

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Library Philosophy and Practice (e-journal)

Libraries at University of Nebraska-Lincoln

April 2021

Ontology of Library and Information Science Theses of the Central Universities in India

syamili C Dr.

University of Calicut, csyamili@gmail.com

Rekha RV Dr.

Pondicherry University, rekhatvm@gmail.com

Follow this and additional works at: <https://digitalcommons.unl.edu/libphilprac>



Part of the [Cataloging and Metadata Commons](#), and the [Scholarly Publishing Commons](#)

C, syamili Dr. and RV, Rekha Dr., "Ontology of Library and Information Science Theses of the Central Universities in India" (2021). *Library Philosophy and Practice (e-journal)*. 5197.

<https://digitalcommons.unl.edu/libphilprac/5197>

Ontology of Library and Information Science Theses of the Central Universities in India

Abstract

The study attempted to design the ontology of Library and Information Science theses of central universities of India to visualize the doctoral research output in Library Information Science. The study also attempts to develop the ontology using 'Methontology', which is an ontology engineering methodology to explore the method. This Ontology contains 247 concepts, 244 Individuals, 125 Object property assertions, 85 Data property assertions. The developed ontology was evaluated and made sure that it meets the information need of its users and it has been found that Ontology could answer the majority of the questions after evaluation. The evaluation results are satisfactory and the model is valid. This ontology will act as a guide for everyone who is looking at India's research productivity in the field of Library and Information Science, especially researchers and academicians all over the world. The ontology is currently published in web protégé since that work will be used for an application of e-learning applications.

Key words: Ontology, Protégé, Methontology, Theses, Shodganga, Library Science, Information Science semantic web

1. Introduction:

The concept of ontology comes from Philosophy. In Philosophy, ontology is the study of being or existence. Later Gruber (1993) coined the term Ontology as a subfield of Computer Science. Gruber defined ontology as a conceptual modal and this definition is widely accepted. Gruber defined "ontology as an explicit specification of a conceptualization". Ontology is a wide term including a wide range of concepts and ideas. Ontology helps the web to be a more meaningful space and plays a major role in different semantic web applications. In recent times ontology has become a major research area in several disciplines of Information Science well as Computer Science. The ontology contains machine-interpretable concepts and the relationships among the concepts in a particular domain.

Research trends in Library & Information Science changing every day. Influence of the technology in this profession is very visible from the research which is taking place in the subject. The main aim of the study is to develop Ontology for the Library and Information Science theses from Shodganga by the central universities of India. Keyword-based search engines can't retrieve information by making semantic matching. So, the users have to apply too much effort and background knowledge to interpret the information. The semantic web can replace the normal web as it is capable of solving problems with the normal web. Here we develop an ontology that will act as a guide for everyone who is looking at India's research productivity in the field of LIS, especially researchers and academicians all over the world. This ontology can be used by anybody who is working in developing the semantic web of this domain.

2. Related Studies

The ontology of Corona Virus infectious disease was developed by Liu et al (2021). A patent ontology has been developed for 600 patent ontologies in the domain logistics development to answer the user queries (Phan, Nguyen, & Nguyen, 2018). An ontology called EOHD has been developed for Historical description using basic ontology model by Zou and Park (2018) for representing historical collections⁴. Another application ontology which assists the teachers in the task of classroom assessment has been developed by Ajetunmobi and Daramola (2017). Aroua, Baazaoui-Zghal and Abed (2017) developed a model for public transportation. In domain like public transportation, heterogeneous and complex data has to be combined together to deal users need. This can be easily solved with the help of ontology. This study put forward a visual framework for personalized itinerary need of the user based on two ontology, one is namely user profile ontology and other domain public transport ontology. Uma and Muneeswaran (2017) developed ontology to retrieve the tag-based information. This study proposes a collaborative tagging based on BAPS method to retrieve. TCMLS is a large ontology of traditional Chinese medicines which contain 30,000 terms and 120,000 concepts. 1,270,000 semantic relations. This ontology is successful in applications such as knowledge visualization and text mining.

3. Methodology

The present study demonstrates the development of the ontology of Library and Information Science theses from central universities in India to visualize the doctoral research output in Library and Information Science. The selected domain for the study is the research in Library and

Information Science. The domain narrowed down into the Library and Information Science theses deposited in Shodganga by the central universities in India. The data has been collected from Shodganga during the 1st week of November 2019. The data may not cover the theses which is uploaded after 10th November 2019 in Shodganga. The metadata of 122 theses from Shodganga has been extracted manually and combined together. These metadata include the title of the thesis, the research guide, the researcher, the year of submission, and the keywords etc. There are several methodologies are there in developing ontologies such as METHONTOLOGY, On- To Knowledge methodology, TOVE, SENSUS, Gruninger and Fox, YAMO. The theses ontology was developed using ‘Methontology’, which is an ontology engineering methodology. Figure 1 shows the steps involved in Methontology.

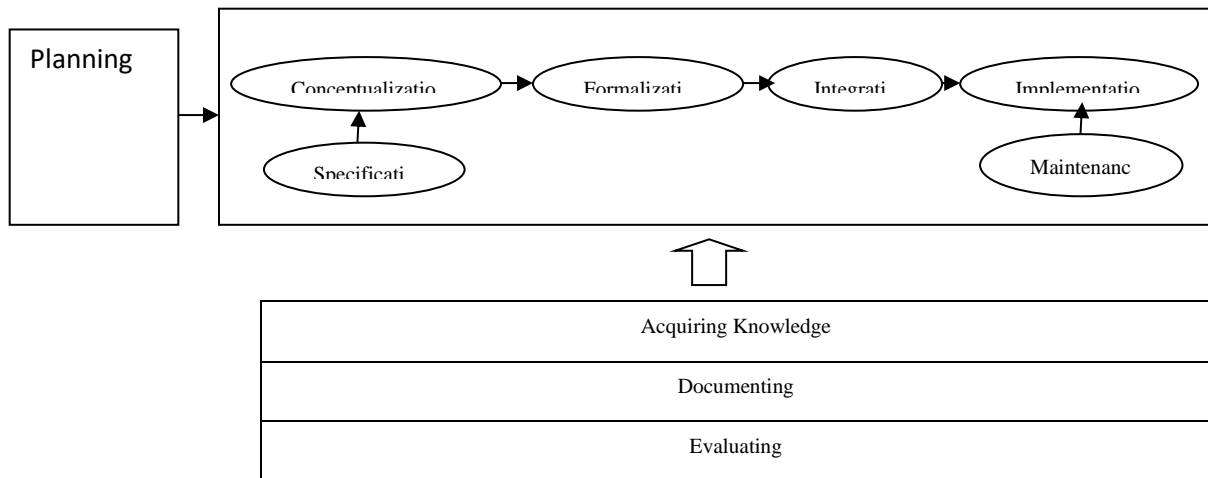


Figure1
States and Activities involved in METHONTOLOGY

The ontology editor Protégé was used to formalize the developed ontology. Protégé is an open-source ontology editing software developed by Stanford University in collaboration with University of Manchester. Desktop version of Protégé (Protégé 4.3.1) used for the present study. The developed theses ontology has been evaluated by a group of scholars from Pondicherry University.

4. Development of Theses Ontology

The steps involved in the development of Library and Information Science theses ontology based on Methontology is explained here.

Step 1 Specification

The intension of the specification phase is to develop a document in natural language about the ontology. METHONTOLOGY defines that this specification document should include the domain, the need of the ontology, expected users, and level of formality, scope and source of data.

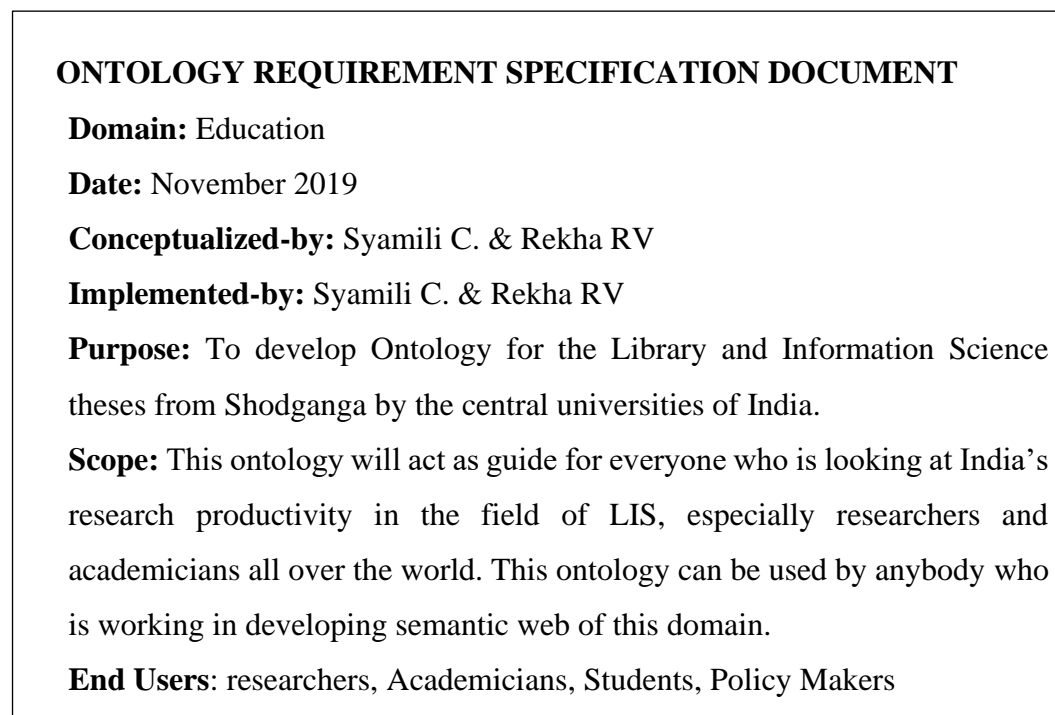


Figure 2

Ontology Requirement Specification Document

Step 2 Conceptualization

Conceptualization step is to identify the various terms and concepts in the domain. The knowledge for this particular ontology is acquired from Shodganga. The domain of the ontology is the Library and Information Science theses in Shodganga by the central universities in India. Out of the 46

central universities in India only 5 universities were uploaded theses in Shodganga. The details are given in the Table 1.

Table 1
Theses in Shodganga

Sl. No	Name of central University with URL	Number of Thesis	Percentage
1	Manipur University http://www.manipuruniv.ac.in	23	18.85
2	Mizoram University http://www.mzu.edu.in	8	6.55
3	Aligarh Muslim University http://www.amu.ac.in	86	70.4
4	Indhira Gandhi National Open University http://www.ignou.ac.in	1	0.81
5	University of Delhi http://www.du.ac.in	4	3.2
Total		122	100

Step 3 Formalization

Formalization of ontology means the structuring of the domain in a conceptual model. Ontology specification and domain vocabulary identification is happening at this stage. After identifying the structure of domain, the developer split the concept into terms. This includes identifying the elements, the attributes. It covers everything irrespective of their hierarchical structure. In defining the class hierarchies there are three approaches. They are top down, bottom up and middle out approach. Top down approach gives good control over the details in context of ontology development. Bottom up requires high degree of effort and among these METHONTOLOGY suggest Middle out approach as a strategy in identifying concepts. It helps in identifying primary concept first. Later one can move on either to specialization or to generalization.

Step 4 Integration

Arranging complete list of terms, relations and characteristic in the previous phases are done in this step. This classification of the facets in their logical and hierarchical order creates knowledge.

It is equally important that what others done in this domain earlier. If so, the ontologist can either refine or extend the sources. There are many ontologies available in electronic form and these can be imported and modified. For the searching of ontology there are a number of ontology libraries available. Some of them are Ontolingua, DAML ontology library, Rosetta Net DMOZ etc. Here the researcher conducted a thorough literature search and found that there is much ontology in education domain, there is no such ontology for theses and dissertations. Ontology in education domain is needed and can be used in e learning applications. The primary tree of the developed ontology is shown in Figure 3.

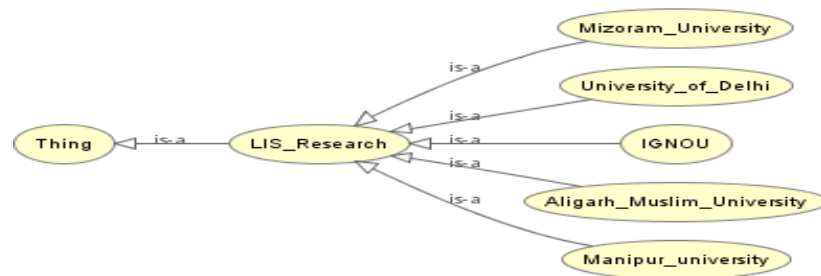


Figure 3

OWL Viz Plugin Visualization of the developed ontology

The relationship between entities in the theses ontology is represented in Figure 4.

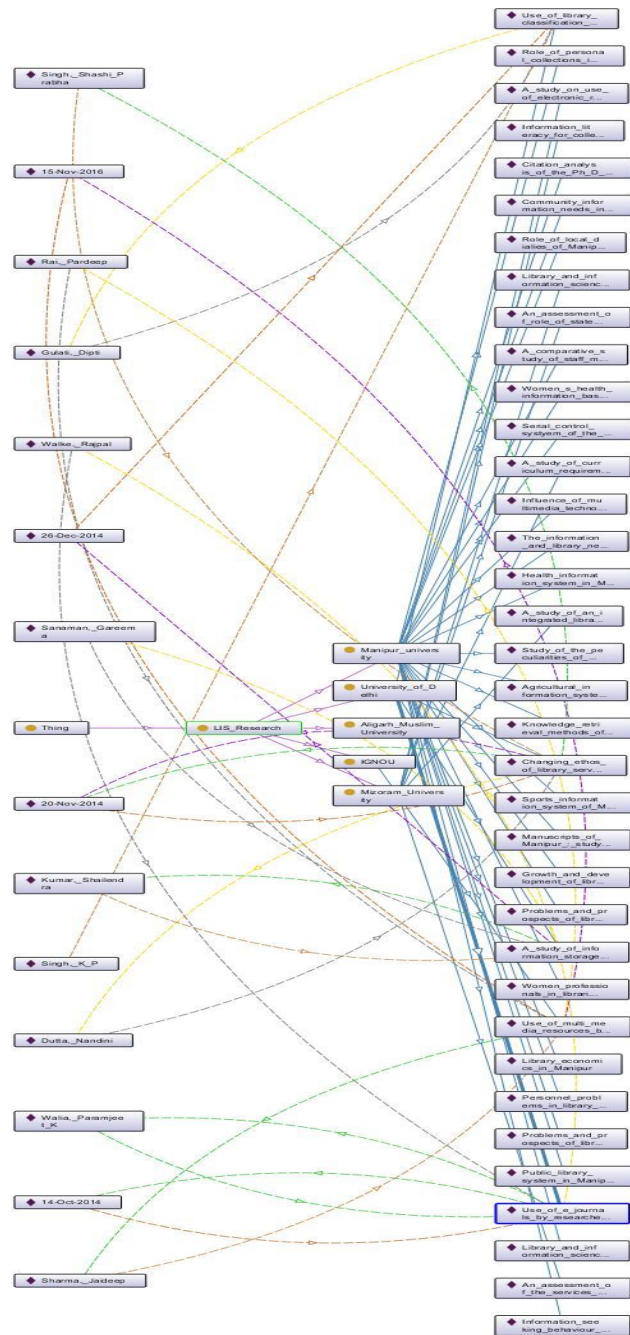


Figure 4

The Screenshot of the Graphical Representation of relations between Entities in the Ontology

Step 5 Implementation

There are a number of Ontology Editors available such as Protégé, SWOOP, Onto Edit, Altova Semantic Works, Onto Studio etc. Among these Protégé is most widely used by professionals, programmers, Researchers and others. So Protégé is using for the present study to formalize the developed ontology.

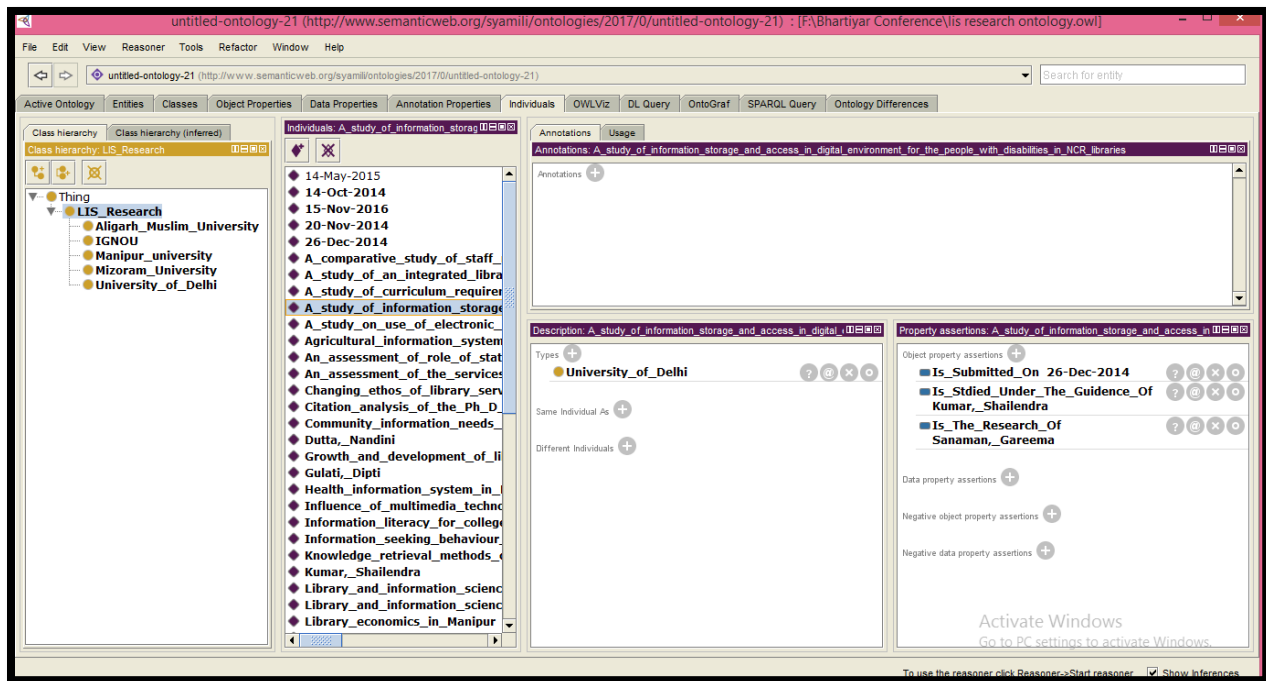


Figure 5

Theses ontology development using Protégé

For each group, abstracting a set of similar featured terms, a standard term is selected to label the group while all other synonymous terms are recorded separately. The process of adding relations among the entities in order to achieve semantic richness is shown in Figure.ie. Rai Pardeep *is the researcher of* the study Use of multimedia resources by the students of distance education in IGNOU: A study of Delhi region. The same entities are also connected using the inverse relation Use of multimedia resources by the students of distance education in IGNOU: A study of Delhi region *is conducted by* Rai Pardeep



Figure6
Relation among two entities in the ontology

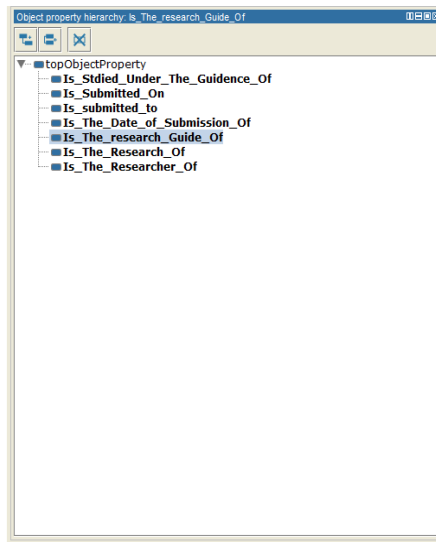


Figure 7
The Properties & Relations followed in the Ontology

Step 6 Maintenance and Evaluation

After formalizing the developed ontology in Protégé, Ontology evaluation has to be carried out. There are different methodologies to evaluate ontology. Evaluation using user’s point of view is one among them. The developed theses ontology has been evaluated by a group of scholars from Pondicherry University. This team includes both men and women from different disciplines. All these evaluators have used Shodganga at least once in their research time. The age groups of the scholars are in between 24 and 40. Interview took place during the span of two month.

The entire ontology was then redrawn on a chart paper. Different colored pens were used to denote relationships, attributes and instances. The researcher introduced the ontology to the

participants. Participants were requested to write down their query and find the answers manually by making search paths. The participants were also asked to mark their answers as satisfactory, partially satisfactory and unsatisfactory.

Table 2
Evaluation of Theses Ontology

Competency Level of Evaluator	Evaluator	Level of Satisfaction			Total Queries
		Satisfact ory	Partially Satisfactory	Unsatisfacto ry	
Evaluators from within the domain	Evaluator 1	5	2	1	8
	Evaluator 2	4	1	0	5
	Evaluator 3	7	3	2	12
	Evaluator 4	11	4	2	17
	Evaluator 5	6	2	1	9
Evaluators from outside the domain	Evaluator 6	5	4	3	12
	Evaluator 7	3	0	2	5
	Evaluator 8	6	1	1	8
	Evaluator 9	9	3	2	14
	Evaluator 10	8	4	2	14
Total		64	24	16	104

The result is diagrammatically represented in the Figure 8. The theses ontology could answer 62 percentages of the questions, i.e. 64 out of the 104 questions. The remaining questions were either partially answered or unanswered questions. These questions were analyzed in detail for further development of the ontology.

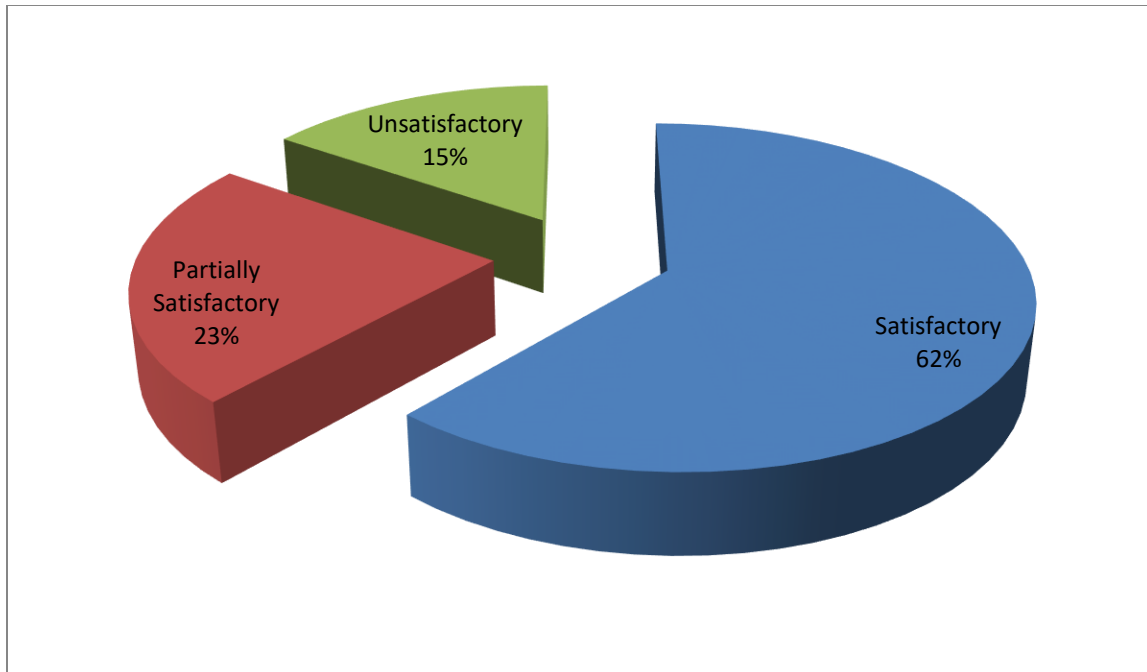


Figure 8
Satisfaction Level of Users

Some questions were omitted as they do not come under the scope of the domain. Remaining questions were analyzed again to extract the missing concept and then fed the key terms into the ontology. The participants' query analysis thus further developed the ontology. After the analysis and addition of new entities This ontology could answer majority of the questions. Mostly asked questions and their key terms are listed in Table 3.

Table 3

List of selected questions and keywords identified

What are the studies related to Digital library?	<Keyword, Digital Library>
Who is the most Productive guide in the field of LIS?	<Guide, number of thesis>
Which is the most productive University in terms of number of produced theses?	<University, number of theses>
Which is the latest submission in Delhi University	<University of Delhi, date of submission>
Who is the author of thesis Use of multimedia resources by the students of distance education in IGNOU: A study of Delhi region	< Title: Use of multimedia resources by the students of distance education in IGNOU: A study of Delhi region, Author>

The missing concepts were fed into the ontology after the evaluation. Hence the ontology after evaluation could answer the majority of the questions raised in the evaluation. More and more concepts can be added into the ontology and this ontology is not a complete and this will develop further. The Table 4 shows the statistics of the ontology. At present the ontology contains 247 concepts, 244 Individuals, 125 Object property assertions, 85 Data property assertions.

Table 4

Theses Ontology - Statistics

Attributes	Count
concepts	347
Object property assertions	125
Individual count	244
Data property assertions	85

Step 7 Documentation

There is no specific format for documenting ontology. The document available at the end of the development is the code of the ontology. Along with this natural language description of step by step approach adopted during the development should be published. METHONTOLOGY strongly argues that each ontology developer should develop document the process of developing and publish it or archive it for the further usage. The ontology is designed in such a way that it can be further modified in a later time. More concepts can be included. Definitely it will grow with time

5. Findings and Conclusion

The study attempted to design ontology for Library and Information Science theses of central universities in India to visualize the doctoral research output in Library and Information Science. This ontology contains 247 concepts, 244 Individuals, 125 Object property assertions, 85 Data property assertions. Methontology is most widely used methodology in developing ontology irrespective of the time. Hence this study adopted the methodology called Methontology. From the study it is identified that Methontology mainly focus on design and implementation and there in not much technological support .Ontology editor Protégé, Web Ontology Language has been used for coding the ontology. The developed ontology was evaluated and made sure that it meets with the information need of its users and it has been found that the Ontology could answer majority of the questions after evaluation. The evaluation results are satisfactory and the model is valid. The evaluation reveals that this model can provide enough functionality to meet user's need. Furthermore semantic sharing between other applications is also possible. The ontology is currently published in web protégé, since that work will be used for an application of e learning applications. This ontology can be used by anybody who is working in developing semantic web of this domain. Only limited attempts have been done in developing ontology for the databases of theses and dissertations. This work is the first attempt to build ontology for theses and dissertations.

References

1. Ajetunmobi, S. A., & Daramola, O. (2017). Ontology-Based Information Extraction for Subject- Focussed Automatic Essay EvaluationFocussed Automatic Essay Evaluation. *International Conference on Computing Networking and Informatics*. IEEE.
2. Aroua, E., Baazaoui-Zghal, H., & Abed, M. (2017). An Ontology Driven Framework for Personalized Itinerary Visualization. *2017 21st International Conference Information Visualisation (IV)*, 152–157. <https://doi.org/10.1109/iV.2017.51>
3. De, A., And, N., & Missikoff, M. (2016). UPON Lite focuses on users, typically domain experts without ontology expertise, minimizing the role of ontology engineers. *Communications of the Acm*, 59(79), 79–86. <https://doi.org/10.1145/2818359>
4. Gruber, T. R. (1993). A trnslational approach to portable ontology specifications. *Knowledge Acquisition*, 5(2), 199-220.
5. Gruninger, M., and Fox, M. S. (1995). “Methodology for the Design and Evaluation of Ontologies”,. *Workshop on Basic Ontological Issues in Knowledge Sharing, IJCAI-95, Montreal*.
6. Ibrahim, A. M. (2013). Ontology-Driven Information Retrieval for Healthcare Information System : A Case Study. *International Journal of Network Security & Its Applications*, 5(1), 61–69. <https://doi.org/10.5121/ijnsa.2013.5105>
7. Liu, Y et al.(2021) Ontological modeling and analysis of experimentally or clinically verified drugs against coronavirus infection. *Scientific Data*. 8(1) <https://doi.org/10.1038/s41597-021-00799-w>
8. Phan, C.-P., Nguyen, H.-Q., & Nguyen, T.-T. (2018). Ontology-based heuristic patent search. *International Journal of Web Information Systems*, IJWIS-06-2018-0053. <https://doi.org/10.1108/IJWIS-06-2018-0053>
9. Uma, R., & Muneeswaran, K. (2017). OMIR: Ontology-Based Multimedia Information Retrieval System for Web Usage Mining. *Cybernetics and Systems*, 48(4), 393–414. <https://doi.org/10.1080/01969722.2017.1285163>
10. Yu, T., Cui, M., & Jia, L. (2016). Towards a TCM domain ontology: Standardization, ontology engineering, and applications. *Proceedings - 2015 8th International Conference on BioMedical Engineering and Informatics, BMEI 2015*, (Bmei), 500–504. <https://doi.org/10.1109/BMEI.2015.7401556>

11. Zou, Q., & Park, E. G. (2018). Linking historical collections in an event-based ontology. *Digital Library Perspectives*, 34(4), 262–275.
<https://doi.org/10.1108/DLP-02-2018-0005>