Semantic Web as the Next Generation Smart Web

[1]Nikita Malik, [2]Sanjay Kumar Malik
[1]Research Scholar, USIC&T, [2]Associate Professor, USIC&T
Guru Gobind Singh Indraprastha University, Sector-16C, Dwarka, New Delhi-110078, India

Abstract: In the present times, information search on the web is quite significant in our daily lives. The vast information on the huge and decentralized web may be present in structured, semi-structured or unstructured form, due to which the present web 2.0 suffers with certain limitations. During web searches, a lot of irrelevant information also gets displayed to the user which makes the information search unpleasant. This demands the need of an efficient information search on the web. For this, a smart and intelligent web is needed which can understand what the user needs. Semantic web is a dream and an effort which may provide a solution towards this. It is the future web 3.0 , envisioned by Sir Tim Berners Lee, whose objective is to extend the current web 2.0 such that a well-defined meaning and structure is provided to the existing information on the web, leading to a better cooperation of humans with machines and web giving a web with semantics embedded. This paper includes how semantic web may act as the future smart web of the next generation including the various critical aspects briefly. In this paper, first, origin and growth of web has been presented along with various web versions and contributions of Berners-Lee, the father of web; second, the vision, goal, capabilities, life cycle of semantic web data and processes, popular fallacies, significant milestones and applications etc. are presented; third, the foundation architecture for semantic web with its basic technologies and few significant tools have been explored, discussed, compared in tabular form and revisited, followed by the future scope in this direction.

***Index Terms* - Semantic Web, Ontology, RDF, SPARQL, Linked Data, Smart Web, Web 3.0.**

I. INTRODUCTION

1.1. Origin and Growth of Web

1.1.1 History of Internet

Early signs of network architecture can be traced back to more than 50 years- in the 1960s. A U.S. Defence sponsored study led researchers to develop ARPANET (Advanced Research Project Agency Network) in 1969, which originally consisted of nodes at four U.S. research institutes. It was designed as a robust network aimed at allowing scientists to share data and access remote systems. Over the years, more nodes were added to this network, and it developed further with the introduction of new capabilities such as file transfers, emails, etc. With the popularity of these applications, ARPANET became a high-speed, distributed, many-to-many connections system, which formed the basis for the Internet [2].

In the 1970s, Transmission Control Protocol/Internet Protocol (TCP/IP) was developed to facilitate connection of separate networks to form a network of networks, i.e. the Internet (International network). It is a set of network communication protocols that enables connection of hosts to the Internet, and also specifies the framework for a few basic services such as email, remote logon and file transfers over a large number of client-server systems [2].The 1980s saw a huge rise in the industries of personal computers and superminicomputers. With the availability of inexpensive desktops and powerful servers, many companies joined the Internet and began using it for communicating with their customers and with each other. And by 1990, with over three lakh hosts, the ARPANET was decommissioned, in a step to move beyond the Internet [2].

While few companies were investing in proprietary networks, Tim Berners-Lee was working at the European Laboratory for Particle Physics, CERN in Switzerland, on designing a cost-effective and efficient way for providing world-wide access to the huge information stores for free. He introduced the idea of a World Wide Web (WWW) - the system of web pages and sites, based on the Internet, for global sharing of information. The Web can be seen as a service which uses the Internet, just like email, peer-to-peer file transfers etc., and a web browser is required for accessing it. It has been readily adopted ever since its conception.

1.1.2 Evolution of the Web

The Web (or WWW) combined text, pictures, sounds on the web pages in an interactive manner, and developers saw this potential for publishing information with the richness of multimedia in a simple way. Berners-Lee and his collaborators' efforts laid the foundation for designing open standards for the Web and built all the tools required for a working web. Tim Berners Lee is regarded as the father of the Web, since he invented the WWW, as presented in the first International

WWW Conference in 1989. This came to be known as the Web 1.0, where content was created by the producers and presented to the users to read, search and share[1] and the first web page was provided in the open internet by the end of 1990. People from outside of CERN were also invited to join this new web community in 1991.

To make the WWW available publically, Lee posted the WWW Project Summary in 1991. Lee also laid the foundation for W3C (World Wide Web Consortium), which was founded in October 1994, and exists as an industry consortium aimed at realizing the web's full potential. It focuses on promoting standards for the Web evolution and for interoperability between WWW products, by generating specifications and reference software[2]. Web evolution took place from Version 1.0 to Version 3.0 as follows:

Table 1. Various Versions of the Web

Parameter	**Web 1.0**	**Web 2.0**	**Web 3.0**
Definition	The age of static websites with just one-way flow of information i.e. for displaying before the users to read.	It is the era of dynamic, interactive websites which holds good in the present time too.	The extension of present web by associating meaning with information to allow machines to understand and process it automatically.
Web Type	Simple	Social	Semantic
Era	Introduced in 1989; extended from 1989 to 2000	Formally defined in 2004; extends from 1999 till present day	Envisioned in 1999; still under development
Actions Supported	Allows searching and reading of information on the web. Does not allow communication between server and user.	Allows reading and writing of information on the web. Interactivity allowed due to introduction of web applications.	Facilitates reading, writing and understanding of information. Meaning and context based smart applications and functionalities.
Basic Concept	Connects Information (through hyperlinks)	Connects People (using client-server programming)	Connects Knowledge (using concept of ontologies)
Focussed on	Formatting and displaying information; Hyperlinking & bookmarking on pages	Content creativity of users and producers; Sharing content (keyword based search)	Linked data sets; Consolidating dynamic content (meaning based search)
Basic Technologies Associated	HTML, Portals, E-mail, Peer to peer file sharing, Search Engines, File and Web Servers	XML, AJAX and Javascript frameworks, Blogs, Instant Messaging, Wikis	RDF, Ontologies, Personal Intelligent Digital Assistants, Semantic Searching
Target Users	Companies	Communities	Individuals
Examples of Associated Websites	CNN website back in the day, any personal web page	Wikipedia, Google, Twitter, Facebook, YouTube	Dbpedia, iGoogle, GoogleMaps

With the release of server side scripting languages such as PHP in 1994, and client-side scripting languages such as Jscript, VBScript in 1996, a dot com bubble came into being. The websites got popular in this duration, adding to the growth and commercialization of the WWW. The convergence of the major events of introduction of the WWW, graphical browsers and commercialization led to the accelerated development of IT (Information Technology) [2].

Along with this, blogging sites and the social media sites were created in late 1990s and surged in early 2000s. This entire phase of web evolution was termed Web 2.0 and became popular in 2004 after the O'Reilly Media Web 2.0 Conference[3].This is the web version that is currently in use, where content creation and sharing by the users form the core basis.

The next phase of web evolution- Web 3.0 or Semantic Web, was conceptualized by Sir Tim Berners Lee in 1999, and is still in its initial stages of development, and will become the next IT revolution for future extension of the web to a smarter web. Various versions of the web that have evolved over the years have been presented and compared in Table 1 [13].

Web 4.0, also called Symbiotic Web, is still an underground idea which is in progress. It allows read-write-execution-concurrency on the web. The dream behind this web is to enable collaboration between humans and machines through possible powerful interfaces like mind controlled interfaces. In other words, using Web 4.0, it will be possible for machines to cleverly read the web contents and react in the form of taking decisions of what to execute and in which order for fast loading of the website with superior quality performance [13].

1.1.3 Originator and Father of Web: Sir Tim Berners Lee and his Contributions

Sir Tim Berners Lee is presently the Director at W3C and has also been honoured with the Turing Award in 2016. In his efforts towards providing a functional web with open access to information, made the following significant contributions, as presented in Fig.1:

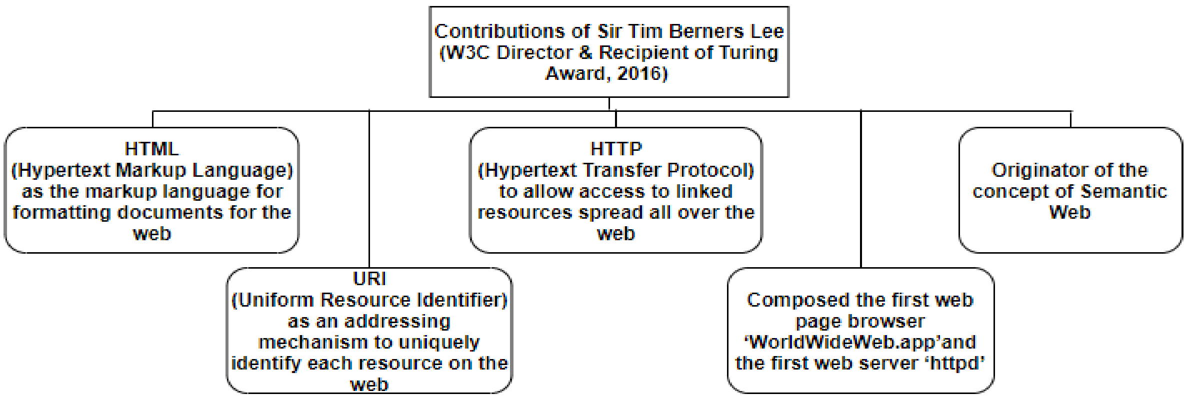

Fig.1. Contributions of the Father of the Web

Sir Tim Berners Lee introduced the concepts of HTML, URI, HTTP, web browser and server. He expressed his vision of an efficient web- the Semantic Web, when he published the roadmap for future web design in 1998, in his book Weaving The Web [3] and his speech at the XML 2000 Conference [5]. His Scientific American article of 2001[1] was a breakthrough in providing the significant components of architecture for semantic web.

1.2. Limitations of Present Web and Need of a Smarter Web

- The present web is HTML based, describing how the content is to be presented, and XML enabling data exchange across the web, none of which says anything about the meaning of the data, nor emphasizes on the machine's understanding and processing capability of the web content [2].
- Expressivity of HTML isn't enough to allow the entities defined in a document to be linked to related entities through typed links [4]. It can't directly be exploited by the techniques involved in information retrieval and extraction processes, so, web information processing remains largely limited to keywords searches. The current web search algorithms have a high recall, that is, are capable of identifying the web pages containing a set of given words, however, they also return results that are irrelevant, i.e. have low precision rate. The web browsers are also constrained to accessing the available information in a keyword-based standard form. As a result, the development of complex networks consisting of meaningful data still remains challenging today, and information search and retrieval has yet to be made efficient [2].
- Moreover, the web is vast and decentralized, with its ever growing size, and the content on the web is a mix of unstructured, semi-structured and structured forms. Therefore, it is difficult to organize and manage all the information on the web.
- Data can have different meanings in different context, and can be perceived in ways varying from user to user. This vagueness and uncertainty in knowing what a user exactly means and needs, is a challenge yet to be overcome in the present age of web.
- Information on the web can be untrustworthy, published by sources that are not validated. Security concerns on the web pose an important issue.

In effect, presently the web has been developed for humans, and there is no focus on the aspect of data that could be processed automatically. The computers are not yet able to process the meaning of the content on the web automatically. So, there is a need to go beyond the basic XML semantics, onto the elements of ontology language, logic connectives, and rule systems, so as to allow interoperable applications, automation and intelligent business logic to be incorporated in the web. User communities including organizations such as W3C and IETF (Internet Engineering Task Force) also emphasized on evolving and implementing methods and techniques for providing foundational roots towards an intelligent web. The Semantic Web introduces these notions, intended to bring about a powerful paradigm shift to fully open the information flow [2].

II. SEMANTIC WEB: THE FUTURE SMART WEB

The new web architecture, termed as the Semantic Web (SW), has advanced from the backgrounds of ontology languages and Artificial Intelligence (AI), and offers the ability to perform automated processing by the machines based on metadata associated with the information on the web [2]. Metadata is about description of sources and relationships, useful for identifying and extracting of information. Ontologies provide meaningful representations of the web content and enable users to work on shared knowledge, and facilitates advanced query search to provide better and more relevant results [23]. Semantic Web agents are softwares that function autonomously and perform tasks on behalf of the users proactively, and can communicate with other web agents, while utilizing metadata, ontologies, and logic [2]. The idea of

Semantic Web is to create a meaningful web system which provides a better platform for representation of knowledge that is globally processable by machines [23]. Smart technologies are the future, and semantic web aims at assisting in the creation of such intelligent environments and their adaptions to provide for humans [24].

2.1. Definitions of Semantic Web

Various definitions for semantic web have been given over time, as illustrated in Fig. 2:

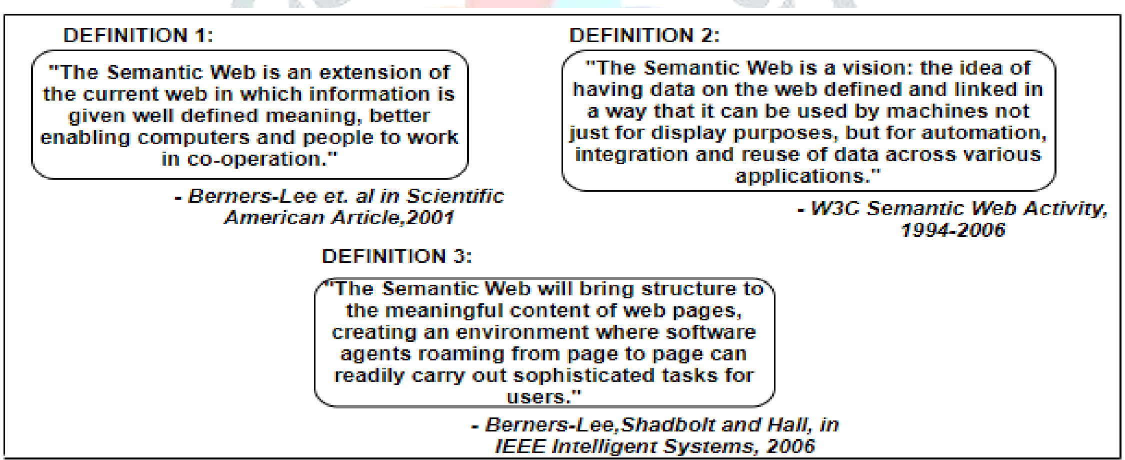

Fig.2. Definitions of Semantic Web

2.2. Vision, Roadmap and Goal of Semantic Web

Vision of the semantic web, as given by Tim Berners Lee in 1999: "I have a dream for the Web [in which computers] become capable of analysing all the data on the Web – the content, links and transactions between people and computers. A 'Semantic Web', which should make this possible, has yet to emerge, but when it does, the day-to-day mechanisms of trade, bureaucracy and our daily lives will be handled by machines talking to machines. The 'intelligent agents' people have touted for ages will finally materialize." [3]
Later, from 2001, Lee presented a broader vision for the semantic web, which included trust, where "The Semantic Web will bring structure to the meaningful content of web pages creating an environment where software agents roaming from page to page can readily carry out sophisticated tasks for users" [1].

Tim Berners Lee gave the roadmap for semantic web in 1998: "The Web was designed as an information space, with the goal that it should be useful not only for human-human communication, but also that machines would be able to participate and help. One of the major obstacles to this has been the fact that most information on the Web is designed for human consumption, and even if it was derived from a database with well-defined meanings (in at least some terms) for its columns, that the structure of the data is not evident to a robot browsing the Web. Leaving aside the artificial intelligence problem of training machines to behave like people, the Semantic Web approach instead develops languages for expressing information in a machine process-able form."[4].

The goal of semantic web is to be "a web talking to machine", that is, a web where machines can effectively and easily take advantage of the web content and provide better help to people. The aims of semantic web are [11]:

- To integrate the semi-structured or structured sources of content over the web with the aid of RDF and RDFS, that is, using RDF format for exposing the web datasets and expressing the intended semantics of these datasets using RDFS [11].
- To improve the current WWW.
- To improve search engines.
- To achieve the dynamic personalization of the websites.
- To achieve semantic enrichment of existing web pages [11].

2.3. Linked Data

Underpinning the evolution of web from a space of global information in the form of linked documents to a universal database that contains a world-wide network of semantic propositions is a suite of practices, best suitable for connecting and publishing data in a structured way on the web, known as Linked Data [4].
For achieving and creating Linked Data, transformations or on-the-go access to prevailing databases is needed, for which technologies should be available for a common format (RDF)[5]. It is based on certain rules referred to as the Linked Data Principles [18]:
1. URIs are used as things' names.
2. For users to be able to look up the things' names, use HTTP URIs.
3. Using the RDF, SPARQL standards, provide useful information for when a URI is looked up by someone.
4. To enable discovering of more things, links to other URIs should be included [18].

The extension of web with the adoption of the linked data, i.e. a universal data space holding millions of statements, this Web of Data has enabled new types of applications since it is possible to automatedly relate heterogeneous data and deduce implicit information from the evident information. Some characteristics of the Web of Data are [4]:

- It can contain different types of data and is generic.
- Data can be published by anyone.
- There is unconstrained choice on the vocabularies with which data publishers want to represent data.
- RDF links create a global data graph connecting the entities spanning across various data sources, enabling their discovery.
- Presentational and formatting aspects are strictly separated from the data.
- If a data is described using a vocabulary that is unfamiliar, the linked data application can identify the vocabulary terms by dereferencing the URIs, and can find their definition. Therefore it can be said that the data is self-describing.
- In comparison to web APIs (application program interface), data access is simplified here by the usage of RDF and HTTP as the standard data model and data access mechanism respectively.
- New data sources can be discovered by applications at runtime by following RDF links, therefore implying that the web of data is open [4].

Linking Open Data Project[6], founded in January 2007, is a community effort supported by the W3C's Semantic Web Education and Outreach Group, and the most evident example of the application of Linked Data principles [4]. DBPedia is a classic example of a large linked dataset that makes available the Wikipedia contents in RDF. Its importance is in the fact that it includes links to other data sets on the web[7].

2.4. Semantic Web Data Lifecycle

A semantic web framework offers consistency in syntax and semantics across various concepts and processing methods, and provides a complete, uniform programming environment for the semantic web. The path for exploring semantic web framework, as shown in Fig. 3, comprises of two portions- for semantic web data development and semantic web data management [15]. Steps followed by Semantic Web Data Development life cycle for direct data manipulation are [15]:

1. Storage-acquire or reference available space in memory or a database for storing emantic web data.

2. Population- populate the referenced storage space with the retrieved semantic web data from files etc.
3. Combinations- combine from multiple places, the referenced semantic web data, to create unions, intersections etc. and test for equality between referenced locations.
4. Reasoning- perform external and internal reasoning on the semantic web data that will produce additional information based on inferencing.
5. Interrogation- through navigation, queries and searching, investigate the semantic web data.
6. Export- employ methods to provide export of semantic web data in various standard formats.
7. De-allocation- free any allocated resources and clear out the referenced storage.

Administration and associated processing of the semantic web data is dealt by the Sematic Web Data Management, which provides [15]:

- Information- specifies the characteristics, size and capabilities of the semantic web data.
- Events-indicates the occurrence of different actions to enable event-driven programming.
- Concurrency- manages the multiple users and threads that concurrently manipulate the same semantic web data.
- Customization-for specialized uses, allows custom substitutes.

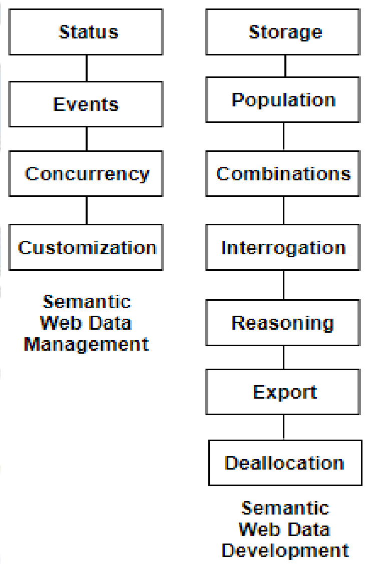

Fig.3. Path for Semantic Web Data Development and Management *(Source: Hebeler et al., 2009)*

2.5. Lifecycle of Semantic Web Processes

Semantic web services facilitate the automated annotation, discovery, invocation of logic across various applications, making the web interoperable. The functionality of web services must be combined to harness their power fully, creating web processes. Web processes represent the complex workflows and interactions on the web, and in all stages of their lifecycle, semantics can play a significant role [22]. As shown in Fig. 4, the various stages which are important for the accomplishment of semantic web processes' lifecycle are:
1. Description/Annotation- ontologies are used for describing the operations of web services, 2. Advertisement- required so that services can be discovered, 3. Discovery- of all the appropriate services based on syntactic information as well as matching the functional, data, quality of service (QoS) semantics required, 4. Selection- locating the service which best matches the requirement, 5. Composition- refers to realizing the web services by efficiently composing into web process and, 6. Execution- involves message sequences, flow of actions and conversation patterns of web services [22].

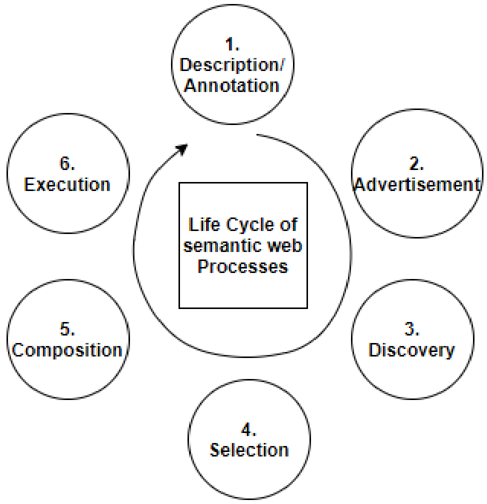

Fig.4. Life cycle of Semantic Web Processes *(Source: Cardoso & Sheth, 2005)*

2.6. Few Popular Fallacies about Semantic Web

Ill-informed criticism based on false assumptions has led to some wrong beliefs about semantic web to become popular [11]:
Fallacy 1- Semantic web tries to enforce, through its RDFS and OWL standards, the meaning on users.
Fallacy 2- Semantic web enforces the subscription to a single predefined meaning for the terms used.
Fallacy 3- Semantic web requires the understanding of complex details of formal knowledge representation for the users.
Fallacy 4- Semantic web requires the impossible task of markup of all the existing web pages manually [11].

2.7. Significant Milestones

The significant milestones towards the achievement of a new generation smart web i.e. Web 3.0 are presented in Table 2 below.

Table 2. Significant Milestones for Semantic Web

Year	Milestone	Explanation
1994	W3C founded	Lee laid the foundation for W3C and it exists ever since as an industry consortium aimed at realizing the web's full potential.
1999	Semantic Web Vision	Lee presented the vision for semantic web [3].
1999	RDF standardized	W3C recommended RDF as the standard for metadata.
2000	Semantic Web Architecture	Lee gave a hierarchy of layers that provide a full visualization of the architecture of semantic web [5]. This was the first version of semantic web architecture by him, followed by three more in 2003, 2005 and 2006.
2001	Scientific American Article	The article talks about the evolution of web from a web consisting primarily of human-readable documents, to a web that includes information for the machine to understand and process. Notions of query languages, inference rules and proof validation beyond metadata were also introduced in it. However, relating data to a particular use context, which was presumed to be a simple task, is still difficult to achieve in present web [1].
2004	RDFS and OWL standardized	W3C recommended RDFS and OWL as the standards to be used for ontology.
2007	Jorge Cardoso's Survey	Cardoso surveyed the usage or market share of different ontology editors, ontology languages and the domains for which ontologies are developed. Table 3 presents the results obtained from that survey. According to him, Protégé is the most commonly utilized ontology editor, OWL is the most

		widely used ontology language and, maximum work in this direction is being done in academia in education domain [20].

Table 3. Jorge Cardoso's Survey's Results *(Source: Cardoso, 2007)*

Usage of different Ontology Editors					
Protégé	Swoop	Onto Edit	Text Editor	Altova SW	Oil ED
68.20%	13.60%	12.2%	10.3%	10.3%	7.3%
Usage of different Ontology Languages					
OWL	RDF(S)	DL	DAML+OIL	Flogic	
75.9%	64.9%	17%	12%	11.8%	
Usage of ontologies in different Domains					
Education	Computer Science	Government	Business Service	Life Sciences	Communications
31%	28.5%	17%	17%	16.5%	13%

2.8. Capabilities of Semantic Web

The semantic web connects facts, that is, instead of linking a piece of information to an application or document, the specific information contained in that particular document or application can be directly referred, such that if the information is ever updated, the advantage of that update can automatically be taken [21]. So, similar to the way Web 1.0 abstracts away the physical and network layers, the document and the application layer involved in information exchange are abstracted away by the semantic web [19].
Semantic web allows web applications to integrate data and programs and work together, without needing intervention by humans, in a decentralized system, catering to a huge business market. Five identified capabilities of semantic web are [14]:

- Formal and Flexible Semantic Modelling: Conceptual or semantic modelling refers to the ability of capturing a universe of discourse. Similar to other modelling approaches like UML (Unified Modelling Language), using their modelling primitives of classes, instances, and properties, ontologies also allow capturing a universe of discourse. They facilitate clear knowledge representation in a formal and flexible way, using conceptual models. The formality of logic based knowledge representation languages used by ontologies assign the modelling constructs with unambiguous meaning. The flexibility aspect of semantic modelling with ontologies is the characteristic of agile schema development, which enables ontology development on instance and schema levels at runtime, and the direct interaction by means of generic tools, which allows populating, creating and changing semantic models [14].

- Web-scale, Intelligent Integration of Knowledge: Semantic web technologies can tackle the challenge of heterogeneity of engineering knowledge by solving, the complex problem of large-scale data integration, in an automatic and flexible manner. This knowledge integration is enabled by the capability to define the transformations and links between ontologies, using techniques such as ontology matching [14].

- Distributed Dataset Exploration and Browsing: For efficient data access on the web and for wide-acceptance with users, browsing and exploring of semi-structured, distributed data sets in a user-friendly manner is a core requirement. Semantic web technologies offer support for such knowledge access through navigation of linked data, supporting sense-making and browsing capabilities across multiple disciplines, providing increased productivity [14].

- Better Search Results- Semantic Web includes more involved questions, relationships and trust while performing search actions. Instead of exact word/pattern matching, related items and new relationships between data can be shown by web. This translates to high recall as well as precision for web search algorithms.

- Assurance of Knowledge's Quality: The mechanisms of semantic web, that enable the formal representation and interlinking of knowledge, allow their interpretation using reasoning. This reasoning capability can therefore be used to support and automate various quality assurance tasks, like through consistency checks. This facilitates creation of flexible, intelligent and reliable applications that detect inconsistencies and increase productivity by virtue of reasoning mechanisms [14].

- Sharing and Reuse of Knowledge: The semantic models of knowledge represented by ontologies form the basis for reuse and sharing of knowledge among domains and different applications. This reuse of existing knowledge supports data integration and increases the productivity of various engineering processes. Multi-disciplinary applications can be constructed on top of these shared and reusable semantic reference models [14].

- Interoperability: Lack of interoperability among web services is an obstacle in realizing all the web's capabilities. The semantic web's aim is to provide an interoperable web, where the machines are able to take advantage of the freely accessible resources. The requirement is for the resources to not be hidden in a proprietary system or a specific tool-oriented format, but those which can be invoked and published in an open, structured format.

2.9. Opportunities for Semantic Web Applications

Semantic web technologies (SWTs) are being used in a broad set of web and enterprise-based applications today, which include the following:

- Consuming and publishing data on the web.

- Search Engines and E-Commerce: By accessing the metadata, search engines are able to return better and relevant results to search queries. Using the notion of annotations, search results are directed to e-commerce sites that follow the standards of annotating data on their web pages, and hence observe an increase in their hits and conversions. This translates to benefits in business to such e-commerce web sites. Schema.org, GoodRelations ontology etc. are frameworks that are currently being adopted by websites for enhancing a user's web experience and profiting on their business [19].

- Semantics based Enterprise Applications: Semantic web technologies offer a technology stack that can be chosen within enterprises for implementing applications and solutions, for various reasons such as:

a) Agile Data Integration-In place of launching large data warehouse projects for integrating data, semantic web technologies offer a flexible and cheaper approach to provide employees with unified views of the information they need [19].
b) Information Classification- The humongous amounts of organizational data requires classification to facilitate easy location of key bits of information. Using the expressivity of ontologies, for retrieval based on tags or automated information discovery, and the support for reasoning/ inferencing, semantic web helps tackle the information representation and discovery [19].
c) Dynamic Content Management-Semantic web technologies help provide a more interactive and dynamic site by supporting richer content at various levels of aggregation, and automating the tagging process. Other applications also focus on question answering, search etc. [19].

- Semantic Bio-informatics-Semantic web concepts are widely applied in the field of bio-informatics. The increasing quantities of biological information can be automatically integrated using semantic concepts- metadata can be developed for biological information, stored data can be annotated, bio-ontologies can be developed for knowledge linking, performing reasoning and having shared definitions of the domain [2].

- e-Learning- By providing a common standard of communication syntax, using semantic concepts and ontologies, the educational content and its usage can be integrated on a large-scale. Semantic web offers extraction of far more information from unified, structured data space, supports collaborative work and discussions, video-conferencing or work across distributed locations to achieve interoperability in the learning community or educational system [2].

- Processing of Knowledge Bases- Methods of semantic web are applicable to tasks of knowledge management. Use of knowledge graphs expressed using RDF, and the type logics attached to these, referred to as ontologies, are central to this field. Models and abstractions are provided by semantic web for the distributed processing of knowledge bases [17].

- Neural Networks- Trained neural networks act like black boxes- they don't indicate directly as to why an output was given, reason being the distributed nature of information that the weighted connections of the network encode. This unsatisfactory situation can be addressed using semantic web background knowledge and technologies which can explain the classification behaviour of artificial neural networks using non-propositional logics. Semantic web methods are used to explain how the outputs are reached, and can be applied to sets of hidden neurons for understanding what their activations indicate [17].

- Software Engineering- Software engineering (SE) and semantic web work closely. The semantic web's ability to process distributed knowledge bases is direly needed for maintenance, implementation and software analysis in SE. Conversely, SE's decades of experience in engineering large application frameworks can be used when developing ontologies [12]. In the context of SE, ontologies can be used in various approaches, by applying ontologies in various stages of the SE development lifecycle. Advantages of this application include ease and flexibility in information integration, facilitates easy information sharing, reuse and interoperability, and ontologies provide formal description of domains and encourage a broader usage of domain models throughout the whole SE lifecycle [16].

III. SEMANTIC WEB STACK, TECHNOLOGIES AND TOOLS

3.1. Semantic Web Architecture/ Layered Cake by Lee

An architecture for a system describes its components' structure, properties, and interfaces and relationships with each other. A reference architecture for Semantic Web was first given by Tim Berners Lee in 2000 [5], as shown in Fig. 5.

His architecture development led to disagreement among some researchers who argued that an architecture should depict functionality of the system rather than the technologies as depicted by Lee. For this reason, Lee's semantic web architecture is termed as Stack or Layered Cake also. Later versions of this layered cake were presented by Lee in the on-going activities of W3C as part of the SIIA Summit in 2003 [6-8], termed version 2 architecture, followed by version 3 at WWW 2005 [9], and finally introduced version 4 at AAAI's keynote address in 2006 [10].

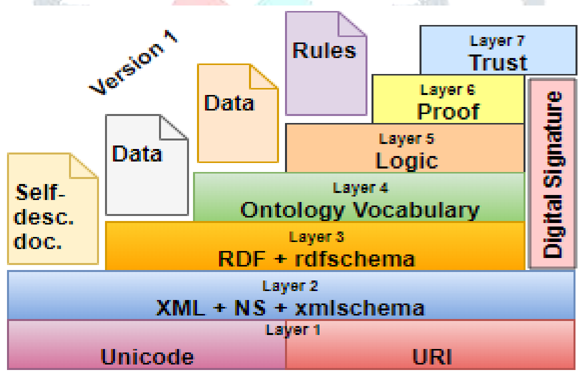

Fig.5. Version 1 of Tim Berners Lee's Semantic Web Layered Cake *(Source: Lee, 2000)*

3.2. Semantic Web Technologies

Lee's semantic web stack illustrates the language hierarchy in which every layer represents a technology and exploits and utilizes the capabilities of the layers beneath it. The technologies used in semantic web can be categorized into Hypertext web technologies, Standardized semantic web technologies, Unrealized semantic web technologies and other significant semantic web technologies, as explained in Table 4.

Table 4. Semantic Web Technologies

Hypertext Technologies	
URI (Uniform Resource Identifier)	An address string to identify each resource on the web uniquely.
Unicode	Provides character set for encoding web documents across different formats.
XML (Extensible Markup Language)	A markup language which provides syntax for storing and exchanging structured data in documents on the web. It includes schema (XMLS), namespace and a query language (XQuery) too.
Standardized Technologies	
RDF (Resource Description	Serves as a conceptual model for publishing and linking data, including metadata

Framework)	on the web. Directed acyclic graphs (DLG) represent the statements in the form of triples of subject-predicate-object as nodes and arcs between them.
RDFS (RDF Schema)	An ontology development language that provides vocabulary to the RDF and a mechanism to group related resources as classes, with their properties and relationships, in a taxonomical structure.
OWL (Web Ontology Language)	A language which captures the ontologies and using the expressive and reasoning power of description logic (DL), acts as a representation tool for the understanding and sharing of a common knowledge.
SPARQL (SPARQL Protocol and RDF Query Language)	It's a standard language for querying data expressed in RDF form, including graph patterns. Also, it acts as a protocol for accessing the ontologies and RDFS data.
RIF (Rule Interchange Format)/ SWRL (Simple Web Rule Language)	RIF allows the exchanges among rule systems and SWRL covers the specification of all kinds of rules, which may be used for reasoning and
Unrealized Technologies	
Logic	Its functioning is based on principles of first order predicate logic, for ensuring the accuracy of information displayed on the web.
Proof	It includes the representation of proofs along with proof validation, justifying why the information on the web should be believed by the user.
Trust	It involves evaluation of credibility of information obtained from different sources and the trustworthiness of the producers of information on the web.
Cryptography	It establishes security of information on the web through mechanisms of digital signature and encryption.
Other Significant Technologies	
Semantic Search and Browsers, Semantic Annotation, Information Retrieval and Web Mining, Semantic Web of Things, Semantic Social Networks, Semantic Sentiment Analysis, Semantic Web Services, Intelligent Web Agents & Softbots , Semantic Search, Semantic Web browsers , Web page segmentation, opinion mining and so on.	

Smart systems, such as smart homes, smart grids etc. are the future of collaboration of technology with humans. They allow sensors, wireless access and databases to cooperatively sense the environment, reason and adapt to it, and provide support for the users within the environment. Ontologies prove helpful and may be used for modelling the contextual information of the intelligent environment and for inferencing on data and drawing conclusions from rules [24].

3.3. Tools

Various tools/frameworks/softwares for ontology designing, reasoning, querying, RDF storing and knowledge management, semantic social network analysis, semantic wiki, linked data, semantic search, and semantic web application development are presented in Table 5.

Table 5. List of Significant Tools for Semantic Web

S. No.	Tool Name	Current version with Year of Release	Purpose	Features	Platform	Online Source
1.	Protégé	Protégé 5.2.0 in 2017	Ontology Editor and Framework for Knowledge base	Feature rich, open source ontology editing environment with full OWL 2 support and direct in-memory connections to Pellet and HermiT reasoners.	Java	http://protege.stanford.edu
2.	Vitro	Vitro v 1.9.3 in 2017	Unified Ontology Editor environment and Semantic Web Application	Web based general purpose instance and ontology editor with customizable public browsing.	Java	http://vitro.mannlib.cornell.edu/

3.	Fluent Editor	Fluent Editor 2015 version 3.6.10.287 10 in 2016	Ontology Editor	Free, comprehensive tool that uses Controlled Natural Language to edit and manipulate complex ontologies by understanding their actual meaning.	Java	http://www.cognitum.eu/Semantics/FluentEditor/
4.	OntoStudio	Ontostudio 3.2.0 Build 353 in 2012	Ontology Development Environment	Most widespread commercial modelling environment. Provides intuitive ontology modelling, mapping tools to quickly match heterogeneous structures, and ability to import many schemas and structures.	Eclipse	http://www.semafora-systems.com/en/products/ontostudio/
5.	NeOn Toolkit	NeOn Toolkit version 2.5.2 in 2011	Ontology Editor	Multi-platform open source ontology editor supporting ontology development in F-Logic and OWL/RDF.	Java/Eclipse	http://neon-toolkit.org/wiki/Main_Page.html
6.	FaCT++	FaCT 1.6.5 in 2017	Reasoner	Open source, tableaux-based reasoner which supports OWL DL and (partially) OWL 2 DL. Used as one of the default reasoners in Protégé OWL editor.	C++	http://owl.cs.manchester.ac.uk/tools/fact/
7.	HermiT	HermiT 1.3.8 in 2013	Reasoner	Free Reasoner for ontologies written using the OWL.	Java	http://www.hermit-reasoner.com/
8.	Pellet	Pellet 2.30 in 2011	Reasoner	Open source reasoner for OWL 2 providing functionality for ontologies' consistency check, computing classification hierarchy, answering SPARQL queries and inference explaining.	Java	https://github.com/stardog-union/pellet
9.	OntoWiki	OntoWiki 1.0.0 in 2016	Semantic data wiki, Linked Data publishing engine	Free and open source web based application serving as both an ontology editor and a knowledge acquisition system, using either a Virtuoso triple store or a MySQL database.	PHP	ontowiki.net/
10.	Semantic MediaWiki	Semantic MediaWiki 2.5.6 in 2018	Semantic Wiki Software	Free, open-source semantic extension to MediaWiki application that runs web-based wikis. It's also a complete framework that can transform a wiki into a flexible and	PHP	https://www.semantic-mediawiki.org/wiki/Semantic_MediaWiki

				powerful knowledge management system using many spinoff extensions in conjunction.		
11.	Dbpedia	Dbpedia Release 2016-10 in 2016	Semantic Knowledge Base	A project to extract structured content in the form of an open knowledge graph (OKG), which is openly available for on the web, from the information created in various Wikimedia projects. The Dbpedia data is served as linked data that can be navigated and queried.	Scala, Java, VSP	http://wiki.dbpedia.org/
12.	GraphDB	GraphDB 8.4 in 2017	RDF Store	Most scalable semantic repository that comprises inference engine, triple store and SPARQL query engine. Offers configurable reasoning support and performance.	Java, C#	https://ontotext.com/products/graphdb/
13.	Cytoscape	Cytoscape 3.6.0 in 2017	Network Visualizing	Open source software platform for network data integration, analysis, and visualization	Java	http://www.cytoscape.org/
14.	Visone	Visone2.16 in 2016	Network Analysis and Visualization	Research project that provides algorithms and models for advancing and unifying visualization and mathematical analysis of social networks.	Java	https://visone.info/
15.	Jena	Jena 3.6.0 in 2017	Semantic Web Development, RDF Store	Open source framework available for semantic web application building. Provides a set of interacting APIs for processing RDF data, inference engine for reasoning, and many storage strategies.	Java	https://jena.apache.org/
16.	Kngine	Kngine1.3.1366 in 2013	Search Engine	An intelligent search engine with capability to understand, answer questions and perform various actions. Useful for automating customer services, for driving enterprise search and for power voice interfaces.	Java/ Xcode	http://www.kngine.com/
17.	Drupal	Drupal	Knowledge	Free and open source	PHP	www.drupa

		8.4.4 in 2018	Management	content-management framework capable of generating and injecting RDFa annotations within HTML pages, and managing textual and multimedia content of web sites enriched with meta-information.		l.org
18.	CubicWeb	CubicWeb 3.25.0 in 2017	Semantic Web Development	Free, open-source semantic web application framework that supports OWL/RDF, RQL. Useful in large-scale LOD and semantic web applications.	Python	https://www.cubicweb.org/
19.	Stanbol	Apache Stanbol 1.0.0 in 2016	Semantic Content Management	Extends conventional content management systems using semantic services. Can also be directly used from web applications like for text completion or tag extraction/suggestion in search fields.	Java	https://stanbol.apache.org/
20.	Marmotta	Apache Marmotta 3.3.0 in 2014	Linked Data Platform	Provides open implementation for publishing or building custom applications on linked data.	Java	http://marmotta.apache.org/

IV. CONCLUSIONS AND FUTURE SCOPE

The evolution of web from its conception to the limitations that lead to the need of a smarter semantic web have been explored and discussed in this paper. The vision, technologies, tools associated with semantic web, with the aim of providing a future of smarter systems and intelligent agents cooperating with humans, have been revisited and presented such that this work forms a useful resource and a tour for users who are new to the concept of semantic web.

Despite the numerous capabilities and applications of semantic web discussed, substantial room for growth still remains. The vision of a complete interoperable web of data hasn't still been realized as most organizations are not well aware and have not widely adopted semantic web technologies. The learning curve is also steep for users who are new to the concepts. However, it is believed that in the years to come, research issues will be addressed and the dream of semantic web will be achieved towards realizing smart systems and machines, along with an environment of efficient search capabilities.

REFERENCES

[1] Berners-Lee, T., Hendler, J. and Lassila, O. (2001). The semantic web. *Scientific American, 284*(5), 28-37.

[2] Alesso, H. P. and Smith, C. F. (2008). *Thinking on the Web: Berners-Lee, Godel and Turing*. Wiley-Interscience.

[3] Berners-Lee, T. and Fischetti, M. (1999). *Weaving the Web: The original design and ultimate destiny of the World Wide Web by its inventor.* Harper & Brothers Publishers.

[4] Bizer, C., Heath, T., and Berners-Lee, T. (2009). Linked data-the story so far. *International journal on semantic web and information systems, 5*(3), pp. 1-22.

[5] Berners-Lee, T. (2000). Semantic Web – XML 2000.*W3C Web site* (Last accessed 11/8/2006). URL: http://www.w3.org/2000/Talks/1206-xml2k-tbl/slide10-0.html

[6] Berners-Lee T. (2003). The Semantic Web and Challenges.*W3C Web site Slideshow* (Last accessed 10/10/2006). URL:http://www.w3.org/2003/Talks/01-sweb-tbl/Overview.html

[7] Berners-Lee T. (2003).Standards, Semantics and Survival. *SIIA Upgrade*, pp. 6–10.

[8] Berners-Lee T. (2003).WWW Past and Future.*W3C Web site* (Last accessed 10/10/2006). URL: http://www.w3.org/2003/Talks/0922-rsoc-tbl/slide30-0.html

[9] Berners-Lee T.(2005). WWW 2005 Keynote.*W3C Web site* (Last accessed 10/10/2006). URL: http://www.w3.org/2005/Talks/0511-keynote-tbl/

[10] Berners-Lee T.(2006). Artificial Intelligence and the Semantic Web. *AAAI 2006 Keynote, W3C Web site*(Last accessed 12/10/2006). URL: http://www.w3.org/2006/Talks/0718-aaai-tbl/Overview.html

[11] Antoniou, G., & Van Harmelen, F. (2004). *A semantic web primer*. MIT press.

[12] Bernstein, A. (2010). Software Engineering and the Semantic Web: A match made in heaven or in hell?. In *International Conference on Software Language Engineering*, Springer, pp. 203-205.

[13] Aghaei, S., Nematbakhsh, M. A., & Farsani, H. K. (2012). Evolution of the world wide web: From WEB 1.0 TO WEB 4.0. *International Journal of Web & Semantic Technology*, *3*(1), 1.

[14] Biffl, S., and Sabou, M. (Eds.). (2016). *Semantic web technologies for intelligent engineering applications*. Cham, Switzerland: Springer International Publishing.

[15] Hebeler, J., Fisher, M., Blace, R., and Perez-Lopez, A. (2009). *Semantic web programming*. John Wiley & Sons.

[16] Happel, H. J., and Seedorf, S. (2006). Applications of ontologies in software engineering. In *Proc. of Workshop on Sematic Web Enabled Software Engineering (SWESE)* on the *ISWC*, pp. 5-9.

[17] Sarker, M. K., et al. (2017). Explaining trained neural networks with semantic web technologies: First steps. *arXiv preprint arXiv:1710.04324*.

[18] Berners-Lee, T. (2011). Design issues: Linked data (2006). URL: http://www.w3.org/DesignIssues/LinkedData. html.

[19] Cambridge Semantics. (2018). Semantic University, https://www.cambridgesemantics.com/blog/semantic-university/

[20] Cardoso J. (2007). The Semantic Web Vision: Where are We? *IEEE Intelligent Systems*, pp.22-26.

[21] Shadbolt, N., Berners-Lee, T., and Hall, W. (2006). The semantic web revisited. *IEEE intelligent systems*, *21*(3), pp. 96-101.

[22] Cardoso, J., & Sheth, A. (Eds.). (2005). *Semantic Web Services and Web Process Composition: First International Workshop, SWSWPC 2004, Revised Selected Papers*, Vol. 3387, Springer.

[23] Shadbolt, N., Berners-Lee, T., and Hall, W. (2006). The semantic web revisited. *IEEE intelligent systems*, *21*(3), pp. 96-101.

[24] Jih, W. R., & Hsu, J. Y. J. (2009). Agent-based context-aware service in a smart space. In *Agent-Based Ubiquitous Computing* (pp. 131-146). Atlantis Press.

ENDNOTES

[1]http://www.w3.org/History/1989/proposal.html
[2]https://www.w3.org/
[3]http://en.wikipedia.org/wiki/web_2.0
[4]https://www.w3.org/DesignIssues/Semantic.html
[5]https://www.w3.org/standards/semanticweb/data
[6]http://esw.w3.org/topic/SweoIG/TaskForces/CommunityProjects/LinkingOpenData
[7]https://www.w3.org/standards/semanticweb/data