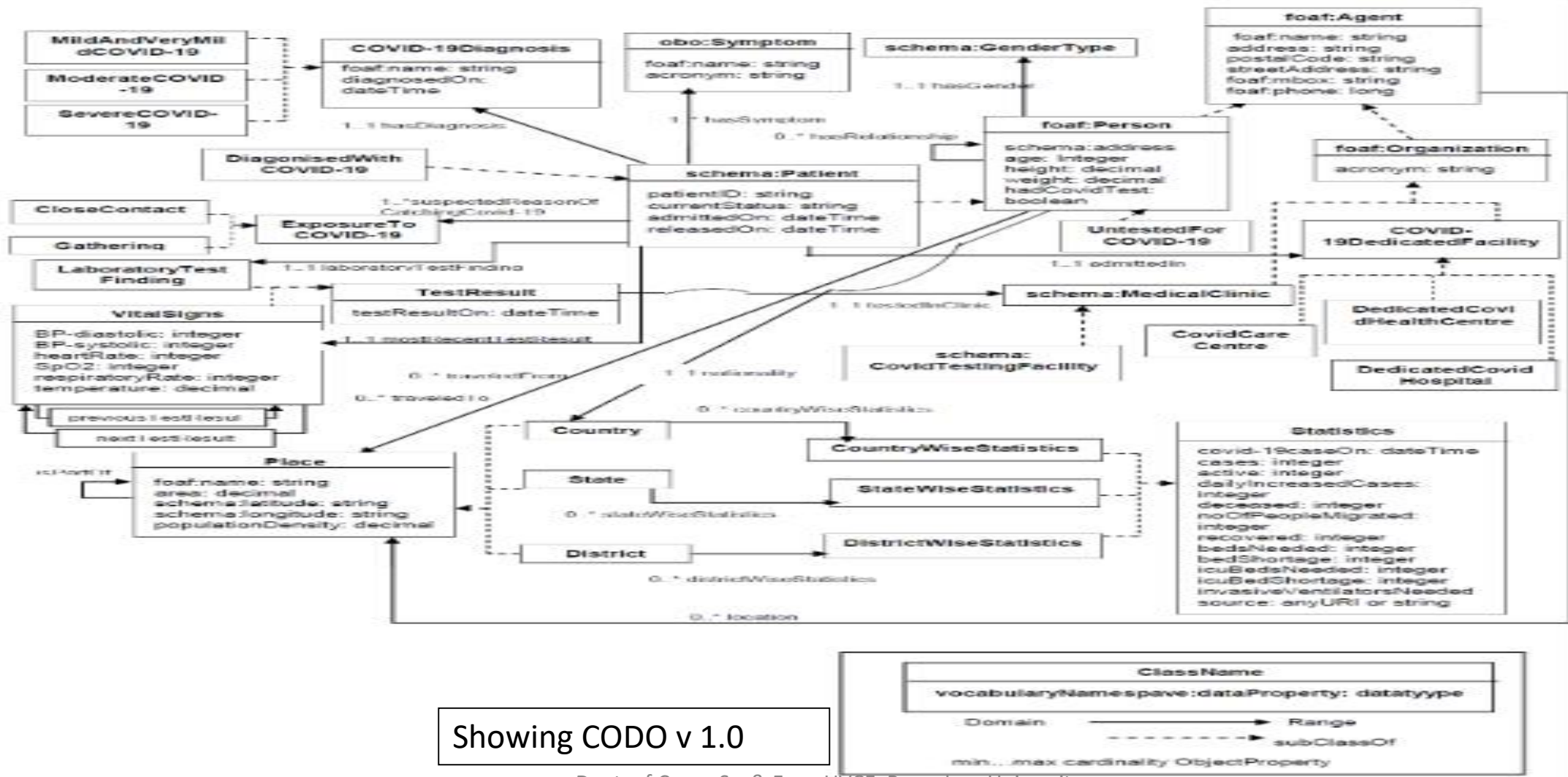
- p000004
- p000005
- p000006
- p000007
- p000008
- p000010
- p000012



Dutta, B. and DeBellis, M.(2020). CODO: an ontology for collection and analysis of COVID-19 data. In Proc. of 12th Int. Conf. on Knowledge Engineering and Ontology Development (KEOD), 2-4 November 2020 (accepted)



# CODO Ontology block diagram




Showing CODO v 1.0


# CODO Knowledge Graph

Primarily with the following two goals:

1. Transforming COVID-19 data as **FAIR Semantic** data
2. CODO ontology evaluation

  
**F**indable

  
**A**ccessible

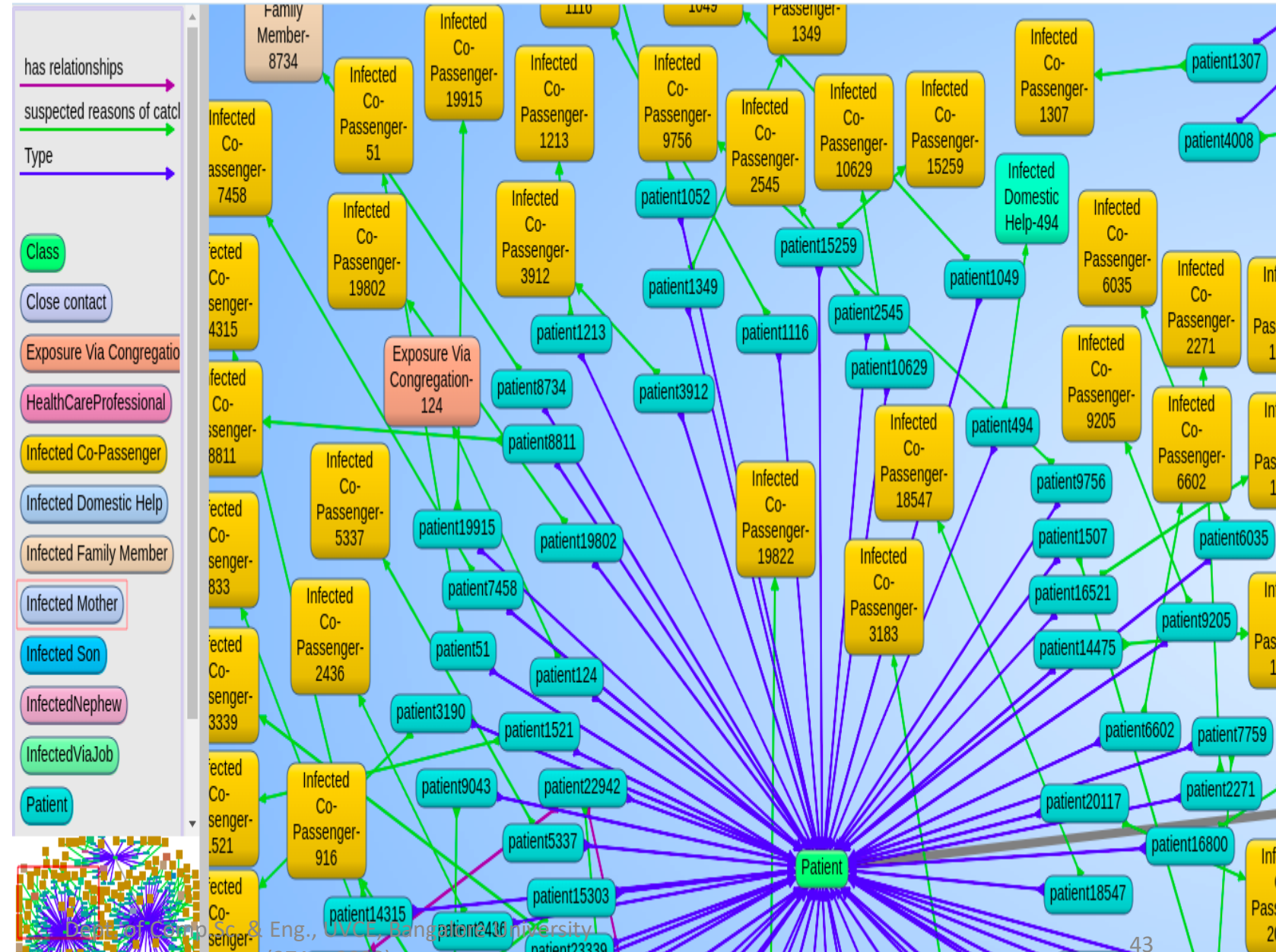
  
**I**nteroperable

  
**R**e-usable

# CODO Knowledge Graph

#Show the patients with the possible reasons of catching COVID-19. Also, display the relationships between the patients, if any.

```
SELECT ?p ?r ?l
WHERE {
?p rdf:type schema:Patient.
?p
codo:suspectedReasonOfCatchingCovid-19 ?r.
OPTIONAL{?p codo:hasRelationship
?l.}
} LIMIT 150
```

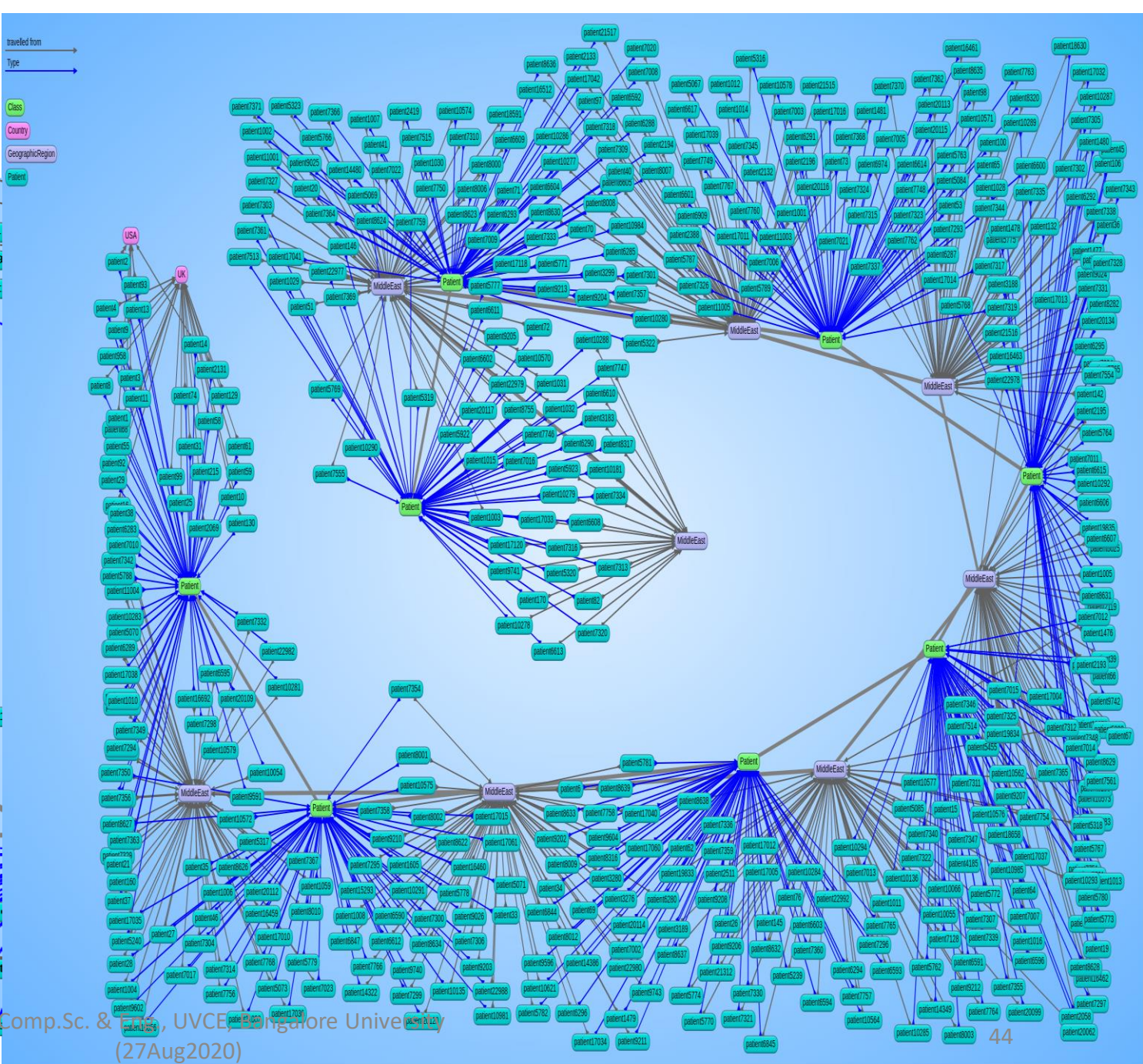
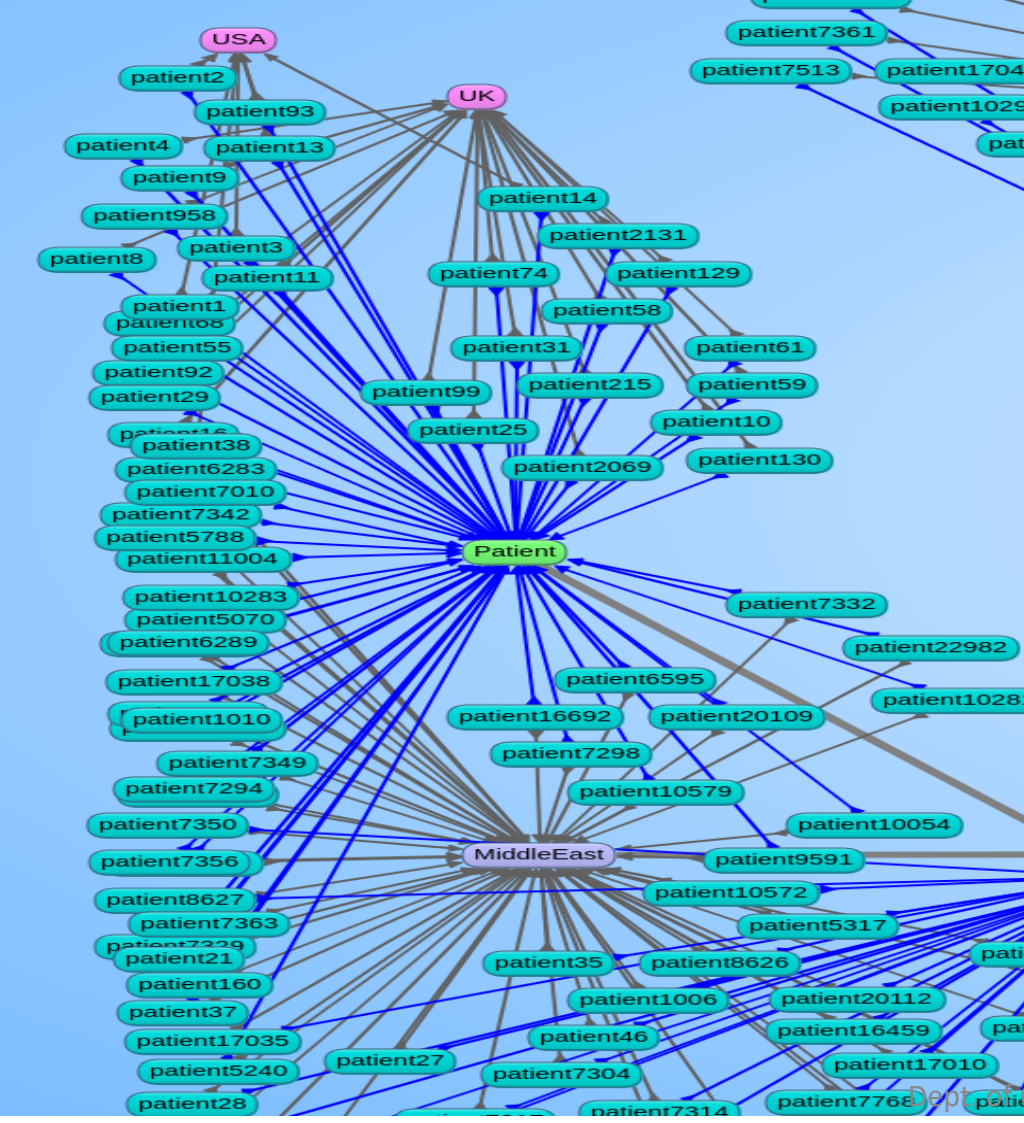




```

SELECT ?p ?tf ?tt
WHERE {
    ?p rdf:type schema:Patient.
    ?p codo:travelledFrom ?tf.
    OPTIONAL{?p codo:travelledTo ?tt.}
} LIMIT 2000

```



# CODO Knowledge Graph consists of ...

- # of axioms: 338977
- # of individuals: 25996
- # of classes: 84
- # of object properties: 73
- # of data properties: 52

## **CODO KG Data source**

<https://www.isibang.ac.in/~athreya/incovid19/data.html>

[https://covid19.karnataka.gov.in/govt\\_bulletin/en](https://covid19.karnataka.gov.in/govt_bulletin/en)

# Experience/ challenges

- Data availability (e.g., clinical data)
- Standard format for data capture and communication
- Most of the datasets are not directly consumable, not suitable to the graph
  - This made the data transformation complicated, time consuming
- Inconsistent data (e.g., sometimes P1342, sometimes 1342. Sometimes P132-P134 and sometimes "P132 and P133 and P134")
- Typo errors
  - Data misplacement
  - Spelling errors (mostly the place names and this complicates the linkage with the external resources)



# Access to CODO

- Both CODO Ontology and CODO Knowledge Graph can be accessed/downloaded from:
  - GitHub (<https://github.com/biswanathdutta/CODO>)
  - Browse CODO Ontology (<https://isibang.ac.in/ns/codo/index.html>)
- Persistent URI for CODO

# Take home message

- Graph data is inevitable
- KG is a powerful way of representing data
- KG can solve many present day data integration and other related tasks
- KG and ML are not technologies that compete with each other but rather solve different problems



# Acknowledgement

- **Michael DeBellis (<https://www.michaeldebellis.com/>)**
  - My friend and colleague for his continuous and active support in making CODO flourish



# References

1. <https://medium.com/virtuoso-blog/linked-data-ontologies-and-knowledge-graphs-a3d0ad6d6f66>
2. Singhal, A. (2012). Introducing the Knowledge Graph: things, not strings.  
<https://www.blog.google/products/search/introducing-knowledge-graph-things-not/>
3. Idehen, Kingsley U. (2020). Linked Data, Ontologies, and Knowledge Graphs.  
<https://www.linkedin.com/pulse/linked-data-ontologies-knowledge-graphs-kingsley-uyi-idehen/>
4. Blumauer, A. and Kiryakov, A. (2020). Knowledge Graphs: 5 Use Cases and 10 Steps to Get There.  
(<https://www.ontotext.com/knowledgehub/webinars/knowledge-graphs-5-use-cases-and-10-steps-to-get-there/>)
5. W3C, 2014. RDF 1.1: Concepts and Abstract Syntax. W3C Recommendation.  
<https://www.w3.org/TR/rdf11-concepts/>
6. W3C, 2014a. RDF Schema 1.1. W3C Recommendation. <https://www.w3.org/TR/rdf-schema/>
7. W3C, 2012. Web Ontology Language Document Overview (Second Edition). W3C Recommendation.  
<https://www.w3.org/TR/owl2-overview/>
8. Gruber, T.R. (1993), “A translation approach to portable ontologies”, Knowledge Acquisition, Vol. 5 No. 2, pp. 199-220.
9. DuCharme, Bob, 2011. Learning SPARQL. O’Reilly.

# References

10. W3C, 2004. SWRL: A Semantic Web Rule Language Combining OWL and RuleML. W3C Member Submission. <https://www.w3.org/Submission/SWRL/>
11. Aasman, J. (2020). Stanford CS 520 Knowledge Graphs. <https://web.stanford.edu/class/cs520/abstracts/Aasman.pdf>
12. Dutta, B., Chatterjee, U. and Madalli, D. P. (2015). YAMO: Yet Another Methodology for Large-scale Faceted Ontology Construction. In Emerald Journal of Knowledge Management. Vol. 19, no. 1, pp. 6 – 24.
13. Graph: Enterprise Knowledge Graph. 2020. (Available from [https://www.youtube.com/watch?time\\_continue=5&v=MJuRnuA0hrM&feature=emb\\_logo](https://www.youtube.com/watch?time_continue=5&v=MJuRnuA0hrM&feature=emb_logo))
14. Bender, Edward A.; Williamson, S. Gill (2010). Lists, Decisions and Graphs. With an Introduction to Probability.
15. Natasha Noy, Yuqing Gao, Anshu Jain, Anant Narayanan, Alan Patterson, And Jamie Taylor (2019). Five diverse technology companies show how it's done. ACM.
16. Dieter Fensel, [...], Alexander Wahler (2020). Knowledge Graphs: methodology, tools and selected use cases. Springer.
17. Andreas Blumauer and Helmut Nagy (2020). The Knowledge Graph Cookbook: recipes that work. Edition Mono/Monochrom, Vienna, Austria.

Thank you!!!

Dr. Biswanath Dutta

Email: [dutta2005@gmail.com](mailto:dutta2005@gmail.com)

[bisu@drtc.isibang.ac.in](mailto:bisu@drtc.isibang.ac.in)