WIKIPEDIA

HTML

The **HyperText Markup Language**, or **HTML** is the standard markup language for documents designed to be displayed in a web browser. It can be assisted by technologies such as Cascading Style Sheets (CSS) and scripting languages such as JavaScript.

Web browsers receive HTML documents from a web server or from local storage and render the documents into multimedia web pages. HTML describes the structure of a web page semantically and originally included cues for the appearance of the document.

HTML elements are the building blocks of HTML pages. With HTML constructs, images and other objects such as interactive forms may be embedded into the rendered page. HTML provides a means to create structured documents by denoting structural semantics for text such as headings, paragraphs, lists, links, quotes and other items. HTML elements are delineated by *tags*, written using angle brackets. Tags such as `<img />` and `<input />` directly introduce content into the page. Other tags such as `<p>` surround and provide information about document text and may include other tags as sub-elements. Browsers do not display the HTML tags, but use them to interpret the content of the page.

HTML can embed programs written in a scripting language such as JavaScript, which affects the behavior and content of web pages. Inclusion of CSS defines the look and layout of content. The World Wide Web Consortium (W3C), former maintainer of the HTML and current maintainer of the CSS standards, has encouraged the use of CSS over explicit presentational HTML since 1997.[2]

HTML
(HyperText Markup Language)

The official logo of the latest version, HTML5[1]

Filename extension	`.html` `.htm`
Internet media type	`text/html`
Type code	TEXT
Uniform Type Identifier (UTI)	public.html
Developed by	WHATWG
Initial release	1993
Latest release	Living Standard (https://html.spec.whatwg.org/) (2021)
Type of format	Document file format
Container for	HTML elements
Contained by	Web browser
Extended from	SGML
Extended to	XHTML
Open format?	Yes
Website	html.spec.whatwg.org (https://html.spec.whatwg.org/)

Contents

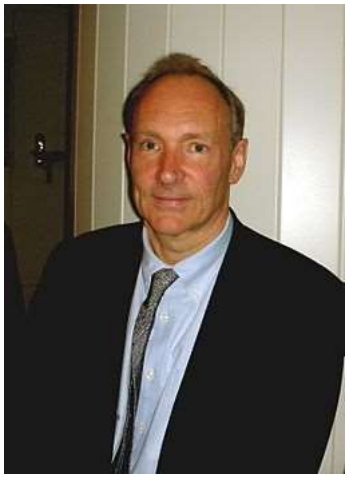
History

Development

In 1980, physicist Tim Berners-Lee, a contractor at CERN, proposed and prototyped ENQUIRE, a system for CERN researchers to use and share documents. In 1989, Berners-Lee wrote a memo proposing an Internet-based hypertext system.[3] Berners-Lee specified HTML and wrote the browser and server software in late 1990. That year, Berners-Lee and CERN data systems engineer Robert Cailliau collaborated on a joint request for funding, but the project was not formally adopted by CERN. In his personal notes[4] from 1990 he listed[5] "some of the many areas in which hypertext is used" and put an encyclopedia first.

The first publicly available description of HTML was a document called "HTML Tags" (http://info.cern.ch/hypertext/WWW/MarkUp/Tags.html), first mentioned on the Internet by Tim Berners-Lee in late 1991.[6][7] It describes 18 elements comprising the initial, relatively simple design of HTML. Except for the hyperlink tag, these were strongly influenced by SGMLguid, an in-house Standard Generalized Markup Language (SGML)-based documentation format at CERN. Eleven of these elements still exist in HTML 4.[8]

HTML is a markup language that web browsers use to interpret and compose text, images, and other material into visual or audible web pages. Default characteristics for every item of HTML markup are defined in the browser, and these characteristics can be altered or enhanced by the web page designer's additional use of CSS. Many of the text elements are found in the 1988 ISO technical report TR 9537 *Techniques for using SGML*, which in turn covers the features of early text formatting languages such as that used by the RUNOFF command developed in the early 1960s for the CTSS (Compatible Time-Sharing System) operating system: these formatting commands were derived from the commands used by typesetters to manually format documents. However, the SGML concept of generalized markup is based on elements (nested annotated ranges with attributes) rather than merely print effects, with also the separation of structure and markup; HTML has been progressively moved in this direction with CSS.



Tim Berners-Lee in April 2009

Berners-Lee considered HTML to be an application of SGML. It was formally defined as such by the Internet Engineering Task Force (IETF) with the mid-1993 publication of the first proposal for an HTML specification, the "Hypertext Markup Language (HTML)" Internet Draft by Berners-Lee and Dan Connolly, which included an SGML Document type definition to define the grammar.[9][10] The draft expired after six months, but was notable for its acknowledgment of the NCSA Mosaic browser's custom tag for embedding in-line images, reflecting the IETF's philosophy of basing standards on successful prototypes. Similarly, Dave Raggett's competing Internet-Draft, "HTML+ (Hypertext Markup Format)", from late 1993, suggested standardizing already-implemented features like tables and fill-out forms.[11]

After the HTML and HTML+ drafts expired in early 1994, the IETF created an HTML Working Group, which in 1995 completed "HTML 2.0", the first HTML specification intended to be treated as a standard against which future implementations should be based.[12]

Further development under the auspices of the IETF was stalled by competing interests. Since 1996, the HTML specifications have been maintained, with input from commercial software vendors, by the World Wide Web Consortium (W3C).[13] However, in 2000, HTML also became an international standard (ISO/IEC 15445:2000). HTML 4.01 was published in late 1999, with further errata published through 2001. In 2004, development began on HTML5 in the Web Hypertext Application Technology Working Group (WHATWG), which became a joint deliverable with the W3C in 2008, and completed and standardized on 28 October 2014.[14]

HTML versions timeline

HTML 2

November 24, 1995

HTML 2.0 was published as RFC 1866 (https://datatracker.ietf.org/doc/html/rfc1866). Supplemental RFCs added capabilities:

- November 25, 1995: RFC 1867 (https://datatracker.ietf.org/doc/html/rfc1867) (form-based file upload)
- May 1996: RFC 1942 (https://datatracker.ietf.org/doc/html/rfc1942) (tables)

- August 1996: RFC 1980 (https://datatracker.ietf.org/doc/html/rfc1980) (client-side image maps)
- January 1997: RFC 2070 (https://datatracker.ietf.org/doc/html/rfc2070) (internationalization)

HTML 3

January 14, 1997

HTML 3.2[15] was published as a W3C Recommendation. It was the first version developed and standardized exclusively by the W3C, as the IETF had closed its HTML Working Group on September 12, 1996.[16]
Initially code-named "Wilbur",[17] HTML 3.2 dropped math formulas entirely, reconciled overlap among various proprietary extensions and adopted most of Netscape's visual markup tags. Netscape's blink element and Microsoft's marquee element were omitted due to a mutual agreement between the two companies.[13] A markup for mathematical formulas similar to that in HTML was not standardized until 14 months later in MathML.

HTML 4

December 18, 1997

HTML 4.0[18] was published as a W3C Recommendation. It offers three variations:

- Strict, in which deprecated elements are forbidden
- Transitional, in which deprecated elements are allowed
- Frameset, in which mostly only frame related elements are allowed.

Initially code-named "Cougar",[17] HTML 4.0 adopted many browser-specific element types and attributes, but at the same time sought to phase out Netscape's visual markup features by marking them as deprecated in favor of style sheets. HTML 4 is an SGML application conforming to ISO 8879 – SGML.[19]

April 24, 1998

HTML 4.0[20] was reissued with minor edits without incrementing the version number.

December 24, 1999

HTML 4.01[21] was published as a W3C Recommendation. It offers the same three variations as HTML 4.0 and its last errata (https://www.w3.org/MarkUp/html4-updates/erratum) were published on May 12, 2001.

May 2000

ISO/IEC 15445:2000[22][23] ("ISO HTML", based on HTML 4.01 Strict) was published as an ISO/IEC international standard. In the ISO this standard falls in the domain of the ISO/IEC JTC1/SC34 (ISO/IEC Joint Technical Committee 1, Subcommittee 34 – Document description and processing languages).[22]
After HTML 4.01, there was no new version of HTML for many years as development of the parallel, XML-based language XHTML occupied the W3C's HTML Working Group through the early and mid-2000s.

HTML 5

October 28, 2014

HTML5[24] was published as a W3C Recommendation.[25]

November 1, 2016

HTML 5.1[26] was published as a W3C Recommendation.[27][28]

December 14, 2017

HTML 5.2[29] was published as a W3C Recommendation.[30][31]

HTML draft version timeline

October 1991

HTML Tags,[6] an informal CERN document listing 18 HTML tags, was first mentioned in public.

June 1992

First informal draft of the HTML DTD,[32] with seven[33][34][35] subsequent revisions (July 15, August 6, August 18, November 17, November 19, November 20, November 22)

November 1992

HTML DTD 1.1 (the first with a version number, based on RCS revisions, which start with 1.1 rather than 1.0), an informal draft[35]

June 1993

Hypertext Markup Language[36] was published by the IETF IIIR Working Group as an Internet Draft (a rough proposal for a standard). It was replaced by a second version[37] one month later.

November 1993

HTML+ (https://www.w3.org/MarkUp/HTMLPlus/htmlplus_1.html) was published by the IETF as an Internet Draft and was a competing proposal to the Hypertext Markup Language draft. It expired in July 1994.[38]

November 1994

First draft (revision 00) of HTML 2.0 published by IETF itself[39] (called as "HTML 2.0" from revision 02[40]), that finally led to publication of RFC 1866 (https://datatracker.ietf.org/doc/html/rfc1866) in November 1995.[41]

April 1995 (authored March 1995)

HTML 3.0[42] was proposed as a standard to the IETF, but the proposal expired five months later (28 September 1995)[43] without further action. It included many of the capabilities that were in Raggett's HTML+ proposal, such as support for tables, text flow around figures and the display of complex mathematical formulas.[43]

W3C began development of its own Arena browser as a test bed for HTML 3 and Cascading Style Sheets,[44][45][46] but HTML 3.0 did not succeed for several reasons. The draft was considered very large at 150 pages and the pace of browser development, as well as the number of interested parties, had outstripped the resources of the IETF.[13] Browser vendors, including Microsoft and Netscape at the time, chose to implement different subsets of HTML 3's draft features as well as to introduce their own extensions to it.[13] (see Browser wars). These included extensions to control stylistic aspects of documents, contrary to the "belief [of the academic engineering community] that such things as text color, background texture, font size and font face were definitely outside the scope of a language when their only intent was to specify how a document would be organized."[13] Dave Raggett, who has been a W3C Fellow for many years, has commented for example: "To a certain extent, Microsoft built its business on the Web by extending HTML features."[13]

January 2008

HTML5 was published as a Working Draft by the W3C.[47]

Although its syntax closely resembles that of SGML, HTML5 has abandoned any attempt to be an SGML application and has explicitly defined its own "html" serialization, in addition to an alternative XML-based XHTML5 serialization.[48]

2011 HTML5 – Last Call

On 14 February 2011, the W3C extended the charter of its HTML Working Group with clear milestones for HTML5. In May 2011, the working group advanced HTML5 to "Last Call", an invitation to communities inside and outside W3C to confirm the technical soundness of the