

Construction of Powerful Online Search Expert System Based on Semantic Web

Yasser A. Nada

*Chairman of Department of computer science,
Faculty of Computers and Information Technology,
Taif University - K S A,
y_nada@yahoo.com*

Abstract— In this paper we intends to build an expert system based on semantic web for online search using XML, to help users to find the desired software, and read about its features and specifications. The expert system saves user's time and effort of web searching or buying software from available libraries. Building online search expert system is ideal for capturing support knowledge to produce interactive on-line systems that provide searching details, situation-specific advice exactly like setting a session with an expert. Any person can access this interactive system from his web browser and get some questions answer in addition to precise advice which was provided by an expert. The system can provide some troubleshooting diagnose, find the right products; ... Etc.

The proposed system further combines aspects of three research topics (Semantic Web, Expert System and XML). Semantic web Ontology will be considered as a set of directed graphs where each node represents an item and the edges denote a term which is related to another term. Organizations can now optimize their most valuable expert knowledge through powerful interactive Web-enabled knowledge automation expert system. Online sessions emulate a conversation with a human expert asking focused questions and producing customized recommendations and advice. Hence, the main powerful point of the proposed expert system is that the skills of any domain expert will be available to everyone.

Keywords— *Expert System, Semantic Web, Online Search, XML*

I. INTRODUCTION

The Semantic Web is an extension of the World Wide Web with new technologies and standards that enable interpretation and processing of data and useful information for extraction by a computer. The World Wide Web Consortium (W3C) recommends XML, XML Schema, RDF, RDF Schema and Web Ontology Language (OWL) as standards and tools for the implementation of the Semantic Web. Ontologies work as the main component in knowledge representation for the Semantic Web. It is a data model that represents a set of concepts and the relationships between those concepts within a domain. Building an ontology starting from scratch is not an easy task since it makes heavy demands on time in addition to expert knowledge related to the domain.

Expert System (ES), also called a Knowledge Based System (KBS), is computer application

programs that take the knowledge of one or more human experts in a field and computerize it so that it is readily available for use. It can also be integrated with textual database which can be used for explanation purposes of basic terms and operations to confirm and to reach conclusion in some situations [1].

A challenge for the Semantic Web is enabling information interoperability between related but heterogeneous ontology [2].

One of the most powerful attributes of expert systems is the ability to explain reasoning. Since the system remembers its logical chain of reasoning, a user may ask for an explanation of a recommendation and the system will display the factors it considered in providing a particular recommendation. This attribute enhances user confidence in the recommendation and acceptance of the expert system [3].

Each time we require certain software or product, we searched the internet twice and more, and maybe 10 times to find what we are searching for, that waste our efforts and time. Many times, we did not even found the software, and then we have to go and buy it from the CD's Shop. Also software fall in hundreds of categories, and sometimes we are very confused what to use and what is the exact features of each one of them. We search through the website of that software and find that all the data is crowded and we just can't differentiate between a software and another and we got confused from that.

That paper intends to solve all of these problems, and serve people really by helping them in searching and downloading software. The system will be a web based expert system for searching and downloading software from all types and usability's.

That expert system will offer classified arrangements for software and powerful searching method through the available software, also information (features, backwards, download instructions, installation instructions... etc), the expert system will be implemented as mentioned above as a web based system, we need to go through several steps to accomplish that goal.

II. RELATED WORKS

In [4] the authors presented the characteristics of disaster management is different in different disaster

relief information, knowledge, standardization of operating procedures and the feasibility of the rescue program. Knowledge-sharing and case retrieval is to develop case-based intelligent decision support system facing the most important issue. Disaster rescue command for decision-making, the use of Web Ontology Language that state the information and knowledge of characteristics of the earthquake disaster rescue, a case-based reasoning and logic to describe the rescue planning business processes.

In [5] presented the main ramification of the idea of “data shouldn’t be dumb” on rules languages which is the language has to be simple and able to cover simple linkage situations. More complex applications of rules should be reserved for the intelligent applications that will populate the web. This isn’t to say that research and standardization on more powerful and flexible rule languages aren’t valuable or shouldn’t be done, it’s not just necessary or even desirable for the rule-based infrastructure of the semantic web. Keeping these things separate – the infrastructure from the residents of the web – lets us have a much simpler rule language as part of the web infrastructure.

This paper [6] has proposed two extensions for the current generation of digital libraries. First, authors proposed to use ontology to represent scholarly information in digital libraries, thus making the libraries be able to share and exchange knowledge in the Semantic Web environment. Second, fuzzy theory is employed to process uncertain scholarly information as the forms of fuzzy ontology and fuzzy queries. A general architecture of digital libraries in the Semantic Web environment has been presented and an experimental system has also been developed to verify our ideas and techniques.

In [7] proposed an ontology-based information retrieval model to improve effectiveness of information retrieval. The ontology embedded in the proposal model is a fuzzy taxonomy generated automatically from the documents.

In [8] presented an approach to level one sensor fusion in the context of the Semantic Web has been presented using a Semantic Web Expert System Shell (SWEEXSYS). Our approach allows for the delegation of sensor fusion tasks to agent-based systems like SWEEXSYS. In our approach, we discussed the merit of having Semantic Web enabled data. By making data Semantic Web compliant, an agent may be capable of disambiguating the information; also, additional information that may be needed in the fusion process can be retrieved without human intervention from the Semantic Web by the fusion agent.

III. KNOWLEDGE IN THE EXPERT SYSTEM

Knowledge acquisition, knowledge representation, knowledge storage and knowledge application constitute the main content of expert system’s work progress [9], whose core is knowledge base and inference engine. The level of an expert system is basically determined by its knowledge base.

Expert’s experience and knowledge are stored in the knowledge base, the more complete and true of the knowledge, the higher level of expert system is [10]. Thus whether it possesses plenty of knowledge is the key of expert system.

There are representation techniques such as frames, rules, tagging, and semantic networks which have originated from theories of human information processing. Since knowledge is used to achieve intelligent behavior, the fundamental goal of knowledge representation is to represent knowledge in a manner as to facilitate inference (i.e. drawing conclusions) from knowledge [11, 12].

We got the following software categories that will be written inside the semantic network knowledge representing as follows:

Internet browser Figure (1), Painting program, Downloading program, Recovery Program Figure (2), Compressors, File sharing Figure (3), Security, Video and audio players, E-books reader, Text Editors, Slides viewers Images viewers Cleaning programs

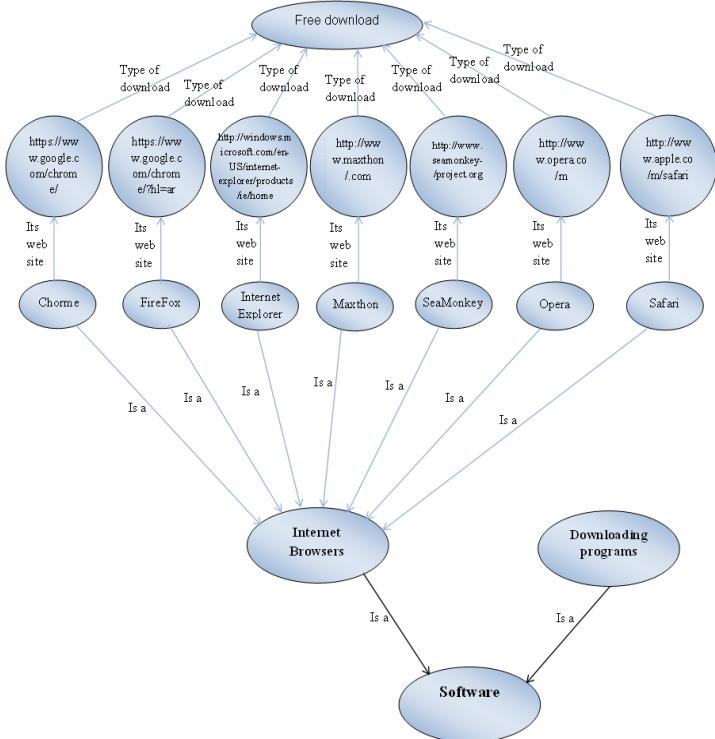


Figure 1: Knowledge Representation for Internet Browser

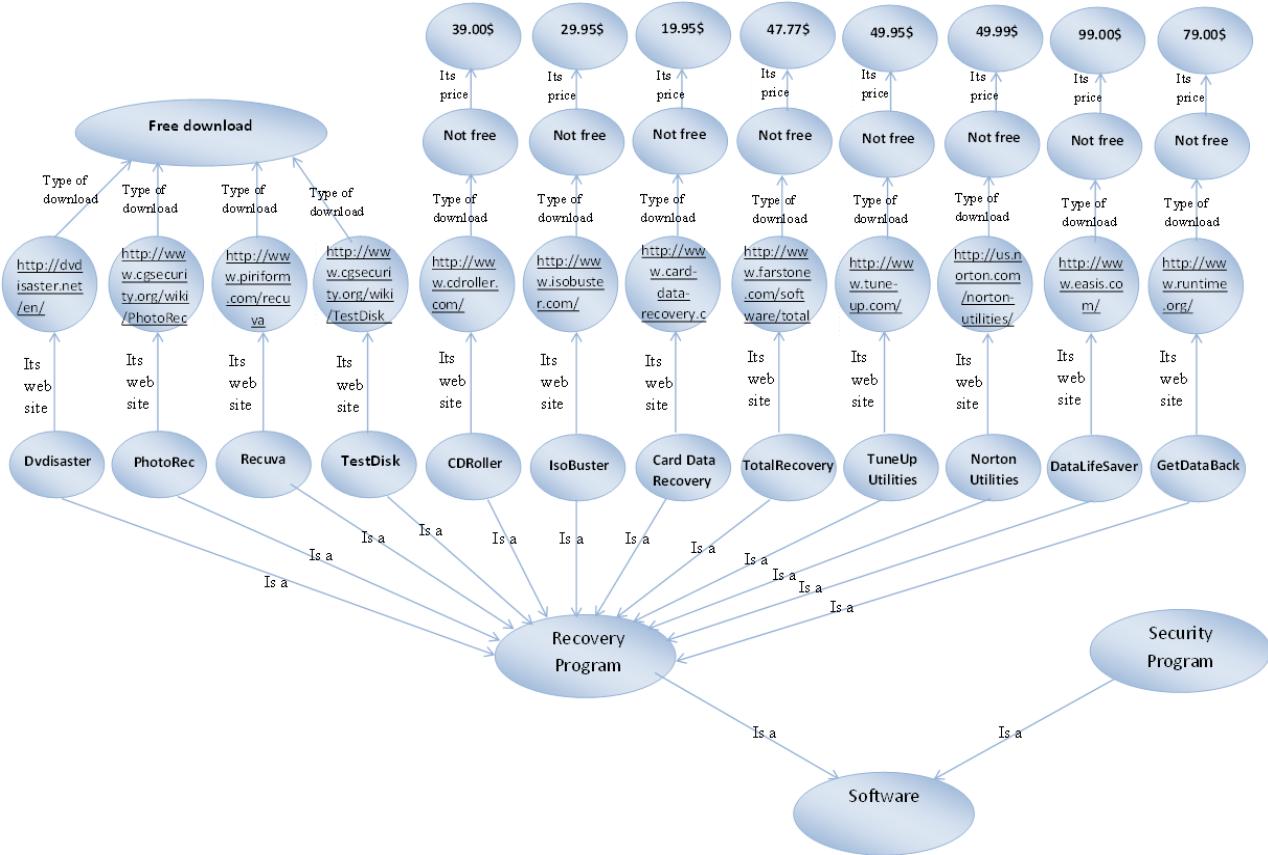


Figure 2: Knowledge Representation for Recovery Program

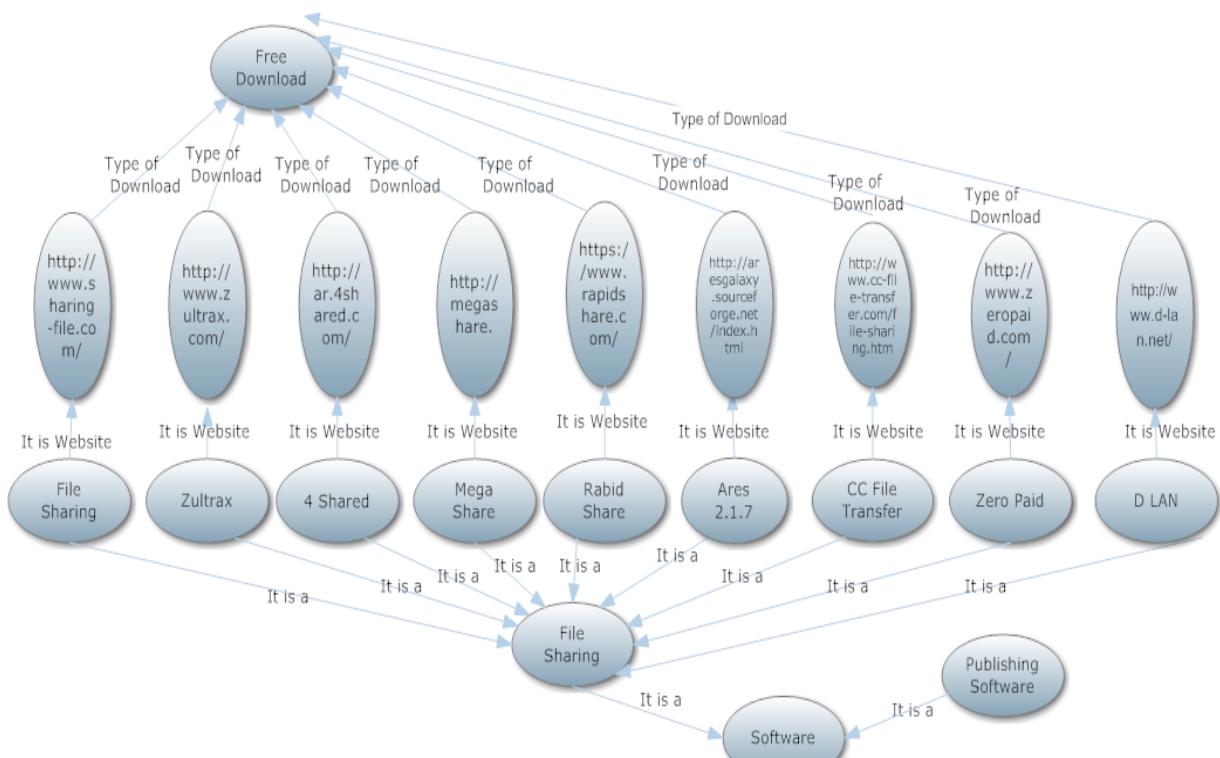


Figure 3: Knowledge Representation for File Sharing Programs

IV. ONTOLOGIES IN SEMANTIC WEB

Ontology is a data model, which can be used to describe a set of concepts and the relationships between those concepts within a domain. Ontology works as the main component in knowledge representation for the Semantic Web. Research groups in both America and Europe developed Ontology modeling languages as The DARPA Agent Markup Language (DAML) and Ontology Inference Layer (OIL). (The W3C Web Ontology Working Group has considered DAML+OIL as the starting point for the introduction of standardized and accepted ontology language for the Semantic Web as Web Ontology Language (OWL) [13]. OWL has three sublanguages: OWL Full, OWL DL and OWL Lite [14]. Existing Semantic Web ontology can be grouped into the following four major categories: metaontology, comprehensive, upper ontology, systematic domain specific ontology, and simple specialized ontology[15].

V. SYSTEM ARCHITECTURE

The system further combines aspects of three research topics (XML, Semantic Web, and Expert System) to facilitate the finding of semantic web services for a client Figure (4).

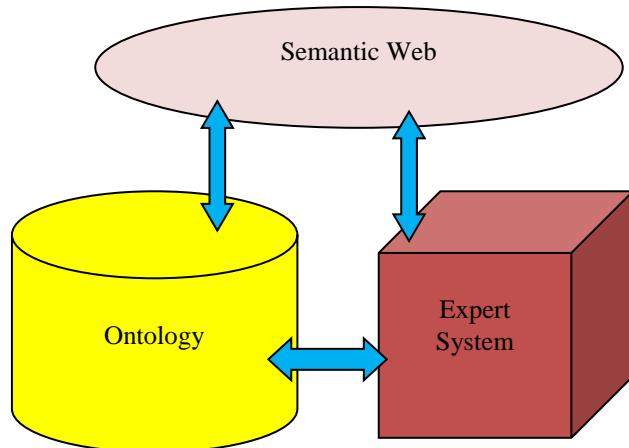


Figure 4: System Architecture

A. XML

Extensible Markup Language (XML) is a set of rules for encoding documents in machine-readable form. It is defined in the XML 1.0 Specifications produced by the W3C, and several other related specifications, all gratis open standards.

The design goals of XML emphasize simplicity, generality, and usability over the Internet. It is a textual data format with strong support via Unicode for the languages of the world. Although the design of XML focuses on documents, it is widely used for the representation of arbitrary data structures, for example in web services.

Many application programming interfaces (APIs) have been developed that software developers use to process XML data, and several schema systems exist to aid in the definition of XML-based languages.

Currently, web services interact by passing XML data, with data types specified using XML Schema. Simple Object Access Protocol [SOAP] can be used as the communication protocol [16], and the I/O signatures for web services are given by Web Services Description Language [WSDL] [17]. UDDI stands for Universal Description, Discovery and Integration [18] and provides the means to publish and discover web services through a UDDI registry.

Today, XML is invading the world of computers and occupying most of its fields. It is widely spreading over the internet, networks, information systems, software and operating systems, DBMS, search tools, web development and services, communication protocols and other fields. As a result, XML data are floating within and between different applications and systems all over the internet and intranets.

A core data representation format for semantic web is Resource Description Framework (RDF). RDF is a framework for representing information about resources in a graph form. It was primarily intended for representing metadata about WWW resources, such as the title, author, and modification date of a Web page, but it can be used for storing any other data. It is based on triples subject-predicate-object that form graph of data. All data in the semantic web use RDF as the primary representation language. The normative syntax for serializing RDF is XML in the RDF/XML form. Formal semantics of RDF is defined as well.

B. Semantic Web Architecture

The architecture of semantic web is illustrated in the Figure (5) below. The first layer, URI and Unicode, follows the important features of the existing WWW. Unicode is a standard of encoding international character sets and it allows that all human languages can be used (written and read) on the web using one standardized form. Uniform Resource Identifier (URI) is a string of a standardized form that allows to uniquely identifying resources (e.g., documents). A subset of URI is Uniform Resource Locator (URL), which contains access mechanism and a (network) location of a document. Another subset of URI is URN that allows identifying a resource without implying its location and means of dereferencing it - an example is urn. The usage of URI is important for a distributed internet system as it provides understandable identification of all resources. An international variant to URI is Internationalized Resource Identifier (IRI) that allows usage of Unicode characters in identifier and for which a mapping to

URI is defined. In the rest of this text, whenever URI is used, IRI can be used as well as a more general concept.

RDFS have semantics defined and this semantics can be used for reasoning within ontology and knowledge bases described using these languages. To provide rules beyond the constructs available from these languages, rule languages are being standardized for the semantic web as well. Two standards are emerging - RIF and SWRL.

For querying RDF data as well as RDFS ontology with knowledge bases, a Simple Protocol and RDF Query Language (SPARQL) are available. SPARQL is SQL-like language, but uses RDF triples and resources for both matching part of the query and for returning results of the query. Since RDFS are built on RDF, SPARQL can be used for querying ontology and knowledge bases directly as well.

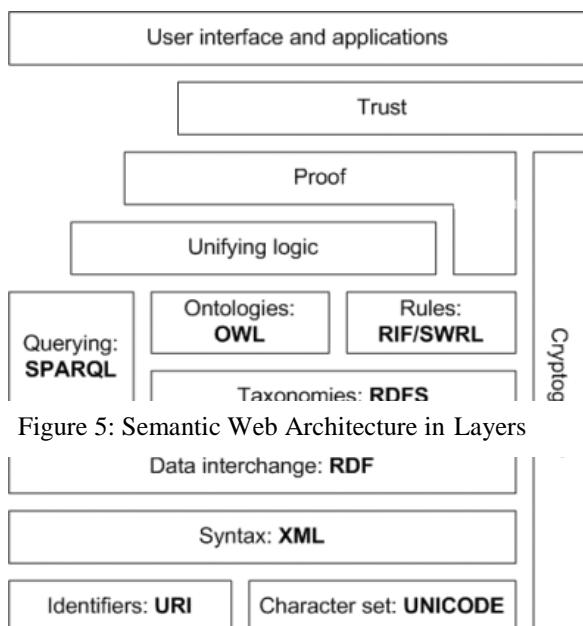


Figure 5: Semantic Web Architecture in Layers

Figure 5: Semantic Web Architecture in Layers
 Note that SPARQL is not only query language; it is also a protocol for accessing RDF data. It is expected that all the semantics and rules will be executed at the layers below Proof and the result will be used to prove deductions. Formal proof together with trusted inputs for the proof will mean that the results can be trusted, which is shown in the top layer of the figure above. For reliable inputs, cryptography means are to be used, such as digital signatures for verification of the origin of the sources. On top of these layers, application with user interface can be built.

C. Expert System

An expert system is a computer program designed to simulate the problem-solving behavior of a human who is an expert in a narrow domain or discipline. An expert system is normally composed of a knowledge base (information, heuristics, etc.), inference engine (analyzes the knowledge base), and the end user interface (accepting inputs, generating outputs). The concepts for expert system development come from the subject domain of artificial intelligence (AI), and require a departure from conventional computing practices and programming techniques.

One of the most powerful attributes of expert systems is the ability to explain reasoning. Since the system remembers its logical chain of reasoning, a user may ask for an explanation of a recommendation and the system will display the factors it considered in providing a particular recommendation. This attribute enhances user confidence in the recommendation and acceptance of the expert system.

All expert systems are composed of several basic components: a user interface, a database, a knowledge base, and an inference mechanism. Moreover, expert system development usually proceeds through several phases including problem selection, knowledge acquisition, knowledge representation, programming, testing and evaluation.

1. Inference Engine

The entire control and operation of the system are done by the inference engine; that is developed using C#; which handles the knowledge in format of XML to get the result from the XML file (Knowledge base) that stores the knowledge rules. Figure (6) shows the block diagram of inference engine. The main roles of the inference engine are summarized as: It applies the expert domain knowledge to what is known about the present situation to determine new information about the domain. The inference engine is the mechanism that connects the user inputs in the form of answers to the questions to the rules of knowledge base and further continues the session to come to conclusions. This process leads to the solution of the problem. The inference engine also identifies the rules of the knowledge base used to get decision from the system and also forms the decision tree.

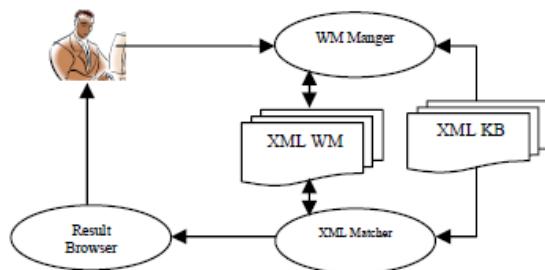


Figure 6: Inference Engine Block Diagram

In this paper the model we took the domain of software is an XML Semantic Web website, we implement the data knowledge based in the form of XML – RDF file to contain all the collected software which we will include in the website, the format of the RDF file as follows:

```
<?xml version="1.0" encoding="UTF-8"?>
<Data>
  - <SoftwareItem>
    <Category>Internet Browsers</Category>
    <Name>Chrome</Name>
    <Link>https://www.google.com/chrome/</Link>
    <Price>Free</Price>
  </SoftwareItem>
  - <SoftwareItem>
    <Category>Internet Browsers</Category>
    <Name>Fire Fox</Name>
    <Link>http://www.mozilla.org/en-US/firefox/new/</Link>
    <Price>20</Price>
  </SoftwareItem>
  - <SoftwareItem>
    <Category>Internet Browsers</Category>
    <Name>Internet Explorer</Name>
    <Link>http://web-development.techweb.com/?kw=sem_ddj_win_goog_</Link>
    <Price>Free</Price>
  </SoftwareItem>
  - <SoftwareItem>
    <Category>Internet Browsers</Category>
    <Name>MaxThon</Name>
    <Link>http://www.maxthon.com/</Link>
    <Price>Free</Price>
  </SoftwareItem>
  - <SoftwareItem>
    <Category>Internet Browsers</Category>
    <Name>Sea Monkey</Name>
    <Link>http://www.seamonkey-project.org/</Link>
    <Price>Free</Price>
  </SoftwareItem>
```

Figure 7: Represented Knowledge in XML.

The knowledge is represented in XML format, because it gives flexibility and being industry standard. Figure (7) shows the part of represented knowledge in XML.

We define 4 attributes:

- Category – Name – Link - Price

VI. RESULTS AND TEST THE SYSTEM

The system was evaluated with different users, including developers, and staff. The system has validated by experts in the domain of online search. Tests of the system were carried out by the developers to make sure the system would work correctly as well as the system is web based system. Figures (8, 9, 10, and 11) shows the snapshots of the developed system.



Figure 8: Home Page

Figure 9: Search Tool Page

All Available Software				
Category	Name	Link	Price	
Recovery Programs	Card Data Recovery	http://www.card-data-recovery.com/	19.95	
Recovery Programs	CD Roller	http://www.cdroller.com/	39.0	
Internet Browsers	Chrome	https://www.google.com/chrome/	Free	
Recovery Programs	Data Life Saver	http://www.easis.com/	99.00	
Recovery Programs	DVDISaster	http://dvddisaster.net/en/	Free	
Internet Browsers	Fire Fox	http://www.mozilla.org/en-US/firefox/new/	20	
Recovery Programs	Get Data Back	http://www.runtime.org/	99.00	
Internet Browsers	Internet Explorer	http://web-development.techweb.com/?kw=sem_ddj_win_goog_Win%7HTML5_adver1_download%20internet%20explorer%209	Free	
Recovery	Iso Ruster	http://www.isohuster.com/	29.95	

Figure 10: All Available Software Page

Figure 11: Search Results Page

VII. CONCLUSIONS

In this paper, an expert system based on semantic web for online search has been proposed. It helps users to find the desired software and read about its features and specifications. The expert system saves users' time and effort consumed in web searching. Users can access and react with the proposed system easily using their web browsers. In addition, the proposed system produces interactive on-line objects that provide searching details and advices exactly like to setup a session with human expert. Furthermore, the proposed system combines three main research topics; Semantic Web, Expert System, and XML. Also, organizations can optimize their most valuable expert knowledge through the proposed powerful interactive Web-enabled knowledge automated expert system. Finally, online sessions which have been constructed by the proposed system emulate a human expert conversation, which receives questions to produce customized recommendations.

REFERENCES

- [1] Khumukcham R., Shikhar Kr. S, "JESS Based Expert System Architecture for Diagnosis of Rice Plant Diseases: Design and Prototype Development", 2013 4th International Conference on Intelligent Systems, Modeling and Simulation, P 674-676, 2013.
- [2] Valerie Cross, Xueheng Hu., "Fuzzy Set and Semantic Similarity in Ontology Alignment", WCCI 2012 IEEE World Congress on Computational Intelligence, Brisbane, Australia June, 10-15, 2012.
- [3] Tanimoto, S. L. , " The Elements of Artificial Intelligence ", Computer Science Press. , 1990.
- [4] Wu Yun, Li Chan, "Semantic Web-based Seismic Disaster Management Expert System", Project 70601011 supported by National Natural Science Foundation of China, 2009.
- [5] Dean Allemang, "Rule-based Intelligence in the Semantic Web ", -or- "I'll settle for a web that's just not so dumb!", Proceedings of the Second International Conference on Rules and Rule Markup Languages for the Semantic Web (RuleML'06), 2006.
- [6] Q. Tho, H. Siu, and C. Tru , "Ontology-Based Fuzzy Retrieval for Digital Library", ICADL 2007, LNCS 4822, pp. 95–98. Springer-Verlag Berlin Heidelberg, 2007.
- [7] C. Been-chian, H. Chih-hung, J. Ming-yi, "Intelligent Information Retrieval Applying Automatic Constructed Fuzzy Ontology", Proceedings of the Sixth International Conference on Machine Learning and Cybernetics, Hong Kong, 19-22 August 2007. 1-4244-0973-X/07, IEEE, 2007.
- [8] Omoju Thomas, David J. Russomanno, "Applying the Semantic Web Expert System Shell to Sensor Fusion using Dempster-Shafer Theory", 0-7803-8808-9/05, IEEE,2005.
- [9] Feng Ding., "Neural Network Expert System", Beijing: Science Press, pp.1-10, 2006.
- [10] Wang Hua, Li Peng-bo. "Fault Diagnosis Expert System of Inertial Navigation System Based on CLIPS and .NET Platform [J]". Journal of Chinese Inertial Technology, Vol.14 No.6, pp.78-80, 2006.
- [11] George F. Luger and William, A. Stubblefield; " Artificial Intelligence and the Design of Expert Systems", The Benjamin/ Cummings publishing Co., Inc., 1989.
- [12] Keith D., "The Essence of Expert Systems", Prentice Hall, 2000.
- [13] G. Antoniou, F. V. Harmelen, "Handbook on Ontologies in Information Systems", Springer-Verlag , pp. 76-92, 2003.
- [14] L. Deborah McGuinness, F. V. Harmelen, "OWL Web Ontology Language Overview", W3C Recommendation , Tech. Rep., 10, Feb, 2004.
- [15] Li Ding, et al., "Using Ontologies in the Semantic Web: A Survey", Department of Computer Science and Electrical Engineering , University of Maryland Baltimore County, Baltimore, Tech. Rep. TR CS-05-07, July, 2005.
- [16] <http://www.w3.org/TR/soap/>
- [17] <http://www.w3.org/TR/wsdl.html>
- [18] <http://www.uddi.org/>

AUTHORS BIOGRAPHY

DR. Yasser Ahmed Nada Was born in Ismailia, Egypt, in 1968. He received the BSc degree in pure Mathematics and Computer Sciences in 1989 and MSc degree for his work in computer science in 2003, all from the Faculty of Science, Suez Canal University, Egypt. In 2007, he received his Ph.D. in Computer Science from the Faculty of Science, Suez Canal University, Egypt. From September 2007 until now, he worked as Assistant Professor of computer science. Chair, Department of computer science, Faculty of Computers and Information Technology, Taif University, KSA. His research interests include Expert Systems, Artificial Intelligence, Object Oriented Programming, Computer Vision, and Genetic.

