Learning Non-Taxonomic Relations of Ontologies: A Systematic Review

Mohamed Hassan Mohamed Ali, Faculty of Science, Alexandria University, Egypt
Said Fathalla, Bonn University, Germany & Faculty of Science, Alexandria University, Egypt
Mohamed Kholief, Arab Academy for Science, Technology, and Maritime Transport, Egypt
Yasser Fouad Hassan, Faculty of Science, Alexandria University, Egypt

ABSTRACT

Ontologies, as semantic knowledge representation, have a crucial role in various information systems. The main pitfall of manually building ontologies is effort and time-consuming. Ontology learning is a key solution. Learning Non-Taxonomic Relationships of Ontologies (LNTRO) is the process of automatic/semi-automatic extraction of all possible relationships between concepts in a specific domain, except the hierarchal relations. Most of the research works focused on the extraction of concepts and taxonomic relations in the ontology learning process. This article presents the results of a systematic review of the state-of-the-art approaches for LNTRO. Sixteen approaches have been described and qualitatively analyzed. The solutions they provide are discussed along with their respective positive and negative aspects. The goal is to provide researchers in this area a comprehensive understanding of the drawbacks of the existing work, thereby encouraging further improvement of the research work in this area. Furthermore, this article proposes a set of recommendations for future research.

KEYWORDS

Non-Taxonomic Relations, Ontology Evaluation, Ontology Learning, Relation Discovery, Systematic Review, Taxonomic Relations

1. INTRODUCTION

The main purpose of using ontologies in recent modern information systems is for representing knowledge about a domain that could be shared between systems. This knowledge representation allows the shared understanding of domain data, semantic processing of information, and interpretation of data, which makes information systems more effective. Ontologies have a crucial role in semantic-based systems, as they are the source of data semantics. Most of the existing ontologies are developed, populated, and maintained manually by domain experts and ontology developers. To tackle this problem, Ontology Learning (OL) as a new field of the research area in ontology engineering has
been proposed. The primary purpose of OL is to automatic or semi-automatic construct and populates ontology from a given corpus of a specific domain (1). A set of semi-automatic or automatic techniques for creating and populate ontologies are proposed, but most of them are focusing on extracting domain concepts and taxonomic relations. Most of the current research in ontology learning ignores the learning process of non-taxonomic relationships. The next decade is likely to witness a considerable rise in Learning Non-Taxonomic Relationships of Ontologies (LNTRO) research. Most studies in LNTRO discovered that the extraction of non-taxonomic relationships from the text, and the further labeling of these relationships, is considered the most challenging task (Maedche and Staab 2000a). In the research area of LNTRO from unstructured data sources, there are two main sub-problems to be considered: identifying the potential semantic relationships between ontology concepts and appropriately labeling these relationships (Ribeiro and Committee 2014).

**Problem statement and requirements.** The Manual development of ontologies is considered a time-consuming and error-prone task, as it needs a lot of effort and time. Therefore, approaches to automatic, or even semi-automatic, development of ontologies from the text are required, especially for LNTRO. The problem of LNTRO is how to automatic or semi-automatic extract non-taxonomic relationships from text.

**Research goal.** In this paper, we present the findings of a systematic review of existing approaches, tools, and evaluation methods that focus on LNTRO from an unstructured data source, such as text corpora and web documents. We would like to point out that a short review paper published in 2012 by Serra et al. (Serra, Girardi, and Novais 2012) focused on discussing only four techniques representing the state-of-the-art LNTRO. Surprisingly, we observed that there are no comprehensive review papers about LNTRO since 2012, except our previous short review (Ali et al. 2017) in 2017. Therefore, we undertook this study, in which we present the results of a systematic review of the state-of-the-art approaches of LNTRO. We qualitatively analyze sixteen of these approaches. The domain, input, output, and language of each approach are presented in Table 1. The solutions they provide are discussed, along with their respective positive and negative aspects. Our goal is to provide researchers with a comprehensive understanding of the existing work in LNTRO and give those willing to use LNTRO tools an overview of the existing one, thereby encouraging further experimentation and new approaches.

**Research question.** We aim at discussing and analyzing existing methodologies, tools, and evaluation methods for semantic relation learning from unstructured data, with a particular interest in LNTRO. To achieve this goal, we aim to answer the following general research question: “How can non-taxonomic relations be learned from unstructured data sources?” we divide this general research question into further four sub-questions: 1) What are the main processes and the techniques used in each process of LNTRO? 2) What are the current approaches for LNTRO from a text of different languages along with the advantages and disadvantages of each approach? 3) What are the evaluation methods used for evaluating LNTRO approaches? and 4) What are the current challenges for LNTRO?

This paper is organized as follows: Section 2 presents the methodology of writing this review. Section 3 presents background information about ontology learning and its related tasks. Section 4 describes the sub-tasks of the LNTRO process. Section 5 presents the state-of-the-art approaches. Section 6 discusses the evaluation techniques for LNTRO approaches. Section 7 presents our discussion with recommendations for future work. Finally, section 8 presents the conclusion and what we emerged from this study.

**2. SURVEY METHODOLOGY**

We conducted this systematic review by following reporting guidelines, such as reporting systematic reviews and meta-analyses of studies evaluating interventions in (Chagnoux et al., 2008; Cimiano & Völker, 2005; Del Corro & Gemulla, 2013). A systematic review can be conducted for several reasons (McDonald et al. 2005), such as 1) the analysis, in terms of positive and negative aspects, of