I.J. Education and Management Engineering, **2016, 3, 9-19**
Published Online May 2016 in MECS (http://www.mecs-press.net)
DOI: 10.5815/ijeme.2016.03.02

Available online at http://www.mecs-press.net/ijeme

Ontology Engineering and Development Aspects: A Survey

[a] Usha Yadav, [b] Gagandeep Singh Narula, [c] Neelam Duhan and [d] Vishal Jain

[a] Research Scholar, Department of Computer Engineering, YMCA University of Science and Technology, Faridabad, India
[b] Research Scholar, M.Tech CSE (IV Sem), C-DAC, Noida
[c] Associate Professor, Department of Computer Engineering, YMCA University of Science and Technology, Faridabad, India
[d] Assistant Professor, Bharati Vidyapeeth's Institute of Computer Applications (BVICAM), New Delhi

Abstract

Ontology can be defined as hierarchical representation of classes, sub classes, their properties and instances. It has led to understanding the concepts of given domain, deriving relationships and representing them in machine interpretable language. Ontologies are associated with different languages that are used in mapping of multiple ontologies. Several applications of ontologies have led towards realization of semantic web. The current web (2.0) is approaching towards semantic web (3.0) that performs intelligent search and stores results in distributed databases.

The paper makes readers aware of various aspects of ontology like types of ontology, ontology development life cycle phases, activities involved in ontology development and ontology engineering tools. Ontology engineering contributes to meaningful search and provides with open source tools for deploying and building ontologies.

Index Terms: Semantic web, Ontology, RDF, OWL and Ontology Engineering.

1. Introduction

Ontology is considered as backbone of semantic web. It is essential to understand meaning of each term along with classes, properties and instances associated with it. It is possible only if we have categorized information. This categorization of information in hierarchical manner is termed as Ontology.

It itself is associated with wide range of concepts like merging, mapping, engineering and development which are considered as individual research areas. Ontology mapping and merging involves integration, aligning and reusing of data so that it can be used with existing web applications. Ontology engineering

* Corresponding author.
E-mail address: usha.yadav.912@gmail.com, gagan.narula87@gmail.com, neelam_duhan@rediffmail.com, vishaljain83@ymail.com

involves use of automatic tools for managing, mapping and integrating ontologies in order to extract knowledge from them. An ontology mapping approach must fulfill given requirements:

- It should be able to generate mappings automatically.
- User should be able to accept, reject and add mappings.
- Mappings should be defined and organized systematically.
- Knowledge must be explored from derived mappings.
- Hierarchical relationships between concepts and instances must be considered

The given paper is categorized into following sections. Section 2 gives brief overview of semantic web in order to relate it with ontology. Section 3 defines various definitions of ontology and its types. It also defines classes, properties and instances of given sentences in OWL format. This section also throws some light on importance of ontology in semantic web. Section 4 describes scope of ontology engineering by introducing METHONTOLOGY approach. Section 5 extends this METHONTOLOGY approach by discussing various activities needed in development of ontologies. Section 6 presents survey of various ontology engineering tools followed by conclusion and references.

2. Semantic Web (Futuristic Web 3.0)

The word semantic conveys "what it means" instead of just focusing on its structure. Semantic web aims to focus on discovering meaning of content that can be understandable by machines as well as humans. It is futuristic web (web 3.0) whose target is to present capabilities of an agent in performing specific tasks. An agent can be entity with/without body who works in an autonomous way by communicating with other agents for performing specific tasks. For instance, one of the agents utilized for extracting data is mining agent while facilitator agent is used for communicating with each other. Few views about nature of semantic web are as follows:

(a) **The machine readable data view:** The idea of semantic web as envisioned by Tim Berners Lee defines data on web and linked in such a way that it can be understandable by machines and utilizing it in various applications [1].
(b) **The intelligent agent's view:** It involves use of agents for retrieving and manipulating data in order to make searching process intelligent.
(c) **The distributed database view:** The information retrieved is stored in databases that need to be mapped to ontologies for knowledge representation of concepts. Semantic web provides flexibility by linking databases and deriving logic rules from them in order to make data more understandable.
(d) **The improved searching view:** Semantic web employs use of agents that performs meaningful search and produces relevant results in context of given query. With the help of Web 3.0, it will become possible to have semantic based search instead of traditional keyword based search.

3. Definitions of Ontology

Guarino and Giaretta (1995) defined seven definitions of ontology (Fig 1) that can be used in different disciplines viz philosophical discipline, formal, specified, conceptualization of system using logical theory and many more. They stated that logical theory helps developers to build ontology as "*A logical theory which gives an explicit, partial account of a conceptualization." [3]*.

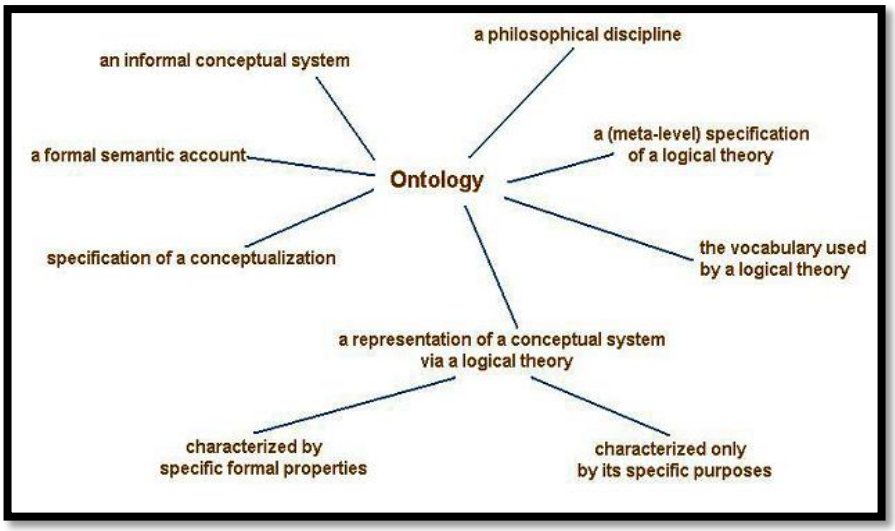

Fig.1. The Seven Definitions of Ontology

In other words, Ontology can be defined as hierarchical representation of classes, sub classes, their properties and instances. It has led to understanding the concepts of given domain, deriving relationships among them and representing them in machine interpretable language [2]. Ontology as philosophical science implies what is structure of objects, events, properties in every area of reality. Ontologies are associated with different languages that are used in mapping of multiple ontologies. Initially RDF and XML were developed in which XML specifies syntax of content rather than its semantics while RDF points to semantics of data. After these startup languages, more expressive and defined languages came into existence. One of the most common ontology languages includes OWL (Web Ontology Language) developed by W3C, DARPA Agent Markup Language (DAML) developed by Defense Advanced Research Projects Agency (DARPA), OIL (Ontology Interface Language) developed by Europeans and DAML+OIL.

- **Below is an example of OWL specification:**

Football and Basketball are sports.
Football is not Cricket.
Football is not Basketball.
< owl: Class rdf: about = "#Football">
<owl: disjointWith rdf: resource "#Cricket"/>
<owl: disjointWith rdf: resource = #Basketball"/>
</owl:Class>
<owl: Class rdf: ID = "Football">
<owl: equivalentClass rdf: resource = "Basketball"/>
</owl: Class>

- **Below is an OWL specification for Property.**

Gagan plays football.
< owl: ObjectProperty rdf: about = "#plays">
<rdfs domain rdf: resource = "#Gagan"/>
<rdfs: range rdf: resource = "#football"/>
</owl: ObjectProperty>

- **Below is an OWL specification for property restriction.**

Basketball is only being played by collegestudents.
< owl: Class rdf: about = "#"Basketball">
<rdfs: subClassOf>
<owl: Restriction>
<owl: onProperty rdf: resource = "#played by">
<owl: allValuesFrom rdf: resource = #collegestudents"/>
</rdfs: subClassOf>
</owl: Class>

3.1. Types of Ontology [6]

Ontologies are generic knowledge bases that are sharable and acceptable. They are of different types:

- **Upper ontology**

Existing well defined structures of given domain point towards upper ontology. IEEE P1600 standard defines standard upper ontology (SUO) which covers 3D and 4D modeling dimensions of multiple ontologies. But 3D and 4D visualization aspects of ontologies have not been reached yet.

- **Heavy-weight and Light-weight ontologies**

Light weight ontologies covers traditional search engines like Google, Yahoo which presents hierarchical structure by taking minor conceptual domains into consideration. Heavy weight ontologies create hierarchical structure by paying attention to each concept, their principles as per philosophical motives and develop semantic relations among classes and subclasses. Example of heavy weight ontologies is upper ontology that builds instance models.

- **Domain and Task ontologies**

Task ontologies are used to specify features of architecture of given knowledge based system whereas domain ontologies specify features of knowledge related to given domain on which various tasks like diagnosis, monitoring, design etc. are performed. Task ontology provides theory of all concepts/vocabularies used in existing structure while domain ontology defines relationship among classes, properties and instances used in designed ontology.

3.2. Construction of Ontologies

- **Building ontologies manually**

According to Noy and McGuiness [4], following are the steps required in the development of ontologies manually.

1. Scope identification: - Ontology is a knowledge base model consisting of classes, properties and their instances (CPI) related to particular domain. So, it is viable to determine scope of given problem and then proceed further.
2. Elaborate resources: - Specify tools and editors used to build ontology and formalize them.

3. Define taxonomy: - After identification of scope and resources, the concepts are organized in hierarchical fashion. It is called as taxonomy. If football is subclass of sport, then all instances in football must be an instance of sport class.
4. Define properties: - Properties needs to be specified from statements. Example: Students read books. Read is their property.
5. Define facets: - It implies use of RDF Schema (rdfs) to express desired ontologies.
6. Define instances: - If book is class, title is subclass and contents in this chapter are instances.
7. Verify for ambiguities: - Ontologies need to be verified so that they cover scope of given domain.

- **Re-using existing ontologies**

Ontologies available from third parties and existing online ontologies can be improved, re-used and modified with the help of web ontology editor tools like Web Protégé, Hozo, Knoodl, Vitro etc.

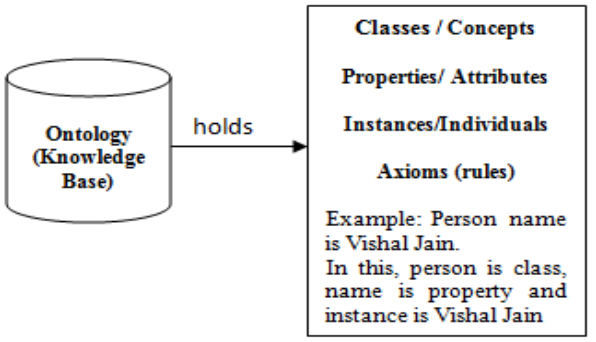

Fig.2. Gruber Ontology Structure

3.3. Role of Ontology in Semantic Web

- Ontology is considered as backbone of semantic web that plays vital role in representing meaningful information from huge amount of unstructured web data.
- Removes word sense disambiguation (WSD),
- Re-use and analyze domain knowledge.
- Helpful in solving reasoning problems, classification and problem solving techniques.
- Helps in achieving interoperability in semantic web.

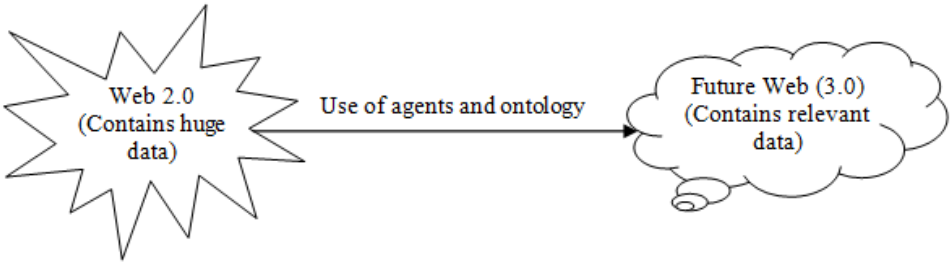

Fig.3. Moving Towards Semantic Web

4. Ontology Engineering

Automated use of tools for developing ontologies is called as Ontology Engineering. It is emerging areas in knowledge management and semantic web development sectors. In other words, ontology engineering is a set of activities that are performed during various phases of ontologies like conceptualization, formulization, design and deployment of ontologies. The main target of ontology engineering is "To build models of every domain and interpret them in machine language". It also includes concepts related to target domain and derives relationships among them. Various ontology editor tools are available for building and deploying ontologies. Some of them are Protégé, OntoEdit, Hozo and many more. These tools consists manuals for development of ontologies and their axioms.

Ontology engineering is time taking task that involves life cycle models. One of the examples of life cycle model is METHONTOLOGY approach. This approach performs development oriented activities (specifying requirements, conceptualization, design, formalization of conceptual model related to particular domain and maintenance of ontologies), support or integral oriented activities (knowledge gathering, integration and documenting ontologies) and management activities [8].

The scope of ontology engineering varies widely but is not limited to philosophy, reusing and sharing knowledge and designing ontologies.

5. Activities Involved In Methontology Approach for Ontology Development

Development of ontology is as complex as measuring quality of software. It requires each and every minute detail of activities and tasks from plinth to paramount. Ontology engineer needs to go through all development methodologies and existing design principles for reaching at some conclusions. Development-oriented activities are sub divided into pre-development, development and post development processes that occur sequentially while Support oriented activities are conducted in parallel manner in conjunction with any of development activities.

5.1. Management Activities

These activities are performed for identifying type of tasks and verifying that designed ontology has covered required specifications.

- Scheduling**:** This activity defines timing of tasks, their dependencies and allocation of resources.
- Quality Assurance**:** This activity assures that the quality of designed ontology and its documentation satisfy user requirements.

5.2. Development-Oriented Activities

These activities are the backbone of ontology development process. It has three sub activities- pre-development, development and post development which are defined below.

- **Pre-development Activities**

Type of environment: It specifies type of application platform to be used for developing ontologies. It also includes selection of ontology editors as tools for defining classes, properties and their instances.

Feasibility study: It checks if given ontology can be built in given environment and complies with user requirements.

- **Development Activities**

Specification phase: - This phase consists of following activities-*domain vocabulary definition, identifying resources, identifying axioms, identifying relationships, identifying data characteristics, applying constraints and verification*. Domain vocabulary definition defines name and properties related to given domain. Identifying resources means grouping different URI's into single class. Identifying axioms are structures that define behavior of concepts. Identifying relationships means relating class with their subclasses with properties and instances. Identifying data characteristics presents features of resources and their relationships. Applying constraints defines restrictions to be used with classes and properties. Verification means checking proposed ontology model for correctness.

Conceptualization phase: - This phase creates a model in context of given domain and presents knowledge at knowledge level. Several strategies are presented for defining conceptualization viz. top down approach and bottom up approach. Top down approach begins with super class that extends to refine ontology structure. This approach is mostly used in philosophical sciences. Bottom up approach begins with databases that are sources of multiple data and then does refinement to develop suitable ontology. This process is followed by information extraction (IE) and ontology learning tools like Text-to-Onto.

Design Phase: - This phase proposes physical structure of designed ontology that is based on RDF model. RDF model consists of three triples-Resource, Property and Value. Example of RDF statement: Potato is cultivated in brown soil with fertilizer value K567. Here potato is resource, brown soil is property and K567 is value.

Formalization phase: - This phase produces ontology as output by using ontology tools. It transforms conceptual model into formal model that can be re-written in suitable syntax.

Implementation phase: - It implements formalized ontology with the help of any of semantic languages (OWL, SPARQL) and executing them using some reasoner like Pellet OWL Reasoner.

- **Post-development Activities**

Maintenance: Ontologies must be updated from time to time so that user gets effective and latest results. It has led to continuous evolvement of ontologies. Ontology maintenance approaches may be centralized or decentralized. Centralized specifies whole control of maintaining ontology is under single entity or group of entities. De-centralized approach provides control to everyone involved in maintenance team. It is faster and cheaper because tasks are divided among several entities. Well defined guidelines must be provided in context of maintenance of ontologies [8].

Re-using existing ontologies: - Ontologies can be re-used by various applications for formalizing knowledge into understandable form. First ontology can re-use second ontology by referencing elements of second ontology in its axioms. The rules or axioms of existing ontologies are adapted to generate new ontology. It is much easier task rather than developing ontology from scratch.

5.3. Support Activities

Knowledge acquisition: This activity gathers knowledge from different sources of information stored in repositories. The multiple sources may include domain expert knowledge, online books and ontologies.

Integration: Combining multiple existing ontologies in order to generate new ontology is called Integration of ontologies. The process holds ontology merging and ontology alignment. Ontology merging obtains new ontology derived from several ontologies belonging to similar domain while ontology alignment is used for identifying mapping between source ontologies. Proper documentation is needed in order to reuse and integrate existing ontologies.

Configuration management: It specifies identification, documentation, recording and reporting of different

versions of ontology. Project management tools and ontology editing environments can be used for performing process configuration management like change request form, control and many more.

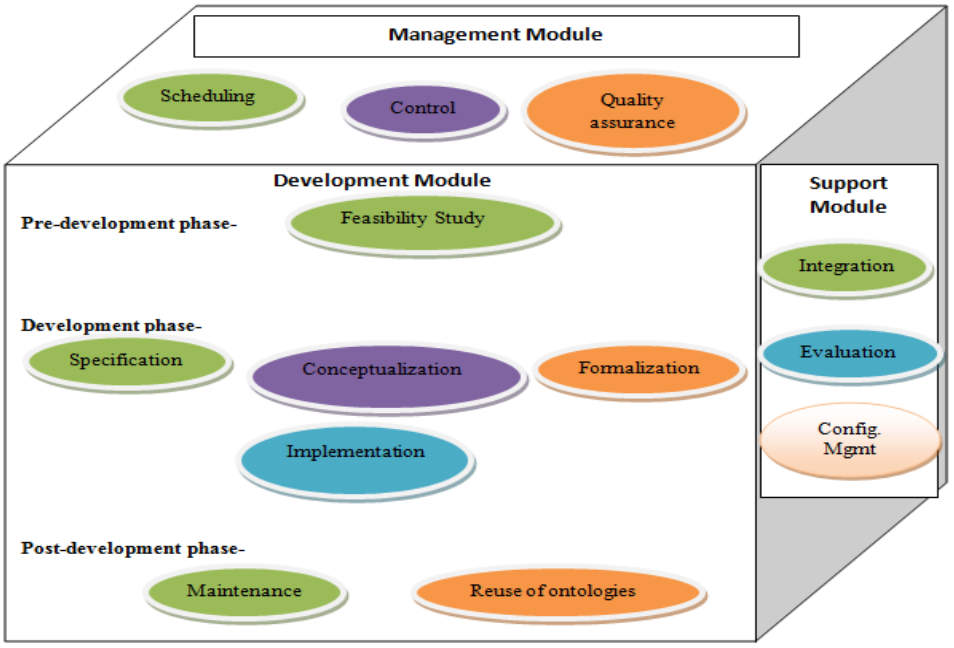

Fig.4. Activities of the Ontology Development Process

Table 1. Comparison of Ontology Engineering Tools

S.No	Tool / Features	Protégé	WebODE	OntoEdit
1.	Collaborative	Yes	Yes	No
2.	Backup recovery	No	yes	No
3.	Querying	Yes	No	Yes/No
4.	Import	RDF(s), OWL	RDF, OWL, DAML+OIL	RDF, DAMIL+OIL
5.	Storage	Files, JDBC	JDBC	Files
6.	Exception handling	No	No	No
7.	Reasoner	Pellet	Prolog	OntoBroker
8.	Base language	OKBC	HTML and Java	Flogic
9.	Merging	PROMPT	ODE Merge	Yes
10.	Implementation language	Java	Java	Java

6. Ontology Engineering Tools

Various tools that support ontology engineering have been identified in Duineveld et al. (2000). Further evaluations on ontology evaluation frameworks have been done in EON 2002 workshop (Sure and Angele, 2002). KAON OImodeller (Bozsak et al., 2002; Motik et al., 2002) falls in category of KAON ontologies. It provides easy integration of ontologies into an enterprise infrastructure. KAON suite consists of KAON24 which is one of semantic web rule languages (SWRL). Protégé´(Noy et al., 2000) was first ontology editor that works with RDF and OWL and is provided with built in plug-ins. WebODE (Arpirez et al., 2001) is an 'ontology engineering workbench' that provides various services for ontology engineering. OntoEdit (Sure et al., 2002, 2003) is an open source framework that integrates and manages services of different extensions of given ontology. It also derives inferences on basis of different extensions of ontologies.

7. Conclusion and Future Scope

Ontology is considered as backbone of semantic web. The paper describes almost every aspect related to ontology. It covers ontology development, ontology engineering, activities needed in METHONTOLOGY approach for development of ontology and ontology engineering tools. The activities cover development, support and management of ontologies. These aspects of ontologies can be used for future development of knowledge bases. Ontologies can be re-used in different domains with the help of ontology engineering tools. Ontology engineering contributes to semantic directed search and helps in solving real world problems.

References

[1] P. Lambrix, "Towards a Semantic Web for Bioinformatics using Ontology-based Annotation", in proceedings of the 14th IEEE international workshops on Enabling Technologies: Infrastructures for Collaborative Enterprises, 2005, pp.3-7.

[2] Sowa, J.F., "Ontology, metadata, and semiotics", Lecture Notes in AI #1867, Springer-Verlag, Berlin, 2000, pp. 55-81.

[3] Gagandeep Singh, Vishal Jain, "Information Retrieval (IR) through Semantic Web (SW): An Overview", in proceedings of CONFLUENCE-The Next Generation Information Technology Summit, 27-28 September 2012, pp 23-27.

[4] Natalya F. Noy and Deborah L. McGuinness. "Ontology Development 101: A Guide to Creating Your First Ontology"'. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, March 2001.

[5] T. Gruber, "Towards principles for the design of ontologies used for knowledge sharing". Int. J. of Human and Computer Studies, 43(5/6):907–928, 1995.

[6] Riichiro Mizoguchi, "Tutorial on ontological engineering", New Generation Computing, 22(2004), 61-96, Ohmsha Ltd and Springer Verilag, 2004.

[7] Dayal U, Kuno H, "Making the Semantic Web Real", IEEE Data Engineering Bulletin, Vol.26, No.4, pp 4-7, 2003.

[8] Rudi Studer, Stephan Grimm, and Andreas Abecker (Eds.), "Semantic Web Services: Concepts, Technologies and Applications", Springer, 2007.

[9] Vishal Jain, Dr. Mayank Singh, "Ontology Based Information Retrieval in Semantic Web: A Survey", International Journal of Information Technology and Computer Science (IJITCS), MECS publishers, 2013, 10, 62-69.

[10] Amjad Farooq and Abad Shah, "Ontology Development Methodology for Semantic Web System", Pakistan Journal of Life Social Sciences, Vol.6 No.1, May 2008, pp 50-58.

[11] J. Mayfield, "Ontologies and text retrieval", Knowledge Engineering Review, 2007.

[12] Vishal Jain, Dr. Mayank Singh, "Ontology Development and Query retrieval using Protégé tool", International Journal of Intelligent Systems and Applications (IJISA), MECS publishers, 2013, 09, 67-75
[13] Kaushal Giri, "Role of Ontology in Semantic Web", DESIDOC Journal of Library & Information Technology, Vol.31 No.2, March 2011, pp 116-120.
[14] S. Luke, L. Spector, D. Rager and J. Hendler, "An Introduction to Ontology", "In Proceedings of the First International Conference on Autonomous Agents (Agents 97)", pages 59-66, 1997.
[15] M. Preethi, Dr. J. Akilandeswari, "Combining Retrieval with Ontology Browsing", "*International Journal of Internet Computing (IJIC), Vol.1, Issue-1*", 2011.
[16] Abdeslem DENNAI, Sidi Mohd., "Semantic Indexing of web documents based on domain ontology", International Journal of Information Technology and Computer Science (IJITCS), MECS publishers, 2015, 02, 1-11.
[17] Berners Lee, J.Lassila, "Ontologies in Semantic Web", "*Scientific American*", May (2001) 34-43.
[18] Gagandeep Singh, Vishal Jain, Dr.Mayank Singh, "Ontology Development Using Hozo and Semantic Analysis for Information Retrieval in Semantic Web" in 'ICIIP-2013 IEEE Second International Conference on Image Information Processing ', Jaypee Univ. Shimla, 9-11 Dec 2013.
[19] Xin Shi, Shurong Tong, "Ontology mapping of design process knowledge based on classification", International Journal of Education and Management Engineering (IJEME), MECS publishers, 2011, 5, 17-25.

Authors' Profiles

Usha Yadav received her B.E. in Information Technology with Honors from Maharshi Dayanand University, Rohtak in 2009 and M.Tech with Honors in Computer Engineering from YMCA University of Science and Technology, Faridabad in 2011. She is pursuing her Ph.D in Computer Engineering from YMCA University of Science and Technology, Faridabad. She is currently working as a Project Engineer in CDAC, Noida and has three years of experience. Her areas of interest are search engines, social web and semantic web.

Gagandeep Singh Narula received his B.Tech in Computer Science and Engineering from Guru Tegh Bahadur Institute of Technology (GTBIT) affiliated to Guru Gobind Singh Indraprastha University (GGSIPU), New Delhi. Now, he is pursuing M.Tech in Computer Science from CDAC Noida affiliated to GGSIPU. He has published various research papers in various national, international journals and conferences His research areas include Semantic Web, Information Retrieval, Data Mining and Knowledge Management. He is also a member of CSI.

Dr. Neelam Duhan received her B.Tech. in Computer Science and Engineering with Honors from Kurukshetra University, Kurukshetra and M.Tech with Honors in Computer Engineering from Maharshi Dayanand University, Rohtak in 2002 and 2005, respectively. She completed her PhD in Computer Engineering in 2011 from Maharshi Dayanand University, Rohtak. She is currently working as an Assistant Professor in Computer Engineering Department in YMCA University of Science and Technology, Faridabad and has an experience of about 12 years. She has published over 30 research papers in reputed

international Journals and International Conferences. Her areas of interest are databases, search engines and web mining.

Vishal Jain has completed his M.Tech (CSE) from USIT, Guru Gobind Singh Indraprastha University, Delhi and doing PhD in Computer Science and Engineering Department, Lingaya's University, Faridabad. Presently, He is working as Assistant Professor in Bharati Vidyapeeth's Institute of Computer Applications and Management, (BVICAM), New Delhi. His research area includes Web Technology, Semantic Web and Information Retrieval. He is also associated with CSI, ISTE.

How to cite this paper: Usha Yadav, Gagandeep Singh Narula, Neelam Duhan, Vishal Jain,"Ontology Engineering and Development Aspects: A Survey", International Journal of Education and Management Engineering(IJEME), Vol.6, No.3, pp.9-19, 2016.DOI: 10.5815/ijeme.2016.03.02