Survey Paper on "Ontology as a Driving Force"

Sumitra Pundlik, Shwetali Jori, Juilee Kapure, Anisha Gaikwad, ShwetaValunj

Abstract—Ontology, a branch of artificial intelligence, is a formal representation of concepts of a particular domain and relationships amongst those concepts. Ontology acts as a powerful tool or a driving factor for many real world applications and this paper presents some of those ontology-based approaches. This paper describes how ontology is modeled, implemented and used in Web Semantics, Business Process Networks and Knowledge and Application Engineering.

Keywords: Ontology, Semantic Web, Ontological Development

I. INTRODUCTION

Aristotle was the first to define what is an Ontology. From then and till now people like Thomas Gruber, Tim Berners Lee and many more have redefined Ontology. From all those definitions we can state that Ontology, a branch of artificial intelligence, is a formal representation of concepts of a particular domain and relationships amongst those concepts. In more simplified words, Ontology is the knowledge representation of a domain of interest. Ontology which is created with different approaches like pattern based extraction, Ontology pruning, data mining etc., acts as a powerful tool in many modern knowledge based systems. It supports natural language processing, information filtering, information retrieval and data access. Ontology acts as a driving force for many real world applications like the Semantic Web, Business Process Networks, Knowledge and Application Engineering and many more. In other words, Ontology acts as an approach for these real world applications. From [1] it can be inferred that Ontology building which is a automatic or semi-automatic process has the following as its fundamentals:

1) Collecting all the data or information related to the domain of interest as a textual corpus
2) Extracting relevant terms from the textual corpus
3) Identifying those as concepts.
4) Determining the relationships amongst the concepts.
5) Determining hierarchy in the concepts identified.
6) Determining the attributes, properties, rules and axioms for the concepts.
7) Populating the existing Ontology by mapping new concepts and individuals to it.

These fundamentals remain the same even if the approaches for building the Ontology vary. The classification generated from the target domain, termed as Ontology, is then implemented in various real world applications, which are discussed in further sections of this paper.

Manuscript Received on November 2014.

Prof. Sumitra Pundlik, Department of Computer, Pune University, Pune, India.
Shwetali Jori, Department of Computer, Pune University, Pune, India.
Juilee Kapure, Department of Computer, Pune University, Pune, India.
Anisha Gaikwad, Department of Computer, Pune University, Pune, India.
Shweta Valunj, Department of Computer, Pune University, Pune, India.

II. ONTOLOGY AND WEB SEMANTICS

According to Tim Berners Lee (1999) the Semantic Web is an extension of the current web in which information is given in a well defined meaning. And that, according to his vision, should enable the machines to "understand" the semantics of the web resources and, therefore, to have a more "intelligent" behavior in their activities of search. To state Tim Berners Lee's definition in a simplified manner we can say that Web Semantics is a process of making web surfing more intelligent or smarter. And it is realized using Ontology. According to [1] Ontology designs a unique path that the information required through queries must follow to arrive at the web resources containing the desired information. Also [1] says that Ontology is expected to provide structured vocabularies that give the relationships between different terms, allowing intelligent agents (and humans) to interpret their meaning. Therefore, [1] helps us to conclude that when Ontology is implemented within a web resource and when a user queries the web resource he/she gets the expected information. This is how Ontology acts as a driving force for Web Semantics or in simple words makes the searching activities smarter.

III. ONTOLOGY AND WEB CONTENT MANAGEMENT SYSTEM

According to [2] these days information over the web is growing at a tremendous rate and it is becoming extremely important to process this pure or raw information. The reason being this 'processed knowledge' has a lot of potential in search engines, product recommended system. Therefore, a tool like Collaborative Content and User-based Web Ontology Learning System holds a strong place. [2] adds that this system integrates two supervised learning approaches- Content-based Learning and User-based Learning Approach. The Content-based Learning Approach extracts Ontology concepts, constructs an Ontology Graph (OG) through the automatic learning of web documents and extracts relevant terms through text mining. The User-based Learning Approach extracts the subset of the Ontology Graphs using features analysis methods and builds a personalized Ontology by using intelligent agent approach so that user reading habit and preferences are captured through their semantic navigation and search over the Ontology-based web content. This system creates collaborative Ontology learning by doing a continuous Ontology matching and refinement process on the Ontology which is created from content-based learning and user-based learning. The specialty of this system is that it gives feedback and personalization along with the Ontology semantic web. [2] Elaborates the two approaches in the following manner:

Content-Based Learning Approach:
This Content-based Ontology Learning Process is comprised with 4 main steps. They are-
1) Textual Analysis
2) Concept Selection
3) Ontology Learning and
4) Ontology Validation.

[2] Depicts a figure involving above four steps

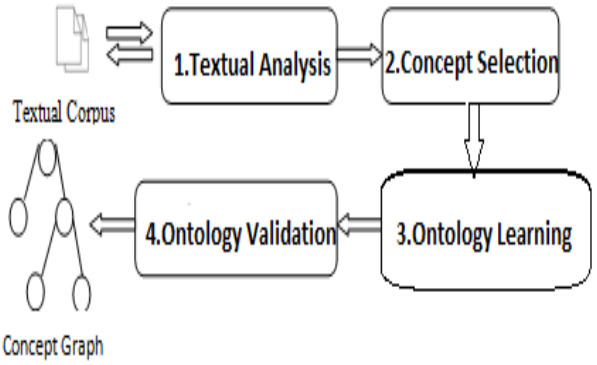

Figure:Four Content Based Ontology Learning Process [2]

Further [2] adds to Content-Based Learning Approach that it defines the Ontology Graph(OG) which is actually the outcome of Ontology learning.[2] also depicts a figure of OG which defines different types of knowledge units according to their level of complexity to comprise knowledge where a knowledge unit is any objects in the Ontology Graph that give semantics expression.

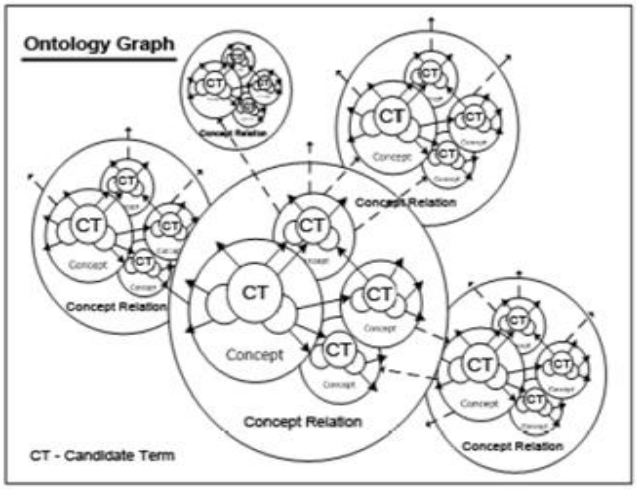

Figure:Ontology Graph [2]

The four components of the OG are described in [2] in the following manner:
1. Candidate Term (CT) – the smallest units that extracted in the form of a sequence of Chinese characters (for an illustration), those are meaningful words in human perspective.
2. Concept (C) – one or more candidate terms groups together with explicit relations to other knowledge unit, it is the basic knowledge unit in the Ontology graph.
3. Concept Relation (CR) – the weight direct relations between two concepts. That defines how two concepts relate to each other.
4. Ontology Graph (OG) – The entire knowledge unit created by groups of concepts, representing a comprehensive knowledge of the domain of a web channel (where web channel is a semantic web system which contains many general concepts with a set of specified domains. It basically includes a content management system, with advance search engine.).

User-Based Approach:
The OG defined in the former approach is used to give the personalization OG output which is accompanied by user reading preferences and each user's Personalization Ontology Search Agent learning result. Actually this approach makes use of a system that is based on the user reading habits and it captures preferences. These preferences are stored in XML and then the XML acts as an input to the user's Personalization Ontology Search Agent. The XML is for every user and is stored on his/her PC. User's Personalization Ontology Search Agent responds to learn user preferences, collect the user feedback, assist the user to search their interests, and report learning result to Ontology learner for further learning.

[2] Depicts OG in second phase i.e. the User Based Approach

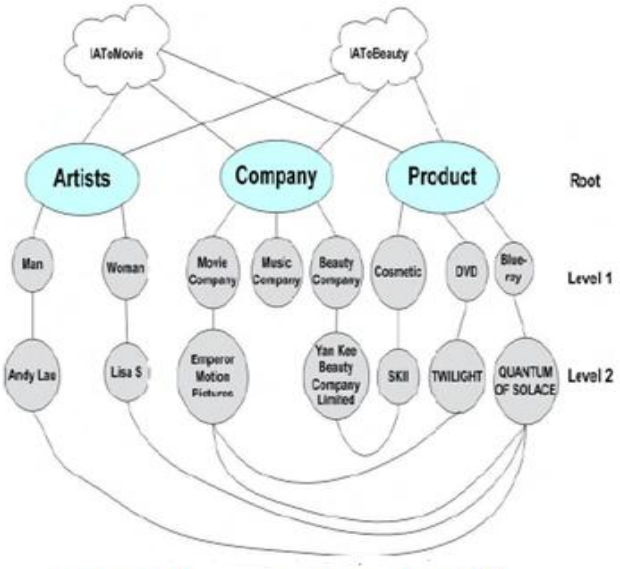

Figure:The Example Structure Of OG [2]

[2] Has given BuBo as the example of such a collaborative system.[2] then has elaborated the architecture of BuBo with the help of the following figure.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

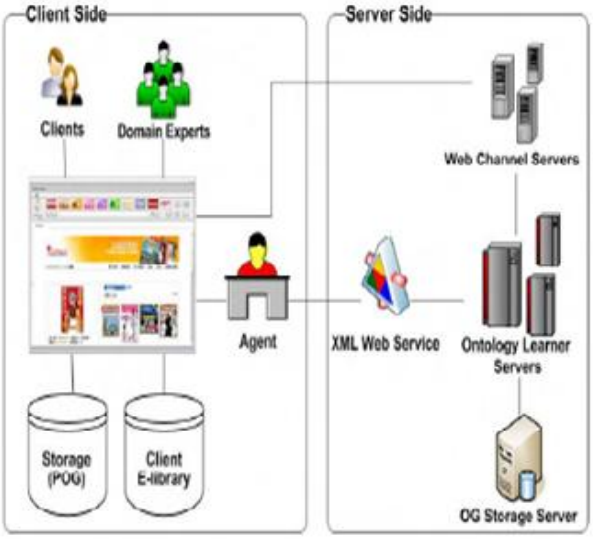

Figure:Ontology Learning Server And Web Channel Server [2]

IV. ONTOLOGY AND BUSINESS PROCESS NETWORKS

As the advance development in research field of AI planning and software engineering is increasing day by day, there is a huge need of process Ontologies. The Ontology can illustrate the need to link declarative and procedural process representations, for managing processes in development of applications and for integrating information from disparate information sources in multiple domains. Ontology supports the organization's wide decision- making and information management process in this modern era. In the near future, automated support will be required for such large scale design and analysis of the business processes. This automated support can be provided with the help of Ontology. From[3] we say that Ontology is a specification of what exists (or may be conceptualized) and what we can say about the world in a formal or informal manner.[3] further states that the major goal of building Ontology is to capture the knowledge about specialized domain. Process Ontologies collect the knowledge related to process models in various domains and development tools. Process Ontology basically supports both semantic of the domain knowledge and domain independent process management. According to [3] Ontology is illustrated by the development of the Semantic Web, Darpa Agent Markup Language (DAML), and Ontology Web Language (OWL). Recent software development of process Ontologies require such Ontology based languages to span languages such as informal textual process descriptions, structured data-flow programming diagrams, UML diagrams etc. But [3] then mentions that it is very challenging to build such complete Ontology in this open and dynamic environment, which will implement large scale business process model. The two techniques are mostly integrated to develop process Ontologies that support requirement analysis as well as software development. The process modeling is studied in the context of theory of automata and distributed system to define the operational semantics of computing system. These different aspects provide an Ontology for modeling and describing computational system from domain point of view. Likewise from application development point of view, domain modeling for system specification is supported with higher level abstractions such as data-flow diagrams and object-oriented modeling. [3] then states that single system can be described in a multiple number of ways with the help of nonstandard and proprietary Ontologies. Though it is very difficult to standardize Ontology, recent attempts to standardize such Ontology have led to evolution of multiple standards such as XPDL, ebXM and BPEL. [3] Adds that Model based software development promote a model driven system approach to standardize the process Ontologies embedded in the system. Ontology plays vital role in such business process modeling by exhibiting intelligent behavior or interacting with the world by providing decision-making and information-management processes to reason about percepts from the world and execute actions on the world. Process management hence focuses on knowledge representation or modeling the different entities that the agent may interact with, in the real world. [3] Mentions that Ontological development is done for formal knowledge representation for a real world phenomena in particular domain. Ontological approach planning provides a framework to reason about the effects of sequences of actions in a discrete dynamic system. The knowledge representation provides the system analysis which further can help in business processing. [3]then states that the utility of Ontologies in huge systems development is given in design of the CYC system.

V. ONTOLOGY AND APPLICATION AND KNOWLEDGE ENGINEERING

[4] says that ,now in many areas such as knowledge engineering, software reuse, digital libraries , web on the heterogeneous information processing, semantic web, information retrieval and so on Ontology is the base. It is the knowledge representation or a semantic network which focuses on the specific content of the domain. In the software field, any software which has the concept, relations amongst those, the description of the transaction processing, semantic interaction, as well as simple reasoning and logic rules is known as the software of the Ontology. [4] gives a development of Ontology. [4] gives description of the assessment phase. It states that the assessment phase is already developed Ontology which with the corresponding software environment tests the validity of the Ontology. Testing is done to check whether the target Ontology meets the requirements specification and answer questions at the beginning of the project dealing with the analysis of capacity issues. Also testing of the target application environment Ontology is done. [4] mentions that the prototype system should track the users for the concept and relationships according to their navigation or search. [4] then mentions the refinement phase in which the Ontology engineers need to perform several cycles until the Ontology reaches the specified level.

[4] mentions about the maintenance phase in which Ontology specifications often need to change in order to reflect the changes in the real world. In order to reflect these changes [4] states that it is required to safeguard the Ontology. Maintenance of Ontology is considered as a process of organization. For the internal Ontology updates - delete - insert the process [4] states that there should be strict rules. Recommended collection of Ontology changes if implemented and the possible impact of the application is thoroughly tested then it begins to transition to a new version of the Ontology. This model of Ontology thus, helps in decision making.

REFERENCES

1. Prof. Ernesto D'Avanzo, Prof. TsviKuflik, Candidato Antonio Lieto, " Manually vs semiautomatic domain specific ontology building", Corso di LaureaSpecialistica in Comunicazioned'impresa e pubblica, Tesi di Laurea in Informatica per ilCommercioElettronico, Anno accademico 2007-2008.
2. Edward H.Y. Lim, Hillman W.K. Tam, Sandy W.K. Wong, James N. K. Liu and Raymond S. T. Lee, " Collaborative Content and User-based Web Ontology Learning System", IEEE 2009.
3. MadhusudanTherani, " Ontology Development for Designing and Managing Dynamic Business Process Networks", IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, VOL. 3, NO. 2, MAY 2007.
4. QIU Wei School of Computer Science, Jia Ying University, Meizhou City, China, "Development and Application of Knowledge Engineering Based on Ontology", Third International Conference on Knowledge Discovery and Data Mining,2010.

AUTHORS PROFILE

Prof. Sumitra Pundlik, Educational details: SSC, May 1998, 73.20 %, S.B.B.A.K. Vidyalaya Hingoli. HSC, Feb 2000, 65.50 %, PCMB, Adresh Education College Hingoli. B.E., Jun 2004, 64%, Computer *Science and Engineering*, S.R.T.M.University, Nanded India. M.TECH May 2011 72.09%, Vishwakarma Institute of Technology, Pune, India.

Research work:

1. Asavari A. Deshpande, "Service Oriented Architecture in Financial Services", National Level **Conference 2007",** at lonavala.

2. P.S.Dhabe, M.S.Patwardhan. **A.A.Deshpande**," Articulated Entity Relationship (AER) Diagram for complete automation of relational database normalization", *International Journals of Database Management System (IJDMS)* **May 2010**, Volume2, Number2, Available at: http://airccse.org/journal/ijdms/curissue10.htmla

3. M.S.Patwardhan,P.S.Dhabe,**Asavari A.Deshpande**,Sandhya G.Londhe," DIAGRAMMATIC APPROACH FOR COMPLETE AUTOMATION OF RELATIONAL DATABASE NORMALIZATION AT CONCEPTUAL LEVEL", *International Journals of Database Management System (IJDMS)* **November 2010**, Volume2, Number4, Available at: http://airccse.org/journal/ijdms/curissue10.html

4. Prof.S.D.Sirsikar, **Asavari Deshpande** "Steganographic Tools for BMP Image Format", *International Journal of Computer Science and Emerging Technologies(IJCSET)*,**February 2011**,Volume2,Issue1 available at: http://ijcset.excelingtech.co.uk/

5. **Asavari Deshpande,** Anup Patil, Saurabh Joshi and Suraj Bothara. *"DBCrypto: A Database Encryption System using Query Level Approach"*. *International Journal of Computer Applications* 45(8):27-32, **May 2012**. Published by Foundation of Computer Science, New York, USA.

6. **Sumitra Pundlik** *,"SQLIJHS: SQL Injection Attack Handling System",* International Journal of Engineering Research & Technology (IJERT)ISSN: 2278-0181Vol. 2 Issue 6, June – 2013

Research Project: AER Modeler and Validator: Eclipse Plug-in for AER Diagrams [July'09-May'10]
The developed tool AERModeler and Validator is an Eclipse Plugin to draw and model AER Diagram [2] proposed by me as research work. It is an Eclipse Plug-in and thus can be accommodated as a part of the Eclipse Framework. It is implemented by extending the Graphical Editing Framework (GEF) and Draw2d. It follows Model View Controller Software Design Pattern for implantation. The Tool Helps to draw, Edit Validate the AER Diagrams .It also provides the XML export and SQL script generation functionality.
Membership:
1) ACM
2) IEEE PROFESSIONAL MEMBERSHIP
3) LMISTE