ISSN: 2229-6959 (ONLINE)
DOI: 10.21917/ijsc.2016.0187

METHODS, APPROACHES, PRINCIPLES, GUIDELINES AND APPLICATIONS ON MULTILINGUAL ONTOLOGIES: A SURVEY

Ravi Lourdusamy[1] and Joseph Merlin Florrence[2]
Department of Computer Science, Sacred Heart College, India
E-mail: [1]raviatshc@yahoo.com, [2]merlinflorrence@gmail.com

Abstract

Ontologies are used in various domains to provide meaningful knowledge from data. The data can be in different natural languages. Ontologies with a lexical layer in different natural languages are needed essential to access the multilingual data by the users. Multilingual ontologies are needed to overcome this defect. Most of the multilingual ontologies are built up in English language and translated into the desired native languages. Many approaches are proposed to develop multilingual ontology applications. The aim of this article is to review the various approaches, methods for developing multilingual ontologies. The principles and guidelines for developing multilingual lexical layer are discussed. The different techniques adopted in the use and developments of multilingual ontologies are enumerated. The multilingual applications are compared using several dimensions.

Keywords:

Multilingual Ontology, Ontology Mapping, Ontology Matching, Ontology Aligning, Ontology Modeling, Ontology Localization

1. INTRODUCTION

Globalization enforces people to make use of information for their day to day life. Most of the information is available in a specific language which restricts the user from obtaining knowledge through different sources. Though many of the web pages have been developed with the translators, the translation becomes imprecise because most of the translators are translating the source text word by word. Translators cannot give the best results while conceptualizing the knowledge from source language to another language. Ontologies are used to solve the problem in knowledge acquisition while conceptualizing the concepts in different natural languages. Ontology is a formal, explicit specification of a shared conceptualization [1]. Ontology provides a shared vocabulary, which can be used to model a domain. It contains set of terms and concepts supporting the meaning of a subject area that are structured in a hierarchical format. Ontologies are used to improve communication machine and human beings. Internal representation of knowledge and data model can be reused with the use of ontologies.

Ontologies are commonly used in Artificial Intelligence and Knowledge Representation Systems. It establishes relationship between concepts of set of information in various natural languages. In Multilingual Ontologies the terms, concepts, attributes are specified in dissimilar natural languages. Multilingual ontologies are constructed using words that have same meaning across the different natural languages in the world. They are used as the names of concept and these words can be used to develop multilingual dictionaries. Multilingual ontologies are used to access the heterogeneous and distributed collection of digital resources. The aim of this article is to review the various research works carried out on multilingual ontologies and to enumerate different methods adopted in the use and development of multilingual ontologies. The multilingual applications are also compared using several dimensions.

The article is structured in the following order. Section 2 presents ontology development methods and needs to develop multilingual ontologies. The principles and guidelines to develop multilingual ontologies are also described in this section. The methods to develop multilingual ontologies are discussed in section 3. The approaches and tools that are used to develop multilingual ontology applications are also narrated in this section. Multilingual ontology applications developed in various domains are illustrated in section 4. The challenges and future enhancements are proposed in conclusion.

2. MULTILINGUAL ONTOLOGIES

Ontologies are used to represent knowledge in an effective manner. Building ontologies to acquire knowledge involves several methodologies and methods. All the methodologies and methods are classified into five categories: 1) extending the existing ontology, 2) learning relations for an existing ontology, 3) ontology construction based on clustering, 4) ontology construction based on semantic graphs and 5) ontology construction from a collection of news stories based on named entities [2]. The classical methodologies and methods used to build ontologies from scratch are: Cyc method, Uschold and King's method, KACTUS approach, methodology, Grüninger and Fox's methodology, SENSUS method and On-To-Knowledge methodology. All these methods are used to develop monolingual ontologies which are quite difficult to access by the different natural language users.

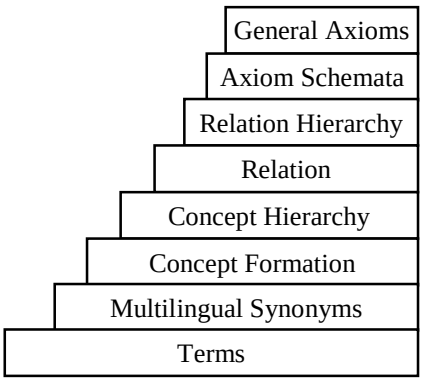

Fig.1. Ontology Learning Layer Cake

Information from the ontology can be extracted in natural languages using entities, descriptions, relations and events. The ontologies that are developed to access different natural languages are called Multilingual Ontologies (MO). It can be constructed by

adopting or extending the ontology construction methods and information extraction techniques. MO are used in: 1) answering information need in any language with respect to semantically structured data available on the semantic web and linked data cloud, 2) verbalizing and access semantically structured data, ontologies or other conceptualizations in different languages, 3) harmonization, integration, aggregation, comparison and repurposing of semantically structured data across languages and 4) aligning and reconciling ontologies or other conceptualizations across languages [3]. The Fig.1 depicts the layer cake for constructing MO, by collecting the terms of sources in different natural languages.

Collecting terms is an initial stage of developing multilingual ontology. Terms that are needed for the domain are collected in different natural languages and are evaluated. Synonyms are to be determined for the collected terms. The concepts are extracted and created using terms and synonyms. With the results of concepts creation, the hierarchy is structured and the relationships among the concepts are established. Rules, axioms and reasoning mechanisms are established to retrieve the information in different natural languages. The principles [4] that to be pursued in developing MO are: 1) user interface should be designed in multilingual and separated from the source code, 2) multilingual applications should support additional languages without reengineering, 3) all elements of the user interface should be identified and the storage and management of these elements should be planned, 4) presentation, processing and saving of data should follow the customs of the locale, 5) the standards such as program libraries for localization can be reused, 6) documentation can be as a part of localization efforts, 7) unicode should be used and 8) the fonts to use and the place of using this font should be considered. The guidelines [5] to be adopted in developing MO are: 1) transparent URI local names should be used, 2) natural language with diacritics should be used when using meaningful URI local names, 3) underscore (e.g. Student_Name) or camel case (e.g. StudentName) for word delimiter should be used when using meaningful URI local names, 4) the singular form for nouns should be used, 5) upper or lower case can be used according to natural language and 6) As many labels as needed for classes and properties should be added. MO applications can be developed for various domain using these principles and guidelines.

3. DEVELOPMENT OF MULTILINGUAL ONTOLOGIES

Multilingual Ontology applications can be developed by adopting different methodologies and methods as follows: 1) ontology mediation, 2) ontology localization and 3) approaches and tools.

3.1 ONTOLOGY MEDIATION

Ontology mediation enables reuse of semantic web application data. It plays an important role in data sharing among heterogeneous knowledge bases [4]. Multilingual ontologies can be developed by adopting the ontology mediation techniques like matching, mapping and merging. Ontology matching consists of generating a set of correspondences between the entities of two ontologies [6]. This process enables the interoperability between heterogeneous ontologies. With the help of ontology matching, the terms from the ontologies that are developed in different natural languages are matched to form multilingual ontologies. In multilingual ontology matching, "the ontologies involved can either share no common natural language or they can share common natural language (s) but at least one ontology contain two or more natural languages within itself" [1]. It can be represented as,

$$\left(L \cap L' = \varnothing\right) or \left(L \cap L' \neq \varnothing\right) \vee (\left|L\right| > 1 \wedge \left|L'\right| > 1) \quad (1)$$

where, L – Shared common natural language, L' – Non shared common natural language

By using the ontology matching techniques the correspondence between different natural languages is matched to form multilingual ontologies. Existing ontology matching techniques are grouped into eight categories [3]. Evaluation of multilingual matching is not mainly designed for evaluation but they are evaluating the ontology based on subsets. Multilingual ontology matching can be implemented in two approaches; they are: direct translation-based and indirect [6]. Direct translation considers direct matching between two ontologies with the help of external resources without intermediary ontologies i.e., translations [6]. Indirect translation is based on indirect alignment of ontologies using the composition of alignments. The translation of ontologies into multilingual ontologies can make use of natural language descriptions, so that the existing matching techniques can be applied [7]. Translation of ontologies can be achieved by using some resources like Machine translators, MultiFarm dataset [8] and OntoFarm project [9]. MultiFarm dataset is "a collection of set of ontologies translated in different languages and the corresponding alignments between them" [8].

Ontology mapping is the task of relating the vocabulary of two ontologies that share the same domain of discourse. Two ontologies in different natural languages are mapped to develop multilingual ontologies [10]. The correspondences between two ontologies are expressed using axioms which are derived from mapping language. Ontology Mapping can be unidirectional or bidirectional. There are three levels of Ontology mapping: 1) conceptual level, 2) instance level and 3) linguistic level. In conceptual level, the concepts from different ontologies described in different languages, can be semantically related by using ontology constructs, either to represent taxonomical relations such as *owl:equivalentClass, owl:sameA, and rdfs:subClassOf*. In instance level, links are established between individuals instead of their associated concepts. In linguistic level of mapping, the links are established between two ontologies associated with their linguistic information. As shown in Fig.2, the correspondences between the terms in different natural languages are expressed. O_T, O_H and O_E are the ontologies developed in Tamil, Hindi and English respectively. All three ontologies are imported into a single multilingual ontology and the similarities between those ontologies are identified and the terms are mapped by specifying the mapping rules.

Multilingual ontology mapping can make use of SOMMO framework to improve results of existing monolingual ontology matching techniques [11]. It is achieved by evaluating the semantics embedded in both the source and target ontologies [11]. Machine translation tools such as the Google Translate, Application Programming Interface (API), SDL Free Translation, Online Translator and Label Translator [12] are used to translate

ontologies. Collection of translation candidates can be generated using existing machine translators. The translated terms should convey the intended meaning. The core ontology is created in the English language, and then each term is translated into the desired natural languages. Ontology Region Mapping (ORM) approach is used to classify the terms that are translated [13]. Ontologies are constructed as semantic networks and the relevant concepts are related. The terms are mapped to the concepts and are related and explored for further related concepts using Graph traversal algorithm [13].

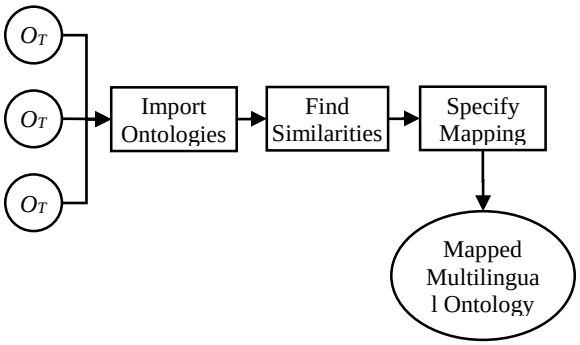

Fig.2. Multilingual Ontology Mapping

Similarities between the terms in different natural language are matched to implement multilingual ontologies. Ontology alignment is automated or semi-automated discovery of correspondences between two ontologies [14]. It creates links between ontologies. Two or more ontologies are developed in different natural languages and links are established between those ontologies to structure multilingual ontologies. Semantic similarity of visual representations of ontology's concepts is adopted to discover the alignments which are used to implement multilingual ontologies [15].

The need to find the equivalent terms of a single concept in several natural languages is an important challenge in the development of MO. Design of multilingual user interface is yet another challenge to be addressed. The process of making terms to express a concept in multi-language has to be identified. While matching, making clear or removing obscurity from the meaning of terms or symbols or expressions are to be considered. Conceptualization mismatches should be considered while performing multilingual ontology matching. Development of methods and tools to evaluate multilingual ontologies is a significant need in MO development phase. There should be clarity while translating natural languages and the ambiguity issues should be removed. While matching multilingual ontologies, the developer should have good knowledge in deemed languages [11].

3.2 ONTOLOGY LOCALIZATION

The adaptation of ontology to the particular language and culture is understood as ontology localization [16]. In this approach the terms that are expressed in one language are translated into another language. Ontology localization is "the process of adapting a given ontology to the needs of a certain community, which can be characterized by a common language, a common culture or a certain geopolitical environment". Guidelines for an automatic localization of ontologies are proposed in generic ontology localization activity. Methodological and technological aspect to evaluate ontology localization activities is implemented to develop MO [17]. The ontology localization activity involves three phases: translation, representation and management.

If the ontology is developed using translation, the developer should identify the exact equivalent terms in different natural languages. Lexical gap or a lack of direct equivalence in the target language should be addressed carefully. Multilingual representation in localization activity can be implemented using three models: i) including the multilingual information in ontology using annotation properties like label, comment and so on., ii) establishing the mapping between the concepts that are developed in different natural languages and iii) relating the external multilingual information to ontology. The type of domain of knowledge represented and the amount of required linguistic information are taken into account before determining these models. Management in ontology localization activity maintains the details of the translated terms. The rules and axioms for changing or translating the terms should be recorded carefully. The process of ontology localization activity is depicted in Fig.3.

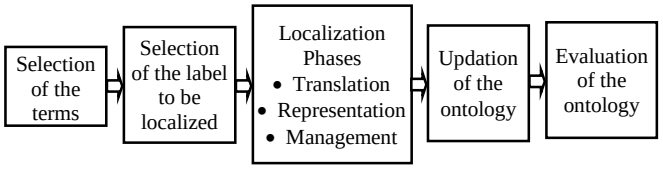

Fig.3. Process of Ontology Localization Activity

The ontology localization takes ontology as input and produces the new ontology as output by extending the label in additional languages. Ontology localization involves two layers: lexical and conceptual layer [18]. The labels, definitions in different natural languages are defined in lexical layer which can be human readable. Conceptual layer converts the labels and definitions in machine readable. The multilingual ontologies developed using localization methods are evaluated in three perspectives [18]: methodological evaluation, performance evaluation and usability evaluation. These evaluation methods produce an eminence of MO. While developing MO using ontology localization, two aspects have to be considered: 1) text classification (TC) and 2) ontology specification. TC is the process of classifying the text documents under same topic or criteria. It can be monolingual and multilingual text classification. Monolingual text classification categorizes the documents which are in one language. In Multilingual Text Classification (MLTC) the documents in different natural languages are arranged by the topics or related criteria [19]. This classification helps the developer to localize the terms into the desired languages. The ontology specification is the method of localization to develop the cross-lingual semantic interoperability of a specific subject field.

MO applications can be implemented by adopting ontology mediation and ontology localization techniques. Before developing multilingual application, modelling ontology is an important aspect. Modelling is the process representing something as a class or an attribute, a lexical or non-lexical object type. It is used to represent the prototype of ontologies. A knowledge management model (KMM) is used to support multilingual ontology applications [20]. The model is based on a global ontology that is manually designed for a specific domain and local contexts, associated with ontology concepts [20]. The

single ontology systems with associated concepts in multiple languages are proposed using a KMM framework. This model captures the cultural and lingual differences using contexts. It also includes the user interface design and labelling the terms in multilingual.

3.3 APPROACHES AND TOOLS

This section discusses the other approaches, tools and platforms that are proposed by various authors to develop multilingual ontologies. These approaches and tools are presented in Table.1.

MO can be designed using Multilingual Ontologies Knowledge Model (MOKM) to capture the lexical semantic of entities [21]. The ontology knowledge model is developed with the set of two terminologies for lexical specification. This model is designed as a language dependant and separate class is defined for each language. Lexical attributes of terms such as description, synonym, etymology and reference are gathered using dictionaries in different languages and created as slots to collect semantic information of each term. The terms can be retrieved using a Software Document Knowledge Management (SDKM) with multilingual ontology model [22]. The terms can also be retrieved using context based multilingual translations in three phases namely: 1) document processing, 2) translating process and 3) retrieval of required document. Linguistic Information Repository (LIR) model associated to the ontology meta-model allows terminological layer localization [23]. LIR provides linguistic elements for each language while localizing ontology elements. Homogeneous access to linguistic information can be distributed in various resources with different granularity using LIR. LIR establishes relations between linguistic elements. LIR overcomes the problem in conceptualization mismatches with various cultures [23].

MO applications can be developed automatically or semi-automatically by using tools such as DODDLE-OWL, DAG, RMRS and Natural OWL. DODDLE-OWL (Domain Ontology rapiD DeveLopment Environment - OWL) [24] is a domain ontology development tool used to construct ontology in different natural languages by applying translation techniques. It is used to develop interactive domain ontology for various natural languages. DAG (Direct Acyclic Graph) is used to build multilingual domain ontologies [25]. It contains the concepts and its relations of a domain. The concepts of the domain are arranged in a hierarchical structure and the related concepts are grouped into categories. The terms are grouped till it reaches the single root. Robust Minimal Recursion Semantics (RMRS) [26] system is used to construct ontologies automatically by extracting knowledge from dictionary definition sentences. RMRS brings out the relation between ontologies in greater depth and precision. Using this tool ontologies are constructed from two different lexicons and these ontologies are linked with the existing crafted ontologies by aligning them. The quality of relations is evaluated by representative evaluation. This system combines various levels of analysis of parsers and semantic structures are generated using this system.

Natural OWL [27] system is developed to generate multilingual natural language. This system adopts Natural Language Generation (NLG) technique over the Semantic Web. It also provides the ability for annotating OWL ontologies with linguistic resources and user modeling information [27]. An important level of automation for building MO can be reached using interlingual meaning representation approach between natural and formal languages. Universal Networking Language (UNL) is used to serve as an intermediate meaning representation, for building MO. Orgonic. Lingua provides cross lingual facility services. It enables the user to find multilingual resources and manage meta data of information resources [28]. Core Ontology Construction Algorithm (COCA) [29] is also used to generate MO automatically using English-Chinese bilingual term bank. The mapping strength of a selected term is computed in this algorithm using WordNet synset associated with Suggested Upper Merged Ontology (SUMO). A graph model is used to measures the strength which is integrated with mapping features. It is possible to find hypernym, part-of-speech of the terms using this algorithm. Developed MO applications can be evaluated using Multi-Role Scenario-Based methodology (MRSB) [28]. Performance of MO is evaluated using Layer wise evidence detection and Comparison with Gold standard [29].

Suite of tools and libraries are used in the development of MO from language resources access and management [31]. One of the MO development language resources is Ontological Linguistic Watermark (OLW) library provides a facility to evaluate the quality of algorithms which are used to develop the ontologies in different natural languages. New knowledge structure [31] is used to implement the bilingual ontologies by planning, conceptualizing, constructing and evaluating phases. Integration of data or ontology can be done by identifying the correspondences between attributes in multilingual schema [32]. A computer-aided system is developed to deal with the data integration. A prototype has been implemented to illustrate the effectiveness of the integration. MO can be developed using Universal Knowledge Core (UKC) which is a collaborative platform to develop multilingual ontologies [33]. This platform developed for UKC is an effective collaborative ontology development enrichment that allows for knowledge engineering in multiple languages. Controlled Natural Language (CNL) [34] is developed by adopting two-level translation approach to build MO. It is enabled by the Grammatical Framework and used to develop multilingual and multi-dialect interface to OWL. Interlingua can be changed relatively by mixing together or using different CNLs in parallel. Natural languages can be generated using three layer OWL based ontology framework [35]. It supports dynamic population of application-neutral ontologies and allows the codification of text planning relevant aspects as part of the communication ontology [36].

K-partite graph learning algorithmic program is used to extract ontology from unstructured text [37]. The terms to develop ontology are collected and separated into different partitions. The ontology is constructed by establishing the relationships between the terms. Terms and information can be retrieved using ontology based approach [38]. The approach has been established using semantic graphs which is a network of heterogeneous nodes and links to support domain ontologies.

Linguistic information extraction algorithms based on repeated segments calculation has been used to index new documents and update the ontology**.** A framework is developed to deliver multilingual learning content and it retrieves learning resources in a language independent manner [39]. This framework

is composed of separate independent component systems. Multilingual information is delivered by implementing the annotation of learning objects in different natural languages. Learning Object Storage (LOS) is developed to store multilingual information. Multilingual learning content is formulated with the help of LOS.

4. MULTILINGUAL ONTOLOGY APPLICATIONS

Ontology applications are engineered for a specific use or application focus [39]. All the ontology applications are classified into the following types: i) Application Ontology, ii) Domain-Task Ontology, iii) Task Ontology and iv) Domain Ontology [40]. Application ontologies are application independent that describes vocabulary related to a specific application. Domain-Task ontologies represent the vocabulary for a task within a specific domain. Task ontologies describe vocabulary relevant to a generic task or activity. The task can be applied in any domain. Domain ontologies are independent of any specific task or application which can be reusable within a specific domain; they represent knowledge and provide vocabulary for concepts and activities within a specific domain [41].

Multilingual dictionary [42], Syllabus tool [43] and MOLTO system [44] are multilingual ontology applications developed for European languages. Multilingual dictionary is developed for electronic commerce with ontological support [42]. The dictionary consists of data in English, Spanish, Finnish, Portuguese, German, and Polish. The ecommerce ontology is developed in multiple languages by creating a plain text file that contains all terms from the central ontology. Each term is translated into other language and separate ontology is created in respective languages. Central version of ontology is developed in English language using ontology editors. The ontologies that are developed in different natural languages are related with the central ontology. Multi-lingua ontology based tool named Syllabus [43] is designed to improve the applications of the European Union Directives (EUD). The tool is developed to support the judiciary terms of Italian, German and European countries and are linked to form a multilingual ontology based conceptual dictionary. Multilingual Online Translation (MOLTO) system [44] is developed to support the EUSFP (European Union Seventh Framework Program). Multilingual ontology application is developed for museum information on the Web. The Museum Reasonable view is an assembly of independent datasets which reduces the cost and risk of using linked data set. They are used as a single body of knowledge with respect to reasoning and query evaluation [44].

Frameworks are proposed for developing MO by integrating the knowledge and thesaurus. The framework that is developed by integrating knowledge from Social Web to support Customer Knowledge Management (CKM) is used to develop multilingual ontology application for the hotel domain [45]. It supports English, French and Portuguese. In this framework, the ontology developed in English acts as the central element. Then the terms are translated into other languages. The framework that is developed by integrating thesaurus is used in Islamic Portal Multilingual Ontology [46]. Thesaurus acts as a tool to develop multilingual ontologies. This framework provides a mechanism to retrieve the Islamic web in four languages, English, Arabic, Urdu/Hindi and Malay. Multilingual ontology is developed by transforming the ontology into relational database for this web. Another framework is implemented to manage the development of ontologies for food, health and nutrition [47]. This framework is used to support different languages which could be mapped during the extraction or query processing time to an actual instance of ontology and allows the same facts to be answered in different languages. The framework provides various systems' components with the extracted information which the search effective.

Many MO applications are developed in education domain to learn and to retrieve educational resources in different natural languages. Semantic ontology-based model is used to index and retrieve the educational resources [48]. Web learning portal prototype is used in this model to perform trilingual searching [Arabic, English and French] for online learning resources. The model is structured using Learning Domain Ontology (LDO) application profile. Resources related to OOPs in Java are stored in database and they are indexed using semantic ontology-based model. Multi- Language Information Retrieval (MLIR) approach for E-learning [49] is developed with the combination of Thesaurus-based Approach and Corpus-based Approach to retrieve the educational resources. Multilingual Course/lecture retrieval system [49] has been developed by using this approach. Multilingual Search engine is also developed based on Apache Lucene by using scoring algorithm for Information Retrieval (IR) library. Multilingual Question Answering over Linked Data (QALD) system [50] and ontology-based dialog system with multiple services are used to improve the efficient learning. QALD is developed with multilingual question answering and ontology lexicalization. Multilingual question answering introduces the multilingualism with the basic structure of the previous challenges. Ontology lexicalization includes the methods that create lexicalizations of ontology concepts semi-automatically. Ontology-based bilingual dialog system is implemented using developed with ontology-alignment algorithm [51]. Island-driven algorithm is used to extract ontology from the specific bilingual ontology.

Different MO systems are developed in medical domain, festival domain, hydrographical domain and earth observation system. Infectious disease surveillance system [52] is a multilingual ontology application which is developed to support infectious disease surveillance. Conceptual structures are formed using multilingual terminological resource which focused on priority pathogens [52]. Bilingual information retrieval system developed for the festival domain [53] is used to retrieve the Indian festival cultures. Domain ontological tree is constructed in Tamil and English languages. The information is retrieved in the users' desired languages based on their requests. Naïve algorithm is used for information retrieval. Keywords are determined using Part-of-speech [POS] tagger. This system contains 200 documents on different festivals for both English and Tamil language. Multilingual ontology for hydrographical domain [54] is an extension of Linguistic Information Repository model (LIR). Two versions of multilingual ontology have been developed for hydro-ontology. The first version of labeling system supports RDF and OWL programming languages. Multilingual information are included the ontology. LIR is used in the second version to relate multilingual information to ontology concepts [54]. Earth observation system is developed by adopting a spatial

data infrastructure based on open geospatial standards, employs an algorithm for semantic search across the multiple multilingual ontologies aligned using the W3C Simple Knowledge Organization System (SKOS) [55].

Approaches and dictionaries have proposed for multilingual text classifications are used in the development MO. Ontology based Hybrid Approach, centroid based classification and Bayesian classification are proposed to classify Punjabi documents [56]. Text classification is carried by removing stop words, punctuations marks, and special symbol. Relevant documents for the text classification are extracted using statistical approach and by using linguistics approach, dictionary is formed. Another model named Classic Vector- space model is used for classification of Telugu documents [57]. Concept based model is developed to analyze the terms on the sentences and documents. This classification is used to overcome the constraints while performing keyword based searching methods. Multilingual dictionaries can be developed by utilizing POS and word structured. The hierarchy of ontology developed with Part-Of-Speech (POS) and Word structure is used to develop ontology based Manipuri-English dictionary [58]. Noun and verb are classified under POS and word structure contains simple word and compound word. The dictionary is implemented with look up and data entry modules. Look up provides searching methods of vocabularies in Manipuri language with the user interface. Data entry can be accessed by Administrator to add Vocabularies.

Wiki-based collaborative tool 'MoKi' is developed for modeling multilingual ontologies [59]. The tool addressed the three important aspects of Linguistic Resources (LRs) modeling [58]. Both the manual and automatic translation is enabled in this tool to translate the labels and descriptions of the entities into ontologies and vocabularies that they are modeling. SmartSlog, "a software development tool is used for programming Smart-M3 agents in Knowledge Processors (KPs)" [60]. The code generation approach has been applied in SmartSlog and it produced the ontology library.

Multilingual ontology applications are compared using several dimensions by identifying the common features adopted in those applications. They are: i) methods that are followed to develop multilingual ontology applications which are discussed earlier, ii) User/Domain of multilingual ontology application, iii) The natural languages that are used in the multilingual ontology applications, iv) Tools that are used to develop the ontology application, v) Schema/Programming Language which are used in developing multilingual ontologies,, and vi) type of multilingual ontology application. By using these dimensions the applications are compared as depicted in Table.1.

Table.1. Approaches and Methods to Develop Multilingual Ontologies

Authors	Year	Method/Approach	Purpose
M. Espinoza and colleagues	2012	Localizing ontology method	To create multilingual ontologies
Montiel-Ponsoda and colleagues	2008	Linguistic Information Repository (LIR) model	Terminological layer localization
E. Andaroodi and colleagues	2014	Ontology design	To capture the lexical semantic of entities
D. Galanis and colleagues	2007	NaturalOWL - Multilingual natural language generation system	To adopt NLG techniques on the Semantic Web, and to annotate owl ontologies with linguistic resources and user modeling information
J. Cardeñosa and colleagues	2008	Interlingual based approach	Automation for ontology building
C. Yirong and colleagues	2008	Graph based Core Ontology Construction Algorithm (COCA)	English-Chinese bilingual term bank
M. T. Pazienza and colleagues	2010	Ontological Linguistic Watermark (OLW) library	Evaluating the quality of algorithms for linguistic enrichment of ontologies
M. Dragoni and colleagues	2013	Multi-Role Scenario-Based methodology (MRSB)	To evaluate multilingual ontology
E. S. Alatrish and colleagues	2014	Semi-automatic procedure	To create ontologies for different natural languages
N. Chrizman and colleagues	2010	DAG (Directed Acyclic Graph)	Automatic construction of multi-lingual domain-ontologies
A. Tawfik and colleagues	2014	Universal Knowledge Core (UKC)	A collaborative platform to develop multilingual ontologies
N. Gruzitis and colleagues	2011	Grammatical Framework	To develop a Controlled Natural Language (CNL)

P. Yu anleagues	2011	Multilingual ontology based method	To integrate Enterprise data
H. Achour and colleagues	2014	Semantic ontology-based model	Multilingual indexing and retrieving the educational resources of a web learning environment
D. G. Korzun and colleagues	2011	SmartSlog, a software development tool for programming Smart-M3 agents	To support generation of multilingual ontology libraries
B. Bhatt and colleagues	2012	K-partite graph learning algorithmic program	To extract ontology from unstructured text
Priyadharshini	2014	Software Document Knowledge Management (SDKM)	To retrieve multilingual documents
Nichols and colleagues	2006	Ontology extraction system	To constructs ontologies by extracting knowledge from dictionary definition sentences using Robust Minimal Recursion Semantics (RMRS)
Aliane and colleagues	2006	Ontology based approach	To retrieve multilingual information
Paulins	2014	High-level framework for multilingual learning content delivery	To annotate and retrieve learning resources in a language independent way

5. CONCLUSION

A survey is presented on the Methods, Approaches, Principles, Guidelines and Applications for developing Multilingual Ontologies. The need for developing multilingual ontologies is discussed. The survey on principle and guidelines specifies some of the rules to be kept while developing multilingual ontology. A developer can follow these principles and guidelines in order to develop multilingual ontology. Further classification of various methods used for developing multilingual ontology is made. The approaches and applications that are proposed by different authors on designing multilingual ontologies are also tabulated.

It was found from the survey that different methods are adopted to develop multilingual ontologies. These methods use translators, ontology matcher to make the existing ontology as multilingual ontologies. The methods focus on matching and mapping ontologies existing in two different languages rather than proposing a method for building a multilingual ontology. Various approaches for developing multilingual ontologies are categorized and the purpose of the approach is highlighted. Each of this approach adopts a different procedure to develop multilingual ontologies. These approaches can be adapted or reused to develop multilingual ontologies. In the survey on applications it was found that very few attempts have been made to develop ontologies in various languages. Multilingual applications that are developed for various domains are compared using several dimensions.

Despite the undeniable advances in creating multilingual ontologies there are a lot of opportunities in this area. The future scope in the field of multilingual ontologies is to find a standardized method for developing multilingual ontologies. Moreover, a multilingual ontology editor can be developed in order to build multilingual ontologies as present editors support only English language. It will avoid the configuration of a system to suite a particular language. Another opening in this field is to develop evaluation techniques and tools to evaluate the multilingual ontologies developed.

ACKNOWLEDGEMENT

This research is financially supported by the University Grants Commission (UGC) of Government of India and Grant MRP-5765/15 (SERO/UGC).

REFERENCES

[1] Thomas R. Gruber, "A Translation Approach to Portable Ontology Specifications", *Knowledge Acquisition*, Vol. 5, No. 2, pp. 199-220, 1993.

[2] John Davis, Rudi Studer and Paul Warsen, "*Semantic Web Technologies. Trends and Research in Ontology Based Systems*", Wiley, 2006.

[3] Cassia Trojahn , Bo Fu, Ondrej Zamazal, Dominique Ritze, "*State-of-the-Art in Multilingual and Cross-Lingual Ontology Matching*", Towards the Multilingual Semantic Web, Springer, pp. 119-135, 2014.

[4] United Nations, "The Principles and Guidelines on Building Multilingual Application for Official Statistics", *United Nations Economic Commission for Europe*, pp. 1-13, 2012.

[5] Elena Montiel-Ponsoda, Daniel Vila-Suero, Boris Villazon-Terrazas, Gordon Dunsire, Elena Escolano Rodriguez and Asuncion Gomez-Perez, "Style Guidelines for Naming and Labeling Ontologies in the Multilingual Web", *Proceedings of the International Conference Dublin Core and Metadata Applications*, pp. 105-115, 2011.

[6] C.T. Dos Santos, P. Quaresma and R. Vieira, "An API for Multilingual Ontology Matching", *Proceedings of 7th Conference on Language Resources and Evaluation*, pp. 3830-3835, 2010.

[7] Haytham Al-Feel, Ralph Schafermeier and Adrian Paschke, "An Inter-lingual Reference Approach for Multi-Lingual Ontology", *International Journal of Computer Science*, Vol. 10, No. 2, pp. 497-503, 2013.

[8] Christian Meilicke *et al.*, "MultiFarm: A Benchmark for Multilingual Ontology Matching", *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 15, pp. 62-68, 2012.

[9] Onto Farm Project, Available at: http://nb.vse.cz/~svatek/ontofarm.html

[10] Cassia Trojahn, Paulo Quaresma and Renata Vieira, "A Framework for Multilingual Ontology Mapping", *Proceedings of the International Conference on Language Resources and Evaluation*, pp. 1034-1037, 2008.

[11] Bo Fu, Rob Brennan and Declan O'Sullivan, "Multilingual Ontology Mapping: Challenges and a Proposed Framework", *Proceedings of Workshop on Matching and Meaning*, pp. 33-35, 2009.

[12] G. Falquet, C. Metral, J. Teller and C. Tweed, "*Ontologies in Urban Development Projects*", 1st Edition, Springer, 2011.

[13] Gerard de Melo and Stefan Siersdorfer, "Multilingual Text Classification using Ontologies", *Proceedings of 29th European Conference on IR Research*, pp. 541-548, 2007.

[14] G. Ganapathy, and R. Lourdusamy, "Ontology Merging and Matching Using Ontology Abstract Machine", *Proceedings of 5th International Conference Knowledge Management*, pp. 654-660, 2010.

[15] Srdan Mihic and Dragan Ivetic, "Multilingual Ontology Alignment based on Visual Representations of Ontology Concepts", *Proceedings of 5th International Conference on Advances in Computer-Human Interaction*, pp. 101-105, 2012.

[16] Mari Carmen Suarez-Figueroa and Asuncion Gomez-Perez, "First Attempt towards a Standard Glossary of Ontology Engineering Terminology, *Proceedings of 8th International Conference on Terminology and Knowledge Engineering*, pp. 1-15, 2008.

[17] Mauricio Espinoza Mejia, Elena Montiel-Ponsoda, Guadalupe Aguado de Cea and Asuncion Gomez-Perez, "Ontology localization", *Ontology Engineering in a Networked World*, pp 171-191, 2011.

[18] P. Cimiano, E. Montiel-Ponsoda, P. Buitelaar, M. Espinoza and A. Gomez-Perez, "A Note on Ontology Localization", *Applied Ontology*, Vol. 5, No. 2, pp. 127-137, 2010.

[19] Gerard De Melo and Stefan Siersdorfer, "Multilingual Text Classification using Ontologies", *Advances in Information Retrieval*, pp. 541-548, 2007.

[20] Aviv Segev and Avigdor Gal, "Enhancing Portability with Multilingual Ontology-Based Knowledge Management", *Decision Support Systems*, Vol. 45, No. 3, pp. 1-38, 2008.

[21] Elham Andaroodi, Frederic Andres, Kinji Ono and Pierre Lebigre, "Protege Tool and Development of Multilingual Ontology for Architectural Corpus, Design Process, Shortcomings", *Proceedings of 8th International Conference on Protege*, pp. 1-3, 2005.

[22] M.R. Priyadarshini, "An Ontology Framework for Context-Based Multilingual Document Retrieval", *International Journal of Computer Science and Engineering Technology*, Vol. 5, No. 3, pp. 178-181, 2014.

[23] Elena Montiel-Ponsoda, Guadalupe Aguado De Cea, Asunction Gomez-Perez and Wim Peters, "Modelling Multilinguality in Ontologies", *Proceedings of 22nd International Conference on Computational Linguistics*, pp. 67-70, 2008.

[24] E.S. Alatrish, D. Tosic and N. Milenkovic, "Building Ontologies for Different Natural Languages", *Computer Science and Information Systems*, Vol. 11, No. 2, pp. 623-644, 2014.

[25] Nitsan Chrizman and Alon Itai, "How to Construct Multilingual Domain Ontologies", Technical Report, Computer Science Department, Israel Institute of Technology, pp. 4345-4350, 2014.

[26] Eric Nichols, Francis Bond, Takaaki Tanaka, Sanae Fujita and Dan Flickinger, "Multilingual Ontology Acquisition from Multiple MRDs", *Proceedings of 2nd Workshop on Ontology Learning and Population*, pp. 10-17, 2006.

[27] Dimitrios Galanis and Ion Androutsopoulos, "Generating Multilingual Personalized Descriptions from OWL Ontologies on the Semantic Web: The NaturalOWL System", *Proceedings of the 11th European Workshop on Natural Language Generation*, pp. 1-25, 2007.

[28] Mauro Dragoni, Chiara Di Francescomarino, Chiara Ghidini, Julia Clemente and Salvador Sanchez Alonso, "Guiding the Evolution of a Multilingual Ontology in a Concrete Setting", *Proceedings of 10th International European Semantic Web Conference*, pp. 608-622, 2013.

[29] Chen Yirong, Lu Qin, Li Wenjie and Cui Gaoying, "Chinese Core Ontology Construction from a Bilingual Term Bank", *Proceedings of International Conference on Language Resources and Evaluation*, pp. 2344-2351, 2008.

[30] Brijesh Bhatt and Pushpak Bhattacharyya, "Domain Specific Ontology Extractor for Indian Languages", *Proceedings of the 10th Workshop on Asian Language Resources*, pp. 75-84, 2012.

[31] Maria Teresa Pazienza, Armando Stellato and Andrea Turbati, "A Suite of Semantic Web Tools Supporting Development of Multilingual Ontologies", *Intelligent Information Access*, Vol. 301, pp. 109-135, 2010.

[32] Awatef Al Azemi, Samia Nefti, Umar Manzoor and Yacine Rezgui, "Building a Bilingual Bio-Ontology Platform for Knowledge Discovery", *International Journal of Innovative Computing Information and Control*, Vol. 7, No. 12, pp. 7067-7075, 2011.

[33] Pen Yu and Hongding Wang, "A Multilingual Ontology-based Approach to Attribute Correspondence Identification", *Proceedings of International Conference on Electrical and Control Engineering*, pp. 1896-1900, 2011.

[34] Ahmed Tawfik, Fausto Giunchiglia and Vincenzo Maltese, "A Collaborative Platform for Multilingual Ontology Development", *International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering*, Vol. 8, No. 12, pp. 3930-3939, 2014.

[35] Normunds Gruzitis and Guntis Barzdins, "Towards a More Natural Multilingual Controlled Language Interface to OWL", *Proceedings of 9th ACM International Conference on Computational Semantics*, pp. 335-339, 2011.

[36] Nadjet Bouayad-Agha, Gerard Casamayor, Simon Mille, Marco Rospocher, Horacio Saggion, Luciano Serafini and Leo Wanner, "From Ontology to NL: Generation of Multilingual User-Oriented Environmental Reports", *Natural Language Processing and Information Systems*, pp. 216-221, 2012.

[37] H. Aliane, "An Ontology Based Approach to Multilingual Information Retrieval", *Proceedings of 2nd International Conference on Information and Communication Technologies*, pp. 1732-1737, 2006.

[38] N. Paulins, I. Arhipova, and S. Balina, "Multilingual Information Delivery Based on a Domain Ontology", *Proceedings of the 15th International Conference on Computer Systems and Technologies*, pp. 430-436, 2014.

[39] James Malone and Helen Parkinson, "*Review of Reference and Application Ontologies*", Available at: http://ontogenesis.knowledgeblog.org/295.

[40] Maha Al-Yahya, Remya George and Auhood Alfaries, "Ontologies in E-Learning: Review of the Literature", *International Journal of Software Engineering and Its Applications*, Vol. 9, No. 5, pp. 67-84, 2015.

[41] Gomez-Perez, Asuncion, Fernandez-Lopez, Mariano and Corcho Oschar, "*Ontological Engineering with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web*", 1st Edition, Springer, 2004.

[42] Johannes Schwall, "Creating an Ontology for a Multilingual E-Commerce Dictionary", Available at: http://www.schwall.de/dl/20071221_comed.pdf.

[43] G. Ajani, G. Boella, L. Lesmo, M. Martin, A. Mazzei and P. Rossi, "A Development Tool for Multilingual Ontology-Based Conceptual Dictionaries", *Proceedings of 5th International Conference on Language Resources and Evaluation*, pp. 1-6, 2006.

[44] Dana Dannells, Mariana Damova, Ramona Enache and Milen Chechev, "Multilingual Online Generation from Semantic Web Ontologies", *Proceedings of the 21st International Conference Companion on World Wide Web*, pp. 239-242, 2012.

[45] Marcirio Silveira Chaves and Cassia Trojahn, "Towards a Multilingual Ontology for Ontology-Driven Content Mining in Social Websites", *Proceedings of 9th International Semantic Web Workshops*, pp. 1-10, 2010.

[46] Juhana Salim, Siti Farhana Mohamad Hashim and Shahrul Azman Mohamad Noah, "Symbiosis of Thesaurus, Domain Expert and Reference Sources in Designing a Framework for the Construction of a Multilingual Ontology for Islamic Portal", *International Journal on Electrical Engineering and Informatics*, Vol. 4, No. 1, pp. 120-133, 2011.

[47] Tarek Helmy, Ahmed Al-Nazer, Saeed Al-Bukhitan and Ali Iqbal, "Health, Food and User's Profile Ontologies for Personalized Information Retrieval", *Proceedings of 5th International Conference on Sustainable Energy Information Technology*, Vol. 52, pp. 1071-1076, 2015.

[48] Marwa Hendez and Hadhemi Achour, "Keywords Extraction for Automatic Indexing of E-Learning Resources", *Proceedings of World Symposium on Computer Applications and Research*, pp. 1-5, 2014.

[49] Leyla Zhuhadar, Olfa Nasraoui, Robert Wyatt and Elizabeth Romero, "Multi-Language Ontology-Based Search", *Proceedings of 3rd International Conference on Advances in Computer-Human Interactions*, pp. 13-18, 2010.

[50] Philipp Cimiano, Vanessa Lopez, Christina Unger, Elena Cabrio, Axel-Cyrille Ngonga Ngomo and Sebastian Walter, "Multilingual Question Answering Over Linked Data (QALD-3): Lab overview", *Proceedings of 4th International Conference of Conference and Labs of the Evaluation Forum Initiative*, pp. 321-332, 2013.

[51] Jui-Feng Yeh, Chung-Hsien Wu and Ming-Jun Chen, "Ontology-Based Speech Act Identification in a Bilingual Dialog System Using Partial Pattern Trees", *Journal of the American Society for Information Science and Technology*, Vol. 59, No. 5, pp. 684-694, 2008.

[52] Nigel Collier, Ai Kawazoe, Lihua Jin, Mika Shigematsu, Dinh Dien, Roberto A. Barrero, Koichi Takeuchi and Asanee Kawtrakul, "A Multilingual Ontology for Infectious Disease Surveillance: Rationale, Design and Challenges", *Language Resources and Evaluation*, Vol. 40, pp. 405-413, 2006.

[53] S. Saraswathi, M. Asma Siddhiqaa, K. Kalaimagal and M. Kalaiyarasi, "Bilingual Information Retrieval System for English and Tamil", *Journal of Computing*, Vol. 2, No. 4, pp. 85-89, 2010.

[54] G. Aguado de Cea, A. Gomez-Perez, E. Montiel-Ponsoda and L.M. Vilches-Blazquez, "Riviere or Fleuve? Modelling Multilinguality in the Hydrographical", *Proceedings of 1st Workshop on the Multilingual Semantic Web*, Vol. 571, pp. 21-28, 2010.

[55] Kristin Stock and Claudia Cialone, "An Approach to the Management of Multiple Aligned Multilingual Ontologies for a Geospatial Earth Observation System", *Proceedings of 4th International Conference on GeoSpatial Semantics*, Vol. 6631, pp. 52-69, 2011.

[56] Nidhi and Vishal Gupta, "Domain Based Classification of Punjabi Text Documents using Ontology and Hybrid Based Approach", *Proceedings of 3rd Workshop on South and Southeast Asian Natural Language Processing*, pp. 109-122, 2012.

[57] A. Kanaka Durga and A. Govardhan, "Ontology Based Text Categorization-Telugu Document", *International Journal of Scientific and Engineering Research*, Vol. 12, No. 9, pp. 1-4, 2011.

[58] S. Ningombam, S.P. Meitei and Bipul Syam Purkayastha, "Building Manipuri-English Machine Readable Dictionary by Implementing Ontology", *International Journal of Engineering Science and Technology*, Vol. 3, No. 10, pp. 7682-7689, 2011.

[59] Alessio Bosca, Matteo Casu, Mauro Dragoni and Andi Rexha, "Modeling, Managing, Exposing, and Linking Ontologies with a Wiki-based Tool", *Proceedings of International Conference on Language Resources and Evaluation*, pp. 1668-1675, 2014.

[60] Dmitry G. Korzun, Alexandr A. Lomov, Pavel I. Vanag, Sergey I. Balandin and Jukka Honkola, "Multilingual Ontology Library Generator for Smart-M3 Information Sharing Platform", *International Journal on Advances in Intelligent Systems*, Vol. 4, No. 3, pp. 68-81, 2011.

[61] S. Shanawaz Basha and L. Sunitha Rani, "A Novel Approach for Text Classification", *International Journal of Computer Science and Technology*, Vol. 4, No. 1, pp. 649-654, 2013.