

# Metric Based Ontology Quality Evaluation

Ravi Lourdusamy, Antony John



**Abstract:** *Ontology helps semantic web to process and understand large amount of data available in Internet. Ontology uses concepts and their relationship with each other to represent knowledge within a domain. The represented knowledge can be analysed, inferred and reused to make decisions and to derive new knowledge. The developed ontology has to be assessed for quality before using or reusing it. Evaluation becomes a key factor to determine the quality of ontology. Different approaches and methods are used to ensure the quality desired by the user. This article identifies various aspects of ontology, provides a framework for metric based ontology evaluation, elucidates components in the framework and develops a tool based on the framework. The framework checks the syntax, structural and semantic measures of ontology. While a reasoner takes care of the syntax and parser errors, the structural metrics analyses the taxonomy of ontology. Semantic measures deal with the distance of concepts in ontology. Further, competency questions are used to do custom based quality checking of a particular domain. This article provides a systematic way to identify and measure the quality of ontology based on metrics.*

**Index Terms:** *Ontology Evaluation, Ontology Quality, Semantic Measures*

## I. INTRODUCTION

Internet contains a large amount of data and is used for searching, advertising, publishing and information retrieval. All these activities are driven by human intelligence and this make the task tedious as data in internet is growing at a very fast rate day by day. The amount of data available in the internet is so vast that sometimes it becomes difficult to get the relevant information intended by the user. Semantic web offers a solution to this problem by making the web machine understandable [1]. The idea of semantic web is to focus on the meaning of data, making it accessible and processable by machines. Machines should process the data based on the query and infer the knowledge contained in the data. In order to achieve this goal machines have to identify the meaning of data, relationship existing between entities and query data to fetch knowledge and deliver results. Semantic web uses ontology to structure data in a way understandable by machines. Ontology uses concepts, axioms and relationships to represent knowledge of a domain. Ontology defines a way to standardize the meaning of concepts and allows knowledge

management and information extraction. Ontologies developed are reused and shared on various platforms[2].

Ontology defined has to be correct to derive the meaning and knowledge from a domain. Wrong representation of data in ontology can bring illogical conclusion. Quality has to be ensured before using, reusing or sharing ontology. In order to ensure quality, ontology has to be evaluated. While evaluating ontology it should be checked whether it is built based on the design principles of ontology and it meets the requirement of the user. The ontology should be evaluated for its syntactic correctness, structural strength and semantic similarities. Hartmann et al. describes ontology evaluation as “assessment of the quality and the adequacy of ontology or parts of it regarding a specific aim, goal or context.” [3] J. Brank et al., defines ontology evaluation as “the process of deciding on the quality of an ontology in respect to a particular criterion with the view of determining which in a collection of ontologies would best suit a particular purpose.”[4] Ontology evaluation can assure building the ontology with some verified standards, ensure the competency, consistency and completeness of the working domain and fulfill the proper implementation of ontology in an application. A systematic approach to ontology evaluation can guarantee correctness while modeling ontology and ensure reuse and knowledge representation.

Ontologies can be built using various methodologies. Ontology evaluation is an important step in the life cycle of the development of ontology. Whichever methodology we adopt to develop ontology it should contain evaluation in its process of development. Ontology evaluation should make a check on the quality and reliability while doing the ontology engineering process. It should also check whether it is built based on the standards set for ontology development. Moreover, it should also see whether the requirements are met by ontology. The evaluation of ontology can be done at various stages using different approaches. Ontology has to be checked for the correctness of syntax, data consistency and constraints. Evaluation unfolds the inconsistencies and redundancies present in the ontology and gives more information about the structure and semantic aspects [5].

This article presents a metric based ontology evaluation. A framework is proposed in this article to check the quality of ontology. The objective of the framework is to check the syntactic, structural and semantic measure of ontology. The framework provides a way to ensure the quality of ontology by checking syntax, data inconsistencies, redundancies, strength and similarity of concepts in ontology. The framework also includes competency questions that will assess specific aspects of ontology. A tool is developed based on the framework to evaluate ontology.

Revised Manuscript Received on October 30, 2019.

\* Correspondence Author

**Ravi Lourdusamy\***, Department of Computer Science, Sacred Heart College, Tirupattur, Tamilnadu, India.

**Antony John**, Department of Computer Science, Sacred Heart College, Tirupattur, Tamilnadu, India..

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

The rest of the article is organized as follows. Section 2 gives fundamental idea about ontology and its structure. Section 3 discusses the related works in ontology evaluation. Section 4 introduces the proposed framework explaining each of its features. Section 5 elaborates the working of the tool developed based on the framework. Last section draws conclusion of the work.

### II. PRELIMINARIES

The primary objective of ontology is to accomplish a shared common knowledge which can be diffused between people and applications in a machine processable format. Common understanding of a domain provided by ontology can be communicated among heterogeneous people, applications and systems. Ontology represents data in a formal and expressive way and defines semantics of concepts and relationships between them. This helps machines to recognize and infer information in meaningful ways and find new knowledge from the existing information.

Ontology is used to describe a specific domain by means of set of concepts which are arranged hierarchically. Ontology consists of concepts, arbitrary relations, axioms, constraints and functions which are used to create knowledge of the domain. It assembles a vocabulary of the fundamental terms and relations of a specific domain and furthermore gives standards to consolidating terms and relations. "Ontology is a formal, explicit specification of a shared conceptualisation." [6]

Ontologies share many structural similarities, irrespective of the language in which they are expressed. Ontologies are built with the following components: classes, attributes, relations, individuals, functions, restriction, rules, axioms and events. Classes, also known as concepts, are abstract types which have set of objects. Attributes define the characteristics or parameter of the classes or objects. Relations explain the way in which classes and instances are related with one another. Individuals or instances are the basic components. Restrictions, rules and axioms enable to create constraints and assertions in a logical form to validate the inputs. Ontologies are widely used in knowledge sharing and reusing. It expresses the knowledge in present in a domain in a formal way. This enables the retrieval and assessment of data by machines in a meaningful way.

Ontology evaluation performs a check on ontology either by comparison or using some criteria to find its quality. There are various approaches, methods and tools to validate and evaluate ontology. Different approaches and methods are used to ensure the quality desired by the user. Though all of them check the suitability of an ontology the main difference between each of them is how they evaluate, the criteria used, motivation behind evaluation and the need of the user. They basically differ in the way it evaluates and on the basis of what it evaluates. While some approaches use comparison, others concentrate on the semantic, qualitative or quantitative aspect of the ontology.

Over the years different approaches were used to evaluate ontology. One of the approaches was Gold standard approach. This approach compared the ontology modeled to a gold ontology created by domain experts [7]. It checked how far the ontology is close to the gold ontology. It assessed the lexical and taxonomical levels of ontology. If it is found to be closer to the gold ontology it is considered to be good. But the

problem here is how to specify the validity of gold ontology. The gold ontology used for comparison is manually created. Who has given the validity of gold ontology that is used for evaluation? There are possibilities of errors and inconsistencies in gold ontology. The result of evaluation can go wrong if the gold ontology is not correct.

Application or task-based evaluation approach shows how well the ontology performs when assigned a task. Mainly three types of errors such as insertion errors, deletion errors and substitution errors are analysed. The effectiveness of ontology when used with application is also measured. The ontology is used in an application and the result is evaluated. Data Driven approach compares the ontology against corpus of a domain [8]. The terms of corpus is collected by domain experts. While doing the evaluation the terms that are present in the corpus and not in the ontology and terms present in ontology and not in corpus are identified. Based on the result the quality of ontology is determined.

User based approach makes use the feed back of a user to evaluate ontology. User satisfaction level is checked. Though there will be subjectivity in the results this is also considered as important [9]. Metric Based Approach evaluates ontology quantitatively based on some measures. Metric based approach tries to find the structural quality, functional quality, analytical quality, pragmatic quality, syntactic quality, cognitive quality, semantic quality, social quality and practical quality of ontology [10]. Structural Evaluation gives the information about the number of classes, individuals, properties and relationships of ontology. Using these values various other metrics can be calculated. Semantic based approach evaluates the cognitive behavior of ontology and analyses the accuracy of ontology in knowledge-based systems. Based on these approaches various frameworks and methods were built to evaluate ontology.

### III. RELATED WORKS

Many efforts were done to suggest various methods and tools to evaluate ontology. Brewster et al., created an architecture to evaluate ontology using data driven approach. It checked the degree of the structure of ontology against the content of the document [11]. Burton-Jones et al., gave a semiotic framework consisting of six criteria for quality evaluation of ontology [12]. Gangemi et al., provided a framework to evaluate ontology from three aspects such as structural, functional and usability-profiling. It gave a set of criteria to select the ontology in a particular context [13]. Yang et al., provided a metric to find the complexity of ontology checking the concepts and their relationship in the hierarchy [14].

Alm et al., checked the quality of ontology based on metrics. The design of ontology was checked to find out the discrepancies in the structure of ontology [15]. Sanchez et al., tried to quantify the semantic dispersion of ontology [16]. Poveda et al., created a tool called OOPS to detect the pitfalls while modeling ontology [17]. Jimborean et al., provided a method to measure the structural and semantic aspects of ontology [18]. Hooi et al., proposed a criteria selection framework for ontology evaluation [19]. Khan et al., created a framework to evaluate bio medical ontologies based on properties [20].

Tovar et al., proposed a metric for automatic evaluation of restricted domain ontologies. The metric is defined in terms of the evaluation of different lexico-syntactic, statistical and semantic approaches automatic evaluation of ontologies of restricted domain, by means of natural language processing, extraction of information and linguistic tools [21].

All these approaches required a certain amount of technical knowledge to evaluate the ontology. Moreover they do not specify what is needed to be corrected and where the problem lies. Mostly they evaluated ontology from one particular perspective. A study of the related works shows that there is a need to develop a framework taking into account the structural, syntactic and semantic aspects of ontology. It is noted that most of the frameworks or methods evaluated only one or two aspects leaving the other aspects unattended. A framework that can give an all round quality check will be handy for users and developers. So a user friendly application which implements the framework with structural, syntactic and semantic aspects will give a better picture of quality of ontology.

#### IV. ONTOLOGY QUALITY EVALUATION FRAMEWORK

Keeping in mind the contributions made already in the field of ontology evaluation, a framework is designed for ontology quality evaluation wherein a user can upload ontology and get a result. The design of this framework has three parts. Each part is distinct from each other. The first part which facilitates the view allows the user to upload ontology for evaluation. The uploaded ontology is passed on to MOQA and the various components of MOQA will evaluate the uploaded ontology based on the metrics of each components. The view component outputs the result of evaluation to the user. Framework uses Java and Jena for implementing the architecture. The flow control is managed by the classes available in Jena and Java is used for viewing and controlling.

##### A. Framework Components

The framework uses reasoners, structural metrics, semantic metrics and competency questions to check the quality of ontology. While a reasoner takes care of the syntax and parser errors, the structural metrics analyses the taxonomy of ontology. Semantic measures deal with the relatedness and distance of concepts in ontology. Further, competency questions are used to do custom based quality checking of ontology.

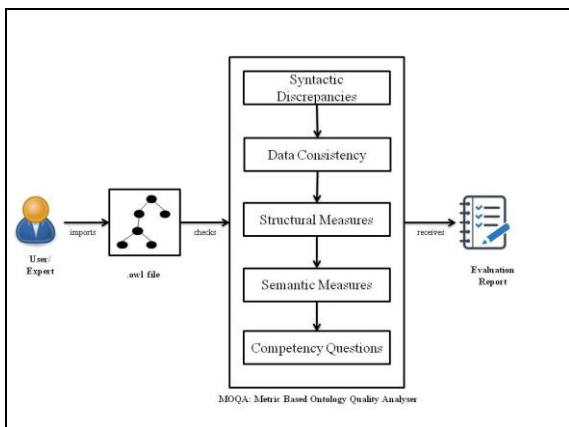


Figure 1: Framework for Ontology Quality Evaluation

##### B. Syntactic Evaluation

A good ontology should avoid errors, inconsistencies, anomalies, redundancies and pitfalls. Ontology should be built in a consistent and coherent manner. Since ontology can define classes, specify properties, express value restrictions, logical constraints and inherit properties of upper of level classes sufficient care should be taken to avoid inconsistent data. There should be right representation of disjoint and exhaustive knowledge. Redundancies should be avoided. The first section of the framework does the syntactic evaluation of taxonomy of ontology. It checks for inconsistencies, errors and redundancy.

Reasoners are used to assess the quality and correctness of ontology. It mainly evaluates the consistency and certainty of concepts in ontology. Unless there is a consistent ontology there can be contradictions in the way concepts and ideas are interpreted. There is possibility of inconsistency coming in the ontology being built and it need to be addressed. Reasoners use various strategies to check different features of ontology. It checks for the validity of inferences from the hypothesis available [22]. The proposed framework includes a reasoner to check for the correctness and consistency of ontology.

##### C. Structural Evaluation

Structurally seen, ontology is a graph whose nodes represent concepts and arcs represent relationships. Knowledge of the structure of the ontology and the concepts used, can give an advantage while reusing or making semantic interpretations. Structural measures mainly refer to the taxonomical structure of ontology. Ontology consists of large number of concepts, relations, individuals and axioms. Structural metrics gives a view of the concepts, relations and individuals used in ontology. It also gives the statistical metrics of ontology and based on the result the quality can be improved. It shows the number of classes, properties, individuals, siblings, depth of classes, classes with subclasses and classes without definition. Further, it specifies the number of axioms and properties used in ontology. This framework gives the quantitative measure of the structure of ontology. Three main details assessed by this framework are the size, name and strength of ontology. Size refers to the number of entities in a module and name refers to name of concepts and entities present in ontology. In order to identify the strength of ontology the following formulae is used [23].

$$\text{Strength of Ontology} = \frac{1}{n} \sum_{i=1}^n M_i$$

where  $M_i$  is the strength of  $i^{\text{th}}$  module of ontology and  $n$  is the number of modules

##### D. Semantic Evaluation

While the structural metrics deals with the taxonomy of ontology, semantic metrics focus on the similarity and distance of concepts in ontology. Semantic metrics are useful in ontology mapping, finding relationships and to avoid redundancies. Approaches used to find the semantic metrics are Structural approach, Information theoretical approach, Feature based approach and Hybrid approach. Structural approach considered the structure of concepts as graphs and used various graph traversal techniques to measure semantic similarity.

Equal weight was given to all the nodes while measuring the similarity. Information theoretical approach considered not only the weight of nodes to find the similarity but also considered the content of information [24]. In feature based approach concepts were seen as a collection of features. These features are considered while measuring the similarity. The concepts are compared for the features they share and what are their distinct features. This approach took into consideration the knowledge defined in taxonomy. Sometimes one approach is not sufficient to find the semantic measures between two concepts. Hybrid approach used the techniques available in other approaches to find the similarity. It took into account the weight of nodes, information content and various features of concepts [25]. This framework assesses the similarity of two concepts in ontology. Semantic similarity and semantic distance of concepts are assessed. Wu and Palmer suggest the following measure to find the similarity between two concepts C1 and C2 [26].

$$Sim(C1,C2) = \frac{2 * N3}{N1+N2+2*N3}$$

N1 and N2: Distance from the specific common concept to concept C1 and C2  
 N3: Depth of the least common subsumer.

**E. Competency Questions**

Competency questions are used in the development process of ontology in order to get the requirements that are specific to the ontologies. CQs represent the requirements expected from ontology. Competency questions are queries formulated based on the ontology requirement [27]. Competency question reveal how far the ontology is able retrieve and reveal the knowledge present in the knowledge base. CQs contains questions that are expected to be answered by ontology. This will help us to know whether the ontology has sufficient axioms to answer questions. Custom based questions can be asked based on a particular domain to check for quality. SPARQL queries and DL queries can be used for asking competency questions. In this framework SPARQL is used to query the CQs to check whether the ontology is built to answer CQs meaningfully.

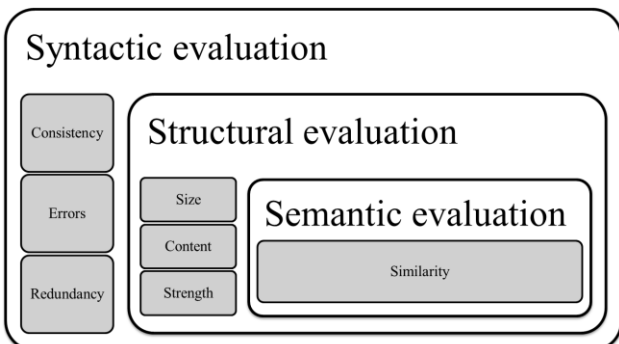


Figure 2: MOQA Components

**V. MOQA: METRIC BASED ONTOLOGY QUALITY ANALYSER**

A tool named MOQA: Metric Based Ontology Quality Analyser is built based on the above mentioned framework using Java and Jena. User can submit ontology in .owl format and the ontology will be evaluated and the result will be

displayed. MOQA is a web based tool which is developed to evaluate a given ontology in different aspects. This tool is developed keeping in mind three aspects viz., syntactic, structural and semantic evaluation. User can upload \*.owl file and choose the evaluation that is to be performed on the ontology.

The application has mainly four tabs; Summary, Taxonomy, Structural Evaluation, Semantic Evaluation and Syntactic Evaluation. The summary tab gives a general description about the ontology as in Figure 3.1. The taxonomy section displays the hierarchical structure of ontology. The result will display all the concepts in ontology in a hierarchical order.

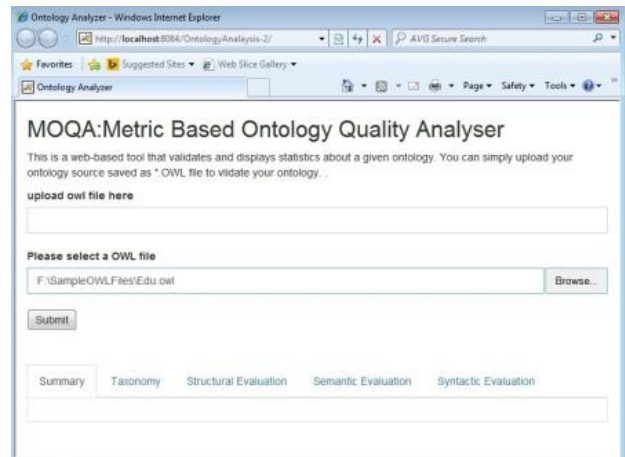


Figure: 3.1

The structural evaluation tab displays the statistical details such as number of classes, individuals, data properties, object properties, axioms and so on as depicted in figure 3.2.

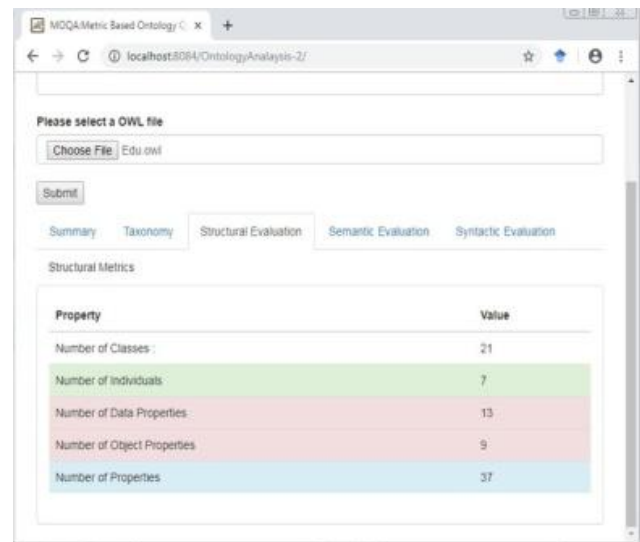


Figure 3.2

In semantic evaluation module, the similarity between the concepts is identified. User can choose two concepts present in the ontology and the similarity measure is checked. The result is displayed as in Figure 3.3.

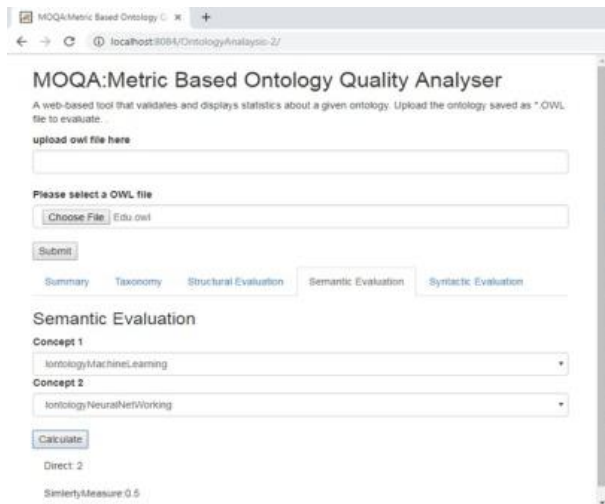


Figure 3.3

MOQA evaluates the syntactic correctness using reasoners. A reasoner is used in this tool to validate the syntax of the concepts. The web application tool was successfully tested for its working. MOQA evaluates the syntactic correctness, structural measures and semantic similarity of concepts in ontology.

## VI. CONCLUSION AND FUTURE WORK

The proposed framework for Metric Based Ontology Quality Evaluation evaluates the quality of ontology to enhance the knowledge representation structure and semantic information retrieval of data. A framework is proposed for the evaluation of ontology. This framework creates uniqueness by the fact that it contains different levels of quality check on various aspects of ontology. The framework helps to understand the internal and external aspects of ontology in the following ways. First of all, the framework proposes a user friendly web application to check the quality of ontology. Secondly, it gives a provision to check the syntax and data consistency using reasoners so that the foundation of the ontology is set right. Thirdly, a view of the structural aspects of ontology gives an idea of the taxonomical description and enables us to know the various components used to model the ontology. Fourthly, it evaluates the semantic metrics of ontology specifying the similarity and distance of concepts in ontology. Lastly, it gives a provision to use custom based competency questions to query ontology of a particular domain. A tool is built based on the framework proposed. For the future development, the framework can be enhanced with more metrics. Secondly, the given framework evaluates the ontology and does not give any corrective suggestions. Knowing how to improve a system is a big advantage. A solution based, intelligent, corrective system can be added to the framework which will inform what to do in order to improve the quality of ontology. Metric Based Ontology Quality Evaluation allows users to learn more about the structure, relations and quality of ontology and to select ontology while modeling, reusing and aligning ontologies.

## REFERENCES

- Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. *Scientific American*, 284(5), 28-37.
- Noy, N. F., & McGuinness, D. L. (2001). Ontology development 101: A guide to creating your first ontology.
- Hartmann, J., Palma, R., & Gómez-Pérez, A. (2009). Ontology repositories. In *Handbook on Ontologies* (pp. 551-571). Springer, Berlin, Heidelberg.
- Brank, J., Grobelnik, M., & Mladenic, D. (2005). A survey of ontology evaluation techniques. In *Proceedings of the conference on data mining and data warehouses (SiKDD 2005)* (pp. 166-170). Citeseer Ljubljana, Slovenia.
- Sure, Y., Daelemans, W., Reinberger, M. L., Guarino, N., & Noy, N. F. (2004). Why evaluate ontology technologies? Because it works!. *IEEE Intelligent Systems*, (4), 74-81.
- Studer, R., Benjamins, V. R., & Fensel, D. (1998). Knowledge engineering: principles and methods. *Data & knowledge engineering*, 25(1-2), 161-197.
- Dellschaft, K., & Staab, S. (2006). On how to perform a gold standard based evaluation of ontology learning. In *International Semantic Web Conference* (pp. 228-241). Springer, Berlin, Heidelberg.
- Hlomani, H., & Stacey, D. (2014). An extension to the data-driven ontology evaluation. In *Proceedings of the 2014 IEEE 15th International Conference on Information Reuse and Integration (IEEE IRI 2014)* (pp. 845-849). IEEE.
- Ouyang, L., Zou, B., Qu, M., & Zhang, C. (2011). A method of ontology evaluation based on coverage, cohesion and coupling. In *2011 Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)* (Vol. 4, pp. 2451-2455). IEEE.
- Lourdusamy, R., & John, A. (2018, January). A review on metrics for ontology evaluation. In *2018 2nd International Conference on Inventive Systems and Control (ICISC)* (pp. 1415-1421). IEEE.
- Brewster, C., Alani, H., Dasmahapatra, S., & Wilks, Y. (2004). Data driven ontology evaluation.
- Burton-Jones, A., Storey, V. C., Sugumaran, V., & Ahluwalia, P. (2005). A semiotic metrics suite for assessing the quality of ontologies. *Data & Knowledge Engineering*, 55(1), 84-102.
- Gangemi, A., Catenacci, C., Ciarmita, M., & Lehmann, J. (2006). Modelling ontology evaluation and validation. In *European Semantic Web Conference* (pp. 140-154). Springer, Berlin, Heidelberg.
- Zhe, Y. A. N. G., Zhang, D., & Chuan, Y. E. (2006). Evaluation metrics for ontology complexity and evolution analysis. In *2006 IEEE International Conference on e-Business Engineering (ICEBE'06)* (pp. 162-170). IEEE.
- Alm, R., Kiehl, S., Lantow, B., & Sandkuhl, K. (2013). Applicability of Quality Metrics for Ontologies on Ontology Design Patterns. In *KEOD* (pp. 48-57).
- Sánchez, D., & Batet, M. (2013). A semantic similarity method based on information content exploiting multiple ontologies. *Expert Systems with Applications*, 40(4), 1393-1399.
- Poveda-Villalón, M., Gómez-Pérez, A., & Suárez-Figueroa, M. C. (2014). Oops!(ontology pitfall scanner!): An on-line tool for ontology evaluation. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 10(2), 7-34.
- Jimborean, I., & Groza, A. (2014, September). Ranking ontologies in the ontology building competition boc 2014. In *2014 IEEE 10th International Conference on Intelligent Computer Communication and Processing (ICCP)* (pp. 75-82). IEEE.
- Hooi, Y. K., Hassan, M. F., & Sharif, A. M. (2015, May). Ontology evaluation—a criteria selection framework. In *2015 International Symposium on Mathematical Sciences and Computing Research (ISMSC)* (pp. 298-303). IEEE.
- Khan, S., Qamar, U., & Muzaffar, A. W. (2015, December). A Framework for Evaluation of OWL Biomedical Ontologies Based on Properties Coverage. In *2015 13th International Conference on Frontiers of Information Technology (FIT)* (pp. 59-64). IEEE.
- Tovar, M., Pinto, D., Montes, A., & González-Serna, G. (2018). A metric for the Evaluation of Restricted Domain Ontologies. *Computación y Sistemas*, 22(1), 147-162.
- Dentler, K., Cornet, R., Ten Teije, A., & De Keizer, N. (2011). Comparison of reasoners for large ontologies in the OWL 2 EL profile. *Semantic Web*, 2(2), 71-87.

23. Lourdasamy, R., & John, A. (2019). A metric suite to analyse strength of modularised ontologies. *International Journal of Information Technology*, 1-8.
24. Lin, D. (1998, July). An information-theoretic definition of similarity. In *Icml* (Vol. 98, No. 1998, pp. 296-304).
25. Rodríguez, M. A., & Egenhofer, M. J. (2003). Determining semantic similarity among entity classes from different ontologies. *IEEE transactions on knowledge and data engineering*, 15(2), 442-456.
26. Wu, Z., & Palmer, M. (1994, June). Verbs semantics and lexical selection. In *Proceedings of the 32nd annual meeting on Association for Computational Linguistics* (pp. 133-138). Association for Computational Linguistics.
27. Ren, Y., Parvizi, A., Mellish, C., Pan, J. Z., Van Deemter, K., & Stevens, R. (2014, May). Towards competency question-driven ontology authoring. In *European Semantic Web Conference* (pp. 752-767). Springer, Cham.

### AUTHORS PROFILE



**Ravi Lourdasamy**, received his Ph.D from Bharathidasan University, Tamil Nadu, India in 2012. Currently he is HOD and Associate Professor in the Department of Computer Science, Sacred Heart College, Tirupattur, India. His research areas include Software Engineering, Ontology and Knowledge Engineering. He has presented and published over 24 papers in international conferences and journals.



**Antony John**, is a Research Scholar in the Department of Computer Science, Sacred Heart College, Tirupattur, India. His research areas include Semantic Web, Ontology and Knowledge Engineering.