Chapter

Taxonomy and Ontology Management Tools: A General Explanation

Sukumar Mandal

Abstract

The World Wide Web is the result of a radical new way of thinking about sharing information. This idea seems familiar now, as the web itself has become pervasive. But this radical new way of thinking has even more profound ramifications when it is applied to a web of data like the semantic web. These ramifications have driven many of the design decisions for the semantic web standards and have a strong influence on the craft of producing quality semantic web applications. Until several years ago, the semantic web was primarily in a research phase. The new implementations of it were mainly to demonstrate the potential of the idea. While much of the activity related to semantic technology still takes place within the academic community, there are now real-world examples of the technology to use as a model. This paper has selected the matured level open-source tools for management of taxonomy and ontology in semantic web environment. Apart from this it also presents some important snapshot available in an online environment for designing and developing the taxonomy and ontology. This is very helpful to the users in using the concept of science and technology in cloud computing and online and offline environment.

Keywords: ontology, semantic web, TemaTres, visual vocabulary, and taxonomy

1. Introduction

This is the age of technology where digital information resources are increasing. It is a good concept for creating both the present and next-level automated and digital library system because it can highly be performed in the semantic web environment for executing the functional activities of libraries and any information resource centers. It helps the researchers for reviewing their literature. Here users can easily access and understand the terms and their relation from these interfaces for the better management of library and information services. It manages the big data in semantic web-linked data environment. It is also possible to manage the bibliographic link data in the visual vocabulary format, so the users can easily download the content as well as information. This is a web-based architecture in the Internet. Semantic web is one of the important aspects in cloud computing [1]. A lot of linking of web resources are available in an online environment [2]. The idea of a web of information was once a technical idea accessible only to a highly trained elite of information professionals: IT administrators, librarians, information architects, and other resources [3]. It is known as big data concept in semantic web level. It is based on open-source standards and formats including SPARQL, RDF, JSON, HTML,

XHTML, DTD, METS, MODS, etc. for managing the different web resources in any institutions or in any library [4]. These formats help to draw the visual graph of each uniform resource identifier for maintaining and managing the big data in the Internet [5]. Visual vocabulary is also an important concept in semantic web because it stores the large number of data and their linking also [6]. It is possible to manage the relations of different subjects by their specific identities, properties, and entities [7]. Big data is easily managed by using the open-source software and open standard formats in an online environment [8]. Users can easily access their necessary information by using this semantic web software in an Internet environment because it can manage the controlled vocabularies and their relationships of each element of any subject fields [9]. So, obviously it can save the time of reader and library professionals also [10]. It can manage the N-Triples, a format for storing and transmitting data, and Terse RDF Triple Language (Turtle), and XML provides an elemental syntax for content structure within documents yet associates no semantics with the meaning of the content contained within [11]. XML is not at present a necessary component of semantic web technologies in most cases, as alternative syntaxes exist, such as Turtle [12]. The World Wide Web contains many billions of pages. The SNOMED CT medical terminology ontology alone contains 370,000 class names, and existing technology has not yet been able to eliminate all semantically duplicated terms [13]. Any automated reasoning system will have to deal with truly huge inputs [14]. The objective of this chapter is very simple to highlight some popular taxonomy, ontology, and visual vocabulary through open-source tools for the users as well as library professionals. The important tools have been discussed in the next section. It also shows how it works in the web environment.

2. Objectives

The essential objectives of this study are explained as below:

i. To explore the modern tools and techniques those are available in an online digital environment for easy generation of taxonomy and ontology in different aspects.

ii. To highlights and visualize some terms by using these open source tool.

iii. To show the taxonomy concepts in an academic and research environment of different relations both technically and graphically.

iv. To design an integrated framework in an offline and online environment by using the open source tool and technique for each an every concept and their relations for the management of information resources.

3. Methodology

The process and methods of this study are one of the important aspects for designing and creating taxonomy and ontology. In this original research paper, the matured level open-source tools and techniques have been selected for the creation of visual vocabulary. Apart from this it also explores how to construct a visual thesaurus in an offline and online environment on Ubuntu operating system for easy constructing of the different terms and relations such as narrower term, broader term, related term, used for, scope note, and so on. So, this paper consists of two parts such as online and offline for term relations.

4. Visuwords

Visuwords is an online visual graphical dictionary as well as controlled vocabulary tool. Actually it provides the visual interface of each word and its respective meaning on the basis of English grammar. This is very user-friendly for definition of any words from any subjects can be represented in a separate color dashboard. Wheel and zoom are possible by using the mouse. It explores the word synonym, derivation, and antonym to extend the meaning of any facets available in an web-based environment (https://visuwords.com/library). **Figure 1** represents the Visuwords interface in an online environment for visual graphic dictionary. It is possible to easily identify the noun, verb, adjective, and adverb for construction the thesaurus against in a word.

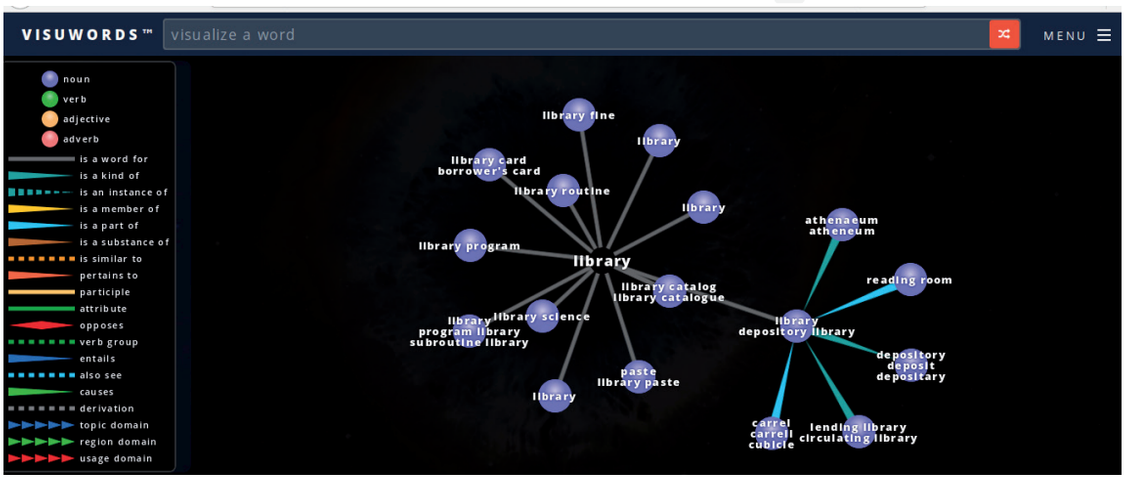

Figure 1.
Visuwords interface for visual graphic dictionary (https://visuwords.com/library).

5. Linked open vocabularies

Linked Open Vocabularies is a new concept in semantic web ontology. It consists of large amount of dataset around 660 vocabularies as on March 10, 2019. It fully supports the OWL and RDF framework for better management and identification of the right link available in an online database. SPARQL endpoint is easily accessed by using this interface. Properties and classes can be done on the basis of ontology concept that is subject, predicate, and object (https://lov.linkeddata.es/dataset/lov/). Here all the dataset belongs to the vocabulary content or element type. **Figure 2** represents the Linked Open Vocabularies interface in semantic web for different elements and their relationship.

6. WebVOWL

WebVOWL is a web-based ontology visualization open-source tool and is very interactive. This is written by high-level programming language named as Java which implements the visual notation and graphical interface of different nodal points in web ontology language. It explores the creation of ontology and combined forced graph in different formats such as JSON, RDF, and XML and other relevant file formats in an integrated semantic web environment (http://visualdataweb.de/webvowl/). **Figure 3** represents the WebVOWL interface in semantic web for linking a custom ontology based on URL and URI.

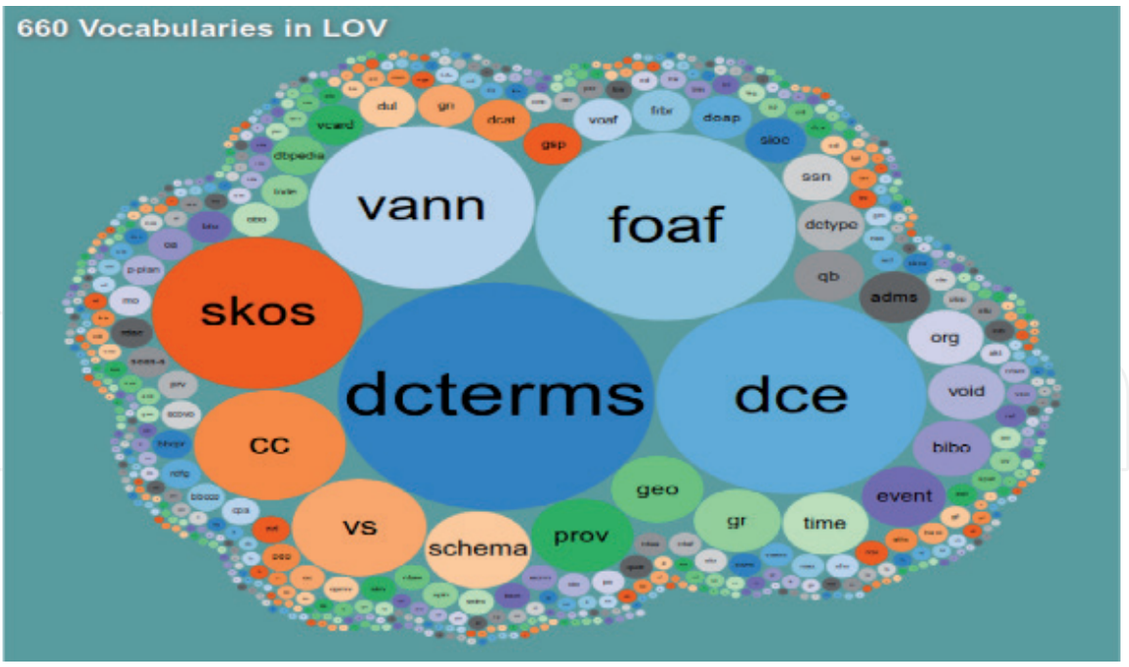

Figure 2.
Linked open vocabularies interface in semantic web ontology (https://lov.linkeddata.es/dataset/lov/).

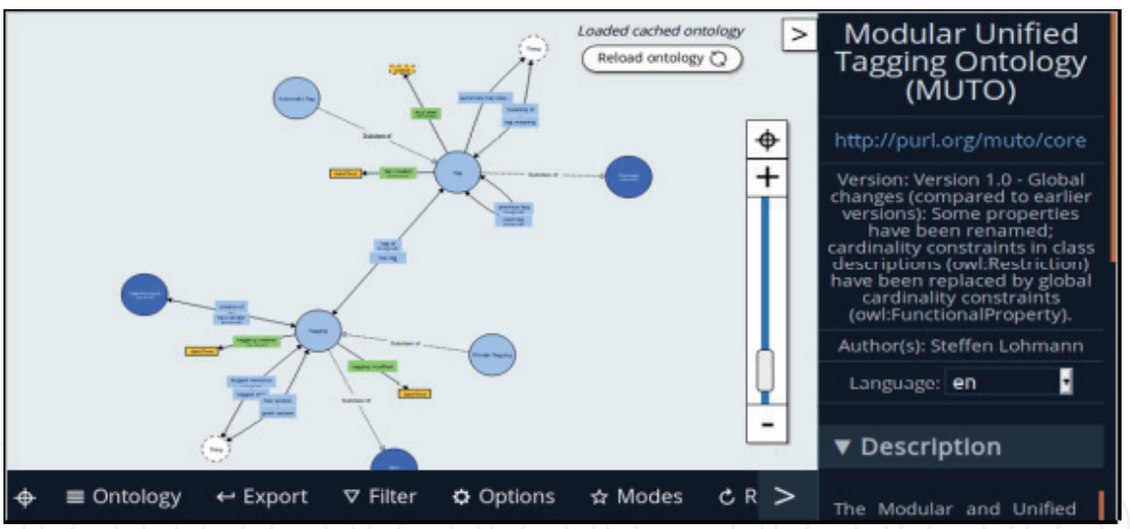

Figure 3.
WebVOWL interface in semantic environment (http://visualdataweb.de/webvowl/).

7. NavigOWL

NavigOWL is a graph visualization open-source Java-based tool in web ontology environment. This tool is a very interactive high-performance graph layout in semantic web environment. Here all ontologies are structured based which facilitate the patrons in thinking the mental map in an ontology environment. Protégé can be graphically represented by using this plug-in for better management of link and their elements (http://home.deib.polimi.it/hussain/navigowl/index. html). **Figure 4** represents the NavigOWL interface in web ontology environment. It loads and uploads the RDF/OWL ontology file for creation of graph and taxonomy in different nodes and edges.

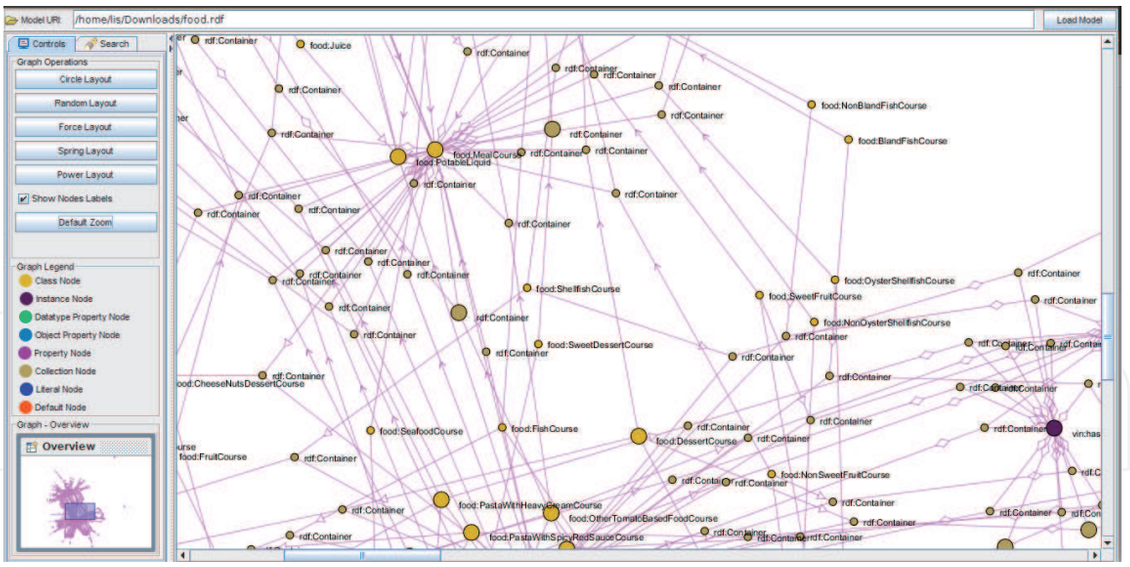

Figure 4.
NavigOWL interface in web ontology.

8. Unilexicon vocabulary

Unilexicon is a visual online controlled vocabulary and thesaurus construction server. This is a high-level open-source software for designing taxonomy and classification. It fully supports the faceted navigation and tagging for document and knowledge classification on the basis of simple knowledge organization system (SKOS). Institutional digital repository can easily be designed and developed by using the metadata and semantic search tool. This is also known as record management tool (https://unilexicon.com/). **Figure 5** represents the Unilexicon vocabulary interface. This is independent of content management system by using tag and facet count concept.

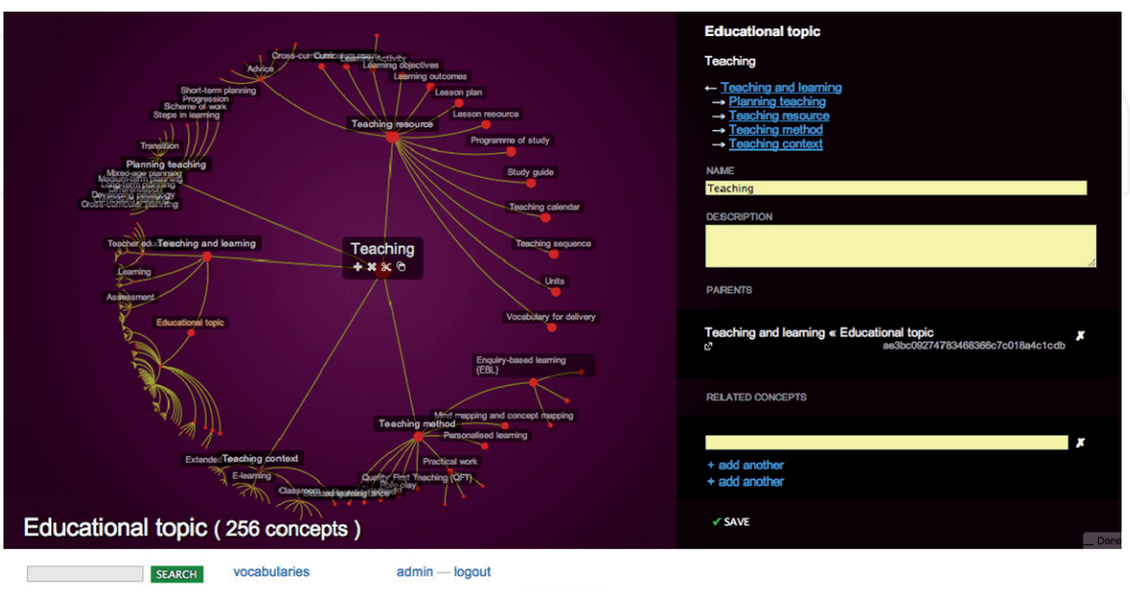

Figure 5.
Unilexicon vocabulary interface in taxonomy (https://unilexicon.com/).

9. Visual vocabulary TemaTres

TemaTres is a web-enabled open-source thesaurus construction software and written in PHP programming language. It is also known as vocabulary server for management formal representations of knowledge, thesauri, taxonomies, and multilingual vocabularies [15]. It has many features including SPARQL Protocol and RDF Query Language, Meta-terms (define facets, collections or arrays of terms), support for multilingual thesaurus, expose vocabularies with powerful web services, search terms suggestion, display terms in multiple deep levels in the same screen, search expansion, vocabulary harmonization, relationship between terms (BT/NT, USE/UF, RT), no limits to number of terms, alternative labels, levels of hierarchy, etc., systematic or alphabetic navigation, complete export in XML format (Zthes, TopicMaps, MADS, Dublin Core,VDEX, BS 8723, SiteMap, SQL), complete export in RDF format, complete export in txt, scope notes, historical and bibliographical notes, user management, terms and user supervision, duplicates terms control, free terms control, quality assurance functions (Duplicates and free terms, ilegal relations), multilingual interface, easy install, utility to import thesauri from tabulated textfiles, unique code for each term, terminology mapping with multilingual, term reports for editors, workflow like candidate, accepted and rejected terms, allow to create user-defined relationships, allow to define published and hidden labels, relationships between terms and web entities, export to WXP (WordPress XML) and import and export data in Skos-core (**Figure 6**).

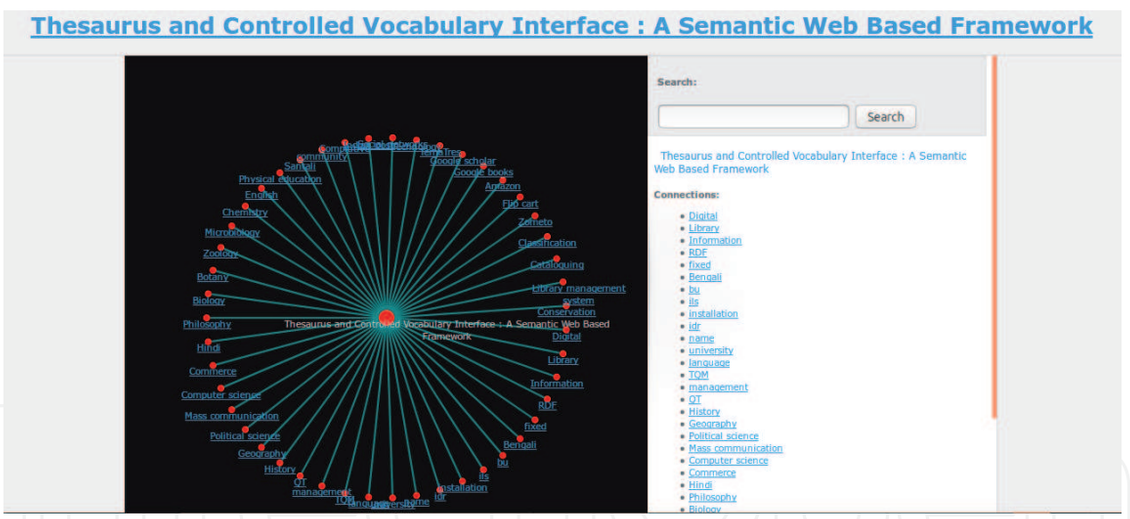

Figure 6.
Visual vocabulary interface in TemaTres.

10. Conclusion

Vocabulary creation is important for the users in any library including academic, public, and research libraries. It can manage and maintain the narrower terms and broader terms in respect to one term of a specific subject like philosophy, Bengali, history, chemistry, etc. Search interfaces of visual vocabulary are very sophisticated, and here quick search facilities are available also. The alphabetical index part contains each and every term, including synonyms, quasi-synonyms, and antonyms, occurring in the systematic part, along with its narrower terms. It controls the different terms used in indexing, providing a means of translating the natural language authors, indexers, and enquirers into a more constrained language used for indexing and retrieval. Visual vocabulary is one of the important concepts

of formulation of different words and their associated terms. All the terms and correlated terms against one specific subject are to be appeared in a visual vocabulary interface. It can manage the different terms including broader terms, narrower terms, and both preferred and non-preferred terms. In this section overview, the concept of a visual vocabulary is a strategy that draws inspiration from the text retrieval community and enables efficient indexing for different terms of a specific item types. Since the occurrence of a given word tends to be sparse across different documents, an index that maps words to the files in which they occur can take a keyword query and immediately produce relevant content. Obviously, it can conclude that the above tools highly overview the taxonomy and ontology in semantic web environment for easy access and download of metadata as well as full-text resources and their measures by graph. Thesaurus, taxonomy, and ontology are properly managed by using these tools for better management and retrieval of linked and resources in the web environment.

Author details

Sukumar Mandal
Department of Library and Information Science, The University of Burdwan, Golapbag, West Bengal, India

*Address all correspondence to: sukumar.mandal5@gmail.com

References

[1] Alfaries A, Bell D, Lycett M. Motivating service re-use with a web service ontology learning. International Journal of Web Information Systems. 2013;**9**(3):219-241. DOI: 10.1108/IJWIS-12-2012-0035

[2] Benslimane SM, Malki M, Bouchiha D. Maintaining web application: An ontology-based reverse engineering approach. International Journal of Web Information Systems. 2009;**5**(4):495-517. DOI: 10.1108/17440080911006225

[3] Bygstad B, Ghinea G, Klæboe G-T. Organisational challenges of the semantic web in digital libraries: A Norwegian case study. Online Information Review. 2009;**33**(5):973-985. DOI: 10.1108/14684520911001945

[4] Calaresu M, Shiri A. Understanding semantic web: A conceptual model. Library Review. 2015;**64**(1/2):82-100. DOI: 10.1108/LR-09-2014-0097

[5] Chi Y-L, Chen H-C. Ontology and semantic rules in document dispatching. The Electronic Library. 2009;**27**(4):694-707. DOI: 10.1108/02640470910979633

[6] Esserhrouchni OEI, Frikh B, Ouhbi B, Ibrahim IK. Learning domain taxonomies: The TaxoLine approach. International Journal of Web Information Systems. 2017;**13**(3):281-301. DOI: 10.1108/IJWIS-04-2017-0024

[7] Khan SA, Bhatti R. Semantic web and ontology-based applications for digital libraries: An investigation from LIS professionals in Pakistan. The Electronic Library. 2018;**36**(5):826-841. DOI: 10.1108/EL-08-2017-0168

[8] Ko YM, Song MS, Lee SJ. Construction of the structural definition-based terminology ontology system and semantic search evaluation. Library Hi Tech. 2016;**34**(4):705-732. DOI: 10.1108/LHT-08-2016-0090

[9] Lausen H, Ding Y, Stollberg M, Fensel D, Hernández RL, Han S-K. Semantic web portals: State-of-the-art survey. Journal of Knowledge Management. 2005;**9**(5):40-49. DOI: 10.1108/13673270510622447

[10] Llanes-Padrón D, Pastor-Sánchez J-A. Records in contexts: The road of archives to semantic interoperability. Program. 2017;**51**(4):387-405. DOI: 10.1108/PROG-03-2017-0021

[11] Navarro-Galindo JL, Samos J. The FLERSA tool: Adding semantics to a web content management system. International Journal of Web Information Systems. 2012;**8**(1):73-126. DOI: 10.1108/17440081211222609

[12] Nguyen H-M, Nguyen H-Q, Tran K-N, Vo X-V. GeTFIRST: Ontology-based keyword search towards semantic disambiguation. International Journal of Web Information Systems. 2015;**11**(4):442-467. DOI: 10.1108/IJWIS-06-2015-0019

[13] Nguyen Q-M, Cao T-D. A novel approach for automatic extraction of semantic data about football transfer in sport news. International Journal of Pervasive Computing and Communications. 2015;**11**(2):233-252. DOI: 10.1108/IJPCC-03-2015-0018

[14] Salvadori IL, Huf A, Oliveira BCN, Mello RS, Siqueira F. Improving entity linking with ontology alignment for semantic microservices composition. International Journal of Web Information Systems. 2017;**13**(3):302-323. DOI: 10.1108/IJWIS-04-2017-0029

[15] Mandal S. Developing thesaurus construction through Tematres for the college libraries under the University of Burdwan. International Journal of English Language, Literature in Humanities. 2016;**4**(6):302-316