Ontology Construction from Text: Challenges and Trends

Abeer Al-Arfaj *abeerarfaj@yahoo.com*
Department of Computer Science
King Saud University
Riyadh, Saudi Arabia

AbdulMalik Al-Salman *salman@ksu.edu.sa*
Department of Computer Science
King Saud University
Riyadh, Saudi Arabia

Abstract

Ontology is one of the most popular representation model used for knowledge representation, sharing and reusing. In light of the importance of ontology, different methodologies for building ontologies have been proposed. Ontology construction is a difficult and time-consuming process. Many efforts have been made to help ontology engineers to construct ontologies and to overcome the bottleneck of knowledge acquisition. The aim of this paper is to give a brief overview of ontology learning approaches and to review some of ontology extraction systems and tools followed by a summarizing comparison of them. Also some of the current issues and main trends of ontology construction from texts will be discussed.

Keywords: Ontology Engineering, Ontology Learning, Ontology Learning Systems, Ontology Evaluation**.**

1. INTRODUCTION

Ontology is a backbone technology for the Semantic Web. It plays an important role in supporting knowledge based applications in the Semantic Web. Since ontology is a representation model which defines domain knowledge with explicit specifications that solve interoperability between human and machine, therefore it is used for knowledge representation, sharing and reusing. Ontology has been used in wide applications like knowledge management, information retrieval, information integration, bioinformatics and e-learning [1,2].

Due to the importance of ontologies in these areas, different methodologies for building ontologies have been proposed. However, the manual building of ontologies requires much time and many resources. Ontology learning, which extracts ontological knowledge from various forms of data automatically or semi-automatically, can overcome the bottleneck of knowledge acquisition and help ontology engineers to construct ontologies.

The ontology term has been adopted from philosophy, where it is defined as the “theory of existence”. Ontology is a well-known term in the field of AI and knowledge engineering. The most popular definition of ontology in information technology and the AI community made by Gruber [3], which states that: “An ontology is a formal, explicit specification of a shared conceptualization”. According to Studer et al [4], conceptualization refers to an abstract model of phenomena in the world by having identified the relevant concepts of those phenomena. Explicit means that the type of concepts used and the constraints on their use are explicitly defined. Formal refers to the fact that the ontology should be machine-readable. Shared reflects that an ontology should capture consensual knowledge accepted by different communities.

The question now is how to build the ontology? The process of constructing ontology is an engineering activity. From the Ontology engineering point of view, there are several methodologies for constructing ontologies from scratch. To reduce the costs involved in the activity of engineering ontologies, ontology learning systems such as Text-To-Onto, Text2Onto, OntoLearn and OntoGen have been developed to extract concepts, relations between concepts, and axioms on relations from domain specific documents.

In this paper, we present a survey on ontology engineering and especially ontology learning from texts. The paper is organized as follows. After the introduction, Section 2 explains the field of ontology engineering. Ontology learning approaches are reviewed in Section 3. In Section 4, the state-of-the art systems and tools that use various approaches to support ontology learning from text is discussed and the comparison of them is summarized in Section 5. Finally, Section 6 concludes the paper.

2. ONTOLOGY ENGINEERING

Ontology engineering is a growing research area that has received much attention in many fields. Ontologies are extensively used in different domains like knowledge engineering, artificial intelligence, natural language processing, e-commerce, intelligent information integration, information retrieval, database design and integration, bio-informatics and etc. In order to support the development of ontologies several methodologies have been proposed to date, facilitating the process of ontology development or ontology engineering [5].

Ontology Engineering is formally defined as “the set of activities that concern the ontology development process, the ontology life cycle, and the methodologies, tools and languages for building ontologies" [6,7]. Most of the ontology engineering methodologies consist of at least the phases shown in Figure 1: feasibility study, requirements analysis, conceptualization and finally deployment, evaluation and maintenance of the ontology. These phases are partitioned into subphases. Conceptualization can be divided into development of the domain model, formalization of the model and its implementation in a certain ontology language. In this phase, ontology learning techniques can be applied [8].

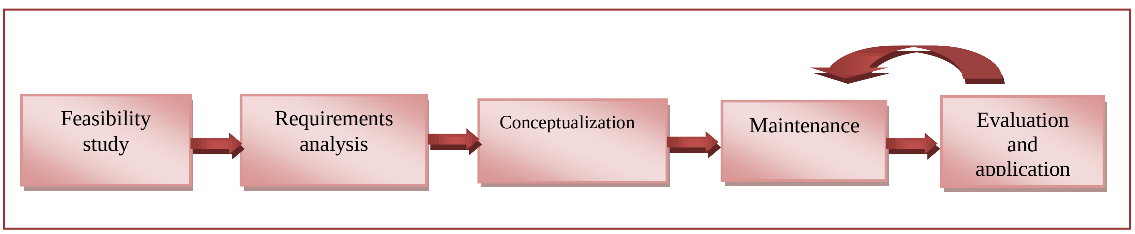

FIGURE 1: Ontology Engineering Process (Source [8]).

There are several methodologies for building ontologies from scratch and for the collaborative and cooperative construction (Seven-Step method [9], knowledge Engineering method which is described in [10], METHONOLOGY [11], TOVE [12], Software Engineering [13] and NeON [14]). Although these improve the ontology development process, building ontologies manually is subjective, very hard, non-evident, time-consuming, error-prone, and can cost very much. To overcome these difficulties, research area known as ontology learning is generated.

3. STATE OF THE ART IN ONTOLOGY LEARNING

Ontology learning is defined as the set of methods and techniques used for building an ontology from scratch, enriching, or adapting an existing ontology in a semi-automatic fashion using several sources. Ontology learning techniques rely on methods from various fields such as machine learning, knowledge acquisition, Natural Language Processing (NLP), statistics, and information retrieval. Such techniques facilitate and support the construction of ontologies by the ontology engineer. This is the reason why ontology learning frameworks have been developed in the last years and integrated with standard ontology engineering tools. Ontology learning can be applied to unstructured, semi-structured and fully structured data to support semi-automatic and cooperative ontology engineering [2,8,15,16].

3.1 Approaches for Ontology Learning

Ontology learning approaches have been classified according to several main dimensions (Figure 2):

The type of knowledge resources for which to learn ontology:

- Structured such as already defined knowledge models include existing ontologies and database schema.
- Semi-structured data designate the use of some mixed structured data with free text such as Web pages, Wikipedia, dictionaries and XML documents.
- Unstructured data is related to any textual content.

The level of automation:

- Semi Automation with user intervention.
- Full Automation with system takes care of all construction process.

The learning targets:

- In addition to concepts and relations, learning targets could be definition that describe the concepts and axioms that constraints interpretation of concepts and relations.

The purpose:

- Ontology can be created from scratch or updating an existing ontology.

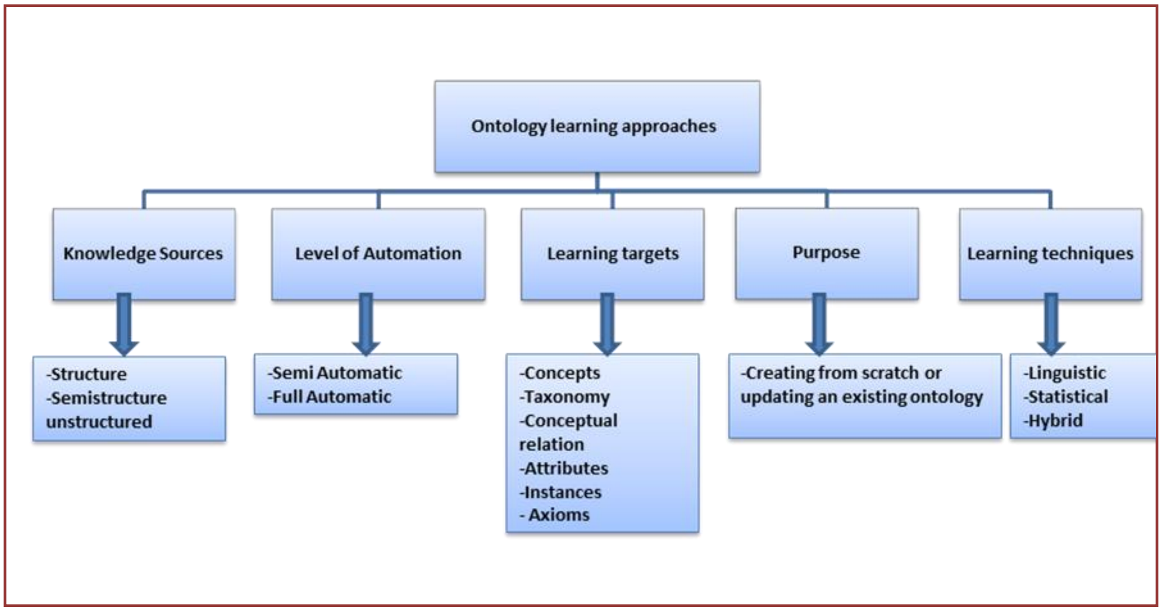

FIGURE 2: Classification of Ontology Learning Approaches.

There are three main approaches to learning ontologies from text: Linguistic, Statistical, and Hybrid approaches [1, 15, 17-20].

3.1.1 Linguistic Approach

Linguistic information is used to extract information from text. Linguistics-based techniques are applicable to almost all tasks in ontology learning and are mainly dependent on natural language processing tools. Some of the techniques include part-of-speech tagging, sentence parsing, syntactic structure analysis and dependency analysis. Other techniques rely on the use of semantic lexicon, lexico-syntactic patterns. For example, the concept linguistically represented using compound, multi-word terms [20]. Also, the conceptual relationships in ontology can be identified by matching to predefined rules to extract the relationships between terms. The most common symbolic approach is to use lexico-syntactic pattern (LSP) such as Hearst Patterns [21] which are used to extract relationships between terms in order to find hyponym and hypernym relations. For example, the pattern “X such as Y” frequently implies Y ISA X. Text2Onto [22] and [23] use this patterns to extract taxonomic relationships. Also [24-26] use syntactic patterns to extract relations between concepts. Another symbolic approach is to use the internal syntactic structure of components terms. This technique is applied in the ontology learning studies by [27] and others. The crucial problem of this method is data sparseness. So, most of the proposed systems overcome this problem by applying the combination methods of linguistic and statistical approaches.

3.1.2 Statistical Approach

The various statistics-based techniques for accomplishing the tasks in ontology learning are mostly derived from information retrieval, machine learning and data mining. Some of the common techniques include clustering, latent semantic analysis, co-occurrence analysis and association rule mining [20].

Clustering approaches based on a similarity measure, cluster similar words based on Harris' distributional hypothesis [28], which states that words that occur in the same contexts tend to have similar meanings. Some approaches assign labels to the clusters, treating the labels as concepts and the terms in the cluster as its instances, which can help in identifying concepts and their synonyms [29]. The hierarchical relationships of concepts can be also extracted by using the similarity measurements.

Co-occurrence analysis attempts to identify terms that tend to occur together to extract related terms and to discover implicit relations between concepts. More recently, Association rules are mined to discover the semantic relations between terms [25, 27, 30]. [31] and [32] further discovered associated concept pairs and verbs, and then employed the verbs to label semantic relations. [33] utilized the distributions of co-occurring concepts and verbs as significance measures for identifying verbs as semantic labels. [34] propose novel fuzzy domain ontology acquisition algorithm. The knowledge construction mechanism constructs fuzzy concept and relation based on concepts.

3.1.3 Hybrid Approach

As the name suggests, this approach borrows ideas from one or more of the previous approaches. Most systems use combination approaches to learn different components and to enrich the ontology. For example, Text2Onto [22] combines machine learning approaches with basic linguistic processing techniques to extract relations between concepts from the text. Ontolearn [35] uses inductive machine learning to associate the appropriate relations that hold among the concepts of the domain.

3.2 Ontology Evaluation

"Ontology evaluation measures the quality of a learned ontology with respect to some particular criteria, in order to determine the plausibility of the learned ontology for the purposes it was built for". Approaches for evaluating learned ontologies can be distinguished into four major categories [36]:

- Gold standard evaluation: the learned ontology is compared to a predefined gold standard ontology.
- Application-based evaluation: the learned ontology is used in an integrated system and is implicitly evaluated through the evaluation of the complete integrated system.
- Data-driven evaluation: the learned ontology is evaluated through comparison with a data source covering the same domain as the learned ontology.
- Human evaluation: the learned ontology is evaluated by domain experts based on predefined criteria, requirements, standards, etc.

Ontologies composed by multiple layers. So, they can be evaluated at different layers, such as:

- Terminological layer: The evaluation here focuses correctness of the terminology.

- Conceptual layer: The evaluation here focuses on how well the extracted terms cover the domain.
- Taxonomical layer: The evaluation of hierarchical relations between concepts.
- Non-taxonomical layer: The evaluation of this layer deals with the relations between the concepts of the ontology.

4. ONTOLOGY LEARNING SYSTEMS

A number of ontology learning systems and tools have been proposed, these tools extract ontological structures from text corpora using various methods and algorithms with the goal of reducing the time and cost for ontology development. In this section, we discuss and compare the major distinguishing factors between the existing ontology learning systems.

DODDLE II [37] is Domain Ontology rapiD DeveLopment Environment. It is a supporting tool to learn taxonomic and non-taxonomic relations. The taxonomic relations extraction based on WordNet and a domain expert. While non-taxonomic relations extraction based on domain specific texts with the analysis of lexical cooccurrence statistics. To evaluate the system, some case studies have been done in the field of law. For taxonomic relations, the precision was 30%. For non-taxonomic relations, the precision was 59%.

Text-To-Onto [38] is a framework for semi-automatic ontology learning from texts which implements a variety of algorithms for diverse ontology learning subtask. It leverages data mining and natural language processing techniques in the ontology development and maintenance task. It proceeds through ontology import, extraction, pruning, and refinement Text-To-Onto was originally integrated into the KAON ontology engineering environment. The advantage of this system is that it has diverse algorithms that support term extraction, taxonomy construction as well as learning relations between concepts. Also it has algorithms for ontology maintenance such as ontology pruning and refinement.

The successor **Text2Onto** [22] was distinguished from the earlier system in three important ways. First, It represented the learned knowledge at a meta-level in the form of instantiated model primitives within a so called Probabilistic Ontology Model (POM), which can then be translated to any expressive knowledge representation language (OWL and RDFS). Second, adding probabilities to the learned structures to facilitate the interaction with the user. Third, by incorporating methods for data-driven change discovery, that selectively updates the POM according to the corpus changes. Which allows a user to trace the evolution of the ontology with respect to the changes in the underlying corpus. Both systems extract taxonomic relationship using Hearst pattern match method and non-taxonomic relations using association rule mining method. They were evaluated in a tourism domain and they achieved 76% accuracy for taxonomic relations; and for non-taxonomic relations, they manually developed a small ontology with 284 concepts and 88 non-taxonomic relationships as the gold standard.

OntoLearn [35] uses a combination of symbolic and statistical methods. It first extracts domain terminology from domain corpora, and then complex domain terms are semantically interpreted and arranged in a hierarchical fashion. Finally, WordNet is trimmed and enriched with the detected domain concepts. It mainly focuses on the problem of word sense disambiguation. In particular, the authors present a new algorithm called structural semantic interconnections relying on the structure of the general ontology for this purpose. The system has been evaluated by two experts. The system was applied in different domains (art, tourism, economy and computer network), and they achieved recall ranging from 46% to 96% and precision ranging from 65% to 97%.

HASTI [41] is a system that learns concepts, taxonomic and non-taxonomic conceptual relations, and axioms, to build ontologies upon the existing kernel. The kernel is domain independent. Therefore, it can be used to build both general and domain ontologies from scratch. The learning approach in HASTI is a hybrid symbolic approach, a combination of linguistic, logical, template driven and semantic analysis methods. It performs clustering to organize its ontology. The authors evaluated HASTI with two cases. In the case of a corpus consisting of primary school textbooks and storybooks, the precision was 97% and the recall was 88%. And with a corpus consisting of computer technical reports, the precision was 78% and the recall was 80%.

OntoGen [29] is a tool that supports the user in building an ontology by extracting possible concepts and relations between them from domain texts using machine learning and text mining algorithms. The system uses supervised methods for concept discovery. It uses improved user interface, in which the user can select a topic through a graphical user interface, and the system automatically suggests some potential subtopics from a set of selected documents. The extracted concepts validated by the user.

TextOntoEx [24] is a tool to construct ontology from natural domain text using semantic pattern-based approach. It analyses natural domain text to extract candidate relations and then maps them into meaning representation to facilitate constructing ontology. TextOntoEx does not discover new relation but discovers instances of known relation. The authors focus in constructing ontology with non-taxonomic relations. They use OWL language to represent the ontology and they applied their system into case study of agricultural domain. The precision ratio is 100% because the irrelevant retrieved is nothing. And the recall ratio is approximately 54%.

OntoGain [30] is a system for unsupervised ontology acquisition from unstructured text which relies on multi-word term extraction. For taxonomic relations discovery, authors exploit inherent multi-word terms' lexical information in a comparative implementation of agglomerative hierarchical clustering and formal concept analysis methods. For non-taxonomic relations discovery, they comparatively investigate in OntoGain an association rules based algorithm and a probabilistic algorithm. The advantage of this system is that it produces a semantically rich ontology of multi-word domain concepts, rather than an ontology of single-word terms and, the extracted ontology transformed into standard OWL statements. OntoGain results are compared to both manual built ontologies, as well as to Text2Onto system, in two different domains: the medical and computer science domains. The evaluation indicated that agglomerative clustering and association rules outperform any other method combination reaching up to 70% precision for identification of taxonomic and non-taxonomic relations respectively in both corpora.

CRCTOL [25] Concept-Relation-Concept Tuple-based Ontology Learning is a system to mine ontologies automatically from domain specific documents. The system adapts a full text parsing technique and employs a combination of statistical and lexico-syntactic methods, including a statistical algorithm that extracts key concepts from a document collection, a word sense disambiguation algorithm that disambiguates words in the key concepts, a rule based algorithm that extracts relations between the key concepts, and also adapts a modified generalized association rule mining algorithm that prunes unimportant relations for ontology learning. The system was evaluated in two case studies: a terrorism domain and a sport event domain. At the component level, quantitative evaluation by comparing with Text-To-Onto and its successor Text2Onto has shown that CRCTOL is able to extract concepts and semantic relations with a significantly higher level of accuracy. The overall accuracy of the system in taxonomic relation extraction is 85.7%, with precision 74.0%. For non-taxonomic relation extraction, the precision is 69.4%. At the ontology level, the quality of the learned ontologies is evaluated by either employing a set of quantitative and qualitative methods including analyzing the graph structural property, comparison to WordNet, and expert rating, or directly comparing with a human-edited benchmark ontology, demonstrating the high quality of the ontologies learned.

OntoCmaps [2] is a domain independent and ontology learning tool that extracts deep semantic representations from corpora. OntoCmaps generates conceptual representations in the form of concept maps. To extract the important elements the system relies on the inner structure of graphs to identify the important concepts. They proposed filtering mechanism based on Degree (number of edges from and to a given term), Betweenness (number of shortest paths that pass through a term), PageRank (fraction of time spent visiting a term) and Hits (ranks terms according to the importance of hubs and authorities) metrics from graph theory. They compared the resulting ontology with the ontology created by Text2Onto and has achieved better results. The learned resources can be unstructured corpus text and other concept maps.

LexOnt [43] is a semi-automatic ontology construction system. It uses the Programmable Web directory of services, Wikipedia, WordNet and the existing ontology to extract relevant terms. LexOnt constructs the ontology iteratively, by interacting with the user. The user can choose, add these terms to the ontology and rank terms. It is a plugin tab for the Protégé ontology editor which interacts with the user to facilitate the ontology creation process. The system accepts unstructured text as input.

Table 1 summarizes the major distinguishing factors between the existing ontology learning systems: The learned elements (concepts, relations, axioms, instances), the main techniques applied for learning (statistical, linguistic-based, pattern matching and hybrid methods), the information sources used for learning, type and amount of user intervention and evaluation results.

System	**Elements learned**	**Main techniques used**	**Learning sources**	**User intervention**	**Evaluation Result**
DODDLE II [37]	Concepts Taxonomic Non-taxonomic relations	Statistics	Dictionaries Domain text documents WordNet	Validates, adapts, defines new domain specific patterns and relations.	The precision was 30% for taxonomic relations, 59% for non-taxonomic relations.
Text-To-Onto [38] Text2Onto [22]	Terms, Synonym Concepts Taxonomic Non-taxonomic Relations Instances	Statistical approach Pruning techniques Association rules	Free text Dictionaries Ontologies	Evaluation	76% accuracy for taxonomic relationship.
OntoLearn [35]	Terms, Synonym Concepts Taxonomic Non-taxonomic relations	Linguistic analysis Machine learning Statistics	Free text WordNet	Evaluation	Recall ranging from 46% to 96% and precision ranging from 65% to 97%.
HASTI [41]	Concepts Taxonomic Non taxonomic relations axioms	Linguistic based Template driven	Free text	Two modes: Automatic and semi-automatic Intervention not necessary	In case1: the precision was 97% and the recall was 88%. Case2: the precision was 78% and the recall was 80%.
Ontogen [29]	Terms Concept Taxonomic relations	Statistical Analysis Clustering	Free text	Evaluation	N/A
TextOntoEx[24]	Instances of known relation	linguistic analysis	Free text	Construct semantic patterns.	The precision was 100%. The recall ratio was approximately 54%.

OntoGain [30]	Concept Taxonomic Non-taxonomic Relations	Clustering, formal concept analysis and Association rules	Free text WordNet	Automatic	The precision was 70% for taxonomic and non-taxonomic relations.
CRCTOL [35]	Concepts Taxonomic Non-taxonomic Relations	Statistical lexico-syntactic Association rules.	Free text WordNet	Edit the learned ontology by adding or removing concepts and relations.	The precision was 74.0% for taxonomic relation. For non-taxonomic relation precision was 69.4%.
OntoCmaps[2]	Linguistic terminology Concepts Taxonomic Non-taxonomic Relations	Pattern-based	Semi/ Unstructured Text	Semiautomatic Expert assess and validate results	For hierarchical relationships the Precision, Recall and f-measure was 81.04 % 45.09 % 57.94 % and 59.23 % 45.00 % 51.14 % for conceptual relation.
LexOnt [43]	Terms Synonyms Taxonomic relations	Statistical analysis Linguistic techniques	Unstructured Text Wikipedia, WordNet Ontology	Semiautomatic user choose and add terms to the ontology and rank terms.	For term extraction the precision was around 4% for Significant Phrases and 3% for TF-IDF terms Recall for both was around 28% For terms matching KB between 30 and 100 percent for Significant Phrases and 11 and 100 percent for TF-IDF terms

TABLE 1: Summary of Ontology Learning Systems from Text.

5. COMPARATIVE ANALYSIS AND DISCUSSION

In the previous section, we presented a brief overview of some ontology construction systems and tools. Table 1 shows that most of the existing ontology learning tools rely on shallow NLP techniques and statistical methods. They employ shallow NLP techniques and focus only on concept and taxonomic relation extraction. For example, Text-To-Onto and Text2onto adopt shallow NLP tools and extract key concepts and semantic relations including non-taxonomic ones from texts. Although, Text2Onto uses information retrieval techniques such as Relative Term Frequency (RTF), TF-IDF, or Entropy to determine the relevance of a given item it is not used to filter out the extracted candidates. Also OntoLearn uses shallow NLP tools.

On the other hand, CRCTOL adopts a full text parsing technique. When compared with Text-To-Onto and Text2Onto, CRCTOL produces much better accuracy for both concept and relation extraction [35]. Also OntoGain adopts a full text parsing technique. It outputs the results of each ontology acquisition step in OWL which allows easier results visualization in any OWL compliant ontology editor, and easier ontology editing, maintenance, reuse and exchange. It outperforms Text2Onto for identification of taxonomic and non-taxonomic relations [30].

We conclude that, we need more than shallow NLP in order to extract rich domain ontologies, involving not only concepts but also semantic relations and axioms. Both linguistic and statistic approaches for ontology construction suffer the limitations. Therefore, domain expert is required to make sense of the results. OntoGain and CRCTOL are examples of the systems that support user involvement.

In order to construct and use of ontologies, there are many challenges that face the ontology engineering community according to [1]:

- Lacking of framework that enables the combination and comparison of different extraction methods. There is also a lack of reusable services for ontology learning, update and evaluation.
- Producing inconsistent or duplicate entries and dealing with these inconsistencies is a particular challenge.
- Absence of full support in the ontology learning tools, regarding many important aspects of ontology engineering especially ontology evolution, reuse, merging, alignment and matching**.**

6. CONCLUSION

In this paper, we have briefly described the ontology engineering field and particularly we have presented an overview of ontology learning. A definition of ontology was provided and the ontology learning approaches were identified. Further, we have also provided a brief overview of some ontology construction systems and tools followed by a summarizing comparison of them. We also discussed some of the current issues and open questions of the ontology construction from texts.

Table 1 shows a summary of all the systems and tools that have been described in this paper. We can make a conclusion that there are many of the research on learning methods consist of term extraction, synonym extraction, taxonomy and semantic relation extraction. But It remains open work how to extract axiom and rules in context of ontology learning from text. A significant challenge for ontology construction from text is the lack of systematic methods for evaluation and ontological gold standards. One clear conclusion that we draw from this literature review is both linguistic and statistic approaches suffer the limitations.
In the further work, we will design, implement and evaluate a method for ontology learning from text that includes human intervention to enhance the existing results.

7. REFERENCES

[1] A. Zouaq and R. Nkambou. “A Survey of Domain Ontology Engineering: Methods and Tools”. In Roger Nkambou, Jacqueline Bourdeau, and Riichiro Mizoguchi, editors, Advances in Intelligent Tutoring Systems, volume 308 of Studies in Computational Intelligence, Springer Berlin / Heidelberg, 2010, pp.103-119.

[2] A. Zouaq, D. Gasevic, M. Hatala. “Towards open ontology learning and filtering”. Information Systems, vol. 36, no.7, pp. 1064–1081, 2011.

[3] T. Gruber.“Towards principles for the design of ontologies used for knowledge sharing”. International Journal of Human–Computer Studies, vol. 43, no. (4–5), pp. 907–928, 1995.

[4] R. Studer, V.R. Benjamins, D. Fensel. “Knowledge Engineering, Principles and Methods”. Data and Knowledge Engineering, vol. 25, no. (1-2),pp.161–197, 1998.

[5] R. Iqbal, M.A.A. Murad, A. Mustapha, N.M Sharef. “An Analysis of Ontology Engineering Methodologies: A Literature Review”, Research Journal of Applied Sciences, Engineering and Technology, vol. 6, no.16, pp. 2993-3000, 2013.

[6] A. G´omez-P´erez, M. Fern´andez-L´opez and O. Corcho, Ontological Engineering. Springer, 2003.

[7] E. Simperl and D. Vrande. “A Methodology for Ontology Learning”, Ontology Learning and Population. Bridging the Gap between Text and Knowledge, P. Buitelaar and P. Cimiano (Eds.) IOS Press, 2008.

[8] P. Cimiano, J. Volker and R. Studer. “Ontologies on Demand? – A Description of the State-of-the-Art, Applications, Challenges and Trends for Ontology Learning from Text”. Information, vol. 57, no. (6-7), pp. 315-320, 2006.

[9] N. Noy and D. McGuinness. “Ontology Development 101: A Guide to Creating Your First Ontology”. Report SMI-2001-0880. Department of Mechanical and Industrial Engineering, University of Toronto, 2001.

[10] Y. Hong-yan, X. Jian-liang, W. Mo-ji, X. jing. “Development of Domain Ontology for E-learning Course”, in proceeding of IEEE International Symposium on IT in Medicine & Education, 2009, pp. 501 – 506.

[11] M. Fernández-López, A. Gómez-Pérez, N. Jurysto. “METHONOLOGY: From Ontological Art Towards Ontological Engineering”, in: Spring Symposium on Ontological Engineering of AAAI, Stanford University, California, 1997, pp. 33–40.

[12] M. Gruninger and M.S Fox. “Methodology for the design and evaluation of ontologies”, in: Proc. Int. Joint Conf. AI Workshop on Basic Ontological Issues in Knowledge Sharing, Montreal ,1995.

[13] A. Nicola, M. Missikoff, R. Navigli. “A software engineering approach to ontology building”. Information Systems, vol. 34, no. 2,pp. 258–275, 2009.

[14] M.C Suárez-Figueroa et al. “Revision and Extension of the NeOn Methodology for Building Contextualized Ontology Networks”. NeOn Deliverable D5.4.3 (2010). NeOn Project , http://www.neon-project.org.

[15] L. Zhou. “Ontology learning: State of the art and open issues”. Information Technology Managemen, vol. 8, no. 3, pp.241–252, 2007.

[16] M. Rizoiu and J. Velcin. “Topic Extraction for Ontology Learning”. Ontology Learning and Knowledge Discovery Using the Web: Challenges and Recent Advances, Wong W., Liu W. and Bennamoun M. eds. (Ed.).2011, pp. 38-61.

[17] P. Cimiano. Ontology Learning and Population from Text: Algorithms, Evaluation and Applications, in Studies in Philosophy and Religion, Springer 2006.

[18] P. Buitelaar, P. Cimiano and B. Magnini. “Ontology Learning from Text: An Overview”. In: Ontology learning from text: methods, evaluation and applications. Breuker J, Dieng R, Guarino N, Mantaras RLd, Mizoguchi R, Musen M, editors. Amsterdam, Berlin, Oxford, Tokyo, Washington DC: IOS Press 2005.

[19] M. Hazman, S. El-Beltagy, A. Rafea. “Ontology learning from domain specific web documents”, International Journal of Metadata, Semantics and Ontologies, vol. 4, no.1, pp. 24 – 33, 2009.

[20] T. Ivanova. “Ontology Learning Technologies - Brief survey, trends and problems”, in proceedings of the International Conference on Information Technologies , 2012, pp.245-255.

[21] M. Hearst. “Automatic acquisition of hyponyms from large text corpora”, in proceedings of the 14th conference on computational linguistics, 1992, pp. 539–545.

[22] P. Cimiano and J. Volker. “Text2Onto – a framework for ontology learning and data driven change discovery”, in proceeding of 10th International Conference on Applications of Natural Language to Information Systems, NLDB 2005, Alicante, Spain, June 15 - 17, 2005, pp. 227–238.

[23] C. Brewster, S. Jupp, J. Luciano, D. Shotton, R. Stevens and Z. Zhang. “Issues in learning an ontology from text”. BMC Bioinformatics, vol. 10(S-5), pp. 1-20,2009.

[24] M. Dahab, H. Hassan and A. Refea. “TextOntoEx: automatic construction form natural English text”. Expert Systems with Applications, vol. 34, no. 2, pp. 1474–1480, 2008.

[25] X. Jiang and A. Tan. “CRCTOL: A Semantic-Based Domain Ontology Learning System”, Journal of the American Society for Information Science and Technology , vol. 61, no. 1, pp.150-168, 2010.

[26] D. Snchez, A. Moreno, L. Vasto-Terrientes. “Learning relation axioms from text: An automatic Web-based approach”. Expert Systems with Applications, vol. 39, no. 5, pp. 5792–5805, 2012.

[27] A. Maedche and R. Volz. “The ontology extraction maintenance framework Text-To-Onto”, in Proceedings of the IEEE International Conference on Data Mining, California, USA, 2001, pp.1-12.

[28] Z. Harris. “Distributional structure”. structural and transformational linguistics, pp.775–794, 1970.

[29] B. Fortuna, M. Grobelnik, D. Mladenic. “Ontogen: semi-automatic ontology editor”, in: Proceedings of the 2007 conference on Human interface: Part II, 2007, pp. 309–318.

[30] E. Drymonas, K. Zervanou and E. Petrakis. “Unsupervised ontology acquisition from plain texts: the OntoGain system”. Natural Language Processing and Information Systems, pp. 277–287, 2010.

[31] I. Serra and R. Girardi. “Extracting non-taxonomic relationships of ontologies from texts”, in proceeding of 6th International Conference on Soft Computing Models in Industrial and Environmental Applications, Spain, 6-8 April, 2011, vol. 87, pp. 329–338.

[32] A. Weichselbrauna, G. Wohlgenannta, A. Scharl. “Refining non-taxonomic relation labels with external structured data to support ontology learning”. Data and Knowledge Engineering, vol. 69, no.8, pp. 763-778. 2010.

[33] J. Punuru and J. Chen. “Learning non-taxonomical semantic relations from domain texts”. Journal of Intelligent Information Systems, vol. 38, no.1, pp.191–207, 2011.

[34] C. Kaladevi, S. Padmavathy, A. Kangaiammal, X. Theetchenya. “Ontology Extraction for E-Learning – A Fuzzy based approach”, in proceeding pf International Conference on Computer communication and Informatics , Coimbatore , Jan.04-06, 2013.

[35] R. Navigli and P. Velardi. “Learning Domain Ontologies from Document Warehouses and Dedicated Websites”. Computational Linguistics, vol. 30, no. 2, pp.151–179, 2004.

[36] G. Petasis, V. Karkaletsis, G. Paliouras, A. Krithara, E. Zavitsanos. “Ontology Population and Enrichment: State of the Art”, in proceeding of Knowledge-driven Multimedia Information Extraction and ontology evolution, LNAI 6050, 2011, pp. 134–166.

[37] T. Yamaguchi. “Acquiring conceptual relationships from domain-specific texts”, in proceedings of IJCAI Workshop on Ontology Learning (OL), USA: Seattle; 2001.

[38] A. Maedche and S. Staab. “Ontology Learning for the Semantic Web”. IEEE Intelligent Systems, Special Issue on the Semantic Web, vol.16, no. 2, pp.72 –79, 2001.

[39] A. Maedche and S. Staab. “Discovering Conceptual Relations from Text”, in: W.Horn (ed.): ECAI 2000. proceedings of the 14th European Conference on Artificial Intelligence, Berlin, Amsterdam , August 21-25, 2000.

[40] A. Maedche and S. Staab. “Ontology learning”. In Handbook on Ontologies, In S. Staab and R. Studer, eds., The Handbook on Ontologies, Springer, 2004, pp. 173–190.

[41] M. Shamsfard and A. Barforoush. “Learning ontologies from natural language texts”. Int. J. Human-Computer Studies, pp.17–63, 2004.

[42] A. Gómez-Pérez and Manzano-Macho. “A survey of ontology learning methods and techniques”. Deliverable 1.5. OntoWeb, 2003.

[43] K. Arabshian, P. Danielsen and S. Afroz. “LexOnt: A Semi-Automatic Ontology Creation Tool for Programmable Web”, In : AAAI Spring Symposium: Intelligent Web Services Meet Social Computing, 2012 ,pp.2-8.

[44] R. Neches, R. Fikes, T. Finin, T.R. Gruber, T. Senator, W. Swartout. “Enabling Technology for Knowledge Sharing”. AI Magazine, vol. 12, no.3, pp. 36–56, 1991.