**W3C: World Wide Web Consortium**

**Terrence Brooks**  
*iSchool, University of Washington, Seattle, Washington, U.S.A.*

**Abstract**  
The World Wide Web Consortium (W3C) is the organization that leads the development of standards for the Web. Sir Tim Berners-Lee, the founder and director of the W3C, envisions a linked network of information resources that guides Web standards development and points the way toward the creation of a Semantic Web. This entry describes the pioneering role of Berners-Lee in the development of the Web, the accomplishments and vision of the W3C, and the development of the Semantic Web.

**INTRODUCTION**

This entry introduces the World Wide Web Consortium (W3C) by outlining the pioneering role of Sir Tim Berners-Lee in the creation of the Web and Web browsers, the development process employed by the W3C community to construct Web standards, and the vision and development of the Semantic Web.

In 2007, the W3C was an international consortium with more than 400 members. Berners-Lee leads a team of about 60 researchers and engineers who are located primarily in three institutions: the Massachusetts Institute of Technology Computer Science and Artificial Intelligence Laboratory (MIT/CSAIL); the European Research Consortium in Informatics and Mathematics, Nice, France; and Keio University, Japan.

The W3C is the leading organization in the development of Web standards and general guidelines for Web development. Its general mission is the development of Web standards and protocols that maximize benefits for the greatest number of people. While the process model of the W3C is consensual agreement and cooperation, it lacks any means to coerce the adoption of its standards and protocols.

Since 1994, the W3C has published more than 90 Web standards called W3C recommendations. By building applications that comply with W3C recommendations, a Web developer supports the community of users and software that promote the sharing of information and platform independence. Berners-Lee and the W3C can justifiably be credited with inventing the World Wide Web, pioneering the development of the Web browser, and developing the protocols that make sharing information possible.

**THE VISION OF THE W3C**

The ultimate motivation of the W3C springs from the vision of Berners-Lee for the Web, and has been variously expressed as the W3C motto: “To lead the World Wide Web to its full potential by developing protocols and guidelines that ensure long-term growth for the Web,” or more simply “Leading the Web to its full potential…” This vision privileges the notions of interoperability and device independency, and suggests a public network where information can be easily shared among strangers, and is not confined to a specific application software suite or a particular vendor’s products. The essence of the W3C vision is that a document or program that follows W3C specifications should work identically across different applications and different computers. To “work identically” in this context means that the essential information or functionality is preserved, while conceding that presentation issues such as the size of the display device, color options, accessibility accommodations, and so on, may introduce variability.

The vision of the W3C is displayed by its future of the Web document[1] which claims that the Web is for:

- **Everyone** (regardless of culture, abilities, etc.).
- **Everything** (applications and data stores, and on devices ranging from power computers with high-definition displays to mobile devices to appliances).
- **Everywhere** (from high to low bandwidth environments).
- **Diverse mode of interaction** (touch, pen, mouse, voice, assistive technologies, computer to computer).
- **Enable computers to do more useful work** (through advanced data searching and sharing).

To complement its advocacy of interoperability and device independence, the W3C also promotes
accessibility. Accessibility refers to the degree that application software is sympathetic to people with disabilities, but also in a more general sense, the degree to which applications can be used in environments degraded by noise, lighting, etc.

THE SEMANTIC WEB

Residing behind the W3C’s vision of the Web as networked personal and corporate information is the more ambitious vision of the Web as a network of semantics. An early statement of this particular vision of the Semantic Web appeared in the Scientific American.[2] This entry sketched a Web network of publicly available documents that possessed semantics structured as metadata. These metadata were to be harvested mechanically and manipulated by inference tools to solve ad hoc everyday problems such as finding the office hours of the closest doctor, the cheapest available tickets for the theater, and so on. The subtitle of the Scientific American article summarizes this vision of a Semantic Web: “A new form of web content that is meaningful to computers will unleash a revolution of new possibilities.”

This original description of the Semantic Web was refined[3] by suggesting the difference between the Web of public HTML documents destined to be displayed in Web browsers, and a web of structured data that would be amenable to mechanical harvest. Theoretically, these two Webs would coexist and even overlap as HTML documents destined for display in Web browsers might also contain structured metadata. An example technology that could be used in this way would be Resource Description Framework (RDF) attributes. Examples of the data that exist on the Web and that could be harvested mechanically might be calendar data, travel arrangements, photograph descriptions, and financial transactions.

The vision of the Semantic Web as a web of linked open data has come to dominate. The result might be called a web of data, where islands of semantics exist; that is, pools of structured documents on the open Web that could be harvested mechanically. Current efforts are now focusing on linking common data points that are related but exist in different semantic islands on the Web. In summarizing the achievements and challenges of constructing the Semantic Web, Ivan Herman[4] gives a voice to the shift of the Semantic Web from a single huge, central ontology organizing all knowledge, which would be unmanageable if not impossible to construct. Instead he described the use of numerous ontologies and vocabularies, such as Friend of a Friend and Dublin Core that would capture local richness and encourage the discovery of new relationships among data. The challenge of such architecture is linking disparate data points together.

The Linking Open Data project[5] centralizes the increasingly common construction of resources that are available for the mechanical harvest of structured information. The following enumerates some of the structured data available on the Web, which are, in effect, islands of semantics that might be linked together:

- IgentaConnect (http://www.ingentaconnect.com/) offers bibliographic metadata storage and has more than 200 million RDF triplets available.
- RDFS/OWL Representation of WordNet (http://wordnet.princeton.edu/) has 150MD of RDF/XML available for download. WordNet facilitates the browsing of meaningfully related words and concepts.
- “Département/canton/commune” is available from the French Statistical Institute and represents statistical surveys on population, employment, wages, prices, business, economy, and French national accounts.
- Geonames Ontology and Data has information on 6 million geographical features.
- RDF Book Mashup makes information about books, their authors, reviews, and online bookstores available on the Semantic Web. Information drawn from sources such as Amazon, Google, and Yahoo can be integrated into the Semantic Web.
- dbpedia is a community effort to extract structured information from Wikipedia and link this information to other sources of structured information.

Semantic Web Case Studies and Use Cases[6] list descriptions of systems that have been deployed in industry that use Semantic Web techniques.

HISTORICAL HIGHLIGHTS OF THE W3C

The invention of the World Wide Web in 1989 was a by-product of Berners-Lee’s solution to a document sharing problem at CERN (European Organization for Nuclear Research). He proposed a global hypertext project of linked documents to help people coordinate their work during large project development. To manifest his vision he created the first World Wide Web server and the first browser/client called “World Wide Web.” This early Web browser was first offered to CERN in December 1990 and then to the Internet at large in the summer of 1991.

During the period of 1991 to 1993, Berners-Lee extended the development of the Web by constructing basic technologies such as Uniform Resource Identifiers (URIs), Hypertext Transfer Protocol, and Hypertext Markup Language (HTML). As an early
model of community development, Berners-Lee’s technological initiatives were modified by feedback from the nascent Web community. In 1994 Berners-Lee founded the W3C at MIT/CSAIL where he is a senior research scientist.

The beginnings of the World Wide Web can be found in the proposal “Information Management: A Proposal” placed by Berners-Lee before CERN’s governing board. Its ambition was to solve the information management problems of large projects by utilizing the idea of a hypertext, which had been pioneered by Nelson. Such a hypertext featured human-readable information linked together in ad hoc and unconstrained ways, thus promoting the discovery of hitherto unrecognized parallels or points of convergence among disparate texts.

In September 1992 Berners-Lee gave an invited presentation to the Computing in High Energy Physics 92 conference, where he claimed that the “W3 project merges networked information retrieval and hypertext to make an easy but powerful global information system...W3 now defines the state of the art in networked information retrieval, for user support, resource discovery and collaborative work.” In 1993 the CERN W3 software suite appeared in the public domain.

The first W3C recommendation, for Portable network graphics, appeared in October 1996. Within several years, major recommendations were issued that were foundational for the growth and development of the Web. These included the recommendation for HTML 4.0 in December 1997, which added tables, scripting, and style sheets. These features permitted authors to create significantly more expressive Web content. And in February 1998, the recommendation for XML 1.0 appeared. XML provided a platform for a host of following recommendations and has evolved into a universal language of the Internet.

The following enumerates some of the important historical achievements of the W3C:

- The portable network graphics recommendation in October 1996 provides a cross-platform graphics format.
- The cascading style sheets recommendation of December 1996 laid the foundation for a uniform strategy for adding style (e.g., fonts, colors, spacing) to Web pages.
- The Web accessibility initiative of February 1997 promoted accessibility of the Web through four primary areas of work: technology, tools, education and outreach, and research and development.
- The HTML 4.0 recommendation of December 1997 added new features to Web publishing such as tables, scripting, style sheets, internationalization, and accessibility.
- The XML 1.0 recommendation of February 1998 introduces the fundamental technology that would become an influential Web standard.
- In August 2000, the scalable vector graphics recommendation introduced two-dimensional graphics and graphical applications in XML.
- Web services activities introduce a fundamental protocol for Web services in January 2002.
- In May 2003, the W3C adopted a royalty-free patent policy and encouraged the development of open standards.
- Two fundamental protocols of the Semantic Web, RDF and Web ontology language (OWL) were introduced in February 2004.
- In March 2004, VoiceXML promoted the use of interactive voice response applications to Web-based development and content delivery.
- The W3C mission of universal access was promoted in February 2005 with the introduction of the character model recommendation. The goal of the character model recommendation was the ease of use of the Web regardless of language, script, writing system, and cultural conventions.
- The mobile Web initiative of May 2005 set the mission of making Web access from a mobile device as simple as Web access from a desktop device.
- In November 2005, the Semantic Web for healthcare and life sciences interest group deployed standardized Semantic Web specifications for the medical industry.

This historical listing illustrates the foundational role of the W3C in designing specifications and protocols that foster the sharing of information. The culmination of the worldwide sharing of information is expressed in the vision of the Semantic Web.

MEMBERSHIP OF THE W3C

The W3C brings together a diverse group of stakeholders to achieve its mission of developing Web standards and protocols. As of 2008 there were more than 400 members of the W3C representing more than 40 countries. Approximately 37% of these members are American and about 9% are from the United Kingdom. The stake holders of the W3C represent many different types of organizations with the largest category being consultants and systems integrators, followed by the university research and development category, and the general software companies’ category. Other categories of members include vendors of technology products and services, content providers, corporate users, research laboratories, standards bodies, and governments.
ACTIVITIES OF THE W3C

The activities of the W3C manifest themselves as various kinds of groups. Working groups focus on technical developments, interest groups focus on nontechnical issues, and coordination groups focus on the communication among related groups.

The following is an overview of some of the leading activities of the W3C:

- XML activities represent a core set of technologies around which orbit a number of working groups that are focused on development in the areas of extensible stylesheet language, efficient XML interchange, XML binary characterization, XML processing model, and XML linking, query, and schema.
- Graphics activities focus on the development of scalable vector graphics (SVG) technologies.
- The HTML activities group works on the evolution of this standard language of the Web. The current goal is to evolve HTML into an XML-based markup language and thereby ease its use with other markup languages.
- Internationalization activities work to ensure that the W3C’s formats and protocols are amenable to all of the world’s languages, writing systems, character codes, and local conventions.
- Math activities facilitate the use and presentation of mathematics on the Web.
- Mobile Web activities are working to overcome fundamental interoperability and usability problems associated with mobile Web access. They attempt to integrate the key players in the mobile area: authoring tool vendors, content providers, handset manufacturers, browser vendors, and mobile operators.
- Multimodal interaction activities target the ability of Web users to dynamically switch to the most appropriate mode of interaction. Ideally, a multimodal Web application would permit users to input data via speech, handwriting, and keystrokes, and receive output via displays, prerecorded and synthetic speech, audio, and various haptic displays.
- Patent policy activities alert the W3C community about developments in the legal and standards environment.
- Rich Web client activities focus on Web-based applications that extend the user experience on the Web beyond static HTML. Often these applications support compound documents that combine multiple formats such as Extensible Hypertext Markup Language, SVG, Synchronized Media Integration Language (SMIL), and XForms.
- Security group activities focus on the security context of Web applications that prevent surfers on the Web from being deceived and defrauded.
- Semantic Web activities focus on the creation of a universal medium for the exchange of data. Ideally both personal information and enterprise applications would be smoothly integrated creating a global sharing of commercial, scientific, and cultural data.
- Style activities focus on style components of Web pages. This group is working to extend style to other types of documents such as XML, SVG, and SMIL.
- Synchronized multimedia activities focus on choreographing multimedia presentations where audio, video, text, and graphics are combined in real time. A Timed Text working group is working in designing formats for streaming text synchronized with other timed media.
- Ubiquitous Web applications activities focus on enabling value-added services and business models for ubiquitous networked devices.
- Voice browser activities aim at technologies for capturing and producing speech and managing the dialog between user and computer.
- Web content accessibility initiative (WAI) focuses on accessibility to the Web resources, while the WAI technical group focuses on the technical aspects in three areas: Web content, user agent, and authoring tool accessibility.
- Web services activities promote the technology of Web services as a standard means of interoperability between different software applications running on a variety of platforms and frameworks. This group is designing the infrastructure, architecture, and core technologies for Web services.
- XForms activity focuses on the use of Web forms for the collection and distribution of information.

New areas of interest and activity are incorporated into the W3C by a recommendation development process.

RECOMMENDATION DEVELOPMENT PROCESS

The W3C strives to create high-quality standards for the Web though the strategy of community consensus. The general public is encouraged to join the members of the W3C in coming to consensus of Web standards and protocols. A new development begins with expressions of interest in a topic, and perhaps a workshop that brings interested persons together for the exchange ideas and information. Next, a proposal for a new activity describes the scope and duration of the project, as well as working groups, interest groups, and coordination groups that will carry out the work. As the development work proceeds, candidate specifications and guidelines cycle in revision and review as they advance to recommendation status. This may be a long process that may
ultimately lead to a success recommendation, or may result in the abandonment of the proposal. As successful proposals near completion, a W3C advisory committee finally examines the mature proposal and issues a recommendation.

CONCLUSION

The W3C has had a major influence on the creation of the World Wide Web, arguably the most important public information utility of our age. It has taken as its mission the development of the Web in the manner of consensus building with the aim of creating the greatest benefit for the greatest number. As the Web grows in size and sophistication, more and more Semantic Web applications are appearing.

REFERENCES