Development of Ontology on Crime Investigation process

Journal of Physics: Conference Series **1716** (2021) 012053 doi:10.1088/1742-6596/1716/1/012053

Development of Ontology on Crime Investigation Process

P V Srimukh[1] and S Shridevi[2]

[1]Student, School of Computer Science and Engineering, Vellore Institute of Technology University, Chennai, India.
[2]Associate Professor, Research Division of Advanced Data Science, Vellore Institute of Technology University, Chennai, India.

[1]srimukh.pampana@gmail.com
[2]shridevi.s@vit.ac.in

Abstract. This paper is an extended version of the research paper "Ontology based crime investigation process" which deals with the working principle and the construction of an ontology based extensively on organized crime. Ontologies on various domains are created in which their assistance has been widely recognized. They are being developed for providing basis for allocation of knowledge in a domain and imparting reasonable information. In this paper we describe the structure of the ontology we created and also by validating the ontology via an online ontology evaluating tool.

1. Introduction

An ontology[14] is a formal description of knowledge as a set of concepts within a domain and the relationships that hold between them. The entitlement of such an elucidation of the knowledge, we need to conventionally enumerate components such as instances of objects i.e individuals, attributes, classes, relations furthermore adding restrictions, rules and axioms. This results that ontologies share information and reiterative knowledge representation but will also append by annexing contemporary knowledge about the domain.

The framework of the ontology can be applied to a compendium of individual facts to create a knowledge representation system which is a multitude of entities, where the types and the relationships between them are expressed by nodes and edges between these nodes, By representing the fabrication of the knowledge in a domain, the ontology forms the stage for the knowledge graph to store the data in it.

There are numerous number of methods that make use of formal stipulations for knowledge representation such as lexicons, classifications, glossary, idea maps, taxonomies and logical models. The ontologies represents relationships between the entities which will provide access to its users to formulate the associations, while the taxonomies or RDBs (relational database schemas) provide relations in their own particular domain.

2. Survey

Categorization of the facts and evidences in a crime scene[15] on the comprehensive and feasible way will form the foundation that will generate a decision support system which establishes the relationships between the evidences that were collected. This gives a logical explanation of the crime that was purported on the scene. In the previous paper we tried to establish a formal way of classifying the knowledge acquired from the crime scene extensively identifying a pattern. A thorough research study was conducted on various knowledge representation of crime domains that were being developed.

3. Review of Various Ontologies

In recent times Ontologies are created in the crime domain that reflects the information needed for the police and other intelligence departments to understand the kind of facts and logical interpretations they need to take as references to solve the cases. Although these ontologies are built in assisting various departments, These ontologies should not be restricted to deal with a certain domain[1,3,7]. Moreover the ontologies were having digital evidence as data Since they lack the appeal to generate enough evidence related to organized crime they are specifically restricted to a certain part of crime investigation process. This system[1] was able to scrutinize evidences via digital forensic report. Other Domains related to situational base[4,5], Information[8] and cyber crime and also on web base[6,9,12] were dealt extensively with respect to crime. [4] A situation based awareness with the culmination of various situations in which the crime happened. Although the system proved to be effective in dealing with the trying of criminal yet the hypothesis postulated for acquiring the biometrics is currently under progress. The novel ontology that we created will have the terminologies related to various domains within the crime jargon. By comparison we increased the scale of domains in the field of crime. The contemporary ontologies were designed to address various relations and attributes that were clinging to one specific domain. Furthermore, These Ontologies illustrates the crimes committed were diminutive and minor. For instance an ontology was created to deal with motorcycle theft[2,11] The crime analysis framework for which the publishers try to do is to store different parameters independently without making any relations between them.

Organized crimes follow a meticulously planned order in committing crimes regardless of whether it was carried out by a single individual or a nexus of people who are a syndicate operating at the highest level of degree. While inspecting a crime scene the forensic scientist explores every part of it and collects all the things present in the field. The facts acquired from the crime scene are carefully analyzed such that a pattern emerges from the crime scene evidence. This pattern is to be imparted into the ontology such that if there are any such patterns that were recorded in the system previously, then based on this parameter the investigation of the case would be solved within no time.

This Novel Ontology that we created have a clearly distinguished types of crimes such as robbery, kidnap, fraud, corruption, murder, drug abuse. We also specified individuals to each and every type of crimes mentioned above. We inculcated various departments as to which type of case will come under the purview of which department. For each specified domain they were completely equipped with the individuals in which they are helpful to form tiniest details. The ontology was designed to include sundry of domains of crime into a single system. The key input to this knowledge representation is the evidence provided by the forensic department analysed[13] from the crime scene and a stack of previously solved cases necessitating to the organized crime. This will serve as a citation to a current case or the system will take the reference to make a scrutiny that whether the crime committed is related to the same culprit or a proxy or the organization behind it.

4. Terminology

The existing ontologies affiliates various tectonic resemblances, nonetheless of the language in which they are demonstrated. Almost all the ontologies that were developed have the similar terminology. They form the basic building blocks for the knowledge representation. This section discusses the elements involved in constructing the ontology.

The following are the components that develop the ontology:

*4.1 Individuals***:** The rudimentary blocks in the development of the ontology. These individuals are also known as instances are used to define various classifications within the individuals themselves. They include objects such as persons, fields, names, animals, plants etc.

*4.2 Classes***:** Classes are typically nothing but classifications in an ontology. Classes are particularly used for denoting that certain categories belongs to one group. In ontology a class can have many sub-classes depends upon the sorting. Thing is a class in which it is the superclass of all classes specified.

4.3 Attributes: They are called Objects. These attributes are utilized for distinguishing relations between objects and sometimes between individuals. Objects form the relation between a certain type of class and certain type of individuals. It provides the link between classes or individuals. The interpretation of an ontology depends upon the objects.
4.4 Relations: Relations are the principal factors that form the meaningful correlations of the objects in the ontologies. In what way the object of one kind is related to the object of other kind is the purpose of relations. These form the true nature of the objects in that specific domain. Mostly the relations can be used in many domains commonly but there are some relations that are specified to utilized only to that particular domain.
4.5 Axioms: Axioms are detailed as the rules specified to create the ontology that forms the entire structure in a more reasonable way. These axioms will illustrate the entire structure of the ontology in which whatever the knowledge that was specified to it has to be reflected
4.6 Events: Events are deployed to change the meaning of the relations between the objects in the ontology. Sometimes the relation between two objects is not making a logical sense. In this perspective these events are used to edit the relations.
4.7 Protégé: Protégé is an open source software which was developed at Stanford university for the generation of ontologies. Protégé is a framework in which a user can develop ontology or tailor an existing ontology for customizing such that it suits his representation of knowledge on a specific domain. It provides a wide range of tools to develop the owl file and it has also visualization tools that can visualize the constructed ontology.
4.8 Resource Description Framework: RDF is a typical facsimile for reciprocation of data on the web. RDF consists of characteristics the expediting the merge of data even if the elementary outline of data contradict. It has the ability to support the evolution of the outline over a period of time without which the developer need not to change the minor discrepancies in the data. RDF broadens the associated structure of the web to use Uniform Resource Identifiers to moniker the connection among the objects as well as both ends of the link which is commonly cited as a triple. This simple schema permits the structured and semi-structured data to be assorted, exhibit, and allocated beyond various applications. This binding structure configurates a directed, labeled graph, where the edges depicts the denominated connection between two objects that is constituted by the graph nodes. This graphical view of the structure is the most simple yet possible mental model for RDF and is often used in easy-to-understand visual .
4.9 SPARQL: SPARQL is a query language which facilitates the developer to write down queries against the RDF data which can also be manipulated to fetch the results as a set of triples i.e in the form of subject-predicate-object. The whole database will be defined as subject-predicate-object. The RDF data will be represented as three columns which comprises of subject column, predicate column and the object column. The entity in a database is similar to that of the subject column in resource description framework. The entity in a conventional SQL database has the fields for a specific object are sorted in various conglomerative columns or laid out exceeding a single table where these entities are pointed out by a unique key.

5. Methodology

A cluster of facts or evidence procured from the scene of crime which was analysed by the forensic department will be etched into the ontology as shown in Fig 1. The ontology provides relationships between the evidences which are relevant to the crime scene. The provided knowledge graph has to be logically correct in which a reasoning system was implemented that will form the semantic web. This data is an inferred data which will be drafted and produced.

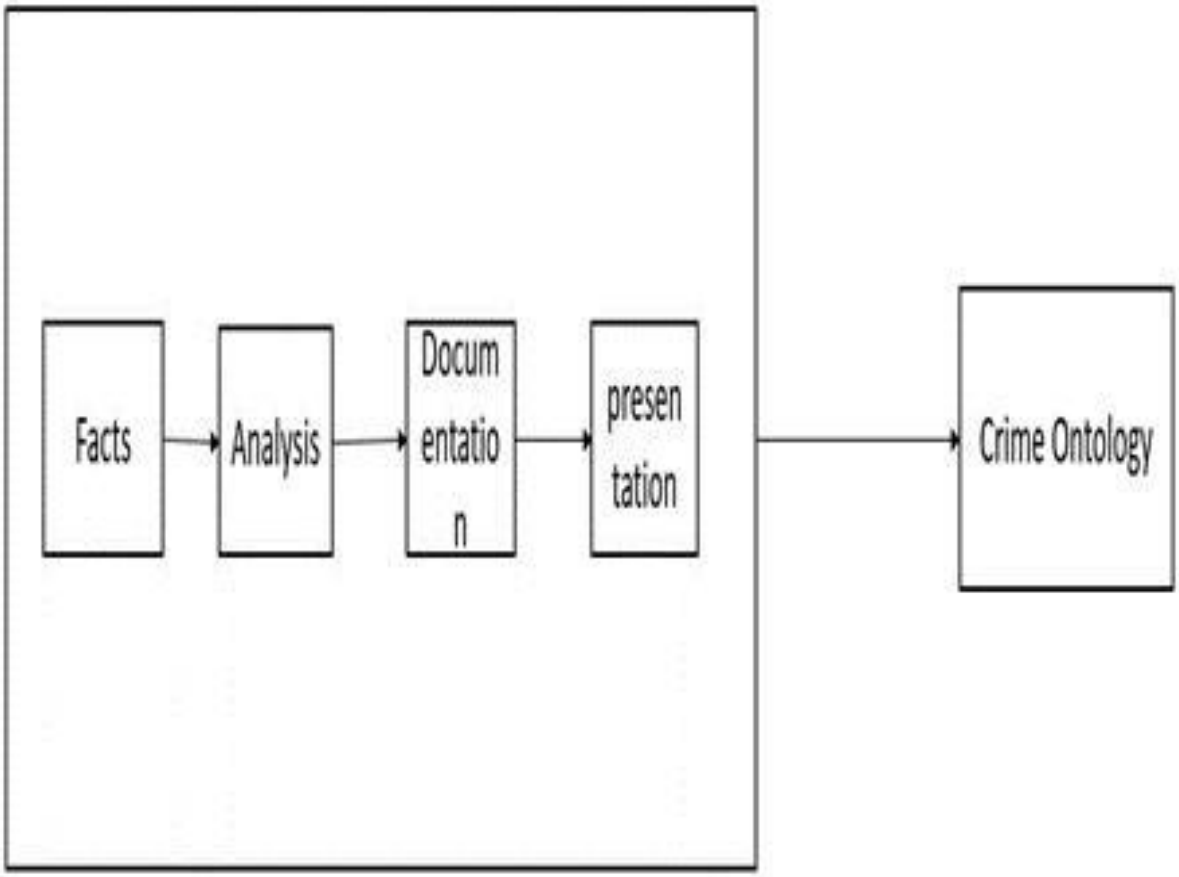


Figure 1. Architecture

6. Implementation

An Ontology based on diverse branches of organized crime was created in this context refer Fig 2. Protégé is an open source software for the creation and editing of ontologies in order to develop a knowledge representation system. It was designed and developed by Stanford University. Protégé-5.5.0 is utilized in the construction of this Ontology. This owl file comprises of different classes related to fields of investigating a certain criminal such as the complete details of his Bio Data. This Bio Data is a Data property of a class person.

6.1 Classes: This Owl file consists of 31 classes related to the criminal, the crime scene, the type of crime, the place in which the crime was committed, which department has jurisdiction on the type of crime and also details on the Organization for which the criminal is working for.

6.2 Object properties: The Object properties which comprises of 14 are generated to relate the classes in between them such that the inferred data will be distinguished relatively to the investigative officer.

6.3 Datatype Properties: 10 properties are bestowed especially for the bio data of the criminal or the accused.

6.4 Individuals: There are 88 individuals that form the meaningful instances in which these are related to all the classes and these instances will interact with the classes.

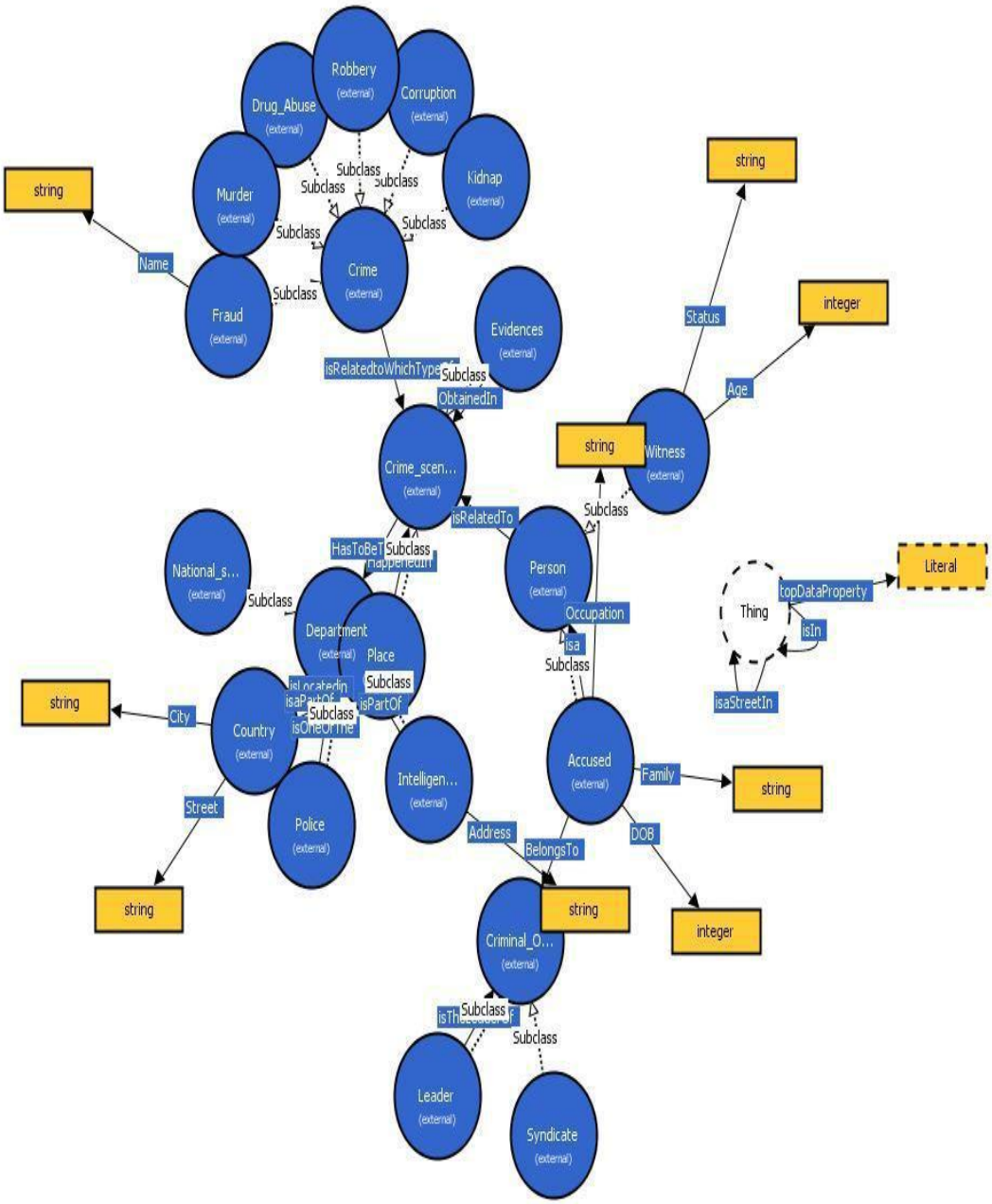


Figure 2. Visual representation of the Ontology

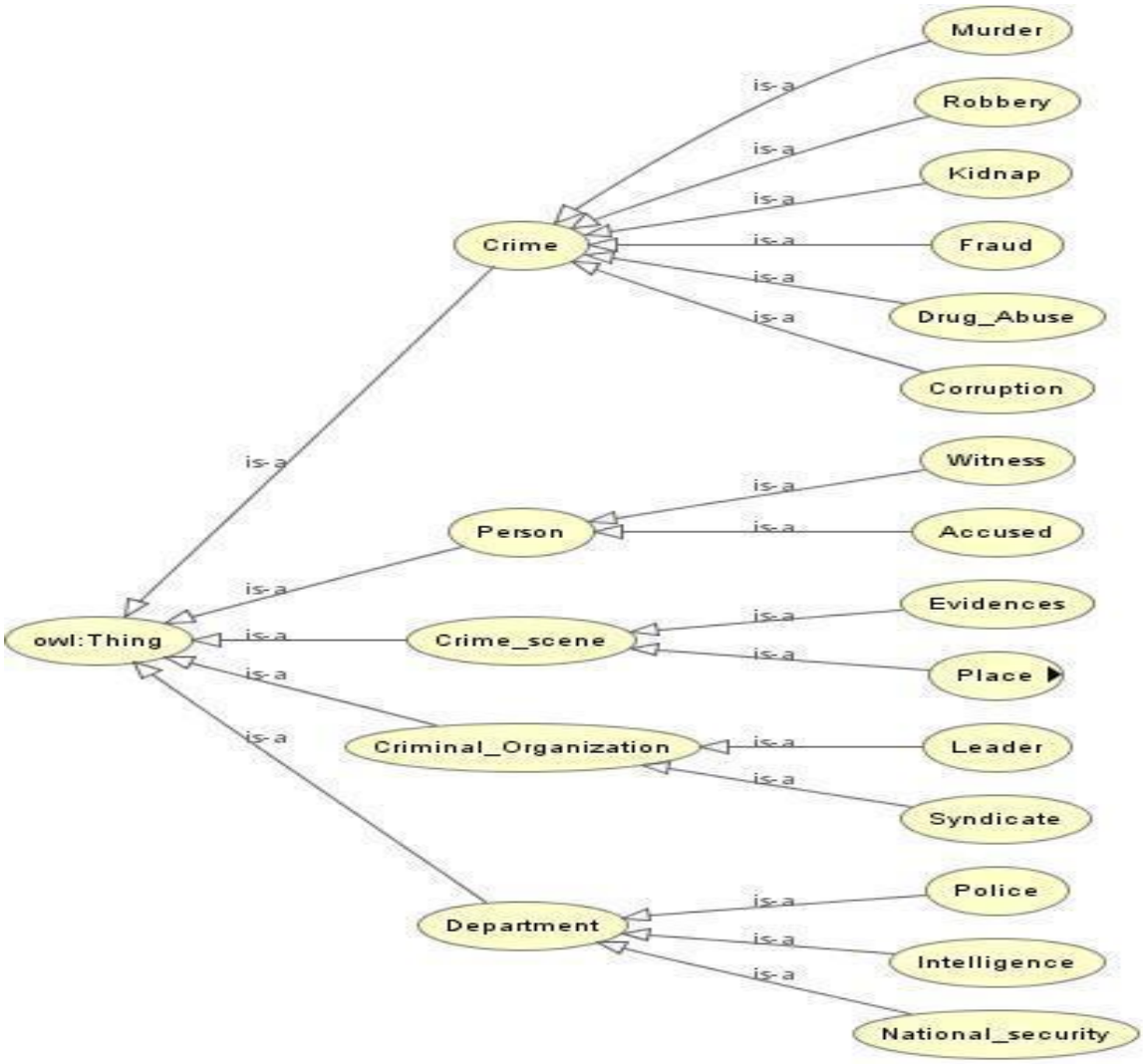


Figure 3. Logical representation of the Ontology

The Above picture gives the visualization of the ontology on crime investigation process. Thing is a superclass to all the classes that were created inside it. Crime class comprises of different types of the domain. Murder, Robbery, Kidnap, Fraud, Drug Abuse, Corruption. Each of these classes have been assigned with individuals relating to that particular type of the domain. For example, Robbery consists of variety of individuals such as burglary, dacoity etc. This was created in such a way that when a person was accused of a crime the ontology will define that person under which part of the domain he belongs to and in that part which individual has to be assigned refer Fig 3.

The most important class that was created in this ontology was the crime scene class. This class has two sub classes the first one is the evidence part and the second one is the Place. The evidence class defines the crime scene by the facts and evidences taken from the crime scene. As the whole point of creating the ontology is to produce a suitable logical inference. On this basis itself the system would suggest that this is a crime scene in which it is a part of the crime domain in the crime class, to which department the case has to be assigned, if a pattern was derived from the evidence of a crime scene using the knowledge the ontology previously acquired it suggests the probability of who might be committing the crime i.e the person class and the criminal organization class. The person class has two sub-classes, one was designated to witness, in which the witness to that crime scene will be recorded. The accused class will record the data belonging to the guilty.

The criminal organization class has two sub-classes leader and syndicate. This class is created such that if the accused have any associations with any organization or the accused was ordered by the organization to do a specific crime. The class Department gives a categorization between different branches of government intelligence institutions. Based on the evidence acquired the respective criminal case will be assigned to respective department. The person class has the datatype properties such as name, age, address, DOB, family, occupation, status. These properties give information on both accused and witness classes in person and not only that, the datatype property "name" is being used in the class criminal organization as well. Individuals were generated based on the class domain. These individuals give meaningful instances between the classes.

The Ontology is appraised using Ontometrics. Ontometrics is an online tool that authenticates and exhibits the statistics about a given ontology. The Ontology is a representation of knowledge that was created in order to acknowledge the mappings and relations between the classes and their properties are logically applied to define the relationship between them as shown in Fig 4.1 to Fig 4.5. This tool checks for the consistency of the ontology based on how accurately it is developed and produces a report. The structure was not complex. This was given in base metrics section in which the DL expressivity was acknowledged as ALH(D) as shown in Fig 4.3.

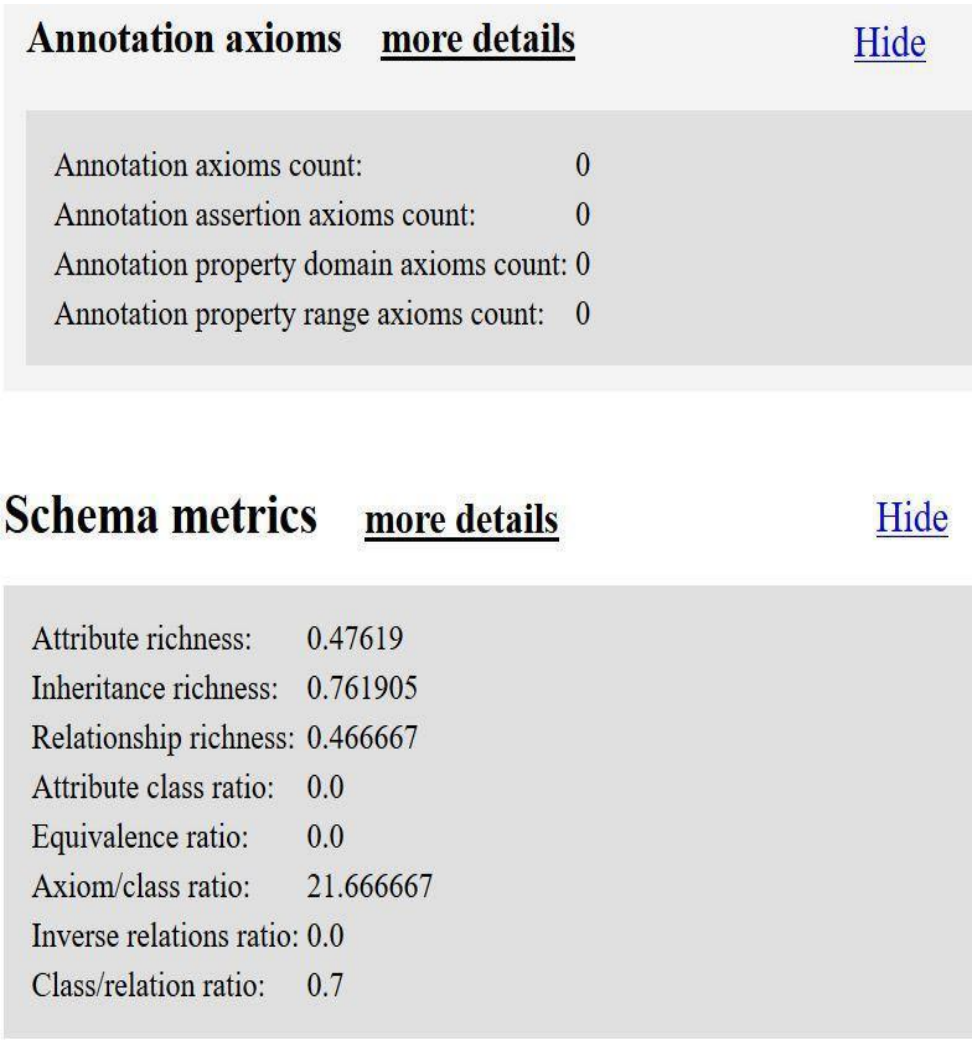
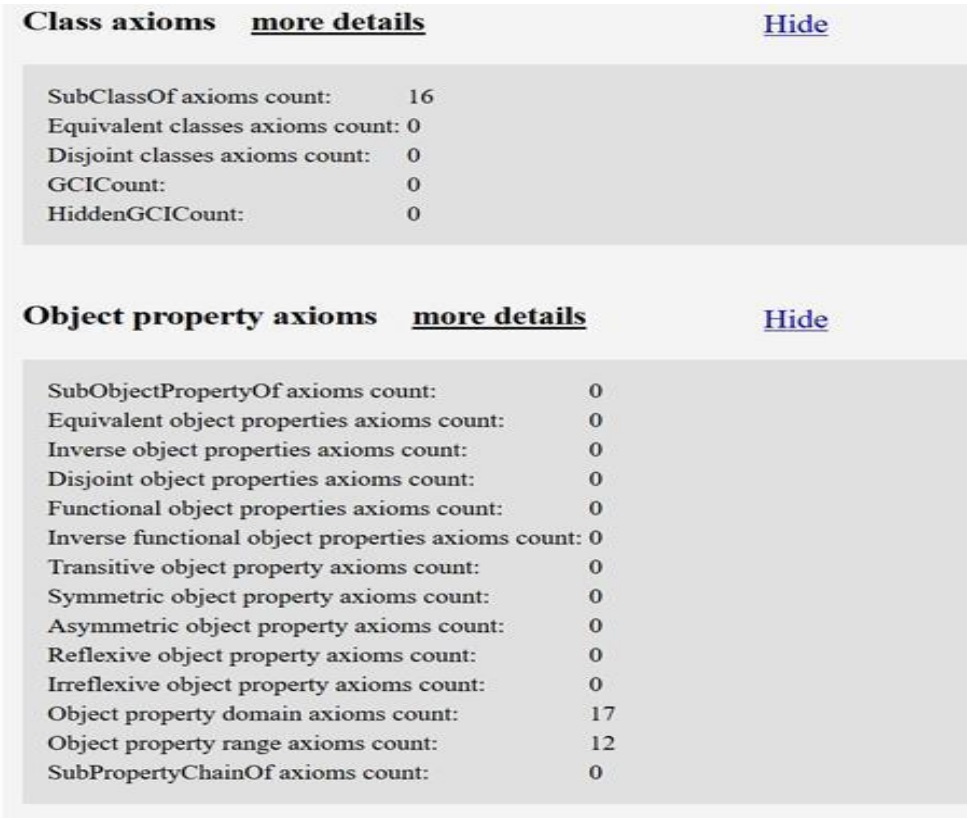


Figure 4.1 Evaluation of schema



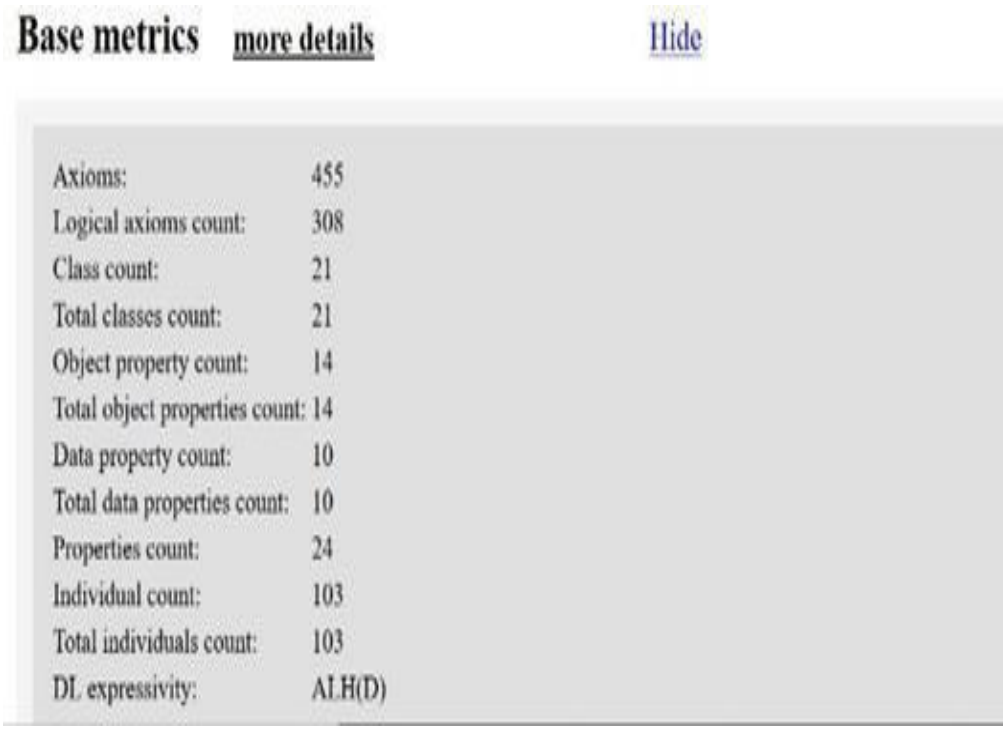
Class axioms more details Hide

SubClassOf axioms count:	16
Equivalent classes axioms count:	0
Disjoint classes axioms count:	0
GCICount:	0
HiddenGCICount:	0

Object property axioms more details Hide

SubObjectPropertyOf axioms count:	0
Equivalent object properties axioms count:	0
Inverse object properties axioms count:	0
Disjoint object properties axioms count:	0
Functional object properties axioms count:	0
Inverse functional object properties axioms count:	0
Transitive object property axioms count:	0
Symmetric object property axioms count:	0
Asymmetric object property axioms count:	0
Reflexive object property axioms count:	0
Irreflexive object property axioms count:	0
Object property domain axioms count:	17
Object property range axioms count:	12
SubPropertyChainOf axioms count:	0

Figure 4.2 Evaluation of Class and Object property



Base metrics more details Hide

Axioms:	455
Logical axioms count:	308
Class count:	21
Total classes count:	21
Object property count:	14
Total object properties count:	14
Data property count:	10
Total data properties count:	10
Properties count:	24
Individual count:	103
Total individuals count:	103
DL expressivity:	ALH(D)

Figure 4.3 Evaluation of base metrics

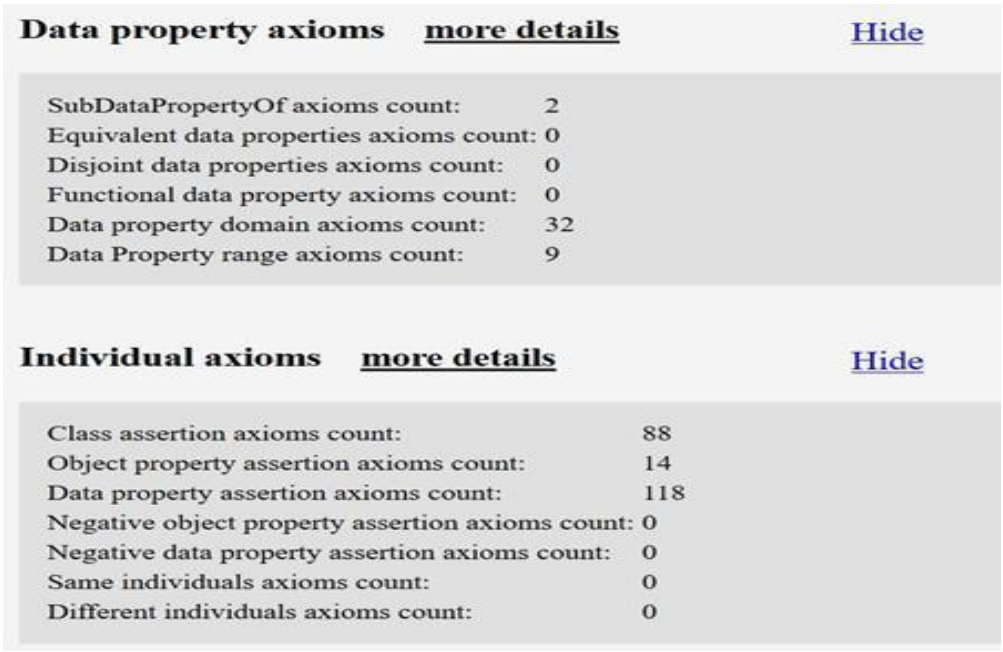


Figure 4.4 Count of various axioms

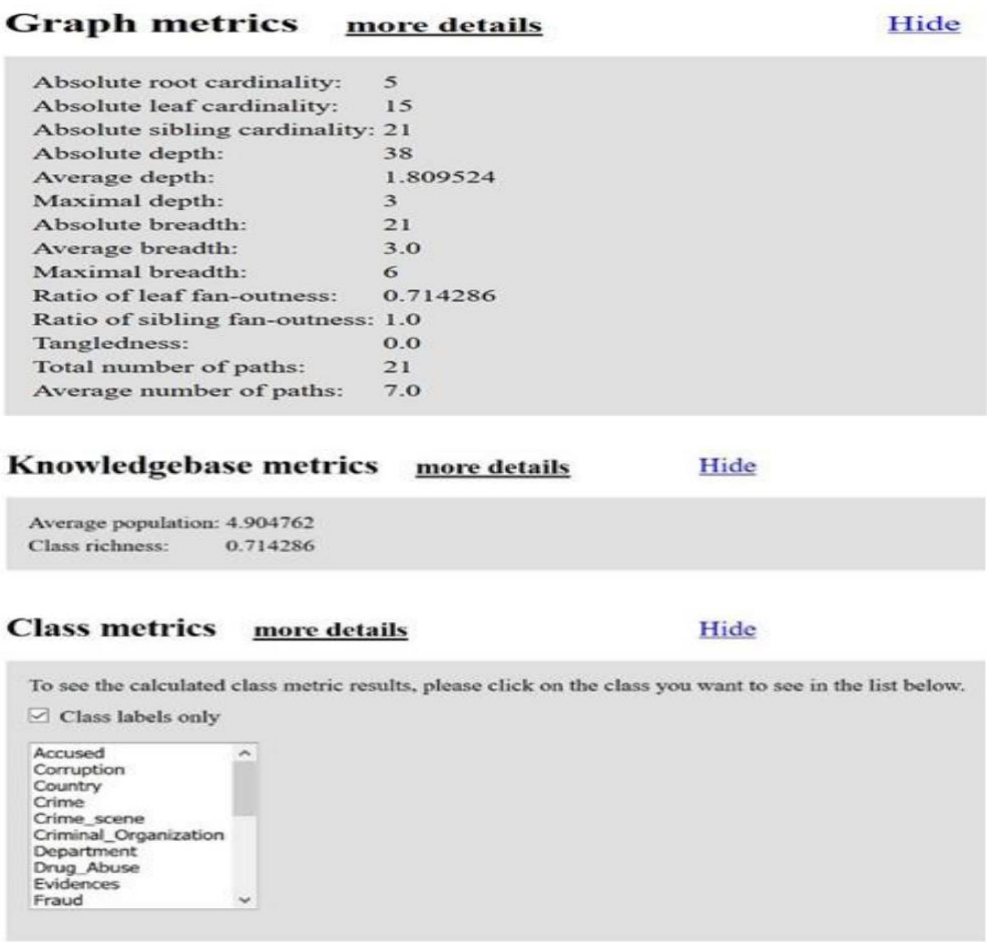


Figure 4.5 Complexity of the Ontology

7. Conclusion

There are numerous Ontologies that enhance miscellaneous problems in crime investigation as mentioned in the above paper. There are a wide range ontologies that are being developed and also are in an inceptive stage in the real-time applications. The above structured ontology will be a first of a kind in which multiple branches criminology is being inculcated. Although this ontology in its commencing phase will be able to allocate with the compendious stage of appraising of crime scene But over a period of time it will metamorphose in dealing with a significant level of solving complex criminal cases. The novelty of this Ontology is that it provides not only for the smaller cases to solve that relates to robbery but also this Ontology deals with bigger issues over nationally and globally. We also want to contemplate the concept of adding this ontology into machine learning such that it will use this system as as an input and predict the logical outcomes more accurately.

References

[1] A M Talib and F O Alomary 2015 Towards a Comprehensive Ontology Based-Investigation for Digital Forensics Cybercrime. *International Journal on Communications Antenna and Propagation (IRECAP)* **5.1** p 263-68.
[2] M A Jalil, C P Ling, Maizura, N M M Noor and F Mohd 2017 Knowledge Representation Model for Crime Analysis. *Procedia Computer Science* **116** p 484-91.
[3] F Amato, G Cozzolino, A Mazzeo and N Mazzocca 2017 Corr-elation of Digital Evidences in Forensic Investigation through Semantic Technologies pp 668-73.
[4] M McDaniel, E Sloan, W Nick, J Mayes, A Esterline 2017 Ontologies for situation-based crime scene identities p 1-8.
[5] J H M Nogueira and W W P Vasconcelos 2007 Ontology for Complex Mission Scenarios in Forensic Computing *The International Journal of Forensic Computer Science* **1** p 42-50.
[6] T Boongoen and N Iam-On 2015 Discovering identity in intelligence data using weighted link based similarities: A case study of Thailand p 327-32.
[7] R Carvalho, M Goldsmith and S Creese 2016 Applying Semantic Technologies to Fight Online Banking Fraud *Revista Brasileira de CiÃªncias Policiais* pp 61-68.
[8] B Onnoom, S Chiewchanwattana, K Sunat and N Wichiennit 2015 An ontology framework for recommendation about a crime scene investigation *14th International Symposium on Communications and Information Technologies*, *ISCIT 2014* pp 176-80.
[9] C Li, Y Hu and Z Zhong 2010 An event ontology construction approach to web crime mining. *Proceedings - 2010 7th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2010* **5** p 2441-45.
[10] V Loia, M Mattiucci, S Senatore and M Veniero 2009 Computer Crime Investigation by Means of Fuzzy Semantic Maps p 183-86.
[11] M A Jalil, F Mohd, C P Ling and N M M Noor 2018 Criminal intelligence analysis based on ontological knowledge representation model. *International Journal of Engineering and Technology(UAE)* **7. 2.28** p 187-90.
[12] E K Elsayed, W Ghonaim and E Zeaar 2018 Semantic Crime Investigation System *Journal of Computers* **13.10** p 1216-26.
[13] D C Harrill and R Mislan 2007 A small scale digital device forensics ontology *Small Scale Digit Device Forensics* J. **1.1**
[14] https://en.wikipedia.org/wiki/Ontology_(information_science)
[15] https://en.wikipedia.org/wiki/Organized_crime