

## **Intelligent Information Retrieval System: A Survey**

**Ankita Sharma**

*Computer Science Department, Banasthali Vidyapith Jaipur  
Jaipur, Rajasthan, India.*

### **Abstract**

In recent years, the rapid use of internet is found in almost every domain therefore it becomes a huge challenge for researchers to retrieve useful information on the World Wide Web. In the traditional Information Retrieval systems users are not able to get semantic description of the information needed by them, so Intelligent Information Retrieval (IIR) systems are used to find out the more relevant information's. In this paper we present a brief survey of Intelligent Information Retrieval Systems and Intelligent agent models based on semantic web and ontology. The performance of such intelligent systems is calculated in terms of Quality of Search, Efficiency, Effectiveness and satisfaction of users according to search results.

**Keywords:** Intelligent information Retrieval, Semantic web, intelligent agent, Ontology, Information Retrieval models, Intelligent Information Retrieval systems.

### **1. Introduction**

The amount of information available on the web grows day by day with the recent development of Internet and web technology. In the web there are more than fifteen billion web pages are available but it is not necessary that we always get good quality of information. To retrieve useful information from bulky websites various links are available on the web, so for this search operation are performed by search engines on the internet like Google, bing, Yahoo, mywebsearch and GoodSearch etc. and there are so many tools are also available on the web for effective search, filtering and retrieve knowledge, retrieve relevant information etc. These tools are performing their work on the basis of keyword search [1] [4] [7] [9]. When a user perform the keyword search in

this case it's very difficult to understand synonyms and polysemy. It is a drawback of existing tools because of its low precision and recall. For this we use Semantic web [18] that provides a solution in the World Wide Web discovery and organizing information.

An IIR [5] is one of the machine learning programs that were used to retrieve the data effectively which was doing earlier by the users manually. IIR provides the high efficiency and personalizing in the recent development of internet. Searching the relevant information has become the most typical task either in public or private networks. So, there were so many challenges came during crawling, indexing of web pages and all these challenges are now days handled by the web search engines. Users are not familiar with search engines as well as query interface and Natural languages queries are used to handle uncertainty, incompleteness etc. Search engine assume that web is a connected graph of nodes and nodes represent the hyperlinks among the web pages. For searching the information on web it uses Meta-search algorithms [1]. The resources are available over the web using the Resource Description Framework (RDF), selection of resources on the basis of the query and then it combine the result came from different search engines. In the resultant set it maintains the documents according to sorted order of relevance with respect to user query.

IIR challenges need to handle access over the internet, heterogeneity, distributed, collaborative and context sensitive retrieval. IR requirement that it must be provide the information for insuring the quality of search in IIR. IR challenges in IIR related to Query representation from user specification, clustering and indexing, classification, Question answering issues, meta-search, distributed information retrieval, matching, ranking the result by relevance, language modeling, performance measure and user feedback in terms of recall and precision. The objective of this article is to provide the brief survey of IIR systems and Intelligent Agent models. We have divided this paper into five sections. In section 1) we introduce a brief introduction of the IIR System, need of IIR system, Searching mechanism, Drawback of keyword search. In section 2) we have given the brief introduction about the IIR Systems and models. In section 3) we describe some future work for semantic web IIR systems and how can we design an effective algorithm to retrieve relevant information on the web.

## **2. Related Work**

In the previous technologies we are not able to provide the specific search results according to user's interests. Mostly searching mechanisms are based on keyword search and the content cannot meet the requirement of IR and these searching mechanisms are not providing the information according to preferences or interests.

Author [10] implements an intelligent multi-agent model which is a combination of conceptual graph (CG) and multi agent model. An intelligent multi agent model can work in active and passive mode also to bring semantic information in retrieval and filtering process. Each node of CG are representation of entities, attributes, states and

maintains the relationship between nodes. An author says that the model returns results in much refined way and quite useful when user have static information need.

An author [8] has given the architecture of multi agent system. In this model, the different types of intelligent agents analysis agent, filter agent, feedback agent and search agent are used for searching the information on the web. Intelligent Agent are computer programs that are tries to learn patterns and the behavior of user and then act of behalf of user. An author ensures that system is much more accurate as compare to traditional search engine.

An author [4] implements search engine intelligent framework based on semantic web and it is a unique identifier for every resource on the WWW. It is used to handle various types of links available resources on the web which handled by computers. This system is humanistic and intelligent and improves the effectiveness of system. It improves the precision and recall rate and allows users to perform the search operation with natural language. An author ensures that retrieval system is more accurate as compare to traditional IR systems.

An author [7] uses semantic query interface and IR server based on the ontology and multi agent. These multi Agent technologies belong in the field of Artificial Intelligence (AI). Multi-agent system [8] in the WWW maintains intelligent agents and Semantic web oriented perfect IIR system. In the framework of ontology it uses Ontology Web Language (OWL), Simplified Object Access Protocol (SOAP). Semantic web oriented IIR system maintains ontology repository is implemented by using clustering algorithm and inference rules in the systemic architecture. An author ensures that the retrieved information can be improved by knowledge maps.

An author [11] implements a Software Reuse System (SRS) based on semantic web and SRS are used to store the information. Future reuse knowledge maintains in the artificial intelligence laboratory (AILab). In SRS semantic web uses RDF, RDF Schema (RDFS) and OWL to represent the knowledge in the knowledge repository. SRS stores the information in the form of Software Development Knowledge Element (SDKE). An author provides an environment in which reuse and sharing of information is possible and it's also provide a mechanism for retrieving relevant information using the user's context.

In this paper an author [9] implements Ontology to describe the semantic information of knowledge base in order to meet the requirement. Ontology can express the knowledge which is too complex and it provides the knowledge sharing and reusing mechanism. Author [2] describes the ontology has the good ability of data sharing, reusing and it understands a domain to overcome differences in technologies. Ontology supports the semantic retrieval and this model is used to improve the search quality.

In the web the available web pages are the combination of text and semantic markup. An author implements [12] a new framework that is marked up statements like DARPA Agent Markup Language and Ontology Interchange Language (DAML+OIL) these are the semantic web language. Author ensures the semantic markup improve the retrieval performance when a web document index, query

processed, retrieval results are returned. Author introduces [14] ontology based IIR system and creates Computer Aided Process Planning (CAPP) application system.

An author implements [3] approach in which extract the information from the web documents. Domain ontology and semantic retrieval approach for retrieving similar words in documents. In this system an author ensures performance increases over other semantic information retrieval techniques.

An author implements [16] a model and an algorithm is implemented by using AI to enhance the efficiency and effectiveness of IR. Search engines containing redundant information and for returning the searching results its take too much time.

**Table 1:** Analysis of IIR models and Systems.

S. No.	Models and Systems	Techniques/Area	Results
1.	An Intelligent Multi Agent Model	Conceptual Graph, Multi Agent model	Intelligent Multi-agent retrieval model
2.	IIR model based on Multi-Agents	AI, KQML/ FIPA-ACL, TCP/IP, WinSock	Technique of IIR based on multi agent
3.	The technology of IIR based on semantic web	AI, OWL, SOAP, RDF, XML, Ontology vocabulary	Search engine framework
4.	Semantic Web Oriented IIR Systems	OWL, Distributed Technologies, SOAP	Biomedical Disease System
5.	A software reuse system based on semantic web	RDF, RDFS, OWL, SDKE	A semantic reuse system
6.	IIR System based on ontology	OWL, Computer Graphics, RDF, RDFS, Jena, SPARQL	Semantic Retrieval system
7.	IR on the semantic web	DAML, OIL, XML, RDFS, SQL query engines	Approach to retrieval of documents
8.	Ontology based IIR system	DAML+OIL, SOAP	IIR CAPP application System
9.	Semantic Retrieval approach for web documents	XML, OWL, AI	Jaundice disease system
10.	IIR System based on three layer agent Structure	AI, Priority Algorithms, Knowledge Library	Three layer agent structures
11.	Comparison of QA system based on semantic web	RDF, Conceptual Graph, AI, URI	Graph Matching algorithm
12.	Extracting and modeling the semantic information	XML, NLA, Semantic document model	NLA techniques

13.	Cognitive Location aware system	Wired/Wireless network, semantic graphs	Cognitive location aware IR semantic matching
-----	---------------------------------	---	---

An author introduces [17] a technique for extracting and modeling the semantic information over the web documents using domain specific ontology and natural language analysis technique. Semantic document model is represented as XML.

Author implements [18] a cognitive location aware system in which intelligent agent operates in both environments it can be wired or wireless networks and retrieve information, then it uses the semantic graphs to classify, organize and match the best suitable information according to users query. Author ensures that the system is faster and accurate as compare to other systems.

### 3. Future work

There are dynamically changes are performed on the web links that are available on the internet and an IIR research has been developed rapidly on the web. In the WWW there are so many further researches is possible and need an improvement, whenever new information is stored and updated. Understanding of user query and returning the right results has become a recent challenge for researchers. This is a task for researchers to identify all the challenges which occur during the selection of relevant documents and classify the IR challenges for handling the complexity of design in the proposed solution.

In the ontology research, semantic document is integrated the various terms synonyms, homonyms and inheritance mechanisms. Ontology systems provide the good mechanisms for reusing the information and data sharing. So there are lots of researches available on ontology based queries and the ontology model must be support semantic based retrieval. For retrieving the relevant information on the web how researchers can design an effective algorithm to retrieve relevant information on the web and how Researcher can do improvement in recall and precision of IIR Systems. In the ontology semantic document model researchers can do their work in the testing and evaluation of find out the semantic web approaches.

The next problem occurs during the communication and sharing mechanisms among the several different patterns. For this researcher can do their improvement in clustering algorithm to mine for personalized the user's patterns and behavior.

### 4. Conclusion

In the World Wide Web (WWW) there are information resources available and for these resources it provides an IIR systems and tools which are used to find more relevant information according to user's interests. We began our discussion of IIR system and its need, searching mechanism, drawback of keyword search. In the semantic web and ontology systems that provides a solution in the World Wide Web

discovery and organizing information, so our paper present a survey on the semantic web oriented system and models. These systems are representing and storing the information Resource Description Framework, Resource Description Framework Schema, Ontology Web Language.

## References

- [1] Mohd Wazih Ahmed, Dr. M. A. Ansari "A survey: Soft computing in Intelligent Information Retrieval Systems," *International Conference on Computational Science and Its Applications*, IEEE 2012
- [2] S. Kalaivani, K. Duraiswamy, "Comparison of Question Answering Systems Based on Ontology and Semantic Web in Different Environment", *Journal of Computer Science* 8 (9): 1407-1413, 2012
- [3] Hany M. Harb, Khaled M. Fouad, Nagdy M. Nagdy, "Semantic Retrieval Approach for Web Documents", *(IJACSA) International Journal of Advanced Computer Science and Applications*, 9, 2011
- [4] Jianguo Jiang, Zhongxu Wang, Chunyan Liu, Zhiwen Tan, Xiaoze Chen, Min Li "The Technology of Intelligent Information Retrieval Based on the Semantic Web" *2nd International Conference on Signal Processing Systems (ICSPS)*, IEEE 2010
- [5] Nicholas J. Belkin "Intelligent Information Retrieval: Whose Intelligence," *Department of Information Studies, University of Tampere*
- [6] LIU Yong-Min, CHENG Shu "Artificial Intelligent Information Retrieval Using Assigning Context of Documents," *International Conference on Networks Security, Wireless Communications and Trusted Computing*, IEEE 2009
- [7] Wenjie Li, Xiaohuan Zhang, Xiaofei Wei, "Semantic Web-Oriented Intelligent Information Retrieval System," *International Conference on BioMedical Engineering and Informatics*, IEEE 2008
- [8] Yi Xiao, Ming Xiao, Fan Jhang "Intelligent Information Retrieval Model Based on Multi-Agents," IEEE 2007
- [9] Pan Ying, Wang Tianjiang, Jiang Xueling, "Building Intelligent Information Retrieval System Based on Ontology" *The Eighth International Conference on Electronic Measurement and Instruments*, IEEE 2007
- [10] Tanveer J. Siddiqui, U. S. Tiwary "Integrating Notion of Agency and Semantic in Information Retrieval multi-agent model", *Proceeding of the 2005 5<sup>th</sup> International Conference on Intelligent Systems Design and Applications (ISDA'05)*, IEEE 2005
- [11] Bruno Antunes, Paulo Gomes and Nuno Seco, "SRS: A Software Reuse System based on the Semantic Web"
- [12] Urvi Shah, Tim Finin, Anupam Joshi, R. Scott Cost, James Mayfield , "Information Retrieval on the Semantic Web"

- [13] Ms. Vandana Dhingra, Dr. Komal Kumar Bhatia, "Towards Intelligent Information Retrieval on Web", *International Journal on Computer Science and Engineering (IJCSE)*
- [14] Wenjie Li, zhilyong, Feng, Yong Li, zhoujun xu, "Ontology-Based Intelligent Information Retrieval System", IEEE 2004
- [15] Tanveer J Siddiqui, "Intelligent Techniques for Effective Information Retrieval ", *ACM SIGIR Forum*, Vol. 40, No.2, December 2006
- [16] Shi-Yi Xie, Jia-Cun Liu, Han Wang , "Research Of intelligent Information Retrieval System based on Three Layers Agent Structure" , *Proceedings of the Second International Conference on Machine Learning and Cybernetics, Xi'an*, IEEE 2003
- [17] Shahrul Azman Noah , Lailatulqadri Zakaria, Arifah Che Alhadi, "Extracting and Modeling the Semantic Information Content of Web Documents to Support Semantic Document Retrieval"
- [18] Eddie C.L. Chan, George Baciuc, s.C. Mak, "Cognitive Location-Aware Information Retrieval by Agent-based Semantic Matching", 8<sup>th</sup> IEEE Int. Conf. of cognitive informatics (ICCI09), IEEE 2009.

