



Science Forum (Journal of Pure and Applied Sciences)

Journal Homepage: <https://atbuscienceforum.com.ng/index.php/jpas>



Semantic Knowledge Retrieval: A Framework for Enterprise Information System Search

Abdulmajid B. Umar¹, Sani Danjuma¹, Shamsuddeen Muhammad Abubakar²

¹Faculty of Computing, Yusuf Maitama Sule University Kano, Nigeria.

²Faculty of Computing, Federal University Dutse, Jigawa State.

Abstract

In today's competitive global market, knowledge is a vital asset for organizational success, making its effective retrieval crucial. However, traditional information systems often struggle to meet user search demands. Semantic knowledge retrieval techniques address this by capturing the meaning behind user queries and knowledge resources, improving search result relevance. This paper proposes a novel semantic knowledge retrieval framework that processes natural language queries. It employs pre-processing techniques like stop-word removal, stemming, and tokenization to refine input data. Query expansion incorporates synonyms to ensure comprehensive results, while a semantic similarity approach compares the original and expanded queries to prioritize relevance and reduce information overload. Domain ontologies are utilized to accurately interpret and execute user requests. The framework is designed to enhance enterprise information system searches by delivering more accurate and context-aware results. By integrating advanced semantic techniques, it aims to improve knowledge retrieval, supporting better decision-making and innovation in a competitive landscape.

Keywords

Semantic framework,
Knowledge retrieval,
Enterprise information
systems,
Information
Search

1. Introduction

Today, businesses are faced with a stark reality: "anticipate, respond, and react to the growing demands of the marketplace, or perish" (Nah, Lau and Kuang, 2001). Intense competitive environment demands innovative ways for survival and actualization of business success. This results in the realization of the importance of knowledge to organizational sustenance. As it is now, only a few would debate validity of the statement that the world is "knowledge-based" (Umar and Danjuma, 2020). Researchers, academicians and industries therefore consider knowledge as a valuable resource that contributes to sustainable competitive advantage, thereby stressing the need to understand its multifaceted concept and multilayered meanings.

Sarker *et al.* (2005) argue that the development of best practices for managing this complex concept of knowledge is the key to riding on today's "competitive wave". Consequently, effective and efficient use of information technology is described as a necessary component or enabler for managing knowledge (O'Dell and Hubert, 2011). In organizations, these technological enablers are called enterprise information systems. They are popularly known as large-scale software applications that support business processes (Gulla, 2004). Like any other information system, information is accessed from these systems through the utilization of search and retrieval mechanisms. Furthermore, this task becomes even more challenging in EIS since they integrate several organizational systems into a single package. Traditional retrieval mechanisms employed by these systems tend to rely on keywords to match user query terms with stored information for retrieval. This results in information overload, irrelevancy of search result (Umar *et al.*, 2018). Knowledge retrieval has been identified as a means of addressing this issue. In this context,

knowledge retrieval deals with identifying and leveraging technologies that facilitate access to information from KMS. Accordingly, several literatures presented knowledge retrieval models of KM and KMS, with some focusing on particular sub-discipline while others provide a more generic view. Though knowledge retrieval has been a topic of recognition in this direction due to its importance, reviewed literatures lack generic formal framework for enterprise information systems.

Semantic web envisions the provision of machine readable descriptions of data, programs, and infrastructure to provide computers the ability to reflect on them (Domingue, Fensel and Hendler, 2013). This idea and technologies of the semantic web opened an opportunity to solve most of the problems related to keyword search. Ontology as a key component of the semantic technology provides an explicit specification of a conceptualization. This offers integration capability while also providing a set of common conceptualization that is easily understood by all components of the EIS. One of the most worthy of note research focus in semantics is semantic knowledge retrieval. It aims at fulfilling user search request through the utilization of the semantic meaning of the query and stored knowledge to improve precision.

In this paper, we propose a semantic knowledge retrieval framework for the enterprise information system, taking into account the distributed nature of these systems and the need for adequate integration. However, we do not aim to discuss much of semantic integration for EIS as the focus is mainly on the retrieval aspect. The rest of the paper is structured as follows: Section 2 presents a review of literatures and related work, followed by presentation and description of the proposed framework in section 3. Section 4 is a brief discussion of future work, and the conclusion is presented in section 5. Finally, section 6 is a presentation of references.

2. Literature Review and Related Work

2.1 Enterprise Information Systems

Enterprise information system (EIS) is a packaged application that supports and automates business processes and manages business data (Gulla, 2004). They come with pre-implemented and customizable modules that reflect best practice for common business operations (Winnemöller, 2009). Business data from different functional areas are integrated and kept consistent across the organization. These systems were developed to integrate disparate legacy systems within an organization. Studies suggests that EIS provide great benefit to organizations by integrating data, providing opportunities to improve business processes, and unifying information support within organizations. The deployment of an EIS

- Rapid availability
- Sound business practice
- Known and verifiable quality
- Low up-front and overall cost
- Inspect-able documentation
- Available maintenance
- Continual research and update
- Varied support and training

Traditionally, enterprise ecosystem is made of dozens and even hundreds of information systems, depending on the enterprise size. These information systems can be stand-alone, interfaced with some others or tightly integrated with information systems inside or outside the enterprise. They can concern various functional domains of the enterprise as it was mentioned in the earlier sections of this paper.

system can involve considerable business process analysis, employee retraining, and new work procedures (Olson and Kesharwani, 2010).

Indeed EIS's been large socio-technical economic systems, are themselves complex and, therefore, comprise the three fundamental components of physical, decision, and information sub-systems. The situation can even be more complicated in the case of networked enterprises or large manufacturing supply chains made of several business entities (e.g. suppliers, manufacturers, assemblers, retailers, etc.) and which become Systems of Systems (SOSs). A characteristic of enterprise systems is their complexities both in terms of business data and in the way they affect the organization's business. According to Robsen (1997), advantages offered by these systems include:

2.2 Semantic Technologies

The semantic web, designed to be an extension of the www envisages a distributed platform of structured information content, better enabling machines to comprehend documents (Berners-lee, Hendler and Lassila, 2001). It uses resource descriptor framework (RDF) to express meaning in a set of triples represented as 'subject, predicate, and object' of an information content. These are written using XML tags, a natural structural way to describe machine process-able data. Berners-Lee suggests a layered structure for the semantic web, so as to realize the vision in an incremental fashion. Figure 1 shows the layers of the semantic web architecture.

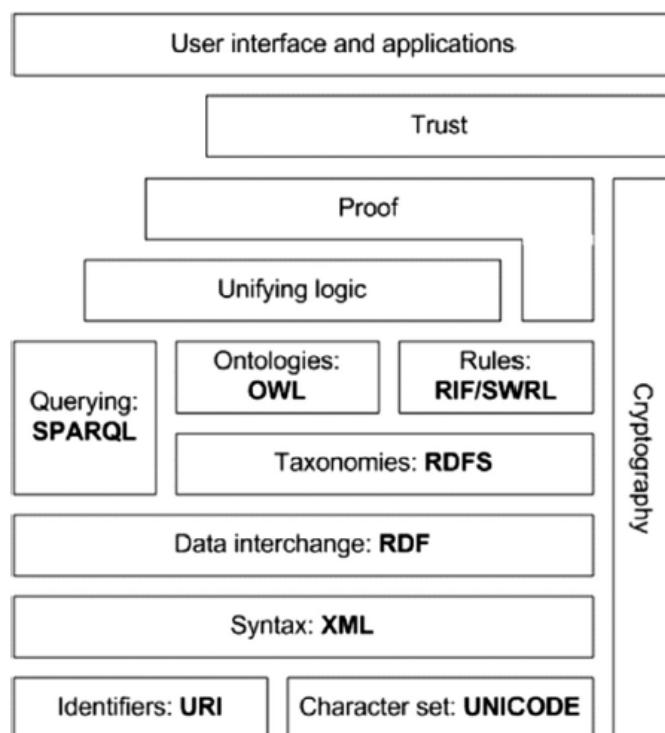


Figure 1: Semantic web architecture

An important component of the semantic web is ontology's, a concept used to formally describe the relationship among terms (Berners-lee, Hendler and Lassila, 2001). Using ontology's, machines could draw inference of the relationship that may exist between concepts. For the formalization of document content to allow for integration and interoperation between distributed data, SW utilizes the resource description framework (RDF). RDF is a graph structured data model that uses the identifier mechanisms previously discussed to permit data integration from disparate sources and enable interoperability (Domingue et al., 2011). Furthermore, an attribute of ontology's that provides a shared specification ensures a level of inter-operability among different knowledge sources. This is considered an important and powerful feature because of its ability to solve

terminology and related problems. It is therefore not surprising that the concept of semantic knowledge retrieval or rather semantic search began to be explored by researchers (Bonino et al., 2004; Kiryakov et al., 2004).

2.3 Semantic Knowledge Retrieval

A very early and popular work in this direction is from (Ding et al., 2002) that uses ontology's in a knowledge system called 'On-to -Knowledge'. An all level ontological approach was used in the architecture which features a web search, information extraction, document exchange, and information maintenance modules over weakly structured data sources. Zidi and Abed (2013) also studied the usage of ontology's to improve information access from transportation systems, by exploiting the defined entities for RDF,

RDFs, and OWL. An entity retrieval model (ERM) based - indexing system was designed to take advantage of the graphical structure of RDF representation of triplets (N-triples) to enhance performance based on the index structure. Similarly, Kara et al. (2012) proposed the idea of using keyword-based querying interface (such as Google's) which they claim are the most user friendly with semantic indexing, for querying of semantically extracted data. They described an ontology-based framework for the extraction and retrieval of information consisting of named entity recognizer (for recognizing and tagging of named entities in narrations) and a two-level lexical analyzer (for the extraction of complete semantic entities based on pre-defined templates).

This section presents the proposed and formulated semantic knowledge retrieval framework for supporting knowledge access in enterprise information systems. The aim of the proposed framework is to improve the access to distributed enterprise knowledge. Furthermore, the framework stands to benefit the EIS users by indirectly reducing the time and effort to find relevant knowledge for organizational business processes wherever and whenever needed, regardless of geographical location of the knowledge seeker. The proposed framework consists of four modules as shown in figure 2, namely; Query module, Knowledge retrieval module, ES services module, Knowledge repository.

3. Semantic Knowledge Retrieval Framework

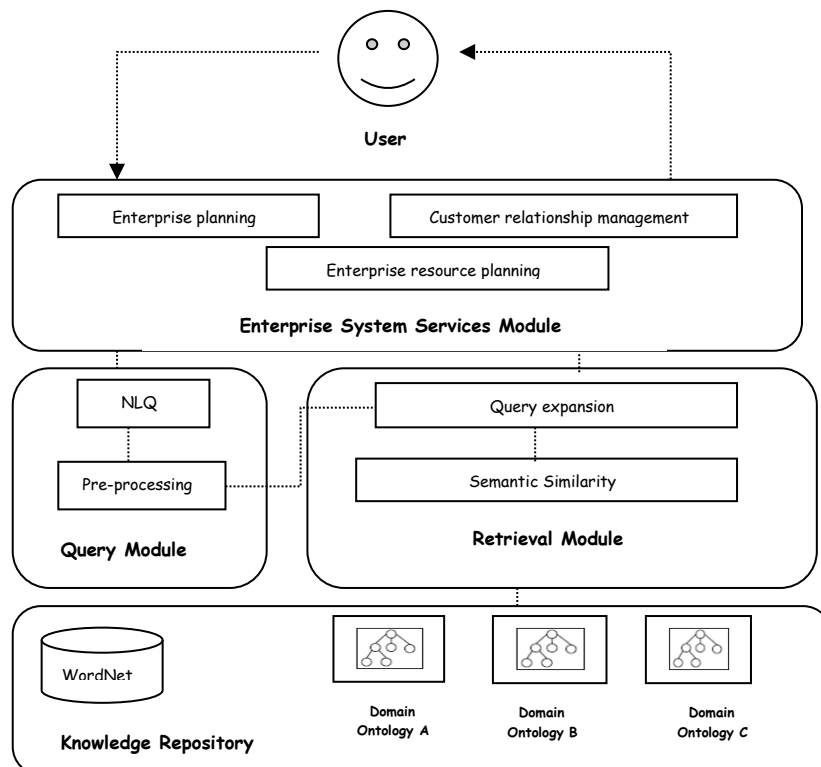


Figure 2: Proposed Semantic Knowledge Retrieval Framework

A. Enterprise System Service Module:

This layer distinguishes the three most common packages included in an enterprise information system, enterprise planning, customer relationship management, and enterprise resource planning. In the proposed framework, it serves as a gateway for users intending to access knowledge from any of the packages initiate the process. Hence, this layer serves as a universal access point to the proposed framework.

B. Query Module:

The layer handles two functions. One is the NLQ component, responsible for accepting natural language query from users. This query is then passed to the next component for pre-processing. At this stage, tokenization is first performed to split the query into pieces (i.e. tokens). This ensures the removal of certain characters (e.g. punctuation). Next is stop-word removal, where words considered to have little or no significant effect to the query are removed (examples of stop-words are "I", "it", "am", etc).

C. Retrieval Module:

This is the heart of the proposed framework. The retrieval module accepts input from the query module and processes it through the knowledge repository. Query expansion is first performed to include synonyms of original query terms into the search log, to increase the completeness of search result. In order to improve precision, similarity of the expanded query is checked against the original query. Were the context or meaning of the original query has been altered due to the

expansion, expanded query is discarded and execution of user query performed.

D. Knowledge repository:

This layer serves as the bedrock of the framework. Here, integration of enterprise's resources is ensured through the use of domain ontology for each package of the enterprise system. Usage of ontology makes communication between several domains easier, and knowledge extraction smatter. WordNet is included in the repository layer to provide a dictionary for query expansion functions.

The complete process outlined in the proposed framework is detailed in figure 3. To provide a simple description of figure 2, a step by step process is presented below:

- *Read NL query*
- *Perform NLP (pre-processing)*
- *Save original query tokens*
- *Expand tokens with synonyms*
- *Compare expanded query with original (similarity)*
- *If context has changed*
 - *discard expansion tokens*
 - *else, replace expanded with original query*
- *Represent in RDF triples*
- *Execute as SPARQL query*
- *Return ranked results*

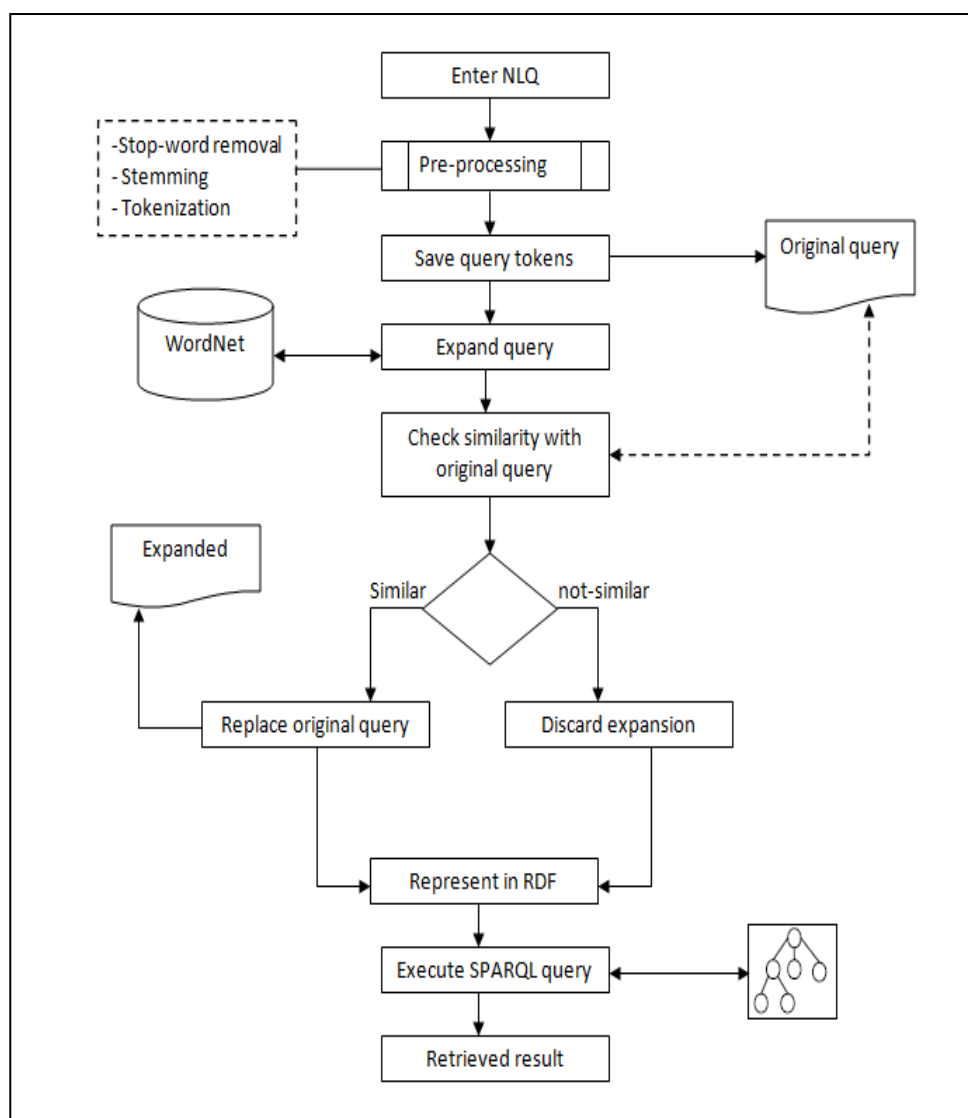


Figure 3: Chart for proposed framework

4. Future Work

Since we presented a framework, pseudo-code and flow chart for semantic knowledge retrieval in enterprise information systems, we aim to extend our work by implementing the proposed framework. Furthermore, issue of evaluation for the performance of our framework is intended to be another research for the future. Particularly, we plan to see how our framework will perform and also compare it with other frameworks to see the resultant effect.

5. Conclusion

In this paper, a framework, pseudo-code and flow chart of semantic knowledge retrieval for enterprise information systems search is presented. The ultimate goal of the research is to support EIS users reach required knowledge within system wherever and whenever needed, and to improve the precision of obtained result for their queries.

6. References

- Berners-lee, T. I. M., Hendler, J. and Lassila, O. R. A. (2001) 'The Semantic Web', *Scientific American*, (May), pp. 1-4.
- Bonino, D. et al. (2004) 'Ontology Driven Semantic Search', *WSEAS Transactions on Information Science and Applications*, 1(6), pp. 1597-1605.
- D. Romero, F. Vernadat, Enterprise information systems state of the art: Past, present and future trends, *Comput. Industry* (2016)
- Ding, Y. et al. (2002) 'The semantic web: Yet another hip?', *Data and Knowledge Engineering*, 41(2-3), pp. 205-227. doi: 10.1016/S0169-023X(02)00041-1.
- Domingue, J., Fensel, D. and Hendler, J. (2013) *Social Semantic Web, Handbook of Semantic Web Technologies*. doi: 10.1017/CBO9781107415324.004.
- Gulla, J. A. (2004) 'Introduction to Enterprise Systems', *Norwegian University of Science and Technology*, 0, pp. 1-42.
- Kara, S. et al. (2012) 'An ontology-based retrieval system using semantic indexing', *Information Systems*, 37(4), pp. 294-305. doi: 10.1016/j.is.2011.09.004.
- Kiryakov, A. et al. (2004) 'Semantic annotation, indexing, and retrieval', *Journal of Web Semantics*, 2(1), pp. 49-79. doi: DOI:%2010.1016/j.websem.2004.07.005.
- Nah, F. F.-H., Lau, J. L.-S. and Kuang, J. (2001) 'Critical factors for successful implementation of enterprise systems', *Business Process Management Journal*, 7(3), pp. 285-296. Available at: <http://www.emeraldinsight.com/doi/abs/10.1108/14637150110392782>.
- O'Dell, C. and Hubert, C. (2011) *The New Edge in Knowledge*. Hoboken, New Jersey: John Wiley & Sons, Inc.
- Olson, D.L., Kesharwani, S.: Enterprise Information Systems: Contemporary Trends and Issues. World Scientific, Singapore (2010)
- Sarker, S. et al. (2005) 'Knowledge transfer in virtual systems development teams: An exploratory study of four key enablers', *IEEE Transactions on Professional Communication*, 48(2), pp. 201-218. doi: 10.1109/TPC.2005.849650.
- Umar, A. B., Jaafar, A., Jabar, M. A., & Murad, M. A. A. Semantic Knowledge Retrieval KMS Components-Validating The Questionnaire Items. *Journal of Computer Science & Computational Mathematics*, 7(3), 85-89.
- Umar, A. B., Ja'afar, A., Jabar, M. A., & Murad, M. A. A. (2018). A success model for semantic technology based-knowledge management systems: an empirical investigation. *International Journal of Engineering and Technology*, 7(4), 151-159.
- Umar, A. B., Danjuma, S. (2020). A Review of Semantic Information Retrieval Approaches for Search Engines in Knowledge Management Systems. *Nigerian Journal of Science and Research*, 1(4).
- Winnemöller, R. (2009) 'Semantic Enterprise Search', *Polibits (Online)*, (39), pp. 11-17.
- Zidi, A. and Abed, M. (2013) 'A generalized framework for ontology-based information retrieval: Application to a public-transportation system', *2013 International Conference on Advanced Logistics and Transport, ICALT 2013*, pp. 165-169. doi: 10.1109/ICAdLT.2013.6568453.