CHAPTER 3
SEMANTIC WEB AND SOFTWARE AGENTS: THE INESCAPABLE RELATIONSHIP

3.1 INTRODUCTION

In the beginning of sixties, the eminent researchers Allan and Elizabeth F had described the new concept of Semantic Network Model [34]. The central aim of their study was to examine, how human brain uses long-term memory to associate things to find the truth of a context. The main result of their work was they categorized things on the bases of their past experiences and made inferences on the basis of categorization to decide whether the statement is true or false [167]. Later on Tim Berners-Lee had described the concept Semantic Web [23] in 2001. The term Semantic Web implies an Intelligent Web [197] that aims to create machines determine the meaning of information on the web pages rather than simply presenting them to users [23]. The idea is to make World Wide Web intelligent and machine readable by providing tools to find, exchange and interpret information with their intended meaning. It proposed an infrastructure where the information obtainable on the web can be appreciably interpreted. It aims to change web development in such a way that machine can understand of words presented on the web pages, easily relate them and produce information more relevant to the subject at ease of surfing.

Tim Berners-Lee [23] indicated that software agents will serve as the core working component in semantic web, since they are able to draw information with the help of ontology from distributed heterogeneous sources, relate them and draw conclusion. Therefore ontology matching is an essential technique to create communication between agents using dissimilar ontologies [51]. The aim can be realized to its goal deploying intelligent agents in semantic web.

The chapter is structured in three major parts. Initial two parts minutiae from the evolution to latest trends in semantic web and software agents respectively and final
part throws light on the noteworthy relationship between two aforementioned components of intelligent web.

3.2 THE SEMANTIC WEB: BACKGROUND DETAILED

The semantic web is a vision which extracts information from the web and crafts it possible to facilitate machines to identify intricate human queries. It brings the proposal of structuring information accessible across the web in a significant way enhancing search technique and thus resulting user satisfaction. The upcoming sub sections outlined the three appraisal phases of the development of the World Wide Web, namely information web, social web and semantic web.

3.2.1 Evolution of Semantic Web

Web is a vision of global repository of the information. It is most suitable approach for extracting and exchanging information over the internet. Nowadays, web is playing significant role in every area in human life, like research, social networking, education and entertainment etc. Since 1990s to today, structure of web is improving towards with more appropriate technologies and services [191]. It has undergone three generation (see figure 3.1) and is ready to proceed towards its fourth generation.

![Figure 3.1: Generations of Web](image1)

3.2.1.1 Information Web

Web 1.0 was an early generation of conceptual evaluation of the web and earliest generation of web, which is called information web. These days were beginning time of web browser. In 1989, Tim Berners-Lee introduced an idea for creating a universal
hypertext model where any network can access information by using Universal Document Identifier (UDI) [2]. The vision behind his work was to change the documentary concept into information based concept where user can search anything in anytime with less human effort and expenditure. It was foundation stone of information era. Web 1.0 represents the information published on the web in static manner. It is also known as “read-only” web where user could only read information and couldn’t contribute anything. The main advantage of this web was that user could find information through search engine and read them same as news paper. The core protocols used in web 1.0 were HTML, HTTP and URL and they all represented the information in static form. However, this was not appealing part of web because it was first time when people could connect with internet for discovering information. The main objective of web 1.0 was to provide lot of features for business communities for promoting advertisements.

3.2.1.2 Social Web

In 1999, term web 2.0 was introduced by Darcy DiNucci [84] and it came to play in dynamic world in 2004 by Tim O’Reilly with O’Reilly Media Web [205]. It described the second generation of web called social web [203]. It enabled the user to use it as a communication medium for sharing information online with each other. It offered “read-write” and interactive web. The vision of social web was to offer an environment which could enhance the social activities and the result of this is present social networking websites like facebook, twitter, Youtube etc. where all friends connect with each other and share their thought, experiences and journey of life [72]. Web 2.0 describes the information in dynamic way and utilized innovative techniques such as RSS, JSON Script and XML languages. However the major drawback of web 2.0 is provides little characteristics of intelligence communities. Table 3.1 describes web development in timely manner.

<table>
<thead>
<tr>
<th>ALIAS</th>
<th>WEB 1.0</th>
<th>WEB 2.0</th>
<th>WEB 3.0</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Information Web</td>
<td>Social Web</td>
<td>Semantic Web</td>
</tr>
</tbody>
</table>

Table 3.1: Developments in Semantic Web
### PARADIGM SHIFT WITH TIME

<table>
<thead>
<tr>
<th></th>
<th>1990-2001</th>
<th>2002-2009</th>
<th>2010-2020</th>
</tr>
</thead>
<tbody>
<tr>
<td>APPROACH</td>
<td>Read-Only</td>
<td>Read-Write</td>
<td>Read-Write-Execute</td>
</tr>
<tr>
<td>CONTENTS</td>
<td>Static</td>
<td>Interactive</td>
<td>Portable Personal Web</td>
</tr>
<tr>
<td>COMMUNICATION</td>
<td>Information Connection</td>
<td>People Connection</td>
<td>Intelligence Connection</td>
</tr>
<tr>
<td>TECHNOLOGY</td>
<td>HTML, HTTP, URL</td>
<td>XML, RSS, AJAX</td>
<td>OWL, RDF, SPARQL</td>
</tr>
<tr>
<td>EXAMPLE</td>
<td>Netscape</td>
<td>Wiki, Blog</td>
<td>Portable Personal Web</td>
</tr>
</tbody>
</table>

#### 3.2.1.3 Semantic Web

The idea of semantic web as third generation of Web was conceived by Tim Berners-Lee in 2001. Lee had proposed the idea with the intention to convert earlier generations of web into semantic web. Semantic web could process the information both for users as well as machines. It made the trade and construal of information more accurate and relevant. Since then the developments in the technology and services behind web is exponential and internet surfing has now become intelligent. The knowledge in semantic web is represented with the help of graphs and such a graphical notation is termed as semantic network [136]. This graphical information can be categorized into a taxonomic interface of nodes and arcs. The nodes characterize the concepts and arcs describe the relationship between concepts. Inheritance plays a major part in semantic network that allow an object to extract the features of several concepts. Figure 3.2 depicts an example of semantic network.
The semantic network illustrates information in the form of intended graph. The nodes of the graph are highlights several concepts, and the edges of the graph provide connections between those concepts. Here, the taxonomies are “Plant”, “Tree”, “Oak”, and so on. The association relations are “is a”, “has” and “used for”.

The details of semantic web are being presented in the upcoming section.

3.3 ARCHITECTURE OF SEMANTIC WEB

The layered architecture of semantic web [25] describes the development of semantic web from its evolution to its present status. Tim Berners-Lee in 2001 and onwards proposed the said architecture and its various versions (see figure 3.3) but unfortunately these were not considered further for W3C recommendation.
The first version V1 comprised of eight layers whereas version V2 had eleven layers. Later, a third version V3 containing twelve layers was proposed, conquering the drawbacks of earlier versions but still it did not fulfil the requirements of Internet users and version V4, having eleven layers again was presented with the idea to enhance the functionality and efficiency of semantic web. However all these versions had inconsistencies and discrepancies leading to uncertainty, as well as conflicting proposals about their meaning.

Later Gerber [65, 66] analyzed the above mentioned architectures and highlighted that all the layers in the layered architecture were not compliant with each other and thus had scope of improvement. The work highlighted few discrepancies and irregularities in the reference architecture (see figure 3.4) such as Side-by-Side layers, Triangular arrangement of the layered architecture, Combining techniques and working descriptions in the naming of layers and Vertical layers.

Figure 3.3: The Semantic Web Reference Architecture [24, 25, 26]
In order to overcome the above highlighted inconsistencies and discrepancies, Gerber [47, 48] described a Comprehensive Functional Layered (CFL) architecture for the semantic web as shown in figure 3.5.

The CFL architecture could conquer the limitations of Lee’s architecture and is discussed below.
3.3.1 Unicode and Uniform Resource Identifier

The layer ensures that the entities and concepts are combined to infer a unique description and refer to an address on the web. Unicode [106] specifies a regular character pre-identified benchmark for the illustration of text for computer processing. It aims that all human languages may be utilized on the web using a consistent format. Uniform Resource Identifiers (URI) [133] is a string in standard form, which allows unique identification of resources. Uniform Resource Locator (URL) [133] is an additional part of URI that provides access technique and location of the web contents. The usage of URI is significant in distributed web as it gives understandable identification of all resources.

3.3.2 Extensible Markup Language (XML)

This layer contains languages essential to compose information obtainable on web pages. XML assists an exact meaning of knowledge representation by providing content, structure and representation information apart. It allows the user to illustrate the mark-up terms enclosed by tags. It is a basic mark-up language for documents describing structured information. XML namespace give specifying dissimilar mark-up vocabularies in one XML file and XML schema provides schema for a specific set of XML documents.

3.3.3 Resource Description Framework (RDF)

This layer is responsible for defining information about resources on the web utilizing XML tags. RDF Schema facilitates semantic interoperability. In order to present resources on the web with structured machine-understandable descriptions of other languages, RDF can be used for representing information about web resources so that information can be exchanged between web applications without loss of meaning. RDF is a technique for describing information about a resource in a graphical shape.
3.3.4 Ontology

Ontology refers to the vocabulary of a domain. To make machines determine sense of several concepts these must be maintained by a few files containing explanation of concepts along with their relationship with all other. It is a file that appropriately describes concepts and their relations [209]. It is expected that ontologies will play a key job in the managing, exchanging and reuse of information among web services. On the semantic web, ontologies should be applied in services entailed to explore across or combine knowledge from various applications. It should have the capability to support conclusion. It is a suitable approach to determine design of conceptualization, where conceptualization is a set of different terms, which describes the objects and their relatedness with other terms [40].

3.3.5 Rules

The main objective of this layer is to support inference drawing and contain rules for transforming a document from one RDF schema into another one.

3.3.6 Logic & Proof Layer

This layer contains properties of writing logic into documents and gives rules for inference of one type of document into another type. Knowledge Interchange Format (KIF) [133] plays an important role to specify logic in this layer. This block also involves proof layer for authentication of individuality. The proof will be a series of declarations and reasoning rules with pointers to all the supporting material.

3.3.7 Trust

Trust layer gives a mechanism to establish the trust levels of obtained information. It also identifies if the source of information is genuine authentic and at the same time ensuring that only approved users can access the information. It is important that the reasoning system must contain signature verification modules.
3.4 SEMANTIC WEB MINING

The present web infrastructure is a large repository of unstructured data that is commonly understood by human beings only. In these conditions, semantic web plays a significant role by giving machine understandable semantics that create the information machine reasonable and web mining [82] handles the semantic interpretability problem by exploring semi-structure approaches to retrieve hidden information from huge quantity of web data. The amalgamation of these two novel techniques create innovative web environment that is called Semantic Web Mining (SWM) [120]. Figure 3.6 depicts the working of SWM. SWM works in three phases as described below:

**Phase I**: This phase used to eliminate the noisy data and provides interpretability between unstructured data such as email, article, discussion forum, and wikis and explores set of techniques to remove the inconsistency.

**Phase II**: Second phase is used to retrieve knowledge from the interconnected hypertext documents. It is a way of determine design of conceptualization and ranking hyperlink documents based on their quality.

**Phase III**: This phase used to extract valuable information from knowledge discovery phase and provides agreeable results.

As shown in figure 3.7, SWM technique can be classified into three forms i.e. Web Content Mining (WCM), Web Structure Mining (WSM) [120] and Web Usage Mining (WUM)[120].

![Figure 3.6: Semantic Web Mining Infrastructure](image)
3.4.1 **Web Content Mining (WCM)**

WCM is a preeminent technique to find valuable contents and documents from the web. The web contents are describing in two forms: text and multimedia contents. Text content further consists of semi-structure content likes HTML data and unstructured text [162]. On other side multimedia content contains of picture, sound,
tape and structured articles that provide semantic description. The current development of WCM technique have encouraged developers to make more intelligent approaches for knowledge accessing, such as Information Retrieval (IR) [11] and Database approach (DB) [74]. IR uses intelligent agent approach [88] to enhance the information searching and extracting the information from the users inferred or solicited profiles. DB uses database approaches to determine the data on the web and integrate them so that more difficult queries could be searched. The rapid growths in WCM techniques have allowed system to increase knowledge deliverance of information through combining of several approaches such as agent based approach and database approach [88].

3.4.1.1 Agent-Based Approach

An agent is an autonomous software program having capability to execute the task on behalf of users. It exhibit key characterises like mobility, proactive, goal oriented and search patterns. It can also determine the topic illustration by a web data and discover web pages across dissimilar corresponding servers. Agent based approach runs with the help of following three agents.

- **Intelligent Information Searching Agent** [30]: various intelligent tools have been made for extracting valuable information and interpret the discovered information. Harvest [28] is used to obtain particular domain information about particular forms of document. ShopBot [117] is explored for retrieving terms information from several merchant websites and utilize it to obtain common information of the product domain.

- **Information Filtering Agent** [74]: Few intelligent approaches use to retrieve, filter and categorize information automatically. An example of such as approach is HyPersuit [35]. Hypersuit approach is explored for creating cluster taxonomy of hypertext documents.

- **Personalized Web Agent** [160]: Personalized Web Agent (PWA) makes a web site with the necessities of particular users, getting benefit of the information obtained from the study of the user’s web surfing behavior in related with other information collected in the web links, queries and user profile login. WebWatcher and Firefly [160, 206] are the general examples of PWA.
3.4.1.2 Database Approach

Database approach is used to managing the semi-structured information into more classified form so that enhanced information retrieval and searching on the web become probable. It also utilizes query languages and data mining techniques to evaluate it. These tools are:

- **Hierarchical Database:** It is a multilevel database technique to extract appropriate data from hypertext document. This technique works on two phases of database i.e. lower hierarchy and higher hierarchy. The lower hierarchy of phase executes with semi-structured information of the hypertext document and higher hierarchy extract metadata knowledge from lower level and arranges them into object oriented database design.

- **Query System:** with the flooding of information on web, it is more complicated to extract knowledge oriented information form web. It has become essential to use some knowledge oriented approaches so that desire information can be extracted. Web mining technique explores standard database query languages such as SQL, W3QL [160] and WEBLOG [206] to extract knowledge from heterogeneity and semi structured information sources.

3.4.2 Web Structure Mining

Web structure mining is the way of discovering valuable information from the interconnected hypertext document on the web. This technique is work with the topology of hyperlinks consisting of web pages as nodes and hyperlinks as edges. It is appropriate technique to calculate the relatedness of each web page. Web structure mining executes on two phases: hyperlink and document structure described as follows.

- **Hyperlink** [157] is a fundamental term that links a web page to multiple locations either inside the same web page or different web pages. The hyperlink inside the web link is called Intra-search hyperlink [104] and hyperlink among dissimilar pages on the web is called Inter-link hyperlink [206]. Hyperlink technique utilizes hyperlink evaluation approach to enhance the search quality method [157]. It also explores various link evaluation tools.
to study the link quality such as PageRank [147], Weighted PageRank [117] and Hypertext Induced Topic search (HITS) algorithms [147].

- **Document Structure** [162] arranges the web contents in form of tree-structure hierarchy and utilizes various HTML and XML tags within the web pages [82]. Its central idea is to automatically retrieve documents object model (DOM) [145] from the all documents. Document structure technique mainly executes in two phases i.e. Structure Based Document Clustering (SBDC) [104] and Discovery of Browsing Pattern [35]. The former technique retrieve right document structure information after clustering process while the later evaluations the hyperlink web pages extracts by dissimilar times and also by detecting the user behaviour.

### 3.4.3 Web Usage Mining

Web usage mining [82] is an innovative approach to automatically identify the user interaction patterns from web services and measures user behavior, when the user works on the web. It helps to identify type of contents in which user are more interested. Today various business firms and e-commerce societies are follows these rules for evaluating life time value of client and gives better link according their browsing behaviors.

Web usage mining retrieves desire knowledge from server log, proxy log, browser log and managed databases. Web server log contains the history of page log and proxy server executes between customer browser and web server. It works on three forms such as data pre-processing, pattern discovery and pattern evolution.

- **Data Pre-processing:** is worked on raw data to convert it into information. Data is first cleaned to remove incompatible or redundant terms from web server logs. Then the user and session log of the user is determined by user session and recognition element and finally the approach executes by referring to cache. However, due to limited caching, proxy servers, agent cache and few significant web usage profiles do not obtain saved in log documents. This might cause problem in examine the browsing activities of users. This difficulty can be solved by inferring cache hits.
• **Pattern Discovery:** Once user transaction has been determined, various data mining approaches can be used to find the user web surfing patterns. These approaches can be statistical evaluation, organization rules, clustering and categorization. Statistical evaluation [30] is a suitable approach to extract information about visitors to a specific websites. It completes statistical evaluation task such as frequency, mean, medium, ranking, session time and length on the source of navigational path. Organization Rules [147] is find of those pages, which are most regularly referenced in a particular server sessions. Clustering [208] is a technique of grouping simultaneously a set of terms having similar features and categorization [208] is to set the information into various predefined classes. This is one of appropriate approach for optimal future website, where researchers create specific content belonging to a specific user group and class.

• **Pattern Analysis:** approach [82] evaluates the information retrieved from pattern discovery method. The right sense of pattern analysis approach is knowledge discovery phase. This approach utilizes several knowledge retrieval tools such as Online Analytical Processing (OLAP) [120]. Knowledge Query Mechanism holds the characteristic of query languages such as SQL and intelligent agent executes much faster and can handle various processes in a given time [207].

### 3.5 Semantic Web Services

Web Services are a software module that offers wide range of service which can be accessed via a network [84]. It involves the automatic selection, measurement and management of appropriate web services to execute some task and provides a high-level view of the task’s objective that makes them machine understandable. The next generation of web promises to provide semantic web services that are self-defined and agreeable to autonomous discovery of information. Some existing semantic web services are shown in figure 3.8 and discussed as follows.
3.5.1 Semantic Web Services for Communication

Simple Object Access Protocol (SOAP) [84] is a messaging protocol that provides a standardized way of exchanging XML-encoded data in a distributed environment such as semantic web. It allows programs to execute on different operating environment (Window and Linux) and communicate using XML and HTML protocol. It plays an important role in heterogeneous environment on web where web application demands a common data encoding protocol and message format. SOAP provides a universal message platform that invokes functions in other applications. It also shares universal understanding of data such that machines can easily identify them. It consists of three parts: first part is SOAP envelope illustrating what's in the message and how to precede it and second is header containing a set of encoding rules and third one is Body consisting of a major part of message.

3.5.2 Semantic Web Services for Description

Web Services Description Language (WSDL) is an XML based web service that uses XML for illustrating web services as collection of communication endpoints that gives the facility of exchanging messages. The basic characteristic of WSDL is that corporations can circulate WSDLs for services and users can access those services using the information in the WSDL. It simply defines services as communications tool in a structured and standardized way. Table 3.2 delineates the components of WSDL message.
Table 3.2: Components of WSDL Message

<table>
<thead>
<tr>
<th>Types</th>
<th>Message</th>
</tr>
</thead>
<tbody>
<tr>
<td>Message</td>
<td>an abstract, which used for communication</td>
</tr>
<tr>
<td>Operation</td>
<td>an action performed by the service</td>
</tr>
<tr>
<td>Port Type</td>
<td>a set of operations that executes with one or more endpoints</td>
</tr>
<tr>
<td>Binding</td>
<td>a combining protocol</td>
</tr>
<tr>
<td>Port</td>
<td>a single endpoint described as an amalgamation of a binding and a network address</td>
</tr>
<tr>
<td>Service</td>
<td>a collection of linked endpoints</td>
</tr>
</tbody>
</table>

3.5.3 Semantic Web Services for Discovery

Universal Description, Discovery and Integration (UDDI) [173] is a standard proposed by companies to find each other and share business information. It gives a mechanism for clients to find web services. The UDDI Business Registry is a universal, open, online directory that describes their services, tool for discovering other companies’ services, which is open to everybody.

3.6 IMPLEMENTATION PLATFORMS

Figure 3.9 represents various platforms on which semantic web is implemented and a description of each one of them are being given in upcoming subsections.

Figure 3.9: Implementation Platforms
3.6.1 pOWL

pOWL is a web based collaborative semantic web platform based on PHP [18], the web scripting language. pOWL works with OWL [209], which is a language for processing web information and gives more rich integration and interoperability of data between ontology domains. It is a vocabulary extension of RDF [65] that defines the RDF graph. pOWL enables parsing, extracting information, searching, manipulating, saving and serializing RDFs and OWL knowledge bases in a collaborative web enabled environment. pOWL exhibits a 4-tier architecture having pOWL store, RDFAPI, pOWL API and User Interface as four layers (see figure 3.10).

Figure 3.10: High level view of pOWL

Here, pOWL Store is a SQL compatible relational database that utilized a database structure, where all resources are attached in full in a table row representing an RDF declaration. The pOWL store is retrieved by RDFAPI where RDFAPI is used for handling RDF, OWL and RDF-Schema. RDFAI extends RDFAPI and OWLAPI class schema with RDF Schema. pOWLAPI contain classes and functions to make web services on top of those APIs. User interface provides a communication interface between a set of PHP pages that retrieves information form a pOWL store.

3.6.2 Letizia

Letizia [112] system was developed by Henry Lieberman at Massachusetts Institute of Technology. It exploits intelligent agent system that work as interface agent on the web browser. The real vision of Letizia system was to build autonomous agent system inferring user interest from browsing behavior. Agent tracks the user browsing pattern with processes such as following link, initiate searching, request for help and tries to
understand the searching pattern. The Letizia system executes in following three steps (see figure 3.11):

- **Interleaving Browsing with Automated Search** is performed when user interacts with web browser and search relevant information. At this instance, Letizia observes user preferences and summarize the browsing history. It is accomplished by information retrieval and information extraction agents.

- **Modelling the User’s Browsing Process** is used to summarize those links, which are most visited by user. The ambition of this process is to automatically perform evaluation and determine user preferences.

- **Inferences from the User Browsing Behaviour** is more important tool for any marketing or e-commerce websites which observes the user interest.

### 3.6.3 Automatic Semantic Matching of Ontologies with Verification (ASMOV)

ASMOV [225] is an ontology matching system that eliminates semantic discrepancy between ontologies. It was introduced by INFOTECH SOFT [225] and University of Miami in Florida [119]. It explores element and structural attributes of two ontologies and calculates relatedness measure between them. On the basis of these mapping, it evaluates the results and ensures that these semantic measures do not lead to semantic inconsistency anymore. ASMOV system computes similarity calculation and semantic verification as described:
• **Similarity Calculation**: It is depends on four major attributes of ontology matching. These attributes are semantic alignment process (SimAlgL) [225], internal process mapping (SimAlgI), external mapping process (SimAlgE) and individual mapping process (SimAlgN) [225] where SimAlgL determines labels and also gives an environment to save user understandable information. SimAlgI describes similarities involving attribute of domain, classes and range of resources. SimAlgE confers connection among parents and child of classes and properties and SimAlgN describes to find similarities between the sets of class members. The similarity is finally computed using equation 3.1.

\[
SimAlg(n^1, n^2) = \frac{\sum_{\gamma \in \Lambda} w^1 \times SimAlg l(n^1, n^2)}{\sum_{\gamma \in \Lambda} w^1}
\]  

(3.1)

Here,

\[w\]: is weight;
\[n^1, n^2\]: are entities;
\[\Lambda\]: are Ontologies;
\[L\]: is lexical similarity;
\[I\]: is internal similarity;
\[E\]: is external similarity;
\[N\]: is individual similarity;

• **Semantic Verification**: Semantic Verification [224] works in five ways:
  
  o Multi-entity correspondences: arises when an alignment process contains several source and destination ontologies.
  
  o Crisscross correspondences: arises when a matching process contains both source and destination ontologies.
  
  o Disjointness: occurs when both ontologies are invalid and therefore cannot be verified.
  
  o Equivalence incompleteness: arises when set of the two axiom is not emphasized in its corresponding ontology


- Domain and range incompleteness: Domain incompleteness arises when this axiom cannot be demonstrated.

3.6.4 KOINOTITES

KOINOTITES [63] is a semantic web platform, which uses the web usage mining techniques to analyze the need of individual or set of users, who have related kind of interest and navigational behaviors. However heterogeneity environment create some problem in this way. In order to address the above mentioned limitation, web personalization [74] and data mining techniques [82] plays a vital role to extract desired information.

KOINOTITES system work as modular architecture consisting of two components, mining component and graphical user interface (GUI) [63]. Mining Component exploits web usage mining technique to extract web surfing pattern and uses data pre-processing, pattern discovery and pattern techniques. GUI allows interactive user interaction with system. Figure 3.12 depicts the working of KOINOTITES system and is further explained.

![Figure: 3.12: Architecture of KOINOTITES System](image)

Having understood the basics of semantic web, the next section now presents the prominent application areas of semantic web.

3.7 APPLICATIONS OF SEMANTIC WEB

The exponential developments in semantic web have enabled the machines to interpret the results with their intended meaning and thus enhance the information interpretability where users can extract knowledge through the combination of several
approaches. The semantic web has the capabilities to extend internet market in more precise way. Currently semantic web is an appropriate technique for achieving goals of information community. The application areas of semantic web is broadly spacious ranging from e-commerce, social media, education, scientific research, medical and many other fields related to human society. The central objective of this section is to outline the application areas of semantic web. Several application areas of semantic web are listed and discussed as follows.

3.7.1 E-Commerce

The concept e-commerce illustrates an electronic business [209]. It is one of the online services running over the web. These services are online marketing, online transaction, and net banking and merchant administration system, just to list a few. The usability of e-commerce is growing swiftly and it provides wide range of services that saves human effort and money. Daily millions of transactions are performed by people attempting to search their needs [101]. The majority of searchers are linked to customer merchandise, where they are searching items for purchasing.

3.7.2 Social Media

Social media are innovative computerized techniques which allow user to make, share information, thoughts and images/videos in virtual communities and networks [110]. The vision of social web was to offer an environment which could enhance the social activities and the result of this is present social networking websites like facebook, twitter, Youtube etc. Semantic web plays an important role in social media for retrieving desired information in web. It allows users to share knowledge with each other in globally.

3.7.3 Education

With the rapid growth of social and semantic web, educational infrastructure are regularly using semantic web technologies that provides a more flexible, personalized and intelligent learning environment [79] thereby enhancing the quality of online educational services such as e-study, video conferencing, web chat and Anytime, Anywhere, Anybody Studding (AAAS) where several of the web approaches are reusable learning objects i.e. intelligent agents.

3.7.4 Scientific Work

The scientific work is a set of methods for finding new trend, retrieving innovative knowledge, or combining and amending earlier work [69]. Semantic Web is designed with the objective to give the necessary services and techniques, which help a user to perform various tasks for finding valuable information and knowledge. The integration of scientific work and semantic web proposed new infrastructure that gives a new direction for technical research and pushing information technology towards making the meaning of information and exploiting some mathematical and probability approaches to automatically retrieve desirable information from the particular system.

3.7.5 Medical

Research in medical necessitates the exploitation of innovative computing techniques for processing large quantities of ambiguous and uncertain biological information. Semantic web has recently emerged as an approach to precisely assign an unambiguous context to an ambiguous term based on ontology and especially known for their capability to make low cost, fast and reasonably intended solutions to complicated GO annotations problems [131]. In the majority of conditions, GO annotation is written as systematic natural language and that is apposite for humans, but not valuable for machine because machine doesn’t have the kind of vocabulary that people have. In that circumstance Semantic Web plays an important role for processing such type of data resources. Role of semantic web in the medical field can be generally partitioned into knowledge based and distributional based techniques. Knowledge based technique explore knowledge with dictionaries, thesauri and semantic networks with path distance and information content (IC) measures [53]. On other side distributional measures used the addition knowledge source, the sharing of concepts within a corpus to calculate similarity such as corpus IC and context vector approach.

3.7.6 Knowledge Representation and Management

Knowledge representation and management (KRM) [40] is an essential application area of semantic web. KRM is correlated with accessing and retaining knowledge
within a specific area. Vast databases have played central role because they help to maintain the internal knowledge and can provide best results, make new example and enhance their competitiveness. Knowledge management [209] is mostly used for online tool where user can extract valuable Knowledge. The role of the semantic web is to facilitate more advanced knowledge enhancement approaches. It aims to signify the information with symbolic and mechanical reasoning systems. However this vision cannot simply obtained. It requires more intelligent tools such as ontology, semantic network and multi-agent system.

Now, turning our attention to software agents and their potential abilities, which make them suitable for semantic web.

**3.8 SOFTWARE AGENTS**

Intelligent agent is an autonomous software entity that has the ability to do the job on behalf of users. It works as a delegated entity that helps to carry out an explicit task and is possessed with higher-level algorithms necessary for managing complex jobs. It attains the information from the environment and responds according to its abilities autonomy, proactivity, reactivity, flexibility and ability to converse, just to list a few. While it may gain knowledge from their own experience, it should be able to carry out its activities without involving constant human guidance. However, whenever it is essential it should communicate and share information with human and other agents by migrating from one place to another over a network.

A software agent can be valuable as standalone application that performs delegated task on behalf of a user but in complicated circumstances, a single agent may not able to gain the desired results. Therefore agents must cooperate with other agents in the system leading to MAS [197]. MAS provide various benefits over single agents like, reliability, robustness, modularity, scalability, flexibility, concurrency, parallelism and dynamism [103]. The software agents can be static and mobile [172] as well. A static software agent is an autonomous software entity that permanently works on a specific host and on other side, a mobile agent is a software agent that has the ability to transport itself from one host to another one over a network. Autonomy and mobility are the most important characteristics that should be embedded in a mobile agent.
3.8.1 Classification of Software Agents

The section classifies agents according to the various characteristics these are designed with. Accordingly, agents have been classified as cooperative agents, autonomous agents, learning agents, interface agents and collaborative agents (see figure 3.13).

![Figure 3.13: Classification of Agents](image)

3.8.1.1 Cooperative Agent

The main goal of cooperative agent is to coordinate with other agents to achieve the common coherent goal. It involves skills for working in a group context to ensure effective joint task completion. The task is performed on five skill elements such as coordinating activities with group participants, sharing knowledge with authority and assertiveness, extracting abilities and supporting others.

3.8.1.2 Autonomous Agent

The intelligent agents perform their tasks without any human guidance. They autonomously decide the series of operations to be performed to achieve their tasks.

3.8.1.3 Learning Agent

Agent has potential to learn about the user’s behavior and adapt themselves to suit the user’s requirements. A search agent retrieves knowledge for the user from the web search engine in precise ways. The search agent should be able to learn about the kind of information the user is most interested in and adapt it to deliver only the appropriate information to the user.
3.8.1.4  **Interface Agent**

Interface agent takes the raw input and ultimately delivers the output by delegating the task to other agents. Interface agent work as a request/response tool, where one agent sends requests to another agent to perform a set of operations.

3.8.1.5  **Reactivity Agent**

Agents have the ability to sense changes occurring in the environment and react according to input.

3.8.1.6  **Proactive Agent**

This is goal oriented behavior of agent, where they do not only react in response to changes in the environment but can also take initiative for change.

3.8.2  **AGENT PLATFORMS**

Agent platform provides an environment where various agent communication interfaces are execute. Details descriptions of agent platform are given as following:

3.8.2.1  **Java Agent Development Framework (JADE)**

JADE [58] is a middleware framework for developing agent-based applications that can interoperate equally in wired and wireless infrastructure. JADE promotes developing MAS with the predefined programmable and testing tools. It provides a set of graphical tools which helps debugging during agent development. It provides the information resources distributed over the network in the form of java compatible devices like cell phones, laptops and desktops etc.
JADE run-time environment is called a container that contains all the agents created in it. Collection of these containers is called a platform that provides a homogenous layer which hides the complication of primary hardware & software from agents and their developers. While JADE offers many advantages such as mobility, scalability, dynamics and communication [47] over competing platforms, the main disadvantage of JADE is that mobility is not a key element in JADE.

### 3.8.2.2 Voyager

Voyager [196] is a distributed computing middleware to manage the remote communication of traditional CORBA and RMI protocols. Voyager provides support to both mobile objects and autonomous agents and enables objects and other agents to send standard Java messages to other agents. Voyager has set of characteristics such as providing a facility of dynamic configuration of CORBA proxies, provides communication among agent and security system, gives flexible life spans for agents and provides variety of span methods. The primary advantage of Voyager is flexible life spans and the main limitation is that it is a commercial term and is not accessible for free, which could prevent various researchers from utilizing it in favor of other alternatives accessible software.
3.8.2.3 Grasshopper

Grasshopper [29] is a mobile based agent framework that has been designed in conformance with MASIF and FIPA standards. Grasshopper was developed by GMD FOKUS and IKV++ [196]. Currently, several international research projects with the European CLIMATE (Cluster for Intelligent Mobile Agent for Telecommunication Environment) are using Grasshopper as an agent platform. Although Grasshopper supports both dynamic and static agents but it is not accessible anymore and latest versions will not be shown in the future.

3.8.2.4 Zeus

Zeus [196] grants a set of software modules that are utilized to design, expand and arrange agent systems. It also provides an integrated infrastructure for the rapid improvement of collaborative agent services. Zeus was developed by Advanced Applications & Technology Department of British Telecommunication labs [196]. It is open source freely obtainable tool executed in Java and complies with FIPA standards. ZEUS describes an intelligent agent system design methodology and defines the methodology for capturing user requirements. It automatically creates the executable source code of the user-defined agents. The main disadvantage of ZEUS is that its documentation is very weak leading to difficulties in creating new applications.

3.9 SEMANTIC WEB AND SOFTWARE AGENTS: THE INESCAPABLE RELATIONSHIP

Software agents are finding their applications in a wide range of discrete optimization problems such as telecommunication network, wireless sensor network, continuous optimization & many more. Their numerous applications motivated the researchers to consider the same as a good solution for semantic web too. With the improving technologies, WWW is transforming from semantic web to Wisdom Web [70]. Semantic web requires several autonomous techniques and standards in order to provide well defined meaningful information to the users. The deployment of software agents in semantic web is probable solution. In fact, available literature [24,
suggest that agent based approaches are well suited to cope with extremely dynamic and huge size character of these environments.

The insight into the available literature [197, 48, 97, 105, 201, 127] highlights that there are many well defined standard in literature on which intelligent agent based techniques can be analyzed. Moreover, the analysis parameters have been different. For instance, Singh et al. in have proposed Multi Agent Framework for Semantic Web (MAFSW) [197] for analyzing the potential benefit of agent based frameworks, exploited in semantic web. MAFSW provides several communication advantages for agents working in semantic cyberspace. Juneja et al. [97] have also proposed an agent based evaluated their works using fuzzy logic. Tamma et.al [201] described the incredible advantage of intelligent-agent framework employed in semantic web. Burstein et.al in [127] proposed Semantic Web Services Initiative Architecture (SWSA) system that made a set of standard protocol that works in backend as an agent and serve as a foundation for Semantic Web services.

Turning our attention on the role of intelligent agent in semantic web, the features like autonomy, mobility, proactive and reactive etc. helps improving the performance of the semantic web. Although there are listless advantages of agents over human users that inspires the employ of agents in semantic web but few of them are listed as follows [197]:-

- Automated methods will help to remove inconsistencies and attaining new knowledge.
- Agents consist of skills for managing resources and for executing required activities to reach goals.
- Agents have abilities for working in a collective context (in any role) to ensure efficient combined task completion.
- Agents can work on abstract task specification using its prior knowledge of general tasks and can choose either of the method available.
- These have the capability to take the decision or select a course of action.
• Advancement is the proficiency of a system to autonomously adjust its components and links in expectation of changes in its environment is also a motivating factor.

Owing to the above advantages that software agents offer to improve the performance of semantic web, it is obvious that semantic web operates hand in hand with these intelligent entities. A discussion on ontology as presented in next chapter also reflects that semantic web is de’ novo equipped with software agents.

### 3.10 CONCLUSION

The chapter provided the detailed insight of semantic web and also presented a description of software agents. The central objective of this chapter was to understand the basics of semantic web and understand the role of agents into the same. The tools and platforms presented highlighted that both the technologies have a long way to go and together these can definitely make the web as wisdom web. Since, the current research work embodies ontology heavily, therefore the next chapter is devoted to ontology mapping and existing algorithms in the domain.