Airo International Research Journal

Volume XV, ISSN: 2320-3714
April, 2018
Impact Factor 0.75 to 3.19

UGC Approval Number 63012

OPTIMIZATION OF ONTOLOGY IN QUERY PROCESSING FOR THE LEGAL DOMAIN IN TEXT MINING APPROACH

M. Rakesh Chowdary
Research Scholar under the faculty of CSE at SSSUTMS, Sehore-MP.
Harishchandra Hebbar
Professor at School of Information Sciences, Manipal Academy of Higher Education, Manipal.

Abstract

By the flow of conceivable outcomes for a ubiquitous content production, the information over-burden problem achieves the level of complexity which can't be overseen by customary demonstrating approaches any longer. Because of their unadulterated syntactical nature conventional information retrieval approaches did not prevail with regards to treating content. This prompts a low usefulness of the consequences of a retrieval procedure for a user's job that needs to be done. As of now, database management systems exclusively depend on correct language structure of questions to recover information. As result query answers often don't meet the user's intention. In this theory we propose an ontology-based query processing approach for database systems. We utilize ontology to change a user query into another query that may give a more important response to the user. For this reason, we characterize and indicate diverse mappings that relate ideas of ontology with those of a basic database and build up an arrangement of algorithms that allow us to discover these mappings in a semi-automatic manner.

Keywords: Legal, information, query, processing, Ontology, etc.

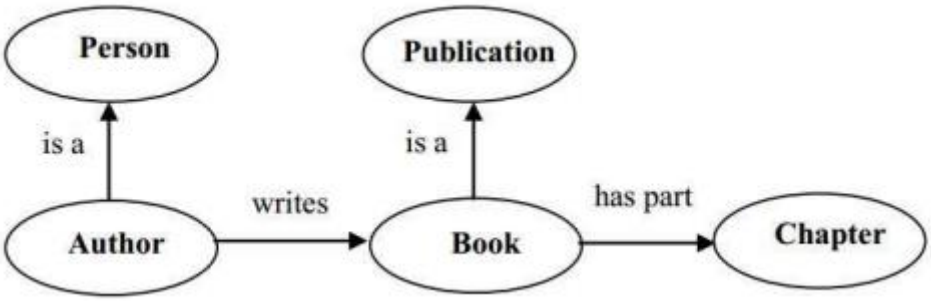
INTRODUCTION

In this paper, we investigate an intriguing issue identified with the Information Retrieval (IR) undertaking of processing of user's query to pick the important judgments. The recovered documents are abridged with term distribution model and the outcomes are exhibited as a user understandable outline. In this work, we depict the development of legal ontology that is helpful in planning a legal knowledge base to answer queries identified with legal cases.

Ontology is an essential building block for the Semantic Web; it's a formal, express specification of a common conceptualization. "Formal" alludes to the way that the ontology ought to be machine understandable. "Express" implies that the sort of concepts utilized and the constraints on their utilization are unequivocally characterized. "Shared" mirrors the thought that ontology catches consensual knowledge, that is, it isn't limited to some individual yet acknowledged by a group. A "conceptualization" alludes to a dynamic model of some phenomenon in the world that recognizes the applicable concepts of that phenomenon [1].

Ontology indicates an arrangement of constraints that announce what ought to essentially hold in any possible world. It used to recognize what "is" or "can be" in the world. It is the intention to fabricate an entire world model for depicting the semantics of information trade. Particularly in the zone of artificial intelligence, ontologies are being

utilized to encourage knowledge sharing and reuse. Ontology is involved concepts, properties, relationships between concepts and constraints. Figure below speaks to a basic ontology likewise called lightweight ontology containing classes and its taxonomical relations.



QUERY PROCESSING ORIENTED LEGAL ONTOLOGY

It is currently broadly perceived that constructing a domain model, or ontology, is an essential advance in the development of knowledge based systems. The recently made ontology follows an inventive structure which covers every single fundamental detail accessible in a legal report. The components of our system were distinguished and introductory ontology has been made based on the dictionary of terms and case ontology made in the beginning time [2]. Case ontology characterizes a hierarchical structure which is made out of the idea depicted in legal cases. To build up this, we had a few discussions with legal communities and likewise on explore different avenues regarding numerous documents identified with various sub-domains for speculation exhibitions. An underlying ontology has six vital top-level components of knowledge in the legal domain – group, person, things, event, facts and

2 acts – which are characterized in Table1 We introduce a top-down approach in the construction of sub-components by utilizing domain knowledge [3]. A top-down approach begins with the meaning of the broadest concepts in the domain alongside subsequent specialization of those concepts. There are various types of relations like is-an, identified with, made out of and so forth, which are utilized as a part of the development of whole ontological structure to depict the relationship between terms with other terms and its semantics. Ontology together with an arrangement of individual instances of classes constitutes a Knowledge Base.

CONSTRUCTING THE DOMAIN ONTOLOGY

Building ontology is the primary theme in this examination; we utilized a top-down approach in building the ontology. Most theoretical concepts are distinguished to start with, then particular into more particular concepts to construct our domain ontology which speaks to the fundamental knowledge in our work. We develop the ontology physically. We have built up the ontology contents for News domain, gathered from various significant research papers[4]. The ontology is executed with Protégé apparatus in OWL organize. As we have already shown the ontology building , the manual development of ontology comprises of the following advances**:**

- Determine the domain and scope of the ontology.
- Consider reusing existing ontologies.
- Enumerate important terms in the ontology.
- Define the classes and the class hierarchy.
- Define the properties of classes (slots).
- Define the facets of the slots.
- Create instances.

CLASSIFICATION APPROACH STRUCTURE

The ontology-based record classification approach comprises of the following parts as appeared in Figure 1[5].

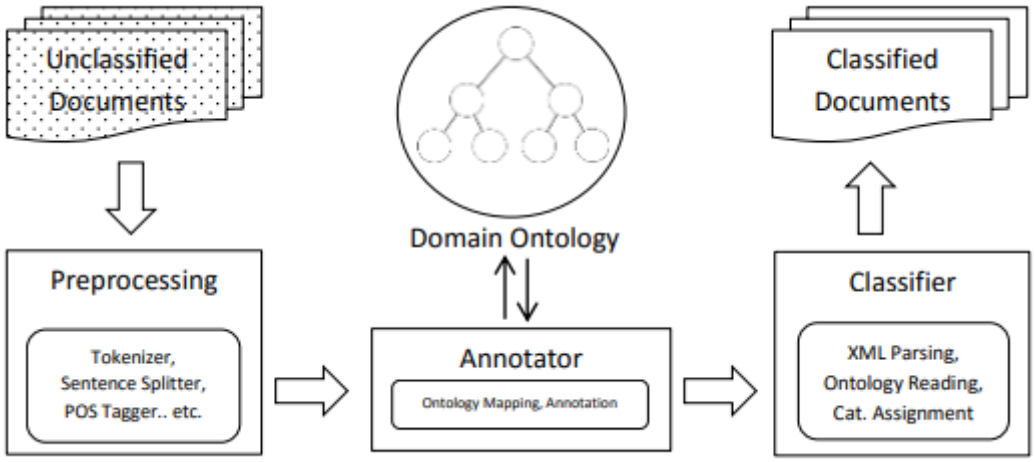


Figure 1 Classification Approach Structure

Document classification components:

- **Preprocessing:** The unclassified documents are passed to the framework, then a preprocessing stage is performed, which incorporate content tokenizing, sentence part, POS labeling, and morphological analyzing.
- **Domain Ontology:** News ontology which utilized as a part of the annotation and classification process.
- **Annotator:** which maps terms to the comparing classes, that is, the sets characterizing the distinctive implications of a term and the linguistic relations from the utilized ontology?
- **Classifier:** which classifies documents relying upon the annotation process?

Table 1: Basic components of legal ontology framework

Components	Descriptions of components
Groups	Sub-domains considered in this study
Person	Two contending parties appear before the court to resolve their conflict; one of the parties will be the subject. The object will be either a thing or a person.
Things	An object can be either a thing or animate beings including humans. The thing can be either corporeal or incorporeal. Subject action on the object or with respect to object has an impact or consequence on the other contending party
Event	The ultimate /last facts in a process (series) of facts that have given rise to the conflict/that have triggered conflict between the contending parties. The conflict is about the duty/obligation (or right) of contending parties.
Facts	This constitutes the process. Subject's actions in relation tot an object consequently constitute the process that has or is likely to have or is apprehended to have an impact on the contending parties. Facts and circumstances of the case that are serious / relevant make up the process
Acts	Courts always deal with application of law to the give facts. The law applied is extracted from the relevant statutes / provisions / rules / regulations / articles / judicial interpretations that are called acts.

A SELECTION OF LEGAL ONTOLOGY

Regardless of the name 'ontology', legal ontologies are quite extraordinary in their approach to legal domains. For illumination purposes, I recognize three distinct approaches. Semantically oriented approaches center around the semantic interpretation of a representation of components and relations in a specific domain. Epistemic ally oriented approaches center around the (securing of) knowledge in a domain [6]. Ontologically oriented approaches pressure the elements and relations that constitute a domain. It's a given that these approaches tend to cover to a specific degree regarding their semantic, epistemic and ontological cases. All things considered, the refinement classifies ontologies based on their accepted relation to the real world: through the relation between language and reality (semantics), between knowledge and reality (epistemology) or through cases with respect to presence (ontology)[7]. What the cases below have in like manner, is that they are based on a 12 Ontologies in the Legal Domain 269 top–down approach, beginning from exceptionally unique concepts, and attempting to apply these on solid domains. There is one striking exemption to this: the LOIS WordNet, which can be considered as a (lexical) ontology, was somewhat worked from the perspective of individual legal concepts [8].

THE PROPOSED SYSTEM ARCHITECTURE

It comprises of two phases: offline and online phases. In the offline phase, the file of the information retrieval system is made and kept up for the corpus based on vector space model. What's more, the domain ontology has composed and executed from the corpus. In the online phase, the user query has expanded utilizing ontology and then the list items are recovered and positioned [9]. These phases can be depicted as follow.

The first phase is offline phase, which comprises of three modules: Documents-Pre-processing, Indexing, and Ontology Building modules. Those modules can portray quickly as follow.

1) *Documents Pre-Processing Module*: This module comprises of three processes: words-tokenization, stop-words Removal, and words-weights. The words-tokenization reaction to break sentences into tokes each record for the corpus.

2) *Indexing Module*: For the terms‟ weights produced by the word-weighting process, this module files the record terms, where the file contains the term weights of words utilizing Vector Space Model (VSM)

3) *Ontology Building Module:* This module reactions to construct the domain ontology, and represents the domain ontology. This ontology is represents by the Web Ontology Language (OWL), which is the standard language for the semantic web.

The second phase is online phase, which comprises of four modules, User-Interface, Query-Pre-processing, Semantic Query Expansion, and Information Retrieval based VSM and Results Ranking. Those models can portray quickly as follow.

1) **User-Interface Module:** the module is offices the query contribution from the end user and presentations the results recovered.

2) **Query-Pre-processing module:** this module preprocesses the info query for tokenization, stop words expulsion, and so on.

3) **Semantic Query Expansion Module:** This module expands the info query based on the domain ontology. For each query words identified with the concepts in the domain ontology, the relations between concepts including that individuals is recovered, and consequently to advance the expanded query.

4) **Information Retrieval based VSM and Ranking Module**: this module coordinates the expanded query vector versus the record vectors to register the comparability between them. In this paper, the cosine comparability is utilized as appeared in condition. Then the documents recovers will rank as indicated by the similitude to the user query[10].

CONCLUSION

Our ontology-based techniques reliably outperform beginning of the workmanship look strategy, notwithstanding when it is forced to work for every one of the three sub-domains considered in this examination. An instinctive clarification for the execution of our ontology-based system is that it gives a knowledge base which had a colossal accumulation of terms and its relationships and other related highlights which is utilized

for better improvements of query terms. Our basic structure can likewise be expanded with the option of terms by including new documents and from various subdomains in future course of time.

REFERENCES

[1] M. Saravanan, B. Ravindran, and S. Raman, Improving legal document Summarization using graphical models, Legal Knowledge and Information Systems, JURIX 2006: The Nineteenth Annual Conference, Paris, 2006, IOS Press, pp.51-60.

[2] A. Valente and J. Breuker, Making Ends Meet: Conceptual Models and Ontologies in Legal Problem Solving, Proceedings of the XI Brazilian AI Symposium (SBIA'94), 1994, pp.1-15.

[3] Asunción Gómez-Pérez. Ontology Evaluation, Handbook on Ontologies, (2004), 251-274.

[4] Nicola Guarino, Formal Ontology and Information Systems, Proceedings of FOIS'98, Trento, Italy, 6-8, June 1998, IOS Press, Amsterdam, 1998, 3 -15

[5] Aggarwal, C., "Data Classification: Algorithms and Applications", Chapman and Hall/CRC, 2014.

[6] Alatrash, E., "Using Web Tools for Constructing an Ontology of Different Natural Languages", Doctoral dissertation, University of Belgrade, 2013.

[7] Aliane, H., Alimazighi, Z., & Cherif, M., "Al-Khalil: The Arabic Linguistic Ontology Project". Language Resources and Evaluation Conference LREC, May 2010.

[8] Aljazeera Website, http://www.aljazeera.net/portal, 2015, January, 12.

[9] Al-Safadi, L., Al-Badrani, M., & Al-Junidey, M., "Developing ontology for Arabic blogs retrieval", International Journal of Computer Applications, vol.19, no.4, pp.41, 46, April 2011.

[10] Al-Shammari, E.T., "Improving Arabic document categorization: Introducing local stem", 2010 10th International Conference on Intelligent Systems Design and Applications (ISDA), pp.385-390, 29 November 2010 - 1 December 2010.