

## **An Overview of Web Intelligence**

**M. Adeeb Ghonaimy**

*Professor Emeritus, Faculty of Engineering  
Ain Shams University, Cairo, Egypt*

*Abstract.* This paper gives an overview of Web intelligence which will enable the current Web to reach the Wisdom level by containing Distributed, Integrated, and Active knowledge. In this case it will be capable of performing tasks like problem solving and question-answering. In addition, it will be capable of processing and understanding natural languages. Web intelligence draws results from a number of disciplines like: Artificial intelligence, Information technology. Mathematics and Physics, Psychology and Linguistics.

The paper covers the following topics: Web evolution and architecture, Topics related to Web intelligence, The Deep Web, Semantic computing and the Semantic Web, The Wisdom Web, Precisiated Natural Language.

### **1- Introduction**

Web intelligence (WI) explores the impact of artificial intelligence and other advanced information technology concepts on the current Web <sup>[1]</sup>. WI will allow better search procedures to return more relevant and precise information from the vast amount of knowledge distributed over the Web. The search space will constitute not only the small portion called the Surface Web, but will go beyond that to search what is called the Deep Web.

#### ***Some of the Main Features of WI are***

The Web should be autonomic, capable of automatically delegating its functional roles to other agents. These agents should be capable of communicating with each other through an appropriate Agent

Communication Language <sup>[2]</sup>. The agent population will change dynamically as some may deactivate and others come in. The intelligent Web agents should be capable of using the Problem Solver Markup Language (PSML) or any other variant to specify their roles, settings, and relationship with other services. They should also be capable of processing and understanding natural language. It must understand and correctly judge the meaning (semantics) of concepts. Web agents must also incorporate a dynamically created source of metaknowledge that deals with the relationships between concepts and knowledge constraints. Finally, the intelligent Web can personalize user interactions.

*From the above, WI draws Results from the following Disciplines:*

Artificial Intelligence, Information Technology, Mathematics and Statistics, Psychology, Linguistics, and Physics.

## **2- Web Evolution and Architecture**

### ***2-1 Web Evolution***

Since its inception in 1990, the World Wide Web has been visualized as a distributed repository of knowledge. In order to search this vast knowledge base, search engines, like Google, were used for that purpose. However, it could search only a small portion of the Web, called the “Surface Web”. The “Deep Web” which contained a huge amount of knowledge remained unfathomed. Recently, some start-up companies used concepts from “quantum linguistics” to search the Deep Web. All this made use of a network architecture called “Client Server” model.

When peer-to-peer networks were introduced to allow social interaction, a new type of Web, called Web-2 was introduced and thus social networking became the next major Web application.

Returning to the original Web, sometimes called Web-1, the search made use of key words which did not in all cases result in the appropriate knowledge needed. Therefore, it was essential to resort to the Semantic Web, which necessitated the construction of Ontologies. These ontologies describe concepts in machine readable form and represent an essential component of the Semantic Web.

There are a number of Ontology Languages, like OWL (Ontology Web Language) and many other supporting tools to help in implementing the Semantic Web, sometime incorporated in what is called Web-3.

The next step in Web evolution is to make use of the distributed semantic knowledge in an “active” way to answer questions and solve problems. This requires the use of a number of “intelligent agents” each one specializing in a certain role and taking into account any constraints relevant to the specific problem. In other words, the Web should be equipped with inferencing capabilities that are performed in an autonomous manner. Those intelligent agents should be capable of using the Problem Solver Markup Language (PSML) to specify their roles, settings, and relationships with other services. Also, they should have the ability to process and understand natural languages. Web intelligence comprises many domains, among them: knowledge networks and management, Web agents, Web mining and farming, and distributed inferencing . Sometimes this conceptual, intelligent Web is called Web-4 or the Wisdom Web.

## **2-2 Web Architecture**

We can visualize the Web as comprising a number of levels as follows <sup>[1]</sup>:

(1) *Internet Level*: This is the infrastructure of the Web <sup>[3,4]</sup>. It has evolved from the ARPANET in 1969 and is developing rapidly toward the New Generation Internet. The top 3 challenges of the Internet are: large-scale support for mobility, efficient content distribution, and security <sup>[5]</sup>. There are many research projects around the world to offer appropriate solutions for the above problems. Three of them are: Japan’s New Generation Network, The U. S. GENI (Global Environment for Network Innovations), and the European Union’s FIRE (Future Internet Research and Experimentation). Most of these projects have established a concept design, are expected to develop and verify the basic technology by 2015, and have the network running by 2020.

Other developments are Quantum Networks for which there are a number of working prototypes around the world <sup>[6-9]</sup> and the Quantum Internet which is still a research activity in a number of academic institutions <sup>[10,11]</sup>.

(2) *Interface Level*: Since the Web functions as an interface level for human-Internet interactions, it requires the following: Adaptive cross-language processing <sup>[12]</sup>, personalized multimedia representation, and multimodal data processing capabilities <sup>[13]</sup> that can benefit from the recent interest in multidimensional translation <sup>[14]</sup>. Web-enabled camera phones can enable queries in the form of an image captured by the phone camera <sup>[15]</sup>.

(3) *Original Web Level*: This is sometimes called Web-1, and it still represents a major level. An important application is searching the Web for pages satisfying a number of keywords. The means for doing that is through a number of search engines, *e. g.* Google. These engines return a list of the pages satisfying these keywords, but ranked in such a way that reflects the importance of each page. Here, it is assumed that pages are hyperlinked to each other. Google uses a PageRank algorithm to rank the different pages. This is, in some respects, similar to citation analysis of scientific papers <sup>[16]</sup>. Although billions of pages amount to this surface(or shallow) Web, there is a huge hidden Web from these search engines, that use only static HTML pages. This hidden Web (called Deep Web) contains scientific databases, library catalogs, and phone books <sup>[17-20]</sup>. A separate section will be devoted to the Deep Web.

(4) *The Social Web Level*: This is sometimes called Web-2. Briefly, this is a Web in which people can contribute as much as they consume. The media coverage concentrated on blogs, video sharing and podcasting. Social networking websites let users build social connections with family, friends, and coworkers <sup>[21]</sup>. In the context of the social Web, user data is composed of three types of information: Identity data, social-graph data which represents who I know on the social Web, and Content data which represents what I have on the social Web <sup>[21]</sup>. The education system can benefit tremendously from social Web services, where students, teachers, administrators, and parents will be more tightly connected.

The scientific community is also making a transition to Web -2 creating what some call Science2.0 <sup>[22]</sup>. An example of that is called OpenWetWare project at MIT where two biological engineering laboratories cooperated together as a wiki (a Web site that can be edited by anyone who has access) and uses the same software as the online encyclopedia Wikipedia (Website-1). In 2007, the National Science

foundation launched a five-year effort to transform this platform into a self-sustaining community independent of its current base at MIT. Other efforts are: the Nature Network which is an online network for scientists to discuss scientific news and events (Website-2), Science Commons which is an online project to aid open-access science on the Web (Website-3).

In the literature, there are some studies on Social Networks and Social Networking that deal with the design, development, and study of social technologies at the level of individuals, groups, and organizations<sup>[23]</sup>. Also, with the development of technologies to locate individuals as they do their daily activities, a new class of location-aware information systems that link people-to-people-to geographical places (called P3 systems) is being created. Such systems can strengthen the relationship between social networks and physical places<sup>[24]</sup>.

The Web-2 technology has also been applied to model Patient-Centered health informatics application. This allows patients to participate in the health care system by sharing qualitative and quantitative information about their care plans, diagnosis, medications, and other relevant information<sup>[25]</sup>.

**(5) *The Semantic Web Level:*** This is sometimes called Web 3.0. In general, the Semantic Web extends the current Web so that information has a well-defined meaning. It has to provide a language that expresses both knowledge and rules for reasoning about knowledge. The Semantic Web consists mainly of three components: XML, RDF, and Ontologies.

The XML language is designed to make information self-describing. It is a metalanguage that enables exchange of information not only between different computer systems but also across national and cultural boundaries since it relies on the Unicode standard<sup>[26]</sup>. In conjunction with Extensible Stylesheets Language (XSL), it is possible to reformat XML into different devices, thus achieving a “write once and publish everywhere”. XML allows users to add arbitrary structures to their documents but says nothing about their meanings.

Meaning is expressed by RDF Resource Description Framework which encodes it in a set of triples, each triple is like the subject, verb, and object of an elementary sentence, These triples can be written using

XML tags, subjects and verbs are each identified by a Universal Resource Identifier (URI) <sup>[27]</sup>.

For the Semantic Web to function, access to structured collections of knowledge, called Ontologies is essential. Ontologies are collections of statements written in a language such as RDF that define the relations between concepts and specify logical rules for reasoning about them. The Web Ontology Language (OWL) is thus the language that powers the Semantic Web.

Some vendors, like Oracle in 2005 offered RDF support in its spatial 10.2g database <sup>[28]</sup>.

There are also protocols for appropriate query languages as the SPARQL (The SPARQL Protocol and RDF Query Language).

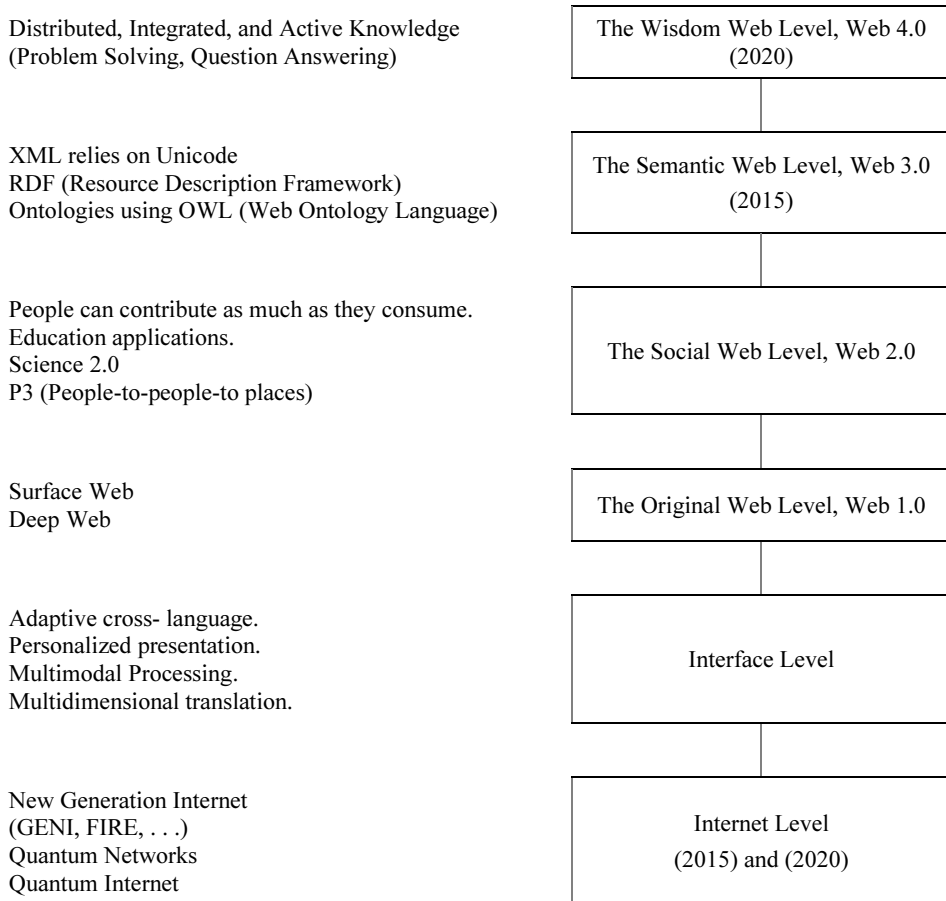
It is possible to look at Web 3.0 as integrating Web 2.0 and the Semantic Web structure as shown in Fig. (1) <sup>[29, 30]</sup>.

<b>Web 3.0</b>	
Web 2.0	Semantic Web (RDF, OWL)
	Linked data (RDF, SPARQL)

**Fig 1. Web 3.0 and its relation to web 2.0 and Semantic Web.**

**(6) *The Wisdom Web Level:*** This is sometimes called Web 4. 0. At this stage we reach the level of Web Intelligence which was indicated briefly in the Introduction. In general, wisdom involves a holistic and integrative understanding of the world. Wisdom is not narrow or specialized knowledge but a broad and a deep knowledge <sup>[31]</sup>. A definition of the Wisdom Web that serves our purpose here is that it deals, with “Distributed, integrated, and active knowledge”. The different ontologies on the Web represents the first property as we have indicated in the previous section. However these Ontologies may be related to different specific domains. If we can integrate some of these ontologies that may cover our domain of interest, then the second property is satisfied. Now we come to the third and most important property, the “active knowledge”. In the context of the Wisdom Web, the distributed and integrated knowledge could be used to solve problems and answer questions. Since knowledge is always updated in a dynamic manner, the solutions and answers could always be updated to reflect the latest state of world knowledge. The ambitious goal could be achieved using

techniques of Web Intelligence. If the Semantic Web and the Web 3.0 research achieve its promise and get stabilized and working by 2015 as some researchers predict, then the Wisdom Web can achieve its promise starting from 2020. It is possible to summarize what we have presented in section (2-2) in Fig. (2).



**Fig. 2. Web Architecture Levels.**

Before leaving this section, we describe some supporting Web services which are a framework of software technologies to support interoperable machine-to-machine interaction over a network <sup>[32]</sup>. Some of these are: SOAP (Simple Object Access Protocol) which transports a message between two points and can include extra information such as routing and the security mechanisms used.

WSDL (Web Services Description Languages) is an XML-based language that provides a description of the message, the protocols used (*e. g.* SOAP), and the address of the Web service.

UDDI (The Universal Description, Discovery, and Integration) specification that is used to quickly find Web services over the Internet. UDDI lists available Web services, gives their description, and provides instructions to use them.

### 3-Topics Related to Web Intelligence

In this section a brief idea is given about some topics related to Web Intelligence <sup>[1],[33-35]</sup>.

(1) **Web Foundations** include the following: Web information description and query languages – The Semantic Web – The Wisdom Web – Web protocols.

(2) **Web Human – Media Engineering** includes: Multimedia information representation and processing-Visualization of Web information-Web based human computer interfacing.

(3) **Web Information Management** includes: Multidimensional Web databases-Multimedia information management-Web knowledge management-Web security, integrity, privacy, and trust.

(4) **Web Information Retrieval** includes: Image retrieval-Multilinguistic information retrieval-Multimedia retrieval-Ontology-based information retrieval.

(5) **Web Agents** include: Dynamics of information sources-Global information collecting-Web-based cooperative problem solving.

(6) **Web Mining and Farming** includes: Data mining and Knowledge discovery-Multimedia data mining-Web-based Ontology engineering-Web farming.

(7) **Web based Applications** include: Business intelligence-Conversational systems-Electronic library-Web-based decision support systems-Web-based distributed information systems-Web-based learning systems.

One of the ambitious goals is to reach human-level intelligence<sup>[36]</sup>. Humans have many remarkable capabilities such as: the capability to



reason, converse, and make rational decisions in an environment of imprecision, uncertainty, incompleteness of information, partiality of truth and possibility. A prerequisite to achievement of human-level intelligence is the mechanization of such capabilities and in particular natural language understanding. The latter necessitates dealing with the precisiation of meaning<sup>[37]</sup>. This will be considered in a separate section.

In order to develop the above applications in an appropriate manner, it was essential to develop the Web Engineering discipline<sup>[38]</sup> which is related to Requirement and Software Engineering but adds the appropriate Web metrics. In addition, it has to take into consideration Security, Legal, Social, and Ethical issues<sup>[39]</sup>.

#### **4- The Deep Web**

The surface Web contains billions of static HTML pages. Two years ago, in 2008, Google had passed a milestone and added the one trillionth address to the list of Web pages it knows about. Beyond this surface Web lies an even vaster hidden Web called the Deep Web behind the forms of searchable databases. It has been estimated that this is 500 times larger than the surface Web<sup>[18]</sup>. These can include financial information, shopping catalogs, flight schedules, medical research and all kinds of other material stored in databases that remain invisible to search engines.

Deep Web searches sometimes use mediated search engines which relies on wrappers that serve as a kind of Rosetta Stone for each data source<sup>[17]</sup>. In order for the Deep Web to be properly searched, its semantic contents should be available for search<sup>[19]</sup>.

An example of a program that provides access to the hidden Web is Deep Query Manager from Bright Planet<sup>[40]</sup>. Google is also presenting a system for Surfacing the Deep Web content. It precomputes submissions for each HTML form and adding the resulting HTML pages into a search engine index<sup>[41]</sup>.

A start-up Company called Dipsie has developed a number of technologies including the Dipsie. bot and algorithms based upon Quantum Linguistics to address problems related to searching the Deep Web. Its programs search a larger portion of the Web as well as analyze

and interpret language in a more human-like manner to derive greater meaning. Based on Quantum Linguistics, its algorithms identify the utility of words and how they interact with, influence and are influenced by one another. Thus, they can predict the semantics of words and phrases within content and also recommend alternative content. Since semantics assumes an important role in Deep Web search, we present a brief account of word semantics and then a brief overview of quantum linguistics and its role in sentence and text semantics in general.

### ***Word Semantics:***

Since meaningful sentences are composed of meaningful words, a logical starting point is to study the semantics of words and their relationships.

A tool that can help in this respect is WordNet<sup>[42-44]</sup>. It has been developed at Princeton University, and has been used together with other lexical operators to improve Web searches<sup>[45]</sup>. WordNet includes the following relations which are summarized in Table (1).

(1) *Synonymy* is WordNet's basic relation, because it uses sets of synonyms (synsets) to represent word senses. Synonymy (syn same, onyma name) is a symmetric relation between word forms.

(2) *Antonymy* (opposing name) is also a symmetric semantic relation between word forms, especially important in organizing the meanings of adjectives and adverbs.

(3) *Hyponymy* (sub-name) and its inverse, *hypernymy* (super-name) are transitive relations between synsets. Because there is usually only one hypernym, this semantic relation organizes the meaning of nouns in a hierarchical structure.

(4) *Meronymy* (part-name) and its inverse, *holonymy* (whole-name), are complex semantic relations. WordNet distinguishes component parts, substantive parts, and member parts.

(5) *Troponymy* (manner-name) is for verbs what hyponymy is for nouns, although the resulting hierarchies are much shallower.

(6) *Entailment* relations between verbs are also coded in WordNet.

**Table (1). Semantic Relations in WordNet.**

Semantic Relation	Syntactic Category	Examples
Synonymy (similar)	N, V, Aj, Av	Pipe, tube Rise, ascend Sad, unhappy Rapidly, speedily
Antonymy (opposite)	Aj, Av, (N, V)	Wet, dry Powerful, powerless Friendly, unfriendly Rapidly, slowly
Hyponymy (subordinate)	N	Suger maple, maple Maple, tree Tree, plant
Meronymy (part)	N	Brim, hat Aluminum, airplane Ship, fleet
Troponymy (manner)	V	March, walk Whisper, speak
Entailment	V	Drive, ride Divorce, marry

Note : N = Nouns, Aj = Adjectives, V = Verbs, Av = Adverbs

With the development of Web 2.0, an increasing amount of user-generated content containing rich opinion and sentiment information has appeared on the Web. Texts containing opinions and emotions are referred to as direction-based texts. Sentiment classification can help determine whether a text contains positive or negative sentiments. SentiWordNet is a lexical resource for sentiment analysis<sup>[46-47]</sup>, and is based on the original Princeton WordNet. The same concepts could be used for other languages.

### ***A Brief Overview of Quantum Linguistics:***

Quantum linguistics is a recent activity that tries to apply quantum mechanics to language problems and in particular, semantics<sup>[48]</sup> in which he relied on the idea of “sign” presented by Ferdinand de Saussure. The sign is composed of “signifier” and “signified” illustrating the duality of symbol and concept. He suggested also a number of postulates for quantum linguistics.

Another domain in which quantum mechanics is applied is Information Retrieval. The pioneering work of Salton<sup>[49]</sup> suggested to represent a document by a vector of terms in the document space which can be considered as an ordinary Euclidean Vector space. Using concepts from quantum mechanics, a document can be represented as a vector in

Hilbert space, and an observable such as “relevance” can be represented by a Hermitian operation. The important notions in quantum mechanics such as: state vector, observable, uncertainly, and superposition translate into analogous notions in information retrieval<sup>[50]</sup>.

Semantic analysis is based on text co-occurrence matrices and data analysis techniques employing Singular Value Decomposition (SVD). Various models provide methods for determining similarity of meaning of words and passages by the analysis of large text corpora<sup>[51]</sup>. In Latent Semantic Analysis one represents words by vectors spanning a finite dimensional space and text passages are represented by linear combinations of such words, with appropriate weights related to the frequency of occurrence of the words in the text. Similarity of meaning is represented by scalar products between certain word vectors.

However, LSA has some problems since it treats a text passage as bag of words in which order is irrelevant. This is a serious difficulty since the syntax is important for evaluating text meaning. As a simple example<sup>[51]</sup>, the sentences “Alice hits Bob” and “Bob hits Alice” cannot be distinguished by LSA.

If we resort to concepts from Quantum Information Theory (QIT) in which a basic object is not a word but a letter with the binary alphabet consisting of 0 and 1 qubits, then the ordering of qubits is obtained by means of the tensor product. Ordering of words can be obtained in the same way. The above reference gives some semantic analysis in quantum notation.

Another research direction is the development of mathematical frameworks that enable us to compute the meaning of a well-typed sentence from the meaning of its constituents. It depends on recasting the Hilbert space formalism in category-theoretic terms to admit a true logic for automating the methods used<sup>[52-53]</sup>. It makes use of the categorical formalism for quantum mechanics introduced by Abramsky<sup>[54-55]</sup>.

The last research direction referred to here is that which demonstrated the presence of quantum structures in language by proving the violation of Bell’s inequality<sup>[51]</sup>. I am not going to give any further details here, but this research gives more evidence for the relation between quantum mechanics and natural language.

## 5- Semantic Computing and the Semantic Web

In Sept.. 2010, the IEEE Computer Society established a technical committee on Semantic Computing. Semantic Computing is in line with Web 3.0 which is characterized by the Semantic Web and the Internet of things, it also includes computing driven by natural language and all computational content such as software, devices, and processes. Semantic computing requires the development of new synergized technologies from natural language processing, data and knowledge engineering, software engineering, computer systems and networks, communication, signal processing, pattern recognition, and other technologies. A number of annual IEEE International Conferences on Semantic computing have already started in 2007. Also, an International Journal of Semantic Computing started to appear in the same year.

Semantic computing tries to match the semantics of computational content and the naturally expressed user intentions to help retrieve, manipulate, or even create content<sup>[56-57]</sup>. The connection between content and the user can be made *via*: semantic analysis, semantic integration, semantic services, service integration, and semantic interface. Figure (3) shows the general architecture of semantic computing adapted from<sup>[57]</sup>.

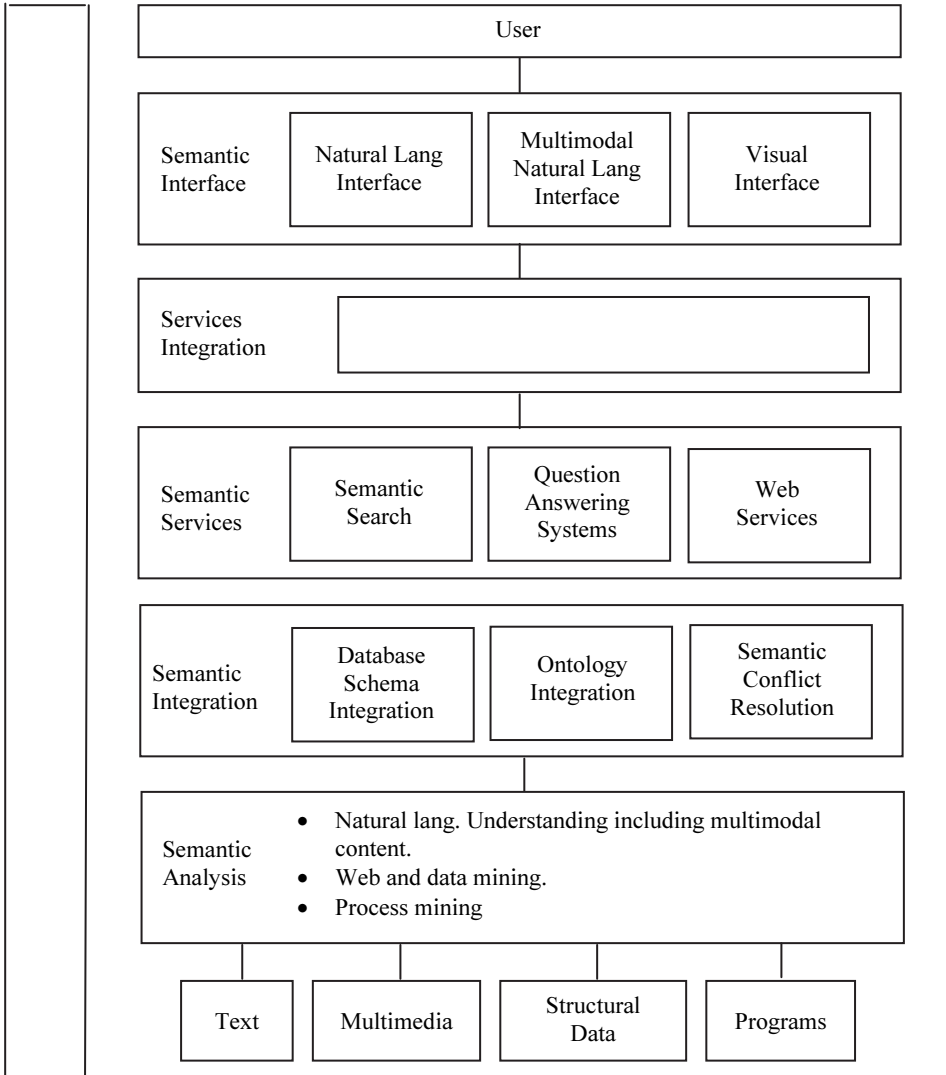
Let us now turn to the Semantic Web and elaborate a little on the presentation given in section 2-2. Figure (4) gives a summary of the Semantic Web layers<sup>[58]</sup>.

A brief exposition is given to terms that were not considered before.

The XML-namespace syntax is used to abbreviate URIs in statements. The XML Schema standard forms a broad base on which developers can build interoperable XML applications. Therefore, XML Schema enables the cross-organizational sharing and verification of documents. The Schema specification consists of two parts: a language to describe the high-level structure of the XML document, and a list of allowable data types that can be used in those documents.

The basic building block in RDF is an object-attribute-value triple, written as A(O,V) that is an object O has an attribute A with value V. Some tutorials are given in<sup>[59-61]</sup>. RDF Schema lets developers define a particular vocabulary for RDF data and specify the kinds of objects to which these attributes can be applied. RDF Schema expressions are also

valid RDF expressions (just as XML Schema expressions are valid XML).



**Fig. 3. Security and Management including Quality of Service.**

In 2004, W3C announced OWL as a standard Web ontology language. It is based on Description Logics which are a family of logic-based knowledge representation formalisms that are descendents of Semantic Networks and other knowledge representation languages, but

that have a formal semantics based on first-order logic. OWL was also based on earlier languages like OIL (Ontology Interchange Language)<sup>[62-63]</sup>. The syntax of description logics consists of:

Description Logics		
Ontology Web Language (OWL)		
Resource Description Framework (RDF)		RDF Schema
XML	Name Space	XML Schema
Unicode	Universal Resource Identifier (URI)	

**Fig. 4. Layers of the Semantic Web.**

A set of unary predicate Symbols that are used to denote concept names.

A set of binary relations that are used to denote role names.

A recursive definition for defining concept terms from concept names and role names using constructors.

Description logics do not make the Unique Name Assumption (UNA) or the Closed World Assumption (CWA). Therefore, they employ the Open World Assumption (OWA). Under OWA, failure to derive a fact does not imply the opposite. This is closely related to the monotonic nature of first order logic which means that adding new information never falsifies a previous conclusion.

The framework for first order logic may be unsuitable for certain situations which require complete knowledge about the world. In this case inference will be non-monotonic, meaning that additional knowledge can invalidate previous conclusions<sup>[64-65]</sup>. Many knowledge modeling constructs are related to CWA and cannot be expressed in first-

order logic. Default rules and constraints depend on CWA. Since many applications require OWA and CWA in parallel, some researchers suggest using Local Closed World (LCW) reasoning. They suggest using Autoepistemic Description Logics (ADL) to achieve that purpose<sup>[66]</sup>. Autoepistemic logic is a formalism concerned with the notion of “knowledge” and “assumption”, *i.e.* to ask what the knowledge base knows or assumes<sup>[67]</sup>.

## 6- The Wisdom Web

Before I give some more information about the Wisdom Web, the following two quotations are given:

Where is the life we have lost in living?

Where is the Wisdom we have lost in knowledge?

Where is the knowledge we have lost in information?

T. S. Eliot, *the Rock*, 1934

To know that we know what we know, and to know that we do not know what we do not know, that is true knowledge”

Nicolas Copernicus (1473 – 1543)

The following definition for wisdom taken from<sup>[31]</sup> is given:

Wisdom is the continually evolving understanding of and fascination with the big picture of life and what is important, ethical, and meaningful. It includes the desire and ability to apply this understanding to enhance the well-being of life, both for oneself and for others.

Many thinkers and philosophers dreamt about that, like H. G. Wells (1866-1946) in his World Brain idea given in his book published in 1938 where he proposed the comprehensive organization of all knowledge<sup>[68]</sup>. Also, Teilhard de Chardin (1887-1955) proposed similar ideas<sup>[69]</sup>.

From the above, it is essential to consider wisdom as combining knowledge with practical applications. So, knowledge is not an end in itself but reasoning about it is useful in understanding protocols in distributed systems since messages can be viewed as changing the state



of knowledge of a system. This is important in cryptography theory, database and knowledge base theories<sup>[70]</sup>.

Therefore, a semantic model for knowledge is needed. This is essential in answering the following questions:

Do we know what facts we know?

Do we know what we don't know?

Do we know only true things, or can something we "know" actually be false?

Sometimes, a possible-worlds semantics is used to model knowledge<sup>[70]</sup>. The idea behind that is that an agent's state of knowledge corresponds to the extent to which he can determine which world he is in. In a given world, we can associate with each agent the set of worlds that, according to the agent's knowledge, could possibly be the real world. Therefore, an agent knows a fact  $\phi$  exactly if  $\phi$  is true in all the worlds in this set, he does not know  $\phi$  if there is at least one world that he considers possible where  $\phi$  does not hold.

The more common axioms that characterize knowledge are summarized in the following<sup>[70]</sup>:

(1) *The knowledge axiom*

$$K_i \phi \Rightarrow \phi, \quad i = 1, \dots, n$$

Which states that only true facts can be known (this is the essential property, that distinguishes knowledge from belief).  $K_i$  means that agent  $i$  knows  $\phi$ .

(2) *Positive introspection axiom*

$$K_i \phi \Rightarrow K_i K_i \phi, \quad i = 1, \dots, n$$

Which states that an agent knows what facts he knows.

(3) *Negative introspection axiom*

$$K_i \phi \Rightarrow K_i \neg K_i \phi, \quad i = 1, \dots, n$$

Which states that an agent knows what facts he does not know.

(4) *Inconsistent facts*

$$\neg K_i (\text{false})$$

Which states that the agent does not know inconsistent facts.

In the following, a brief definition will be given for two terms: common knowledge and distributed knowledge.

*Common knowledge:* In Some situations it is needed to reason about the state of knowledge of a group of agents, *i.e.* we want to reason about facts that everyone in the group knows. At other times we want to add that not only does everyone knows about them, but everyone knows that everyone knows them. These facts are said to be common knowledge.

*Distributed Knowledge:* It is also desirable to reason about the knowledge that is distributed in the group.

For example, if Alice knows  $\phi$  and Bob knows  $\phi \Rightarrow \psi$ , then the knowledge of  $\psi$  is distributed among them, even though it might be the case that neither of them individually knows  $\psi$ . This is called distributed knowledge that corresponds to what a (fictitious) “wise person” would know. Distributed knowledge is a useful notion in describing the total knowledge available to a group of agents in a distributed environment<sup>[70]</sup>.

At this point it may be appropriate to give the following quotation taken from<sup>[71]</sup>:

“Do you think me a learned, well-read man?”

“Certainly” replied Zi-gong. “Aren’t you?”

“Not at all” said Confucious. “I have simply grasped one thread which links up the rest”.

In Section 2-2 (6), we defined the Wisdom Web as the layer that deals with Distributed, Integrated, and Active knowledge. So the Semantic Web with its ontologies should be augmented with extra components in order to reach the Wisdom Web level<sup>[72]</sup>. The first step is to integrate the different distributed ontologies or part of them into wisdom knowledge in the sense that was just explained<sup>[73-75]</sup>.

In this respect, the concept of a knowledge lens may be useful which can synthesize convergent, legitimate perspectives of the desired knowledge while suppressing the irrelevant<sup>[76]</sup>.

To take care of the active component of the wisdom knowledge, we indicated before that a step in that direction is to develop Web services.

Some suggestions to develop coordinated agent-based services was given by<sup>[77]</sup>. To give such services a form of autonomy was considered in<sup>[78]</sup>. Some work was also developed to mining actionable knowledge on the Web as in<sup>[79]</sup>. Distributed problem solving and other activities that need the cooperation of multiple agents were also treated in the literature<sup>[80-81]</sup>. Human-level intelligence is also a research topic that is currently receiving a considerable attention<sup>[82]</sup>.

### 7- Precisiated Natural Language

A prerequisite to having the capability to reason and make rational decisions in an environment of imprecision, uncertainty, and incompleteness of information is to mechanize natural language understanding<sup>[36]</sup>. Concepts of Precisiated Natural Language (PNL)<sup>[83]</sup> and the process of language precisiation is a new approach towards semantic interpretation of natural language<sup>[84]</sup>. It is based on a generalized theory of uncertainty, and language percisiation is a process of adding constraints in order to clearly describe a complicated or ambiguous concept in natural language.

In PNL, the meaning of a proposition,  $P$ , in a natural language may be represented as a generalized constraint of the form:

#### $X \text{ is } r \text{ R}$

Figure (5) illustrates the concept of a generalized constraint. In this figure,  $X$  is the constrained variable,  $R$  is the constraining relation which, in general, is not crisp (bivalent), and  $r$  is an indexing variable whose values define the modality of the constraint.

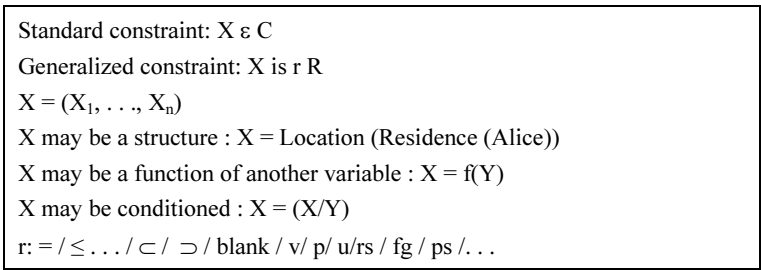


Fig. 5. Generalized constraint.

The principal modalities given by  $r$  are:

Possibilistic ( $r = \text{blank}$ )

Veristic ( $\mathbf{r} = \mathbf{v}$ ).  
 Probabilistic ( $\mathbf{r} = \mathbf{p}$ ).  
 Random set ( $\mathbf{r} = \mathbf{rs}$ ).  
 Fuzzy graph ( $\mathbf{r} = \mathbf{fg}$ ).  
 Usually ( $\mathbf{r} = \mathbf{u}$ ).  
 Pawlak set ( $\mathbf{r} = \mathbf{Ps}$ ).

The set of all generalized constraints together with their combinations and qualifications constitutes the Generalized Constraint Language (GCL)<sup>[37]</sup>.

A concept which plays a key role in PNL is that of a protoform (an abbreviation of “ prototypical form”). Informally, the protoform of a lexical entity such as a proposition, command, question, or scenario is its abstracted summary. For example, the protoform (PF) of

P: Alice is young is A(B) is C,

Where A is the abstraction of age, B is abstraction of Alice, and C is abstraction of young.

Similarly, the protoform of

P: Most Swedes are tall is Count (B/A) is Q

Where A is abstraction of Swedes, B is abstraction of tall Swedes, count (B/A) is abstraction of the relative count of tall Swedes among Swedes, and Q is abstraction of most<sup>[85]</sup>.

Importance of the protoform concept shows the deep semantic structure of the lexical entity to which it applies. So, propositions p and q are PF-equivalent written as PFE(p, q), if they have identical protoforms.

As a simple example, p: Most Swedes are tall, and

q: Few professors are rich, are PF-equivalent.

The above reference gives an application of this approach to question-answering systems.

## 8- Conclusions

Web intelligence explores the impact of artificial intelligence and other advanced information technology concepts on the current Web.

Topics related to Web intelligence include the following: Web foundations including the Semantic and the Wisdom Web-Web human-media engineering-Web information management including security, integrity, privacy, and trust-Web information retrieval-Web agents-Web mining and farming-Web based applications.

The Deep Web search engines are being developed to access information beyond the reach of current search engines. Semantics is a key issue in this respect making use of concepts from quantum linguistics. It extends Latent Semantic Analysis by some ideas from Quantum Information Theory.

The Semantic Web is briefly presented relating it to the new discipline of Semantic computing. Different layers of the Semantic Web are briefly described including the Web Ontology Language (OWL) and the Description Logics forming its theoretical foundation.

The Wisdom Web is also presented together with some basic material including: some knowledge axioms and the definition of the Distributed knowledge which is the foundation for Wisdom knowledge.

Finally, a very brief account for Precisiated Natural Language is presented which models natural language using a Generalized Constraint Language giving some very simple examples.

### ***Acknowledgement***

I appreciate very much the efforts of Professor Hosam Fahmy and Professor Hassan Shehata in this respect.

### **References**

- [1] **Zhong, N., Liu, J. and Yao, Y.** "In Search of the Wisdom Web", IEEE COMPUTER, Nov., PP.27-31 (2002).
- [2] **Bradshaw, J. M.** (Ed). "Software Agents", MIT Press, (1997).
- [3] **Kleinrock, L.** "History of the Internet and its Flexible Future", IEEE Wireless Communication, Feb., pp: 8-18 (2008).
- [4] **Kleinrock, L.** "An Early History of the Internet", IEEE Communications, August, pp: 26-36 (2010).
- [5] **Ortiz, S.** "Internet Researchers Look to Wipe the Slate Clean" IEEE COMPUTER, Jan., pp: 12-16 (2008).
- [6] **Elliot, C.** "Quantum Cryptography", IEEE Security and Privacy, July-August, PP: 57-61 (2004).
- [7] **Elliott, C. et al.,** "Current Status of the DARPA Quantum Network", BBN Technologies, March, pp: 1-12 (2005).
- [8] **Stix, G.** "Best-Kept Secrets", Scientific American, Jan., pp: 64-69 (2005).

- [9] **Curcic, T. et al.**, “Quantum Networks: From Quantum Cryptography to Quantum Architecture”, *ACM SIGCOMM Computer Communication Review*, **34** (5), (October): 3-8 (2004).
- [10] **Lloyd, S., Shahriar, M. S. and Hemmer, P. R.** “Teleportation and the Quantum Internet”, arXiv: quant-ph0003147, (2000).
- [11] **Kimble, H. J.** “The Quantum Internet”, arXiv: quant-ph/0806.4195 (2008).
- [12] **Chung, W.** “Web Searching in a Multilingual World”, *Communications of ACM*, May: 32-40 (2009).
- [13] **Oviatt, S., Darrel, T. and Flickner, M.** “Multimodal Interfaces that Flex, Adapt, and Persist”, *Communications of the ACM*, Jan.; 30-33 (2004).
- [14] **Gottlieb, H.** “Multidimensional Translation: Semantics Turned Semiotics”, *Proc. of the Conference on Challenges of Multidimensional Translation*, pp: 1-29 (2005).
- [15] **Stix, G.** “A Farewell to Keywords”, *Scientific American*, (July): 73-75 (2006).
- [16] **Ma, N., Guan, J. and Zhao, Y.** “Bringing PageRank to Citation Analysis”, *Information Processing and Management*, **44**: 800-810 (2008).
- [17] **Wright, A.** “Searching the Deep Web”, *Communications of the ACM*, (Oct.): 14-15 (2008).
- [18] **He, B., et al.**, “Accessing the Deep Web”, *Communications of the ACM*, (May): 94-101 (2007).
- [19] **Goth, G.** “Reaping Deep Web is a Matter of Semantics”, *IEEE Internet Computing*, (May/June): 7-10 (2009).
- [20] **Handschuh, S., Volz, R. and Staab, S.** “Annotation for the Deep Web”, *IEEE Intelligent Systems*, (Sept/Oct.): 42-48 (2003).
- [21] **Ko, M. N. et al.**, “Social Networks Connect Services”, *IEEE COMPUTER*, (August): 37-43 (2010).
- [22] **Waldrop, M.** “Science 2.0”, *Scientific American*, (May): 46-51 (2008).
- [23] **Churchill, E. F. and Halverson, C. A.** “Social Networks and Social Networking”, *IEEE Internet Computing*, (Sept/Oct): 14-19 (2005).
- [24] **Jones, Q. and Grandhi, S. A.** “P3 Systems: Putting the Place Back into Social Networks”, *IEEE Internet Computing*, (Sept/Oct): 38-46 (2005).
- [25] **Weitzel, M. et al.**, “A Web 2.0 Model for Patient-Centered Health Informatics Applications”, *IEEE COMPUTER*, (July): 43-50 (2010).
- [26] **Bosak, J. and Bray, T.** “XML and the Second Generation Web”, *Scientific American*, (May): 79-83 (1999).
- [27] **Berners-Lee, T., Hendler, J. and Lassila, O.** “The Semantic Web”, *Scientific American*, (May): 28-37 (2001).
- [28] **Lassila, O. and Hendler, J.** “Embracing Web 3.0”, *IEEE Internet Computing*, (May-June): 90-93 (2007).
- [29] **Hendler, J.** “Web 3.0 Emerging”, *IEEE COMPUTER*, (Jan.): 111-113 (2009).
- [30] **Hendler, J.** “Web 3.0: The Dawn of Semantic Search”, *IEEE COMPUTER*, (Jan.): 77-80 (2010).
- [31] **Lombardo, T.** “Wisdom Facing Forward”, *The Futurist*, (Sept.-Oct.): 34-42 (2010).
- [32] **Leavitt, N.** “Are Web Services Finally Ready to Deliver?”, *IEEE COMPUTER*, Nov. PP: 14-18 (2004).
- [33] **Yao, Y. Y. et al.** “Web Intelligence (WI)”. *Proc. Asia-Pacific Conference on Web Intelligence*, (2001).
- [34] **Yao, Y., Zhong, N. and Liu, J.** “Web Intelligence: Exploring Structures, Semantics, and Knowledge of the Web”, *Knowledge-Based Systems*, **17**:175-177 (2004).
- [35] **Zhong, N., Liu, J. and Yao, Y.** “Envisioning Intelligent Information Technologies through the Prism of Web Intelligence”, *Communications of the ACM*, March, PP: 89 – 94 (2007).
- [36] **Zadeh, L.** “Toward Human Level Machine Intelligence-Is it Achievable? The Need for a Paradigm Shift”, *Int. Journal of Advanced Intelligence*, **1** (1), (Nov.): 1-26 (2009).
- [37] **Zadeh, L.** “Precisiated Natural Language (PNL)”. *AI Magazine*, **25**, (3): 74-92 (2004).

- [38] **GINIGE, A.** and **MURUGESAU, S.** “The Essence of Web Engineering”, *IEEE Multimedia*, April-June, PP: 22-25 (2001).
- [39] **DESHPANDE, Y.** and **HANSEN, S.** “Web Engineering: Creating a Discipline Among Disciplines”, *IEEE Multimedia*, April-June, PP: 82-87 (2001).
- [40] **MOSTAFA, J.** “Seeking Better Web Searches”, *Scientific American*, Feb., PP: 51-57 (2005).
- [41] **MADHAVAN, J. et al.** “Google’s Deep-Web Crawl”, *Very Large Data Bases VLDB08*, 24-30 August, PP: 1241 – 1252 (2008).
- [42] **MILLER, G. A.** “*WordNet*” *A Lexical Database for English*, *Communications of ACM*, Nov., PP: 39-41 (1995).
- [43] **FELLBAUM, C.** “*WordNet*”, The MIT Press, (1998).
- [44] **GHONAIMY, M. ADEEB** “A Tutorial on WordNet”, *Proc. The Fourth Conf. on Language Engineering*, Oct., Ain Shams Univ., Cairo, Egypt, PP: 1-28 (2003).
- [45] **MOLDOVAN, D.** and **MILALCEA** “Using WordNet and Lexical Operators to Improve Internet Searches”, *IEEE Internet Computing*, (Jan-Feb.): 34-43 (2000).
- [46] **DANG, Y., ZHANG, Y.** and **CHEN, H.** “A Lexicon-Enhanced Method for Sentiment Classification: An Experiment on Online Product Reviews”, *IEEE Intelligent Systems*, July-August, PP: 46-53 (2010).
- [47] **TUFIS, D.** “*Playing with World Meanings*”, *From Natural Language to Soft Computing: New Paradigms in Artificial Intelligence*, Eds: **ZADEH, L. et al.**, PP: 211-223 (2008).
- [48] **CHEN, J.** “Quantum Computation and Natural Language Processing”, *Ph. D. Thesis, University of Hamburg, Germany*, (2002).
- [49] **SALTON, G.** and **McGILL, M.** “*Introduction to Modern Information Retrieval*”, McGraw-Hill, (1984).
- [50] **VAN RIJSBERGEN, C. J.** “*The Geometry of Information Retrieval*”, Cambridge University Press, (2004).
- [51] **AERTS, D.** and **CZACHOR, M.** “Quantum Aspects of Semantic Analysis and Symbolic Artificial Intelligence”, *Journal of Physics A: Mathematical and General*, **37**,(12) (26 March): 123-132 (2004).
- [52] **COECKE, B., SADRZADEH, M.** and **CLARK, S.** “*Mathematical Foundations for a Compositional Distributional Model of Meaning*”, *Linguistic Analysis* 36, (Ed.) **LAMBECK et al.**, 34 pages (2010).
- [53] **CLARK, S., COECKE, B.** and **SADRZADEH, M.** “A Compositional Distributional Model of Meaning”, *Proc. Conf. on Quantum Interactions, University of Oxford*, PP: 133-140 (2008).
- [54] **COECKE, B.** and **PAQUETTE, E.** “*Introducing Categories for the Practicing Physicist*”, arXiv. Org/0808.1032, 100 pages (2008).
- [55] **COECKE, B.** “*Kindergarten Quantum Mechanics*”, quantum-ph/0510032, Oct., 18 pages (2005).
- [58] **HENDLER, J.** “Agents and the Semantic Web”, *IEEE Intelligent Systems*, March/April, PP: 30-37 (2001).
- [59] **DECKER, S., MITRA, P.** and **MELNICK, S.** “Framework for the Semantic Web: An RDF tutorial”, *IEEE Internet Computing*, Nov./Dec., PP: 68-73 (2000).
- [60] **GHONAIMY, M. ADEEB**, “A Tutorial on the Semantic Web and Ontology Languages”, *Proc. The Fourth Conf. on Language Engineering*, Oct., Ain Shams Univ., Cairo, Egypt, PP: 29-68 (2003).
- [61] **KLAPSING, R., NEUMANN, G.** and **CONEN, W.** “Semantics in Web Engineering: Applying the Resource Description Framework”, *IEEE Multimedia*, April-June, PP: 62-68 (2001).
- [62] **HORROCKS, I.** “*Ontologies and the Semantic Web*”, *Communications of the ACM*, Dec., PP: 58-67 (2008).
- [63] **FENSEL, D. et al.** “OIL: An Ontology Infrastructure for the Semantic Web”, *IEEE Intelligent Systems*, March/April, PP: 38-45 (2001).
- [64] **GENESERETH, M.** and **NILSSON, N.** “*Logical Foundations of Artificial Intelligence*”, Morgan Kaufmann, (1987).

- [65] **Schwarz, G.** "In Search of a "true" Logic of Knowledge: The Nonmonotonic Perspective". *Artificial Intelligence*, **79**, (1), (Nov.): 39-63 (1995).
- [66] **Grimm, S.** and **Motik, B.** "Closed World Reasoning in the Semantic Web through Epistemic Operators", (2006).
- [67] **Levesque, H.** "All I know: A Study in Autoepistemic Logic", *Artificial Intelligence*, **42**, (March): 263-309 (1990).
- [68] **Rossman, P.** "*The Emerging WorldWide Electronic University*", Praeger, (1993).
- [69] **Pelton, J.** "*The Fast-Growing Global Brain*", *The Futurist*, August/September 1999, PP. 24-27.
- [70] **Halpern, J.** and **Moses, Y.** "A Guide to Completeness and Complexity for Model Logics of Knowledge and Belief", *Artificial Intelligence*, **54**, (April): 319-379 (1992).
- [71] **Castells, M.** "*The Rise of the Network Society*", Blackwell Publishers, (1996).
- [72] **Heflin, J.** and **Huhns, M.** "The Zen of the Web", *IEEE Internet Computing*, (Sept./Oct.): 30-33 (2003).
- [73] **Beneventano, D. et al.** "Synthesizing an Integrated Ontology", *IEEE Internet Computing*, (Sept./Oct.): 42-51 (2003).
- [74] **McGuinness, D. et al.** "The Emerging Field of Semantic Scientific Knowledge Integration", *IEEE Intelligent Systems*, (Jan./Feb.): 25-26 (2009).
- [75] **Poole, D., Smyth, C.** and **Sharma, R.** "Ontology Design for Scientific Theories that Make Probabilistic Predictions", *IEEE Intelligent Systems*, (Jan./Feb.): 27-36 (2009).
- [76] **Edgington, et al.** "Adopting Ontology to Facilitate Knowledge Sharing", *Communications of the ACM*, (Nov.): 85-90 (2004).
- [77] **Sycara, K. et al.** "Dynamic Discovery and Coordination of Agent-based Semantic Web Services", *IEEE Internet Computing*, (May/June): 66-73 (2004).
- [78] **Paolucci, M.** and **Sycara, K.** "Autonomous Semantic Web Services", *IEEE Internet Computing*, (Sept./Oct.): 34-41 (2003).
- [79] **Yang, Q., Knoblock, C.** and **Wu, X.** "Mining Actionable Knowledge on the Web", *IEEE Intelligent Systems*, (Nov./Dec.): 30-31 (2004).
- [80] **Drashansky, T. et al.** "Networked Agents for Scientific Computing", *Communications of the ACM*, (March): 48-54 (1999).
- [81] **Guha, R.** and **Lenat, D.** "Enabling Agents to Work Together", *Communications of the ACM*, (July): 127-142 (1994).
- [82] **Beal, J.** and **Winston, P.** "The New Frontier of Human-level Artificial Intelligence", *IEEE Intelligent Systems*, (July/August): 21-23 (2009).
- [83] **Zadeh, L.** "*Web Intelligence, World Knowledge and Fuzzy Logic*", BISC Program, University of California, Berkeley, (2004).
- [84] **Thint, M., Beg, M.** and **Qin, Z.**, "Precisiating Natural Language for a Question-Answering System", 11<sup>th</sup> World Multi Conference on Systems, Cybernetics, and Informatics, 2007.
- [85] **Zadeh, L.** "From Search Engines to Question-Answering Systems-The Role of Fuzzy Logic.", *Progress in Informatics*, No. **1**, PP: 1-3 (2005).

#### Web Sites (WS)

- 1- Website -1, ([www.openwetware.org](http://www.openwetware.org)).
- 2- Website -2, (<http://network.nature.com>).
- 3- Website-3, ([www.sciencecommons.org](http://www.sciencecommons.org)).



## نظرة عامة على ذكاء الشبكة العنكبوتية Web

أ.د. محمد أديب غنيمي

أستاذ متفرغ، كلية الهندسة، جامعة عين شمس، القاهرة، مصر

المستخلص. يعطي هذا البحث لمحة عامة عن ذكاء الشبكة العنكبوتية ويب Web والتي ستمكن ويب الحالية للوصول إلى مستوى الحكمة عن طريق احتوائها على المعرفة الموزعة، المتكاملة، النشطة. في هذه الحالة سوف تكون قادرة على أداء مهام مثل حل المشكلات وإيجاد الحلول والرد على الاستفسارات. وبالإضافة إلى ذلك، سوف تكون قادرة على معالجة وفهم اللغات الطبيعية. إن ذكاء الشبكة نتيجة إلى عدد من التخصصات مثل: الذكاء الاصطناعي وتكنولوجيا المعلومات والرياضيات والفيزياء، وعلم النفس واللغويات.