International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
Volume 12 – No. 13, May 2018 – www.ijais.org

A Review of Ontology-based Information Retrieval Techniques on Generic Domains

Ishaq Oyebisi Oyefolahan
Federal University of Technology, Minna, Nigeria
Information & Media Technology Department

Enesi Femi Aminu
Federal University of Technology, Minna, Nigeria
Computer Science Department

Muhammad Bashir Abdullahi
Federal University of Technology, Minna, Nigeria
Computer Science Department

Muhammadu Tajudeen Salaudeen
Federal University of Technology, Minna, Nigeria
Crop Production Department

ABSTRACT
A promising evolution of the existing web where machine and people are in cooperation is the Semantic Web. That is, a machine's represented and understandable web. This is against the existing web which is syntactic in nature - where meaning of query search and its expected results on the web is mostly understood and interpreted by user not machine. However, the technologies drive behind this goal of semantic web is on one hand ontologies and on the other hand information retrieval techniques. Ontology is a data modeling technique for structured data repository premised on collection of concepts with their semantic relationships and constraints on a chosen area of knowledge. While on the other hand information retrieval technique is a mechanism of retrieving relevant information based on the query search. There are existing techniques for information retrieval processes, which includes that of ontological process. Therefore, this paper aimed to present a review on these existing techniques based on different classifications processes. Also, the analysis and comparison of the review are carried out based on some fundamental criteria which include various ontology's domains, ontological tools, information retrieval techniques along with the weights computation algorithms and different evaluation techniques. Thus, a review of ontology based information retrieval techniques had been carried out and this paper has disambiguates the categorization processes of the techniques and serves as a developer's guide for chosen a technique for any domain.

General Terms
Semantic Web; Ontology; Information Retrieval

Keywords
Semantic Web; Ontology; Information Retrieval; Query Expansion; Semantic Annotation; Weights Algorithms

1. INTRODUCTION
Presently, the volume of information stored on the electronic document repositories such as World Wide Web (WWW) and its demand by users is becoming high and therefore drawing research attentions. The existing WWW which is adjudged as the largest platform for electronic documents is described to be syntactic in nature. That is, the meaning of query search and its expected results on the web is mostly understood and interpreted by user not machine. This shortcoming can be attributed to the fact that the mechanism to retrieve relevant results that satisfy user's needs becomes a challenge as a result of word mismatch and short queries; in other words, ambiguous nature of the natural languages [1]. Thus, in order for the web to be meaningful both in terms of structure and retrieving relevant information, the web must be understood and interpreted by both user and machine hence, the term semantic web.

Semantic web is a promising evolution of the existing web where machine and people are in cooperation [2]. In other words, a web that is machine's represented and understandable. Unlike the existing web that is only human understandable. However, the technologies drive behind this goal is on one hand ontologies and on the other hand the information retrieval techniques [3]. Stressing further, [4]; [5] reported that one suitable means of achieving semantic web is by exposition of knowledge via the means of ontologies.
Ontology is therefore, a data modeling technique for structured data repository premised on collection of concepts with their semantic relationships and constraints on a chosen area of knowledge. [6] described five components to formalized ontologies as classes, relations, attributes, axioms and instances. Area of knowledge otherwise known as domain, sparse across any real life scenario such as agriculture, medicine, law, news, engineering, sports, arts and entertainment and so on. Similarly, one of the popular definitions of ontology is that of Gruber [7] which states that ontology is an explicit and formal specification of a conceptualization.

Information Retrieval (IR) is a mechanism of retrieving relevant information based on the query search of user's intent. There are existing techniques for information retrieval processes, which includes that of ontological process which is the focus of this paper. Generally speaking, Information Retrieval Systems (IRSs) are categorized into two folds: The syntactic search systems, otherwise known as keyword-based systems and semantic search systems also called conceptual-based systems [8]. Word mismatch is a serious challenge of Information Retrieval [9]. It is described as situation of IR system where the concept of query terms of the user is at variance with the concepts of the documents. This is attributed to the vocabulary issues of synonyms, polysemy and the likes.

Typically, when user invoke query, IRS is to choose information that is likely to satisfy user's intent. In order to satisfy this demand, three processes are carried out [10]. In the light of this, [11] described IRS as a function that maps a query Q (user' request) from H_Q (a set of query) to a set of m documents within collection D of all indexed documents.

$$IR : H_Q \rightarrow D^M$$

In recent time, ontology has been progressively used for developing applications for different domains. Researchers have proposed some techniques however, in terms of application for the information retrieval systems, it has not yield a considerable integration with searching methodologies [12]. Therefore, the applications of ontology in semantic information retrieval are diverse in nature and cannot be underestimated. That is, not limited to contextual semantic search only but also to semantic similarity measures [13].

2. WHY ONTOLOGY-BASED INFORMATION RETRIEVAL

Ontology-Based Information Retrieval is becoming an interesting area in the current research trend of ontology and semantic web [14]. The capability and applicability of ontology are diverse in nature but more indispensible in knowledge management and information sharing. To this end, the work of [15] presented a picture for modeling ontology for information sharing, recovery and interoperability amid multidisciplinary areas.

Search on the existing web is by keyword-based approach. While commendable efforts in terms of retrieving relevant documents have shown towards IRS by the researchers, there is still room for further improvement. For instance, ability to explore conceptualization lies in user's request and corpus meanings by using additional knowledge. That is, the capacity of IRS to draw inference or relation between the queries terms [16].

An ontology based search system as a knowledge search and retrieval system developed for any domain has the capacity to address the limitations of the existing keyword approaches [17]. The authors first developed ontology for mechanical domain as the knowledge area and adopted query expansion technique for its knowledge retrieval. It was reported that the proposed method outweighs the performance of keyword based search methods in terms of precision and recall metrics. Similar approach was used by [18] and [19] for engineering domain and technique preparation domain ontologies respectively. An ontology based information retrieval was presented in the domain of soccer [20]. In their research work, three semantic search issues were taken into consideration which includes the retrieval performance issues.

Ontology design and its inherent technologies without developing further algorithm or applying the existing techniques are considered to address the pitfalls associated with information retrieval systems by some literatures. Similar example of this aforementioned approach is the work of [21]. The researchers described an ontology development methodology that would be incorporated with protégé and promised that the proposed method would have the capacity to achieve efficient ontology based information retrieval. Similarly, cash crop ontology developed by [22] is reported to provide precise cash crop farmers market's information in Nigeria.

[23] argued that the traditional information retrieval based on the keyword does not satisfied user's query intent in terms of recall and precision. Therefore, in order to enhance the relevance of results return to user, an information retrieval model based on domesticated plants ontology is developed. The ontology was used to produce an information retrieval system that aid context based search consequently, changing the keyword based search system. Also, the research of [24] described how agricultural information in Nepal became accessible by converting different available agricultural data which are not readily and easily accessible before to easier and accessible structure through the use of Resource Description Framework (RDF). RDF is a triple based ontology representation language.

Similarly, some researches consider the approach of ontology mapping to address the problem of information retrieval. For examples; [25] proposed ontology refinement algorithm to refine ontologies for information retrieval, by mapping some popular biomedical ontologies (MeSH and Gene). [26] proposed hybrid fuzzy ontology whereby concepts which are similar from two ontologies of a domain are retrieved and aligned. They stated that, affirmation of ontology into accessing information and retrieval is an effective approach to implement searching effects of appropriate information that users want. Therefore, in order to have a more efficient information retrieval system, an ontology based IR techniques have to be considered. Also, its accuracy is proportionate to the efficiency of its retrieval techniques [27].

3. TECHNIQUES FOR ONTOLOGY BASED INFORMATION RETRIEVAL

As earlier stated, the sufficiency of ontology alone to efficiently take care of information retrieval becomes a research issues. Therefore, in an attempt to address the research gap of effective information retrieval based on ontology, some techniques are reviewed in this paper. Namely; Query Expansion technique, Semantic Annotations techniques and others.

3.1. Query Expansion (QE) Technique

To retrieve a relevant hit of information based on a search term (query), additional useful terms may be added to the initial search term. Such mechanism is described as Query Expansion or Query Augmentation Technique of IRS. The scheme of adding knowledge to the existing initial query term in order to retrieve relevant information has been catching attentions of researchers in the field of study. From the literatures reviewed, the mechanism of additional useful terms can be achieved manually, interactively or automatically [28]; [29]. Manual QE is achieved based on the skillful decision of user. User takes decision of which term to add to the new query. The interactive (semi-automatic) QE is often described as user-assisted QE. In semi-automatic, once the system produces additional useful terms, user decides which of the term to add to the initial query. For Automatic QE, weights are computed for each of the additional terms and the term with the highest value will be added to the initial query autonomously by the system. There are different weighting functions that give different results. Thus, it is important to note that the efficiency of retrieval system in this scheme depends on how the weighting values have been computed.

The history of automatic query expansion in the field of information retrieval remains indispensible. It is recommended as an efficient approach of solving the word mismatch and short query term issues associated with the

information retrieval systems of the existing web [30]. It was propounded in the early year of 1960 by Maron and Kuhns. However, just about a decade ago it has attained the level of scientific and experimental usefulness, particularly in laboratory scenario such as the Text Retrieval Conference (TREC). Other notable examples of this technique of solving vocabulary issues as a result of natural language ambiguities in terms of synonyms and ploysemy are interactive query refinement, relevance feedback, word sense disambiguation, search results clustering, and the likes [31].

Essentially, with query expansion, the initial query is augmented with an additional meaningful terms and these terms are usually the derivative of knowledge collections or the returned documents of the initial query. There are two classes of knowledge collections, they are: knowledge independent collections and knowledge dependent collections [14]. Query expansion technique can be effectively used both in common knowledge bases (wikipedia) and expert-design knowledge bases such as ontologies [32]. Thus, the new search term produced from the mechanisms described is regarded as contextual information for the initial search term hoping to enhance the retrieval results. The contextual information may be obtained from term co-occurrence, relevance feedback and the latest which can be deduced from knowledge models for instance, ontologies.

3.1.1 Classification of Query Expansion Techniques

The classification of this technique varies in literatures. For instance, Query Expansion is categorized into two types. They are: local analysis and global analysis [33]. The former implies that only a subset of documents that are returned from the initial query term to deduce appropriate words in order to recast the initial query is considered. Two categories of techniques are identified for local analysis; the relevance feedback and the local feedback. The relevance feedback enhances the expanded query by using the search results of the initial query that related to the feedback. Conversely, the local feedback recast the new query or extracted the terms for the expanded query by using the top-ranked document of the initial query results. The latter (global analysis) is described as query-independent, examines all documents for all queries. That is, the expansion terms are extracted from the whole document collection. [34] refer to it as the commonest form of query expansion. Techniques applied for query expansion involves thesaurus, ontology-based [35].

Furthermore, [36] categorized global and local techniques as automatic query expansion while, relevance feedback and manual thesaurus are categorized as manual techniques. The authors classified pseudo-feedback (also called local feedback) as a more topical and recent local technique as at the time of their research. Pseudo feedback drew its strength from relevance feedback. More also, term clustering is described as one of the first global techniques. It works by grouping allied terms into cluster based on their co-occurrences in a collection of documents. Similarly, [12] reported that there are two types of techniques for automatic query expansion. These are: global and local techniques.

In addition, the research work of [14] presented a generic classification. This is classified into three folds. These are: manual, user-assisted and automatic methods. These methods have been earlier explained in the previous section. [1] in their paper described the query expansion techniques as syntactical, semantic or combination of both. Syntactical technique involves statistically analysis the terms dependence. While, analyzes the terms dependences semantically (that is, ontology or thesaurus) portrays semantic approach. Again, combination of the two approaches may be used.

3.1.2 Literatures Review: Ontology-Based Query Expansion Technique of Information Retrieval

Over a period of time, query expansion technique of information retrieval has been continuously reviewed and applied under various expansion mechanisms. This has evidently yielded a good retrieval results. However, since web of knowledge is currently trending, a knowledge-based query expansion techniques is required [1].

[14] maintained that the retrieval mechanisms of an ontology based system are drawing research attentions in the field. They carried out their research on ontology based query expansion technique for the domain of agriculture. The format of agricultural information on the web poses a lot of challenges in terms of knowledge base redundancy and also complexity to manage the different natures of relationship that exist among the concepts of the knowledge base. Therefore, modeling agricultural information ontologically is capable to resolve the research problem.

The research work of [33] focused on query expansion using global analysis to improve the performance of information retrieval of agricultural domain ontology using simple query terms and association rule mining based method for inference. The ontology is based on agricultural expertise retrieval framework (ARGIX) of [37]. It is constructed using Protégé and generated the agricultural OWL files.

[38] in their research examined the use of Relevance feedback and Pseudo relevance feedback techniques of query expansion based on the information from news domain ontology in a probabilistic retrieval system. The sole goal of the research is to evaluate the effect of ontology in terms of recall and precision metrics. The ontology written in XML was developed by [39]. The authors concluded that the results of pseudo-relevance feedback outweigh the performance of the relevance feedback. TREC document collection was selected for evaluation owing to the chosen domain.

An easy access to research information of any domain or field of study remains fundamental for qualitative research development. However, the intent of user's search cannot be semantically interpreted by the conventional search engines. To this end, Aree [40] proposed query expansion-based information retrieval methodology for plant production ontology specifically, rice domain. Thus, the IR technique along with the reasoning components of the ontology were reported to greatly enhance the performance of the ontology based IR. The ontology was evaluated by domain experts and users following how satisfactory the ontology to the competency questions. The query efficiency was measured by precision and recall.

To enhance search over large document repositories, [41] proposed a query expansion algorithm in sports domain for semantic information retrieval. The proposed SIRSD algorithm works by using wordNet and domain ontology to enhance the returned results of query and it was experimented with respect to the traditional search. The research depicted the significance of query expansion using ontology (wordNet) over query expansion using relevance feedback. The sport ontology is developed using protégé tool and its information

represented in RDF. Web Jaccard was used to model and compute the semantic similarity between two words.

Therefore, ontology-based query expansion technique of information retrieval can be viewed, analyzed, or designed in three forms depending on the goals of a research. Namely: query expansion technique using wordNet ontology, query expansion technique using domain-specific ontologies and combination of wordNet and domain-specific ontologies. Primarily, the first two ontologies's considerations have their own potentials and limitations as they have been captured in this review paper by the subsequent sections.

3.1.3 Query Expansion Technique using WordNet

As earlier stated, wordNet is described as knowledge independent collection or a corpus independent knowledge model. That is, a collection of documents that is not restricted to particular domain of knowledge. Generally, WordNet is a lexical data repository that is ontologically design for linguistic purpose, specifically for parts of speech. It has three databases. These are noun, verb, adjective and adverb. The set of synonyms denoted as synset makes available various semantic relations such as synonyms, antonymy, hyponymy, hypernymy, holonymy and meronymy [26]. Despite its limitations, wordNet has played significant role in information retrieval techniques from the literatures reviewed [42]. Besides web search, it has been tested on patent search tasks [43]; all these for query expansion. Beyond query expansion, is its ability to disambiguates the sense of query words [44] all to achieve efficient information retrieval. In contrast, the limitations of this technique includes most of the semantic associations between two terms are not found in the wordnet [45]; [46] and since wordnet has wide coverage, ambiguous terms within the ontology can pose a challenge [28].

[47] presented a query expansion approach using wordNet synonyms and meronyms for retrieving geographical terms in GeoCLEF2005 English monolingual exercise. The popular Lucene search engine was used for indexing and retrieval. However, the results of the proposed method were not appropriate for the exercise. It therefore concluded that, wordNet can be better applied during the indexing stage, by toting up synonyms and holonyms to the index terms.

3.1.4 Query Expansion Technique using Domain-Specific Ontologies

In recent time, ontology has been progressively used for developing applications for a chosen domain and information retrieval systems. Ontology may be developed or reuse in any field of knowledge for instance, agriculture. So, when ontology is used as a knowledge model to cushion the effect of retrieving relevant information in a given domain to form conceptual information based on the initial query, it is described as query expansion using domain-specific ontology. It models terms and concepts for a chosen field of knowledge.

In the literatures of [14]; [33] and [48] query expansion was implored for agriculture domain ontologies. Similarly, [12]; [41]; [20] and [49] for sports based ontology.

3.2 Semantic Annotation (SA) Technique

A variation form of QE is the SA technique where data or concept in ontology is tagged to another data from a structured document (for instance, ontology) or other documents to provide meaningful information. The importance of annotation technique in semantic web development cannot be over emphasized in that the driving technologies behind Semantic web is on one hand ontologies and on the other hand semantic annotation [3]. Ontology as earlier stated; a knowledge representation model for a chosen domain based on a methodology. While, Semantic Annotation the subject matter of this section, is defined as metadata that offers relation between entities that emerge in resources or collections of document and domain concepts created in ontology. It is noteworthy to mention that annotation can be carried out across any multimedia contents. However, for the sake of effective information retrieval, it is paramount to define the scope of content to annotate.

Semantically, ontology-based annotation of textual content of resources involves two major activities: detection of entities of a resource to annotate and tagging of each entity with most suitable class of domain ontology [3]. The former can be achieved via Named Entity recognition technique while on the other hand the latter can also be accomplished by web-based statistical measures or wordNet. The significance of Semantic annotation as another useful tool in carrying out information retrieval tasks has been reechoed in the research work of [50]. The research described annotation in various forms. That is, annotation can be in form of named entities, semantic roles, temporal information and the likes.

Furthermore, the research of [51] presented a good understanding of annotation. They admitted that ontology and annotation strongly form the basis of machine understandable description of documents (semantic web). Therefore, such description can be formed by annotating resources with metadata resulting in the technique called annotations about that resource. An annotation adds data to some other pieces of data. It creates within the given context, an association between the annotated data and the annotating data. There are three types of annotations; these are informal, formal and ontological. This paper is restricted to ontological annotation only.

In addition, the authors further described annotation as a quadruple model (a_s, a_p, a_o, a_c) where a_s is the subject of the annotation (the annotated data) a_o is the object of the annotation (the annotating data) a_p is the predicate (the annotation relation) that defines the type of relationship between a_s and a_o, and a_c is the context in which the annotation is made. As a result, ontological annotation A_s is thus, a formal annotation A_f , where the predicate a_p and the context a_c are an (arbitrarily complex) ontological term, and the object a_o conforms to an ontological definition of a_p [51].

Semantic annotation according to the survey work of [52] is a process of tagging ontology class's instance data and mapped into ontology classes. The research is aimed to examine semantic annotation platforms that can be applied to carry out semi-automatic annotation. Also [53] in their research stated that knowledge management such as semantic annotation partly has to do with a mechanism in which new terms might be obtained from corpus in order to relate them to already existing or new documents. In the context of this paper, the existing or new document refers to ontology. Therefore, we reviewed that the information retrieval technique of ontology can be enhanced via annotation.

Based on literatures reviewed, ontology along with annotation technique serves as a paradigm shift solution towards the challenge of retrieving relevant information from any given document collections. In line with this conviction, [6] developed an ontology-based semantic annotation system to

solve the difficulty in accessing cloud based services suitable to user's intention. Similarly, [54] presented a system that enhance patient's diagnosis and treatment scheme by annotation of clinical images via regular vocabularies from clinical ontologies.

Thus unlike query expansion technique who aimed to argument query term, this technique attaches data to the given context of discourse or domain's entities directly. The domain of discourse includes document annotation, semantic blogs, semantic wikis and tagging. Even though the annotation paradigm supports annotating various kind of multimedia content, the subject granularity or scope for any domain has to be taken into consideration. For instance, the research of [3] focuses on textual content of document annotation.

3.2.1. Classification of Semantic Annotation

Documents can be semantically annotated in three different forms. Namely: manual, semi-automatic and automatic [52]. Manual simply means annotation of document (in this context; ontology's classes) is tagged by human in entirety. While, semi-automatic means that annotation of documents are partly carried by machine because at some points human intervention are required. On the other hand, automatic form of document annotation although a tedious form of annotation [55] is completely handled by machine without any form of human intervention. This third category is attracting research attentions such as the research of [56].

[52] further examined the semantic annotation platforms and presented that the classification of the platform depends on the type of annotation approach applied. Thus, two main types of approach presented are pattern-based semantic annotation platform which includes discovery and rules, and machine learning based semantic annotation platforms such as induction and probability. Also, these two methods of the two platforms can be merged together to form what is called multistrategy semantic annotation platforms.

3.2.2. Literatures Review: Semantic Annotation Technique of Information Retrieval

An architectural framework of Semantic Based Information Retrieval System (SBIRS) for a semantic search was designed and implemented by [57]. The work presented an improved TF-IDF algorithm to retrieve information in a more reliable manner. A semantic indexer component of the SBIRS establishes a weighted semantic annotation and indexing where scores were computed by an adaption of the enhanced algorithm. The system was implemented by C#.net as web-based system in visual studio 2010. Effectiveness of the IR system was measured by precision, recall and F-measure. Comparison of KBIRS and SBIRS in terms of precision and recall were carried out.

A prototype tool (AutoMeta) that enable semi-automatic annotation of document for publishing on the web with the aid of RDFa was presented by [58]. The authors explored the ontology inference capability for semantic annotation and on meta-annotation concepts. The system was developed in Java 1.6 using the APIs: OWLAPI, Pellet-Reasoner, JavaRDFa and Jakarta-commons. Evaluation was carried out with and without the reasoner by taken precision and recall metrics into consideration. However, there were no clear measures to preprocessed initial documents to avoid stop words and the likes. More also, concepts of the domain ontology as terms that are present in the web document are exact match therefore, recall may likely be affected. Synonyms of those terms ought to be taken into consideration. AutoMeta was not evaluated along with existing semantic annotation tools. For example; Gate, Zemanta, Annotea, GoNTogle.

A semantic annotation information retrieval system based on corn plant ontology was proposed in the research of [59]. The semantic annotation technique was RDF triple-based of FCA approach. FCA approach equally involved TF-IDF algorithm to compute terms weight. Terms are chosen as core domain concepts if only their weight values are more than the threshold value. The semantic annotation based on ontology's processes as described in their work involves document preprocessing; extraction of feature words using the algorithm and extraction of semantic triples.

[49] proposed a semantic based system that consists of four modules. Automatic semantic annotation forms an integral component of the system. This is for efficient retrieval of (web) documents. The annotation component of the system was designed based on the algorithm of SPARQL. With the aid of the algorithm, terms related to the domain (sports) would be added to the updated ontology knowledge base and equally filters out irrelevant documents found in the document base (DB).

As the existing web is gradually evolving to web of meaning (semantic web) then, the huge volume of textual web resources becomes practically difficult to annotate manually. To this end, [3] in their article presented a methodology that partly annotates textual content of web resources in an automatic approach without any human intervention. The methodology exploits machine learning algorithms to discover relevant entities of the textual content and relate them to the classes of an input ontology via linguistic pattern approach.

As earlier stated, annotation technique can be carried out in manual form, human-assisted (semi-automatic) form and automatic form. Also, the technique can be implemented on any multimedia contents. Consequently, [55] in their review work canvasses and stressed the need for automatic annotation and using structured vocabularies for annotating image content. The authors concluded that the use of semantic hierarchy for semantic image analysis is attracting the attentions of the researchers as a result of its proven better performance.

Similarly, [60] in their review implore the Automatic Image Annotation (AIA) technique to bridge the semantic gap between low level image features and high level semantic. AIA exploits semantic attributes with the aid of machine learning algorithmic procedures. The article was thus concluded by presenting the difficulties associated with the technique.

Besides these major techniques of information retrieval (query expansion and semantic annotation) reviewed in this paper, the following sections equally presented and discuss other techniques.

3.3.O-A-V Model [61]

Besides, the popularly used query expansion and semantic annotation techniques for information retrieval, [61] proposed an ontology based Object-Attribute-Values (O-A-V) information retrieval model. The proposed system displays search results with a structured insight in the form of O-A-V. The proposed model works based on the designed ranking algorithm. Precision at 10 (P@10) and Mean Average precision (MAP) were the metrics for the system's evaluation

based on TREC technique.

3.4.Keyword Matching Indexing System [26]

A better querying mechanism for information Retrieval that integrates the ontology queries with keyword search was the research presented by [26]. The proposed model used hybrid fuzzy ontology whereby concepts which are similar from two ontologies of a domain are retrieved and aligned. A semantic matching result between input queries and information in the ontology field was developed. Fuzzy ontology was generated with the aid of FOGA (a fuzzy logic). The essence of keyword matching ontology is to primarily take care of keywords that are not specified in the ontology. OWL is used for mapping the generated fuzzy ontology to semantic representation. The ontology evaluations were carried out using the IR metric of precision, recall and f-measure among the standard FOGA, keyword-matching and hybrid FOGA.

4. REVIEW OF THE REVIEWS

The review work of [28] on ontology based query expansion presented a detail accounts on various approaches of augmenting an initial search query with additional term to form contextual based information using the technique. The review discussed the three processes of adding terms which include manual, semi-automatic and automatic. The different approaches of query expansion reviewed equally include query expansion using relevance feedback (both traditional and pseudo approaches), query expansion using corpus dependent knowledge models, query expansion using corpus independent knowledge models, using ontologies for information retrieval tasks. Above all, the review accounted for how the use of ontologies (both wordNet and domain specific) for query expansion outweighs the other approaches. The researchers therefore concluded that using context from ontology to query expansion for the domain of newswire is recommended for further study.

Similarly, [14] in their review of ontology-based query expansion classified the technique into three. These are manual QE, interactive QE and automatic QE. The additional terms for initial query reformation can be derived from either the return documents of the initial query or from the knowledge collections. The review presented two types of knowledge collections. They are: knowledge independent collections (such as thesauri, ontologies, dictionaries) and knowledge dependent collections (such as specific domain ontology). In addition, the review categorized the purpose of query into three forms. They are navigational, transactional and informational queries. It was stated in the research that query expansion technique is best applied when the query is informational. The review on ontology-based query expansion was carried out using agriculture domain ontology and wordNet.

[27] from their survey research on ontology-based information retrieval admitted the tedious task of filtering relevant information by user from the large chunks of collected documents as a result of keyword based search. Thus, the concepts of ontology and its retrieval techniques such as query expansion and indexing are possible panacea to address the shortcomings.

5. DISCUSSION

In this section, query expansion and semantic annotation techniques of information retrieval are discussed. This is due to their popularity judging from the literatures reviewed.

For query expansion technique, different authors presented different categorization process however; they are still related but have some level of ambiguity for researchers to comprehend. That is, the classification of this technique considering the literatures reviewed is not a clear-cut case. For instance, while [14] categorized the technique into manual, user-assisted and automatic methods; [33] categorize it into two approaches as local analysis and global analysis. As earlier explained in Section 3.1.1, the local analysis approach extended the initial query from part of the document that returns either by relevance feedback or local feedback techniques. While on the other hand, global analysis expanded the initial query by considering the whole documents that returns using ontology or thesaurus. In this paper, we used Figure 1 to illustrate this classification.

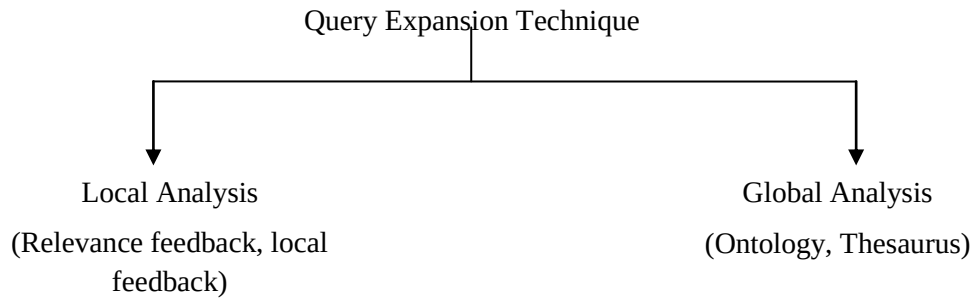

Figure 1: Classification of Query Expansion Techniques [33]

The classification according to [14] is illustrated by Figure 2. Query Expansion Technique is classified into manual, interactive and automatic. The query can be expanded irrespective of these techniques either based on search results or based on knowledge collections (knowledge independent collections or knowledge dependent collections). Ontologies, dictionaries, glossaries and thesauri are examples of knowledge independent collections. While knowledge of specific domains such as agriculture ontology are regarded as knowledge dependent collections.

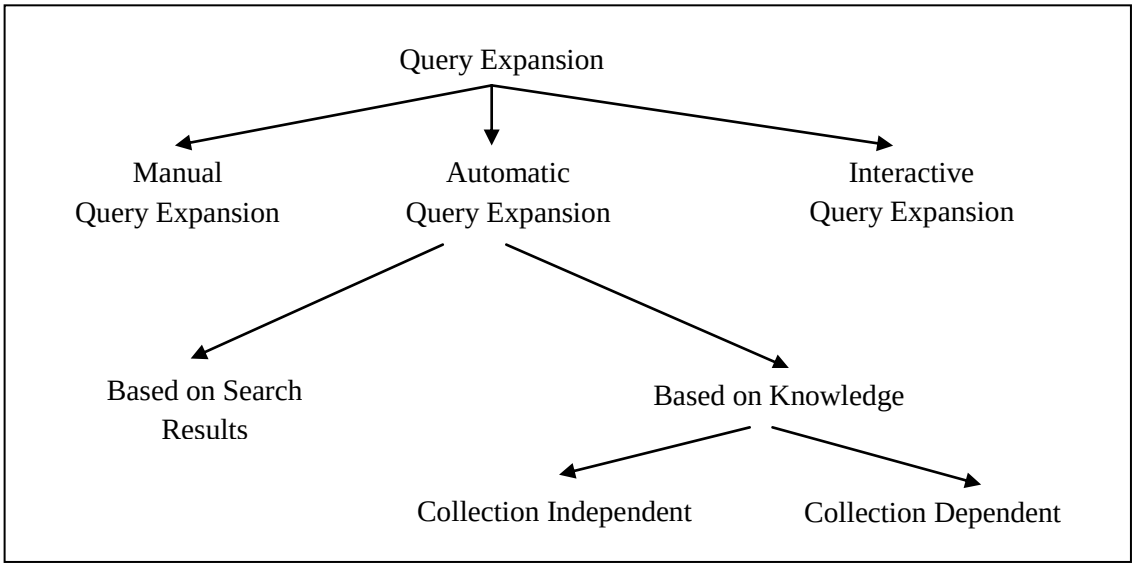

Figure 2: Classification of Query Expansion [14]

Similarly, [36] classified the technique into two. They are automatic and manual query expansion. Just like in the work of [12]; they classified local and global techniques as automatic QE with pseudo-feedback as local technique and term clustering as global technique. Relevance feedback and manual thesaurus are classified as manual query expansion. Figure 3 illustrated it further.

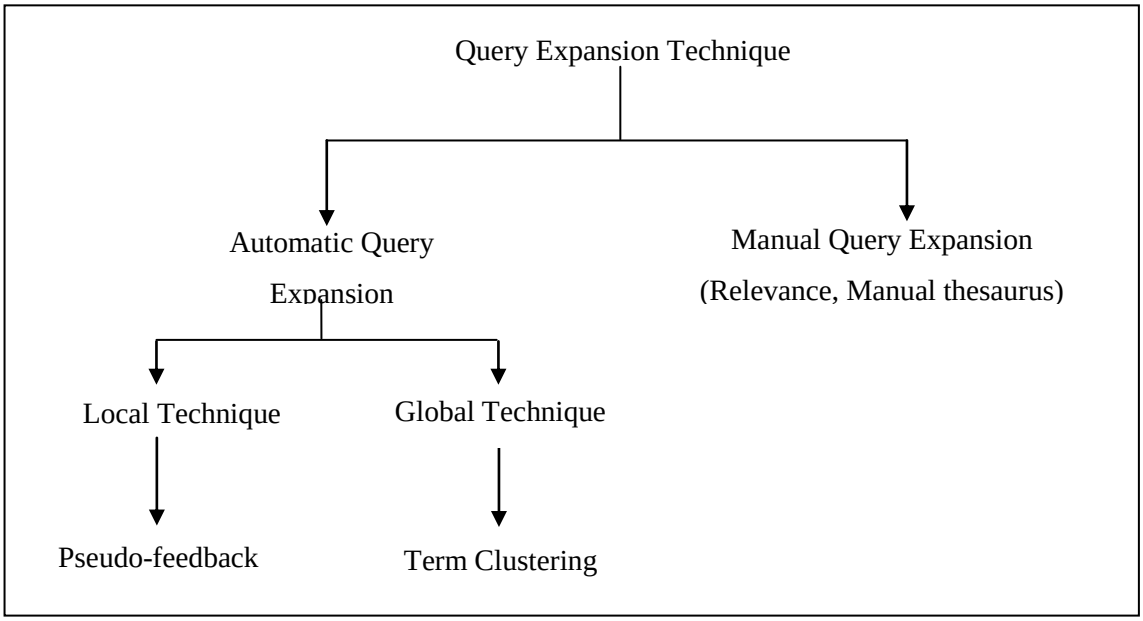

Figure 3: Classification of Query Expansion [36]

More also, [1] classified the technique into two. Namely: syntactic approach and semantic approach. The former is with the aid of statistic analysis while the latter is with ontology or thesaurus. In other hand, the two approaches can be combined. Figure 4 represented this classification.

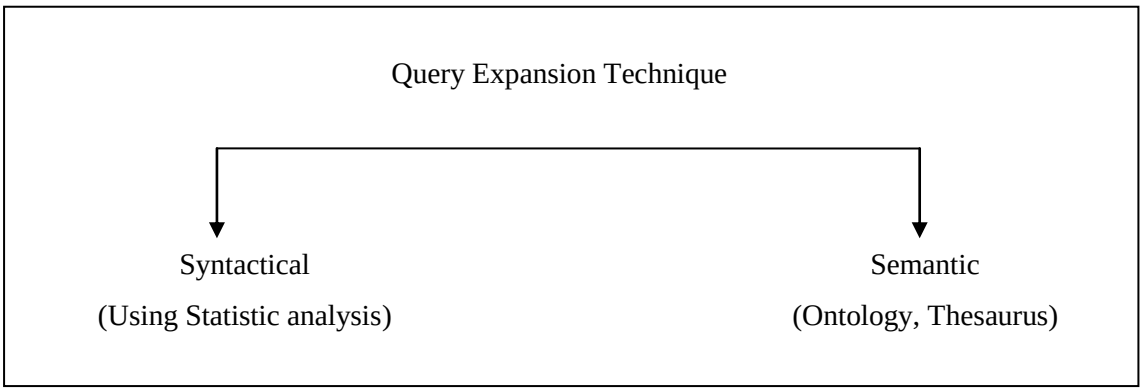

Figure 4: Classification of Query Expansion [1]

Since the primary aim of query expansion technique is to add new meaningful terms to the initial search query and that the process of adding terms can either be manual, user-assisted or automatic [28]; we therefore presented in this review a revised classification of query expansion techniques as clearly shown by Figure 5 in order to disambiguate the categorization process of the technique.

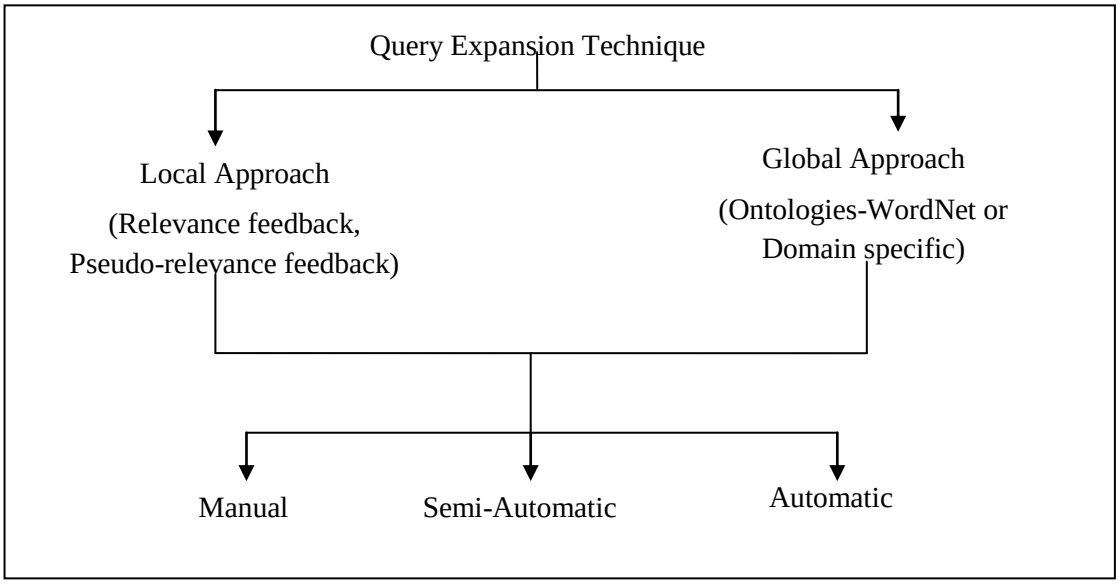

Figure 5: Revised Classification of Query Expansion Technique

Figure 5 in this review paper, analogously classified query expansion technique into local and global techniques. While the global technique consists of ontologies, the local consist of relevance feedback and its alternative called pseudo-relevance or local feedback. The local technique may be described as syntactic approach if the return documents in which relevance feedback or pseudo-relevance feedback would be acted upon is not ontology that is, unstructured document collections. While global technique may be described as semantic based approach because the returns document for expanding the initial query is a structured document collections such as ontologies. This involves generic ontology (wordnet) or domain specific ontology (maize ontology).

Therefore, this paper takes a position that irrespective of the techniques under local or global approach, it can be achieved manually, semi-automatically or automatically. Automatic means the system autonomously add the relevant term to a query. On the other hand, semi-automatic or user-interactive approach means user can judge on the results return by the system. Manual means adding the relevant term to a query is wholly done by user. In any case, it is worth mentioning that the two techniques (global and local) can be hybridized.

For Semantic Annotation, the classification of the technique based on the literatures reviewed is a clear-cut case. That is, the literatures presented a straight forward classification as in the case of [52] in Figure 6.

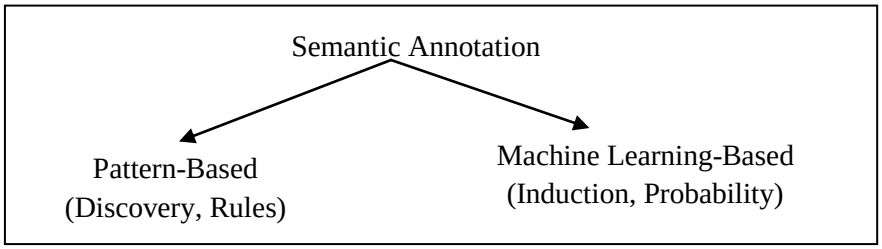

Figure 6: Classification of Semantic Annotation Platform [52]

Figure 6 presented a technique of attaching data to the given context of discourse or domain's entities directly (that is, annotation) in two categories. These are: pattern based technique and machine learning based technique. The former is achieved by imploring rules and axioms or discovery (in form of seed expansion). AeroDAML and Armadillo are examples of semantic annotation platforms that use the technique. The latter has to do with statistical tools to forecast the positions of entities within text. This also includes some induction models. Ont-O-Mat is a clear example of this technique. In any case, a multistrategy semantic annotation platform can still be achieved for optimal performance by combining the two techniques.

[51] classified the technique into three types. They are: informal, formal and ontological annotations as presented by Figure 7 of this paper. However, both informal and formal are not semantic based context. For instance, a handwritten margin annotation in a book is an informal type while, metadata written in formal languages is described as formal type. The only category that has relevance to this paper is the ontological annotation. That is, machine readable annotation written in ontological term such as RDF or OWL.

Annotation technique can be applied in any multimedia contents. However, it is very significance and research effective to define the scope of content to annotate. For instance, if a textual content has to be annotated, scope as; is the annotation about document, paragraph or sentence of

document has to be defined. Semantically, ontology-based annotation of textual content of resources involves two major activities: detection of entities of a resource to annotate and tagging of each entity with most suitable class of domain ontology [3].

Therefore, in this paper; we proposed that ontology-based annotation can be achieved via any proven mechanisms or algorithms include those of pattern and machine learning based that we have earlier reviewed. More also, any chosen mechanism can be manually, semi-automatically or automatically implemented. However, based on the literatures reviewed; semi-automatic and automatic forms of annotations are research trending.

Furthermore, some literatures adopt TF-IDF algorithm for computing weights for entities (concepts) in order to determine their relevance over the seed concept or the query term; in other word, the semantic similarity of the concept(s). The algorithm is not only often used by both query expansion and semantic annotation techniques; it is equally applied by other IR techniques. However, the correctness of setting the arbitrary threshold value for concepts becomes a research issue. For instance, while [12] set 0.4 as threshold value, [59] and [57] set an arbitrary value of t. The research issue for example is that, why 0.4? What consequence does changing the value above or below the set value has on the correctness of concepts? Thus, this forms an open research study. Table 1 presents other criteria for ontology-based information retrieval techniques.

Table 1: Criteria for Ontology-based Information Retrieval Techniques

S/N	Literatures	Information Retrieval Technique.	Algorithm for Weights	Ontology's Domain	Ontology's Tools	Evaluation Technique
1	[57]	Semantic Annotation	TF-IDF	Food, News		Comparison of KBIRS and SBIRS
2	[58]	Semantic Annotation	Ontology Inference Capability	Wine domain	-	TREC approach
3	[38]	Query Expansion	Designed a method	News domain (WNO)	WNO written in XML	TREC approach
4	[41]	Query Expansion	SIRSD Algorithm	Sports domain	RDF, Protégé	Comparison of the semantic search with the traditional search
5	[48]	Ontology techniques	Rules and Inference approach	Citrus	RDFS, Gruff on AllegroGraph	User-based Evaluation
6	[6]	Semantic annotation	-	ICT ontology	OWL2	Expert based evaluation
7	[40]	Query Expansion	-	Rice domain	OWL and AGROVOC CS WB	Expert based evaluation
8	[12]	Query Expansion	Proposed mathematical model	Sports	RDF, Protégé	Expert based evaluation
9	[59]	Semantic Annotation	TF-IDF	Corn	RDF	Comparison of KBIRS and SBIRS
10	[49]	Semantic Annotation	Lavenshtein Algorithm	Updated sports ontology	-	Expert based evaluation
11	[20]	Keyword based Semantic Retrieval Approach	Rules and inference	Soccer	OWL	Expert based evaluation
12	[33]	Query Expansion	Association Rule Mining	Agriculture ontology	Protégé, OWL files	User based evaluation

As clearly presented by Table1, to develop an ontology-based IR system, the following criteria are fundamental.

i. **Ontology:** ontology may be created from scratch or reuse an existing one. [40]; [20]; and [48] created plant production (rice), soccer and citrus ontologies respectively. On the other hand, [58]; [57] and [49] reuse wine, several domain (such as food, news) and sports domain (mapped and updated) ontologies respectively. It is also important not to overlook the indispensible nature of some generic ontologies such as WordNet – a linguistic based database, AGROVOC – an agriculture based database [15]. In order to improve the results return, IR technique can reformulates user's query using wordNet and domain ontology [41]. Similarly, it can be used to take care of concepts' synonyms [12]; [49].

ii. **Ontology's Tools:** ontology is being designed with the aid of some tools. For instance; knowledge representation languages such as OWL [62], RDF/S [63], DAML+OIL [64], XML [2], XTM [65]; and editors such as Protégé [66], AGROVOC CS WB [67], AllegroGraph [48] and TopBraidComposer [68]. Some of these editors like protégé have supportive technologies for important functions. For

example, reasoner modules (pellet, hermit, fact++), JENA API and SPARQL query language.

iii. **Information Retrieval Techniques:** most often based on the literatures reviewed, the capability of ontology for effective information retrieval is further enhanced by incorporating the IR technique such as query expansion or semantic annotation.

iv. **Weights Algorithm:** a certain numbers of related terms to the initial query term would be returned from the document (ontology) under consideration and the top terms are most often ranked as most relevant. Therefore, mechanisms are being put in place to compute and assign weights (values) based on the set criteria as in the case of automatic query expansion or semantic annotation. These mechanisms are what we regarded as weights algorithms in this context. In this paper, while some literatures implored the capabilities of ontology in terms of rules and axioms [20]; [58]; [48], others adopt existing algorithms such as TF-IDF, Lavenshtein Algorithm and Association Rule Mining. On the other hand, some literatures formulated their own such as [41].

v. **Evaluation Technique:** most often metrics for evaluation are precision, recall, f-measure, precision at 10 (P@10) and mean average precision (MAP). Conversely, from the literatures reviewed, there is no single and standardized technique of evaluating the performance of the developed ontology-based information retrieval techniques. However, the following techniques are frequently used.

- TREC approach: the approach is the IR community's yearly evaluation event. It is split into various tracks for various research interests [69]; [70].
- Comparison of the semantic search with the traditional search, in other words, comparison of semantic based information retrieval system (SBIRS) and keyword based information retrieval system (KBIRS) as in the case of [57].
- User based or Expert based evaluation. System evaluation can still be carried out based on some datasets validated by expert(s) of a domain as in the case of [40]. This type of evaluation technique may still be advanced further by establishing comparison with traditional search based on the available datasets as in the case of [20]. The authors established comparison with traditional search and four lucene indices for detail comparison.

6. SUMMARY AND CONCLUSION

This paper reviewed majorly query expansion and semantic annotation techniques of ontology-based information retrieval techniques. Other techniques reviewed in the paper include keyword matching indexing system and O-A-V model. The reviews for the two major techniques were carried out based on generic (wordNet) and domain specific (such as agriculture, sports) ontologies differently.

Query expansion is a prominent technique in ontology based information retrieval. The technique has different classifications forms according to various literatures reviewed but related although with some level of ambiguity. While on one hand query expansion is classified as manual, automatic and interactive; on the other hand, it is classified as local analysis and global analysis. Similarly, another literature classified it as automatic query expansion with local and global techniques as its sub categories and manual query expansion having relevance and manual thesaurus under it. Syntactical approach using statistic analysis and Semantic approach (ontology and thesaurus) is yet another classification of query expansion from a literature in this review paper. Consequently, in this review we revised the technique classification form into: Local approach and Global approach. Local approach includes relevance feedback and pseudo-relevance feedback. Ontologies, generic or domain specific are classified as global approach of query expansion technique. And any of the approach can be implemented manually, semi-automatically, automatically.

Semantic annotation is yet another major technique reviewed in this paper. Annotation as a technique can be classified as formal, informal and ontological. The subject matter of this review limits the work to the third category – ontological. In addition, a literature classified the technique as pattern-based which includes rules, discovery; and machine learning-based which also includes induction or probability. Consequently, semantic annotation can be achieved via any proven algorithms for effective information retrieval. The algorithm can equally be implemented manually, semi-automatically or automatically.

More also, the paper examine ontology-based information retrieval literatures based on the set out criteria such as the ontology itself, domain of applications, tools for developing the ontologies, incorporation of information retrieval techniques, various weights computation algorithms and different evaluation techniques for the system.

Therefore, while ontology based query expansion technique aimed to argument and reformulates query term from the given ontology for relevant information returns, semantic annotation technique on the other hand equally sought to return relevant information by attaching data to the given domain of discourse's entities directly prior to the detection of entity. More also, the importance of generic ontologies such as the lexical database (wordNet) and agriculture knowledge database (AGROVOC) have been duly examined however; we concluded in this paper that better results of effective information retrieval can be recorded if systems are hybridized for instance, wordNet and wordWeb for efficient concept's synonyms representation.

In addition, based on the literatures reviewed, OWL and Protégé technologies as ontology's language and editor respectively outweigh other available technologies owing to their potentialities. Again, it is important to mention that even though information retrieval techniques can be implemented in three forms that is, manual, semi-automatic and automatic; the last two forms of implementations are trending. Finally, as earlier stated the correctness of setting the arbitrary threshold value for concepts becomes a research issue for weights computations algorithms because various weighting functions generate different results thus, it is important to note that retrieval performance largely relies on how the weightings have been computed.

7. REFERENCES

[1] Jiewen W., Ihab, I., and Grant, W. 2011. A Study of Ontology-based Query Expansion Technical Report CS-2011-04

[2] Grigoris, A. and Frank-van, H. 2008. A Semantic Web Primer. The MIT Press Cambridge, Massachusetts London, England.

[3] David, S. David, I. and Miquel, M. 2011. Content annotation for the semantic web: an automatic web-based approach. Knowl Inf Syst (2011) 27:393-418

[4] Wache, H., Vogele, T., Visser, U., Stuckenschmidt, H., Schuster, G., Neumann, H. and Hubner, S. 2001 Ontology-based integration of information—a survey of existing approaches. In IJCAI-01 Workshop: Ontologies and Information Sharing, Seattle, WA, pp. 108–117.

[5] Lutz, M. and Klien, E. 2006. Ontology-based retrieval of geographic information. International Journal of Geographical Information Science Vol. 20, No. 3, March 2006, 233–260

[6] Miguel, A. R-G, Rafael, V-G, Francisco, G-S,and Javier J. S-Z. 2014. Ontology-based annotation and retrieval of services in the cloud. *Knowledge-Based Systems* 56 (2014) 15–25

[7] Gruber, T. R. 1993. A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition,* 5(2): 199-220.

[8] Haav, H., and Lubi, T. 2001. A survey of concept-based information retrieval tools on the web. In: 5th East-European Conference, ADBIS 2001. Vilnius, Lithuania.

[9] Chih-Ping W., Paul J-H. H., Chia-Hung, T., Chun-Neng, H., and Chin-Sheng, Y. 2007. Managing Word Mismatch Problems in Information Retrieval: A Topic-Based Query Expansion Approach. *Journal of Management Information Systems* /Winter 2007-8, Vol. 24, No. 3, pp. 269-295.

[10] Manning, C.D., Raghavan, P., Schutze, H. 2008. Introduction to information retrieval. Cambridge University Press.

[11] Sylvie R., Benjamin, D., Mohameth-Francois, S., Jacky, M., Patrick, A. and Vincent, R. 2013. How ontology based information retrieval systems may benefit from lexical text analysis. New Trends of Research in Ontologies and Lexical Resources, Springer, pp.209-230

[12] Rashmi, C., Rayan, G., Robin, S. and Atul, C. 2013. Domain ontology based Semantic Search for Efficient Information Retrieval through Automatic Query Expansion. 2013 International Conference on intelligent Systems and Signal processing (ISSP)

[13] Suriati, A., Li-Hsing, S., and Rafael, B. 2014. Ontology-based similarity for product information retrieval. *Computers in Industry* 65 (2014) 91–107

[14] Rayner, A., Kim, O. C., Patricia, A., Phang, W. S., Tan, L. I., Leow, C. L. and Gan, K. S. (2014). Ontology Based Query Expansion for Supporting information Retrieval in Agriculture. The 8th International Conference on Knowledge Management in organizations, Springer proceedings in Complexity

[15] Rodrigo, B., Olga, F. N. & Ivo, P. J. 2016. Ontology models of the impacts of agriculture and climate changes on water resources: Scenarios on interoperability and information recovery. *Future Generation Computer Systems.*

[16] Amir, Z. and Mourad, A. 2013. A Generalized Framework for Ontology – Based Information Retrieval. 2013 International Conference on Advanced Logistics and Transport. 29 – 31, May.

[17] Songhua, M. and Ling, T. 2013. Ontology-based semantic retrieval for mechanical design knowledge. *International Journal of Computer Integrated Manufacturing*, 28:2, 226-238

[18] Xutang, Z., Xin, H., Xiaofeng, C., and Ting, Z. 2013. Ontology-based semantic retrieval for engineering domain knowledge. *Neurocomputing* 116 (2013) 382–391

[19] Liu, X., Zhang, X., and Li, Z. (2012). A Domain Ontology-based Information Retrieval Approach for Technique Preparation. *Physics Procedia* 25 1582 – 1588

[20] Soner, K., Ozgur, A., Orkunt, S., Samet, A., Nihan, K. C. and Ferda, N. A. 2012. Ontology-Based Retrieval System using Semantic Indexing. *Information Systems* 37, 294–305

[21] Zhanjun L., Victor R. and Karthik, R. 2007. A Methodology of Engineering Ontology Development for Information Retrieval. International Conference on Engineering Design, Iced'07 28 - 31, Paris, France

[22] Godspower, O. E. and Esingbemi, P. E. (2016). Ontology for Alleviating Poverty among Farmers in Nigeria. INFOS '16, May 09-11, 2016, Giza, Egypt

[23] Ruban, S., Kedar, T., Austin, P. R., and Niriksha, S. 2014.An Ontology-Based Information Retrieval Model for Domesticated Plants. International Journal of Innovative Research in Computer and Communication Engineering Vol.2, Special Issue 5

[24] Suresh, P., Mohamed, A. S., and Jens, L. 2014. Ontology Based Data Access and Integration for Improving the Effectiveness of Farming in Nepal. International Joint Conference on Web Intelligence (WI) and Intelligent Agent Technologies (IAI) – Vol. 02 Pages 319 – 326

[25] Antonio, J-Y., Rafael, B-L., and Dietrich, R-S. 2010. Ontology refinement for improved information retrieval. *Information Processing and Management* 46 (2010) 426 – 435

[26] Uthayan, K. R. and Mala, G. S. A. 2015. Hybrid Ontology for Semantic Information Retrieval Model Using Keyword Matching indexing System. *The Scientific World Journal*, Volume 2015

[27] Vishal, J. and Mayank, S. 2013. Ontology Based information Retrieval in Semantic Web: A Survey. I.J. Information Technology and Computer Science, 2013, 10, 62-69

[28] Bhogal, J., Macfarlane, A. and Smith, P. 2006. A review of ontology based query expansion. Information Processing and Management 43 (2007) 866–886

[29] Francesco, C., Massimo D. S., Luca, G., and Paolo, N. 2014. Weighted Word Pairs for query expansion.

Information Processing and Management (2014).

[30] Hang, C., Ji-Rong, W., Jian-Yun, N., and Wei-Ying, M. 2002. Probabilistic Query Expansion Using Query Logs. ACM 1-58113-449-5/02/0005

[31] Carpineto, C. and Romano, G. 2012. A survey of automatic query expansion in information retrieval. ACM Comput. Surv. 44, 1

[32] Maaike, de B., Klamer, S., and Wessel, K. 2016. Knowledge based query expansion in complex multimedia event detection. *Multimed Tools Appl* (2016) 75:9025–9043

[33] Maleerat, S., Pilapan, P., and Nalinpat, P. 2013. An Ontology-Based Query Expansion for an Agricultural Expert Retrieval System. *iiWAS2013*, 2-4 December, 2013, Vienna, Austria. ACM 978-1-4503-2113-6/13/12

[34] Olga, V. and Murat, K. 2007. Query expansion with terms selected using lexical cohesion analysis of documents. *Information Processing and Management 43*, 849-865.

[35] Alejandra, S. N., Salvador-Sanchez, Elena G-B, and Manuel, P. 2011. An empirical analysis of ontology-based query expansion for learning resource searches using merlot and the gene ontology. Knowledge-Based Systems 24, 15.

[36] Jinxi, X. and Croft, W. B. 2000. Improving the Effectiveness of Information Retrieval with Local Context Analysis. ACM Transactions on Information Systems, Vol. 18, No. 1, Pages 79–112.

[37] Phonarin, P., S. N., and Haruechaiyasak, C. 2012. Agrix: An ontology based agricultural expertise retrieval framework. *Advanced Materials Research 403-408*.

[38] Jagdev, B. and Andrew, M. 2013. Ontology Based Query Expansion with a Probabilistic Retrieval Model. 6th Information Retrieval Facility Conference, IRFC.

[39] Kallipolitis, L., Karpis, V., and Karali, I. (2007). World News Finder: How we Cope without the Semantic Web. In: Devedžic, V. (ed.) Proceedings of AIA 2007, pp. 549–221

[40] Aree, T., Asanee, K., Supamard, P. and Uamporn, V. 2009. Ontology Development: A Case Study for Thai Rice. Kasetsart J. (Nat. Sci.) 43 : 594 - 604

[41] Devi, M. U. and Gandhi, G. M. 2015. Wordnet and Ontology Based Query Expansion for Semantic Information Retrieval in Sports Domain. *Journal of Computer Science*, 2015

[42] Meili L., Xiaobing, S., Shaowei, W., David, L., Yucong, D. 2015. Query Expansion via Wordnet for Effective Code Search. SANER 2015, Montréal, Canada, 978-1-4799-8469-5/15/ 2015 IEEE

[43] Walid, M. and Gareth, J. F. J. 2011. A Study on Query Expansion Methods for Patent Retrieval. *PaIR'11,* October 24, 2011, Glasgow, ACM 978-1-4503-0955-4/11/10

[44] Dipasree, P., Mandar, M., and Kalyankumar, D. 2014. Improving Query Expansion Using WordNet. Journal of the Association for Information Science and Technology

[45] Rila, M, Tokunaga, T. and Tanaka, H. 1998. The use of wordNet in information retrieval. Workshop on Usage of WordNet in Natural Language Processing Systems

[46] Anna, W., Grzegorz, P., Robert, B., and Teresa, P-M. 2013. Association between Text and Ontologies. Intell. Tools for Building a Scientific Information SCI 467, pp. 305-321.

[47] Davide, B. Paolo, R., Emilio, S. A. 2005. A WordNet-based Query Expansion method for Geographical Information Retrieval.

[48] Jianwei, L., Li, L. and Xiaoyan, L. 2015. An Integrated, Ontology-Based Agricultural Information System. *Information Development*, Vol. 31(2) 150–163

[49] Tulasi, R. L., Meda, S. R., Ankita, K. and Hgoudar, R. 2017. Ontology Based Automatic Annotation: An Approach for Efficient Retrieval of Semantic Results of Web Documents. Proceedings of the First International Conference on Computational Intelligence and informatics, Advances in Intelligent Systems and Computing 507

[50] Krisztian, B., Jeffrey, D. Antoine, D. and Yusra, I. 2016. Report on the Eighth Workshop on Exploiting Semantic Annotations in Information Retrieval (ESAIR '15). ACM SIGIR Forum, Vol. 50 No. 1 June 2016

[51] Eyal, O., Knud, H. M., Simon, S., Siegfried, H. and Michael S. 2005. What are Semantic Annotations? http://www.siegfriedhandschuh.net/pub/2006/whatissemannot2006.pdf

[52] Lawrence, R. and Hyoil, H. 2005. Survey of Semantic Annotation Platforms. 2005 ACM Symposium on Applied Computing.

[53] Mohamed A. H. T., Mohamed, B. A. and Abdelmajid, B. H. 2013. A new semantic relatedness measurement usingWordNet Features. *Knowl Inf Syst* DOI 10.1007/s10115- 013-0672-4

[54] Sascha, S., Michael, K., Manuel, M., Saikat, M., Alexander, C., Martin, H., and Dorin, C. 2010. Semantic Annotation of Medical Images. Proc. of SPIE Vol. 7628 762808-1

[55] Anne-Marie, T., Stephane, H., and Jean-Yves, A. 2012. Semantic hierarchies for image annotation: A survey. Pattern Recognition 45 (2012) 333-345

[56] Guido, B., Luigi, D. C., Alice, R. and Livio, R. 2014. Learning from syntax generalizations for automatic semantic. J Intell Inf Syst

[57] Thangaraj, M., and Sujatha, G. 2014. An architectural design for effective information retrieval in semantic web. *Experts Systems with Applications* 41 (2014) 8225 – 8233

[58] Celso, A. F., Maria, C. C. and Ana, M. M. 2013. An Ontology Based Reasoning Approach for Document Annotation. 2013 IEEE Seventh international Conference on Semantic Computing

[59] Hong, Q., Liangliang, Z., and Ying, G. 2010. Semantic Retrieval System Based on Corn Ontology. 2010 Fifth international Conference on Frontier of Computer Science and Technology

[60] Dengsheng, Z., Md, M. I. and Guojun, L. 2012. A review on automatic image annotation technique. Pattern Recognition 45 (2012) 346-362

[61] Vijayarajan, V., Dinakaran, M. Priyam, T., and Mayank, L. 2016. A generic framework for ontology based information retrieval and image retrieval in web data. *Human-centric* Computing and Information Sciences 6:18

[62] Bechhofer, S. 2009. OWL: Web Ontology Language. Encyclopedia of Database Systems, pp 2008 – 2009

[63] Jeen B., Michel K., Stefan D., Dieter F., Frankvan H., and Ian H. (2002). Enabling knowledge representation on the Web by extending RDF Schema. *Computer Networks* 39 609–634

[64] Ian, H. (2002). DAML+OIL: A Reason-able Web Ontology Language. International Conference on Extensible Database Technology, EDBT.

[65] Enesi, F. A. and Adewale, O. S. 2015. A Mechanism for Detecting Dead URLs in XTM-Based Ontology Repository International Journal of Computer Applications (0975 – 8887) Volume 111 – No 12.

[66] John, H. G., Mark, A. M., Ray, W. F., Williams, E. G., Monica, C., Henrik, E., Natalya, F. N. and Samson W. T. 2003. The evolution of Protégé: an environment for knowledge-based systems development. *International Journal of Human-Computer Studies*, Vol. 58, Issue 1, Pages 89 – 123.

[67] Panita, Y., Dussadee, T., Thanapat, S., Asanee, K., Sachit, R., Margherita, S., and Johannes, K. 2008. The AGROVOC Concept Server Workbench: A Collaborative Tool for Managing Multilingual Knowledge. World Conference on Agriculture Information and IT.

[68] Alatrish, E. S. (2013). Comparison Some of Ontology. *Journal of Management Information* Systems.

[69] Ellen, M. V. and Donna, K. H. 2005. TREC: Experiment and Evaluation in Information Retrieval Computational Linguistic, Volume 32 Number 4.

[70] Miriam, F., Ivan, C., Vanesa, L., David, V., Pablo, C. and Enrico, M. 2011. Semantically enhanced Information Retrieval: an ontology-based approach. *Web Semantics: Science, Services and Agents on the World Wide Web*. *Vol.* 9, Issue 4, 434 – 452.