



WIKIPEDIA
The Free Encyclopedia

- [Main page](#)
- [Contents](#)
- [Featured content](#)
- [Current events](#)
- [Random article](#)
- [Donate to Wikipedia](#)

- ▼ Interaction
 - [Help](#)
 - [About Wikipedia](#)
 - [Community portal](#)
 - [Recent changes](#)
 - [Contact Wikipedia](#)

- ▶ [Toolbox](#)
- ▶ [Print/export](#)

- ▼ Languages
 - [Afrikaans](#)
 - [Alemannisch](#)
 - [العربية](#)
 - [Aragonés](#)
 - [Azərbaycanca](#)
 - [বাংলা](#)
 - [Башҡортса](#)
 - [Беларуская](#)
 - [беларуская \(тарашкевіца\)](#)
 - [Български](#)
 - [Boarisch](#)
 - [Bosanski](#)
 - [Brezhoneg](#)
 - [Català](#)
 - [ЧӀвашла](#)
 - [Česky](#)
 - [Corsu](#)
 - [Cymraeg](#)
 - [Dansk](#)
 - [Deutsch](#)
 - [Eesti](#)
 - [Ελληνικά](#)
 - [Español](#)
 - [Esperanto](#)
 - [Euskara](#)
 - [فارسی](#)
 - [Føroyskt](#)
 - [Français](#)
 - [Frysk](#)
 - [Furlan](#)
 - [Gaeilge](#)
 - [Galego](#)
 - [한국어](#)
 - [Հայերեն](#)
 - [हिन्दी](#)
 - [Hornjoserbsce](#)
 - [Hrvatski](#)
 - [Bahasa Indonesia](#)
 - [Interlingua](#)
 - [Íslenska](#)
 - [Italiano](#)
 - [עברית](#)
 - [Basa Jawa](#)
 - [ქართული](#)
 - [Қазақша](#)
 - [Kiswahili](#)
 - [Kurdî](#)

Article [Talk](#) Read [Edit](#) ▼ 

HTML

From Wikipedia, the free encyclopedia

For the use of HTML on Wikipedia, see [Help:HTML in wikitext](#).

HyperText Markup Language (HTML) is the main [markup language](#) for displaying web pages and other information that can be displayed in a web browser.

HTML is written in the form of [HTML elements](#) consisting of *tags* enclosed in [angle brackets](#) (like `<html>`), within the web page content. HTML tags most commonly come in pairs like `<h1>` and `</h1>`, although some tags, known as *empty elements*, are unpaired, for example ``. The first tag in a pair is the *start tag*, the second tag is the *end tag* (they are also called *opening tags* and *closing tags*). In between these tags web designers can add text, tags, comments and other types of text-based content.

The purpose of a [web browser](#) is to read HTML documents and compose them into visible or audible web pages. The browser does not display the HTML tags, but uses the tags to interpret the content of the page.

HTML elements form the building blocks of all websites. HTML allows [images and objects](#) to be embedded and can be used to create [interactive forms](#). It provides a means to create [structured documents](#) by denoting structural [semantics](#) for text such as headings, paragraphs, lists, links, quotes and other items. It can embed [scripts](#) in languages such as [JavaScript](#) which affect the behavior of HTML webpages.

Web browsers can also refer to [Cascading Style Sheets \(CSS\)](#) to define the appearance and layout of text and other material. The [W3C](#), maintainer of both the HTML and the CSS standards, encourages the use of CSS over explicitly presentational HTML markup.^[1]

HTML (HyperText Markup Language)

Filename extension	.html, .htm
Internet media type	text/html
Type code	TEXT
Uniform Type Identifier	public.html
Developed by	World Wide Web Consortium & WHATWG
Type of format	Markup language
Extended from	SGML
Extended to	XHTML
Standard(s)	ISO/IEC 15445 W3C HTML 4.01  W3C HTML5  (draft)

Contents [hide]

- 1 History
 - 1.1 Origins
 - 1.2 First specifications
 - 1.3 Version history of the standard
 - 1.3.1 HTML version timeline
 - 1.3.2 HTML draft version timeline
 - 1.3.3 XHTML versions
- 2 Markup
 - 2.1 Elements
 - 2.1.1 Element examples
 - 2.1.2 Attributes
 - 2.2 Character and entity references
 - 2.3 Data types
 - 2.4 Document type declaration
- 3 Semantic HTML
- 4 Delivery
 - 4.1 HTTP
 - 4.2 HTML e-mail
 - 4.3 Naming conventions
 - 4.4 HTML Application
- 5 Current variations
 - 5.1 SGML-based versus XML-based HTML
 - 5.2 Transitional versus strict
 - 5.3 Frameset versus transitional
 - 5.4 Summary of specification versions
- 6 Hypertext features not in HTML
- 7 WYSIWYG editors
- 8 See also
- 9 References
- 10 External links

- Кыргызча
- Latgaju
- Latviešu
- Lëtzebuergesch
- Lietuvių
- Lumbaart
- Magyar
- Македонски
- മലയാളം
- मराठी
- Bahasa Melayu
- Монгол
- Nederlands
- नेपाल भाषा
- 日本語
- norsk (bokmål)
- norsk (nynorsk)
- Олык марий
- Oʻzbekcha
- ਪੰਜਾਬੀ
- Polski
- Português
- Qaraqalpaqsha
- Română
- Русский
- Shqip
- Simple English
- Slovenčina
- Slovenščina
- Soomaaliga
- كوردی
- Српски / srpski
- Srpskohrvatski / српскохрватски
- Suomi
- Svenska
- Tagalog
- தமிழ்
- తెలుగు
- ไทย
- Тоҷикӣ
- Türkçe
- Türkmençe
- Українська
- اردو
- Tiếng Việt
- ਯਾਰ
- Yorùbá
- 粵語
- Žemaitėška
- 中文

History

[edit]

Origins

[edit]



In 1980, physicist **Tim Berners-Lee**, who was 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.^[2] Berners-Lee specified HTML and wrote the browser and server software in the last part of 1990. In 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^[3] from 1990 he lists^[4] "some of the many areas in which hypertext is used" and puts an encyclopedia first.

First specifications

[edit]

The first publicly available description of HTML was a document called "HTML Tags", first mentioned on the Internet by Berners-Lee in late 1991.^{[5][6]} 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 SGML based documentation format at CERN. Eleven of these elements still exist in HTML 4.^[7]



Tim Berners-Lee

HyperText Markup Language 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 processing; HTML has been progressively moved in this direction with **CSS**.

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: "**Hypertext Markup Language (HTML)**" **Internet-Draft** by Berners-Lee and **Dan Connolly**, which included an **SGML Document Type Definition** to define the grammar.^[8] The draft expired after six months, but was notable for its acknowledgement of the **NCSA Mosaic** browser's custom tag for embedding in-line images, reflecting the IETF's philosophy of basing standards on successful prototypes.^[9] 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.^[10]

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.^[9] Published as **Request for Comments** 1866, HTML 2.0 included ideas from the HTML and HTML+ drafts.^[11] The 2.0 designation was intended to distinguish the new edition from previous drafts.^[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.

Version history of the standard

[edit]

HTML version timeline

[edit]

November 24, 1995

HTML 2.0 was published as IETF **RFC** 1866. Supplemental **RFCs** added capabilities:

- November 25, 1995: **RFC** 1867 (form-based file upload)
- May 1996: **RFC** 1942 (tables)

HTML

- **HTML** and **HTML5**
- **Dynamic HTML**
- **XHTML**
 - **XHTML Basic**
 - **XHTML Mobile Profile** and **C-HTML**
- **Canvas element**
- **Character encodings**
- **Document Object Model**

- August 1996: [RFC 1980](#) (client-side image maps)
- January 1997: [RFC 2070](#) (internationalization)

January 1997

HTML 3.2^[14] 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 in September 1996.^[15]

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](#).

December 1997

HTML 4.0^[16] 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.^[18]

April 1998

HTML 4.0^[19] was reissued with minor edits without incrementing the version number.

December 1999

HTML 4.01^[20] was published as a W3C Recommendation. It offers the same three variations as HTML 4.0 and its last [errata](#) were published May 12, 2001.

May 2000

ISO/IEC 15445:2000^{[21][22]} ("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).ajmal.^[21]

As of mid-2008, HTML 4.01 and ISO/IEC 15445:2000 are the most recent versions of HTML. Development of the parallel, XML-based language XHTML occupied the W3C's HTML Working Group through the early and mid-2000s.

HTML draft version timeline

[\[edit\]](#)

October 1991

HTML Tags,^[5] an informal CERN document listing eighteen HTML tags, was first mentioned in public.

June 1992

First informal draft of the HTML DTD,^[23] with seven^{[24][25][26]} 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^[26]

June 1993

Hypertext Markup Language^[27] was published by the IETF IIR Working Group as an Internet-Draft (a rough proposal for a standard). It was replaced by a second version^[28] one month later, followed by six further drafts published by IETF itself^[29] that finally led to HTML 2.0 in RFC1866

November 1993

HTML+ was published by the IETF as an Internet-Draft and was a competing proposal to the Hypertext Markup Language draft. It expired in May 1994.

April 1995 (authored March 1995)

HTML 3.0^[30] was proposed as a standard to the IETF, but the proposal expired five months later without further action. It included many of the capabilities that were in

- Document Object Model
- Browser Object Model
- Font family
- HTML editor
- HTML element
- HTML Frames
- HTML5 Audio and HTML5 video
- HTML scripting
- Web browser engine
- Tag soup
- Quirks mode
- Style sheets
- Unicode and HTML
- W3C and WHATWG
- Web colors
- Web storage
- WebGL
- Comparison of
 - document markup languages
 - web browsers
 - layout engines for
 - HTML
 - HTML5
 - HTML5 canvas
 - HTML5 media
 - Non-standard HTML
 - XHTML (1.1)

V · T · E ·



Raggett's HTML+ proposal, such as support for tables, text flow around figures and the display of complex mathematical formulas.^[31]

W3C began development of its own [Arena browser](#) as a [test bed](#) for HTML 3 and Cascading Style Sheets,^{[32][33][34]} 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](#) ([link ↗](#)) by the W3C.^[35]

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.^[36]

May 2011

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 specification. The W3C is developing a comprehensive test suite to achieve broad interoperability for the full specification by 2014, which is now the target date for Recommendation.^[37]

XHTML versions

[\[edit\]](#)

Main article: [XHTML](#)

XHTML is a separate language that began as a reformulation of HTML 4.01 using [XML](#) 1.0. It continues to be developed:

- XHTML 1.0,^[38] published January 26, 2000, as a W3C Recommendation, later revised and republished August 1, 2002. It offers the same three variations as HTML 4.0 and 4.01, reformulated in XML, with minor restrictions.
- XHTML 1.1,^[39] published May 31, 2001, as a W3C Recommendation. It is based on XHTML 1.0 Strict, but includes minor changes, can be customized, is reformulated using modules from [Modularization of XHTML](#) [↗](#), which was published April 10, 2001, as a W3C Recommendation.
- XHTML 2.0 was a working draft, but work on it was abandoned in 2009 in favor of work on [HTML5](#) and [XHTML5](#).^{[40][41][42]} XHTML 2.0 was incompatible with XHTML 1.x and, therefore, would be more accurately characterized as an XHTML-inspired new language than an update to XHTML 1.x.
- XHTML5, which is an update to XHTML 1.x, is being defined alongside [HTML5](#) in the HTML5 draft.^[43]

Markup

[\[edit\]](#)

HTML markup consists of several key components, including *elements* (and their *attributes*), character-based *data types*, *character references* and *entity references*. Another important component is the *document type declaration*, which triggers [standards mode](#) rendering.

The following is an example of the classic [Hello world program](#), a common test employed for comparing [programming languages](#), [scripting languages](#) and [markup languages](#). This example is made using [9 lines of code](#):

```
<!DOCTYPE html>
<html>
  <head>
    <title>Hello HTML</title>
  </head>
  <body>
    <p>Hello World!</p>
  </body>
</html>
```

(The text between `<html>` and `</html>` describes the web page, and the text between `<body>` and `</body>` is the visible page content. The markup text `<title>Hello HTML</title>` defines the browser page title.)

This Document Type Declaration is for HTML5. If the `<!DOCTYPE html>` declaration is not included, various browsers will revert to "[quirks mode](#)" for rendering.^[44]

Elements

[\[edit\]](#)

Main article: [HTML element](#)

HTML documents are composed entirely of **HTML elements** that, in their most general form have three components: a pair of **tags**, a "start tag" and "end tag"; some **attributes** within the start tag; and finally, any textual and graphical *content* between the start and end tags, perhaps including other nested elements. The **HTML element** is everything between and including the start and end tags. Each **tag** is enclosed in angle brackets.

The general form of an **HTML element** is therefore: `<tag attribute1="value1" attribute2="value2">content</tag>`. Some HTML elements are defined as *empty elements* and take the form `<tag attribute1="value1" attribute2="value2" >`. Empty elements may enclose no content, for instance, the BR tag or the inline IMG tag. The name of an HTML element is the name used in the tags. Note that the end tag's name is preceded by a slash character, "/", and that in empty elements the end tag is neither required nor allowed. If attributes are not mentioned, default values are used in each case.

Element examples

[\[edit\]](#)

Header of the HTML document: `<head>...</head>`. Usually the title should be included in the head, for example:

```
<head>
  <title>The Title</title>
</head>
```

Headings: HTML headings are defined with the `<h1>` to `<h6>` tags:

```
<h1>Heading1</h1>
<h2>Heading2</h2>
<h3>Heading3</h3>
<h4>Heading4</h4>
<h5>Heading5</h5>
<h6>Heading6</h6>
```

Paragraphs:

```
<p>Paragraph 1</p> <p>Paragraph 2</p>
```

Line breaks: `
`. The difference between `
` and `<p>` is that 'br' breaks a line without altering the semantic structure of the page, whereas 'p' sections the page into paragraphs. Note also that 'br' is an *empty element* in that, while it may have attributes, it can take no content and it may not have an end tag.

```
<p>This <br /> is a paragraph <br /> with <br /> line breaks</p>
```

Comments:

```
<!-- This is a comment -->
```

Comments can help understanding of the markup and do not display in the webpage.

There are several types of markup elements used in HTML.

Structural markup describes the purpose of text

For example, `<h2>Golf</h2>` establishes "Golf" as a second-level [heading](#).

Structural markup does not denote any specific rendering, but most web browsers have default styles for element formatting. Content may be further styled using [Cascading Style Sheets](#) (CSS).

Presentational markup describes the appearance of the text, regardless of its purpose

For example `boldface` indicates that visual output devices should render "boldface" in bold text, but gives little indication what devices that are unable to do this (such as aural devices that read the text aloud) should do. In the case of both `bold` and `<i>italic</i>`, there are other elements that may have equivalent visual renderings but which are more semantic in nature, such as `strong text` and `emphasised text` respectively. It is easier to see how an aural user agent should interpret the latter two elements. However, they are not equivalent to their presentational counterparts: it would be undesirable for a screen-reader to emphasize the name of a book, for instance, but on a screen such a name would be italicized. Most presentational markup elements have become [deprecated](#) under the HTML 4.0 specification, in favor of using [CSS](#) for styling.

Hypertext markup makes parts of a document into links to other documents

An anchor element creates a [hyperlink](#) in the document and its `href` attribute sets the link's target [URL](#). For example the HTML markup,

```
<a href="http://www.google.com/">Wikipedia</a>, will render the word "Wikipedia" as a hyperlink. To render an image as a hyperlink, an 'img' element is inserted as content into the 'a' element. Like 'br', 'img' is an empty element with attributes but no content or closing tag. <a href="http://example.org"></a>.
```

Attributes

[edit]

Most of the attributes of an element are [name-value pairs](#), separated by "=" and written within the start tag of an element after the element's name. The value may be enclosed in single or double quotes, although values consisting of certain characters can be left unquoted in HTML (but not XHTML).^{[45][46]} Leaving attribute values unquoted is considered unsafe.^[47] In contrast with name-value pair attributes, there are some attributes that affect the element simply by their presence in the start tag of the element,^[5] like the `ismap` attribute for the `img` element.^[48]

There are several common attributes that may appear in many elements:

- The `id` attribute provides a document-wide unique identifier for an element. This is used to identify the element so that stylesheets can alter its presentational properties, and scripts may alter, animate or delete its contents or presentation. Appended to the URL of the page, it provides a globally unique identifier for the element, typically a sub-section of the page. For example, the ID "Attributes" in <http://en.wikipedia.org/wiki/HTML#Attributes>
- The `class` attribute provides a way of classifying similar elements. This can be used for [semantic](#) or presentation purposes. For example, an HTML document might semantically use the designation `class="notation"` to indicate that all elements with this class value are subordinate to the main text of the document. In presentation, such elements might be gathered together and presented as footnotes on a page instead of appearing in the place where they occur in the HTML source. Class attributes are used semantically in [microformats](#). Multiple class values may be specified; for example `class="notation important"` puts the element into both the 'notation' and the 'important' classes.
- An author may use the `style` attribute to assign presentational properties to a particular element. It is considered better practice to use an element's `id` or `class` attributes to select the element from within a stylesheet, though sometimes this can be too cumbersome for a simple, specific, or ad hoc styling.
- The `title` attribute is used to attach subtexual explanation to an element. In most browsers this attribute is displayed as a [tooltip](#).
- The `lang` attribute identifies the natural language of the element's contents, which may be different from that of the rest of the document. For example, in an English-language document:

```
<p>Oh well, <span lang="fr">c'est la vie</span>, as they say in France.</p>
```

The abbreviation element, `abbr`, can be used to demonstrate some of these attributes:

```
<abbr id="anId" class="jargon" style="color:purple;" title="Hypertext Markup Language">HTML</abbr>
```

This example displays as [HTML](#); in most browsers, pointing the cursor at the abbreviation should display the title text "Hypertext Markup Language."

Most elements also take the language-related attribute `dir` to specify text direction, such as with "rtl" for right-to-left text in, for examples, [Arabic](#), [Persian](#) or [Hebrew](#).^[49]

Character and entity references

[edit]

See also: [List of XML and HTML character entity references](#) and [Unicode and HTML](#)

As of version 4.0, HTML defines a set of 252 [character entity references](#) and a set of 1,114,050 [numeric character references](#), both of which allow individual characters to be written via simple markup, rather than literally. A literal character and its markup counterpart are considered equivalent and are rendered identically.

The ability to "escape" characters in this way allows for the characters `<` and `&` (when written as `<`; and `&`; respectively) to be interpreted as character data, rather than markup. For example, a literal `<` normally indicates the start of a tag, and `&` normally indicates the start of a character entity reference or numeric character reference; writing it as `&`; or `&`; or `&`; allows `&` to be included in the content of an element or in the value of an attribute. The double-quote character (`"`), when not used to quote an attribute value, must also be escaped as `"`; or `"`; or `"`; when it appears within the attribute value itself. Equivalently, the single-quote character (`'`), when not used to quote an attribute value, must also be escaped as `'`; or `'`; (not as `'`; except in XHTML documents^[50]) when it appears within the attribute value itself. If document authors overlook the need to escape such characters, some browsers can be very forgiving and try to use context to guess their intent. The result is still invalid markup, which makes the document less accessible to other browsers and to other [user agents](#) that may try to parse the document for [search and indexing](#) purposes for example.

Escaping also allows for characters that are not easily typed, or that are not available in the document's [character encoding](#), to be represented within element and attribute

content. For example, the acute-accented e (é), a character typically found only on Western European keyboards, can be written in any HTML document as the entity reference `é`, or as the numeric references `é` or `é`, using characters that are available on all keyboards and are supported in all character encodings. Unicode character encodings such as UTF-8 are compatible with all modern browsers and allow direct access to almost all the characters of the world's writing systems.^[51]

Data types [\[edit\]](#)

HTML defines several **data types** for element content, such as script data and stylesheet data, and a plethora of types for attribute values, including IDs, names, URIs, numbers, units of length, languages, media descriptors, colors, character encodings, dates and times, and so on. All of these data types are specializations of character data.

Document type declaration [\[edit\]](#)

HTML documents are required to start with a **Document Type Declaration** (informally, a "doctype"). In browsers, the doctype helps to define the rendering mode—particularly whether to use **quirks mode**.

The original purpose of the doctype was to enable parsing and validation of HTML documents by SGML tools based on the **Document Type Definition** (DTD). The DTD to which the DOCTYPE refers contains a machine-readable grammar specifying the permitted and prohibited content for a document conforming to such a DTD. Browsers, on the other hand, do not implement HTML as an application of SGML and by consequence do not read the DTD. **HTML5** does not define a DTD, because of the technology's inherent limitations, so in HTML5 the doctype declaration, `<!doctype html>`, does not refer to a DTD.

An example of an HTML 4 doctype is

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN" "http://www.w3.org/TR/html4/strict.dtd">
```

This declaration references the DTD for the 'strict' version of HTML 4.01. SGML-based validators read the DTD in order to properly parse the document and to perform validation. In modern browsers, a valid doctype activates standards mode as opposed to **quirks mode**.

In addition, HTML 4.01 provides Transitional and Frameset DTDs, [as explained below](#).

Semantic HTML [\[edit\]](#)

Main article: [Semantic HTML](#)

Semantic HTML is a way of writing HTML that emphasizes the meaning of the encoded information over its presentation (look). HTML has included semantic markup from its inception,^[52] but has also included presentational markup such as ``, `<i>` and `<center>` tags. There are also the semantically neutral **span** and **div** tags. Since the late 1990s when **Cascading Style Sheets** were beginning to work in most browsers, web authors have been encouraged to avoid the use of presentational HTML markup with a view to the **separation of presentation and content**.^[53]

In a 2001 discussion of the **Semantic Web**, Tim Berners-Lee and others gave examples of ways in which intelligent software 'agents' may one day automatically trawl the Web and find, filter and correlate previously unrelated, published facts for the benefit of human users.^[54] Such agents are not commonplace even now, but some of the ideas of **Web 2.0**, **mashups** and **price comparison websites** may be coming close. The main difference between these web application hybrids and Berners-Lee's semantic agents lies in the fact that the current **aggregation** and hybridisation of information is usually designed in by web developers, who already know the web locations and the **API semantics** of the specific data they wish to mash, compare and combine.

An important type of web agent that does trawl and read web pages automatically, without prior knowledge of what it might find, is the **Web crawler** or search-engine spider. These software agents are dependent on the semantic clarity of web pages they find as they use various techniques and **algorithms** to read and index millions of web pages a day and provide web users with **search facilities** without which the World Wide Web would be only a fraction of its current usefulness.

In order for search-engine spiders to be able to rate the significance of pieces of text they find in HTML documents, and also for those creating mashups and other hybrids as well as for more automated agents as they are developed, the semantic structures that exist in HTML need to be widely and uniformly applied to bring out the meaning of published text.^[55]

Presentational markup tags are **deprecated** in current HTML and XHTML recommendations and are illegal in HTML5.

Good semantic HTML also improves the **accessibility** of web documents (see also **Web Content Accessibility Guidelines**). For example, when a screen reader or audio browser can correctly ascertain the structure of a document, it will not waste the visually

impaired user's time by reading out repeated or irrelevant information when it has been marked up correctly.

Delivery [edit]

HTML documents can be delivered by the same means as any other computer file. However, they are most often delivered either by [HTTP](#) from a [web server](#) or by [email](#).

HTTP [edit]

Main article: [Hypertext Transfer Protocol](#)

The [World Wide Web](#) is composed primarily of HTML documents transmitted from web servers to web browsers using the [Hypertext Transfer Protocol](#) (HTTP). However, HTTP is used to serve images, sound, and other content, in addition to HTML. To allow the Web browser to know how to handle each document it receives, other information is transmitted along with the document. This [meta data](#) usually includes the [MIME type](#) (e.g. `text/html` or `application/xhtml+xml`) and the character encoding (see [Character encoding in HTML](#)).

In modern browsers, the MIME type that is sent with the HTML document may affect how the document is initially interpreted. A document sent with the XHTML MIME type is expected to be [well-formed XML](#); syntax errors may cause the browser to fail to render it. The same document sent with the HTML MIME type might be displayed successfully, since some browsers are more lenient with HTML.

The W3C recommendations state that XHTML 1.0 documents that follow guidelines set forth in the recommendation's Appendix C may be labeled with either MIME Type.^[56] The current XHTML 1.1 Working Draft also states that XHTML 1.1 documents should^[57] be labeled with either MIME type.^[58]

HTML e-mail [edit]

Main article: [HTML email](#)

Most graphical email clients allow the use of a subset of HTML (often ill-defined) to provide formatting and [semantic](#) markup not available with [plain text](#). This may include typographic information like coloured headings, emphasized and quoted text, inline images and diagrams. Many such clients include both a [GUI](#) editor for composing HTML e-mail messages and a rendering engine for displaying them. Use of HTML in e-mail is controversial because of compatibility issues, because it can help disguise [phishing](#) attacks, because of accessibility issues for blind or visually impaired people, because it can confuse [spam](#) filters and because the message size is larger than plain text.

Naming conventions [edit]

The most common [filename extension](#) for files containing HTML is `.html`. A common abbreviation of this is `.htm`, which originated because some early operating systems and file systems, such as [DOS](#) and [FAT](#), limited file extensions to [three letters](#).

HTML Application [edit]

Main article: [HTML Application](#)

An HTML Application (HTA; file extension ".hta") is a [Microsoft Windows](#) application that uses HTML and Dynamic HTML in a browser to provide the application's graphical interface. A regular HTML file is confined to the [security model](#) of the [web browser's security](#), communicating only to web servers and manipulating only webpage objects and [site cookies](#). An HTA runs as a fully trusted application and therefore has more privileges, like creation/editing/removal of files and [Windows Registry](#) entries. Because they operate outside the browser's security model, HTAs cannot be executed via HTTP, but must be downloaded (just like an [EXE file](#)) and executed from local file system.

Current variations [edit]

Since its inception, HTML and its associated protocols gained acceptance relatively quickly. However, no clear standards existed in the early years of the language. Though its creators originally conceived of HTML as a semantic language devoid of presentation details,^[60] practical uses pushed many presentational elements and attributes into the language, driven largely by the various browser vendors. The latest standards surrounding HTML reflect efforts to overcome the sometimes chaotic development of the language^[61] and to create a rational foundation for building both meaningful and well-presented documents. To return HTML to its role as a semantic language, the [W3C](#) has developed style languages such as [CSS](#) and [XSL](#) to shoulder the burden of presentation. In conjunction, the HTML specification has slowly reined in the presentational elements.

There are two axes differentiating various variations of HTML as currently specified: SGML-based HTML versus XML-based HTML (referred to as XHTML) on one axis, and

HTML is precisely what we were trying to PREVENT—ever-breaking links, links going outward only, quotes you can't follow to their origins, no version management, no rights management.

[Ted Nelson](#)^[59]

strict versus transitional (loose) versus frameset on the other axis.

SGML-based versus XML-based HTML

[edit]

One difference in the latest HTML specifications lies in the distinction between the SGML-based specification and the XML-based specification. The XML-based specification is usually called **XHTML** to distinguish it clearly from the more traditional definition. However, the root element name continues to be 'html' even in the XHTML-specified HTML. The W3C intended XHTML 1.0 to be identical to HTML 4.01 except where limitations of XML over the more complex SGML require workarounds. Because XHTML and HTML are closely related, they are sometimes documented in parallel. In such circumstances, some authors [conflate the two names](#) as (X)HTML or X(HTML).

Like HTML 4.01, XHTML 1.0 has three sub-specifications: strict, transitional and frameset.

Aside from the different opening declarations for a document, the differences between an HTML 4.01 and XHTML 1.0 document—in each of the corresponding DTDs—are largely syntactic. The underlying syntax of HTML allows many shortcuts that XHTML does not, such as elements with optional opening or closing tags, and even EMPTY elements which must not have an end tag. By contrast, XHTML requires all elements to have an opening tag and a closing tag. XHTML, however, also introduces a new shortcut: an XHTML tag may be opened and closed within the same tag, by including a slash before the end of the tag like this: `
`. The introduction of this shorthand, which is not used in the SGML declaration for HTML 4.01, may confuse earlier software unfamiliar with this new convention. A fix for this is to include a space before closing the tag, as such: `
`.^[62]

To understand the subtle differences between HTML and XHTML, consider the transformation of a valid and well-formed XHTML 1.0 document that adheres to Appendix C (see below) into a valid HTML 4.01 document. To make this translation requires the following steps:

1. **The language for an element should be specified with a lang attribute rather than the XHTML `xml:lang` attribute.** XHTML uses XML's built in language-defining functionality attribute.
2. **Remove the XML namespace (`xmlns=URI`).** HTML has no facilities for namespaces.
3. **Change the document type declaration** from XHTML 1.0 to HTML 4.01. (see [DTD section](#) for further explanation).
4. If present, **remove the XML declaration.** (Typically this is: `<?xml version="1.0" encoding="utf-8"??>`).
5. **Ensure that the document's MIME type is set to `text/html`.** For both HTML and XHTML, this comes from the HTTP `Content-Type` header sent by the server.
6. **Change the XML empty-element syntax to an HTML style empty element** (`
` to `
`).

Those are the main changes necessary to translate a document from XHTML 1.0 to HTML 4.01. To translate from HTML to XHTML would also require the addition of any omitted opening or closing tags. Whether coding in HTML or XHTML it may just be best to always include the optional tags within an HTML document rather than remembering which tags can be omitted.

A well-formed XHTML document adheres to all the syntax requirements of XML. A valid document adheres to the content specification for XHTML, which describes the document structure.

The W3C recommends several conventions to ensure an easy migration between HTML and XHTML (see [HTML Compatibility Guidelines](#)). The following steps can be applied to XHTML 1.0 documents only:

- Include both `xml:lang` and `lang` attributes on any elements assigning language.
- Use the empty-element syntax only for elements specified as empty in HTML.
- Include an extra space in empty-element tags: for example `
` instead of `
`.
- Include explicit close tags for elements that permit content but are left empty (for example, `<div></div>`, not `<div />`).
- Omit the XML declaration.

By carefully following the W3C's compatibility guidelines, a user agent should be able to interpret the document equally as HTML or XHTML. For documents that are XHTML 1.0 and have been made compatible in this way, the W3C permits them to be served either as HTML (with a `text/html` MIME type), or as XHTML (with an `application/xhtml+xml` or `application/xml` MIME type). When delivered as XHTML, browsers should use an XML parser, which adheres strictly to the XML specifications for parsing the document's contents.

Transitional versus strict

[edit]

HTML 4 defined three different versions of the language: Strict, Transitional (once called

Loose) and Frameset. The Strict version is intended for new documents and is considered best practice, while the Transitional and Frameset versions were developed to make it easier to transition documents that conformed to older HTML specification or didn't conform to any specification to a version of HTML 4. The Transitional and Frameset versions allow for **presentational markup**, which is omitted in the Strict version. Instead, **cascading style sheets** are encouraged to improve the presentation of HTML documents. Because XHTML 1 only defines an XML syntax for the language defined by HTML 4, the same differences apply to XHTML 1 as well.

The Transitional version allows the following parts of the vocabulary, which are not included in the Strict version:

- **A looser content model**

- inline elements and plain text are allowed directly in: `body`, `blockquote`, `form`, `noscript` and `noframes`

- **Presentation related elements**

- `u` (underline) (Deprecated. can confuse a visitor with a hyperlink.)
- `s` (strike-through)
- `center` (Deprecated. use CSS instead.)
- `font` (Deprecated. use CSS instead.)
- `basefont` (Deprecated. use CSS instead.)

- **Presentation related attributes**

- `background` (Deprecated. use CSS instead.) and `bgcolor` (Deprecated. use CSS instead.) attributes for `body` (required element according to the W3C.) element.
- `align` (Deprecated. use CSS instead.) attribute on `div`, `form`, `paragraph` (`p`) and heading (`h1...h6`) elements
- `align` (Deprecated. use CSS instead.), `noshade` (Deprecated. use CSS instead.), `size` (Deprecated. use CSS instead.) and `width` (Deprecated. use CSS instead.) attributes on `hr` element
- `align` (Deprecated. use CSS instead.), `border`, `vspace` and `hspace` attributes on `img` and `object` (caution: the `object` element is only supported in Internet Explorer (from the major browsers)) elements
- `align` (Deprecated. use CSS instead.) attribute on `legend` and `caption` elements
- `align` (Deprecated. use CSS instead.) and `bgcolor` (Deprecated. use CSS instead.) on `table` element
- `nowrap` (Obsolete), `bgcolor` (Deprecated. use CSS instead.), `width`, `height` on `td` and `th` elements
- `bgcolor` (Deprecated. use CSS instead.) attribute on `tr` element
- `clear` (Obsolete) attribute on `br` element
- `compact` attribute on `dl`, `dir` and `menu` elements
- `type` (Deprecated. use CSS instead.), `compact` (Deprecated. use CSS instead.) and `start` (Deprecated. use CSS instead.) attributes on `ol` and `ul` elements
- `type` and `value` attributes on `li` element
- `width` attribute on `pre` element

- **Additional elements in Transitional specification**

- `menu` (Deprecated. use CSS instead.) list (no substitute, though unordered list is recommended)
- `dir` (Deprecated. use CSS instead.) list (no substitute, though unordered list is recommended)
- `isindex` (Deprecated.) (element requires server-side support and is typically added to documents server-side, `form` and `input` elements can be used as a substitute)
- `applet` (Deprecated. use the `object` element instead.)

- **The `language` (Obsolete) attribute on `script` element** (redundant with the `type` attribute).

- **Frame related entities**

- `iframe`
- `noframes`
- `target` (Deprecated in the `map`, `link` and `form` elements.) attribute on `a`, client-side image-map (`map`), `link`, `form` and `base` elements

The Frameset version includes everything in the Transitional version, as well as the `frameset` element (used instead of `body`) and the `frame` element.

Frameset versus transitional

[\[edit\]](#)

In addition to the above transitional differences, the frameset specifications (whether XHTML 1.0 or HTML 4.01) specifies a different content model, with `frameset` replacing `body`, that contains either `frame` elements, or optionally `noframes` with a `body`.

Summary of specification versions [edit]

As this list demonstrates, the loose versions of the specification are maintained for legacy support. However, contrary to popular misconceptions, the move to XHTML does not imply a removal of this legacy support. Rather the X in XML stands for extensible and the W3C is modularizing the entire specification and opening it up to independent extensions. The primary achievement in the move from XHTML 1.0 to XHTML 1.1 is the modularization of the entire specification. The strict version of HTML is deployed in XHTML 1.1 through a set of modular extensions to the base XHTML 1.1 specification. Likewise, someone looking for the loose (transitional) or frameset specifications will find similar extended XHTML 1.1 support (much of it is contained in the legacy or frame modules). The modularization also allows for separate features to develop on their own timetable. So for example, XHTML 1.1 will allow quicker migration to emerging XML standards such as **MathML** (a presentational and semantic math language based on XML) and **XForms**—a new highly advanced web-form technology to replace the existing HTML forms.

In summary, the HTML 4.01 specification primarily reined in all the various HTML implementations into a single clearly written specification based on SGML. XHTML 1.0, ported this specification, as is, to the new XML defined specification. Next, XHTML 1.1 takes advantage of the extensible nature of XML and modularizes the whole specification. XHTML 2.0 will be the first step in adding new features to the specification in a standards-body-based approach.

Hypertext features not in HTML [edit]

HTML lacks some of the features found in earlier hypertext systems, such as **typed links**, **source tracking**, **fat links** and others.^[63] Even some hypertext features that were in early versions of HTML have been ignored by most popular web browsers until recently, such as the link element and in-browser Web page editing.

Sometimes Web services or browser manufacturers remedy these shortcomings. For instance, **wikis** and **content management systems** allow surfers to edit the Web pages they visit.

WYSIWYG editors [edit]

There are some **WYSIWYG** editors (What You See Is What You Get), in which the user lays out everything as it is to appear in the HTML document using a **graphical user interface** (GUI), where the editor renders this as an HTML document, no longer requiring the author to have extensive knowledge of HTML.

The WYSIWYG editing model has been criticized,^{[64][65]} primarily because of the low quality of the generated code; there are voices advocating a change to the **WYSIWYM** model (What You See Is What You Mean).

WYSIWYG editors remains a controversial topic because of their perceived flaws such as:

- Relying mainly on layout as opposed to meaning, often using markup that does not convey the intended meaning but simply copies the layout.^[66]
- Often producing extremely verbose and redundant code that fails to make use of the cascading nature of HTML and **CSS**.
- Often producing ungrammatical markup often called **tag soup**.
- As a great deal of the information in HTML documents is not in the layout, the model has been criticized for its "what you see is all you get"-nature.^[67]

Nevertheless, since WYSIWYG editors offer convenience over hand-coded pages as well as not requiring the author to know the finer details of HTML, they still dominate web authoring.^[citation needed]

See also [edit]

- **Breadcrumb (navigation)**
- **CSS**
- **Dynamic web page**
- **HTML5**
- **HTML decimal character rendering**
- **HTTP**
- **List of document markup languages**
- **Microformat**
- **SGML**
- **XML**

References [edit]

- ↑ **HTML 4 — Conformance: requirements and recommendations** . W3.org. Retrieved on 2012-02-16.
- ↑ Tim Berners-Lee, "Information Management: A Proposal." CERN (March 1989, May 1990). W3.org
- ↑ Tim Berners-Lee, "Design Issues"

4. [^] ^a ^b ^c Tim Berners-Lee, "Design Issues" [↗](#)
5. [^] ^a ^b ^c "Tags used in HTML" [↗](#). World Wide Web Consortium. November 3, 1992. Retrieved November 16, 2008.
6. [^] "First mention of HTML Tags on the www-talk mailing list" [↗](#). World Wide Web Consortium. October 29, 1991. Retrieved April 8, 2007.
7. [^] "Index of elements in HTML 4" [↗](#). World Wide Web Consortium. December 24, 1999. Retrieved April 8, 2007.
8. [^] Tim Berners-Lee (December 9, 1991). "Re: SGML/HTML docs, XBrowser (archived www-talk mailing list post)" [↗](#). Retrieved June 16, 2007. "SGML is very general. HTML is a specific application of the SGML basic syntax applied to hypertext documents with simple structure."
9. [^] ^a ^b Raymond, Eric. "IETF and the RFC Standards Process" [↗](#). [↗](#) [The Art of Unix Programming](#). http://www.faqs.org/docs/artu/ietf_process.html [↗](#). "In IETF tradition, standards have to arise from experience with a working prototype implementation — but once they become standards, code that does not conform to them is considered broken and mercilessly scrapped. ...Internet-Drafts are not specifications; software implementers and vendors are specifically barred from claiming compliance with them as if they were specifications. Internet-Drafts are focal points for discussion, usually in a working group... Once an Internet-Draft has been published with an RFC number, it is a specification to which implementers may claim conformance. It is expected that the authors of the RFC and the community at large will begin correcting the specification with field experience."
10. [^] "HTML+ Internet-Draft — Abstract" [↗](#). "Browser writers are experimenting with extensions to HTML and it is now appropriate to draw these ideas together into a revised document format. The new format is designed to allow a gradual roll over from HTML, adding features like tables, captioned figures and fill-out forms for querying remote databases or mailing questionnaires."
11. [^] Berners-Lee, Tim; Connolly, Dan (November 1995). "RFC 1866: Hypertext Markup Language – 2.0 – Acknowledgments" [↗](#). Internet Engineering Task Force. Retrieved 1 December 2010. "Since 1993, a wide variety of Internet participants have contributed to the evolution of HTML, which has included the addition of in-line images introduced by the NCSA Mosaic software for WWW. Dave Raggett played an important role in deriving the forms material from the HTML+ specification. Dan Connolly and Karen Olson Muldrow rewrote the HTML Specification in 1994. The document was then edited by the HTML working group as a whole, with updates being made by Eric Schieler, Mike Knezdovich and Eric W. Sink at Spyglass, Inc. Finally, Roy Fielding restructured the entire draft into its current form."
12. [^] Berners-Lee, Tim; Connolly, Dan (November 1995). "RFC 1866: Hypertext Markup Language – 2.0 – Introduction" [↗](#). Internet Engineering Task Force. Retrieved 1 December 2010. "This document thus defines an HTML 2.0 (to distinguish it from the previous informal specifications). Future (generally upwardly compatible) versions of HTML with new features will be released with higher version numbers."
13. [^] ^a ^b ^c ^d ^e ^f Raggett, Dave (1998). *Raggett on HTML 4* [↗](#). Retrieved July 9, 2007.
14. [^] "HTML 3.2 Reference Specification" [↗](#). World Wide Web Consortium. January 14, 1997. Retrieved November 16, 2008.
15. [^] "IETF HTML WG" [↗](#). Retrieved June 16, 2007. "Note: This working group is closed"
16. [^] "HTML 4.0 Specification" [↗](#). World Wide Web Consortium. December 18, 1997. Retrieved November 16, 2008.
17. [^] Arnoud Engelfriet. "Introduction to Wilbur" [↗](#). Web Design Group. Retrieved June 16, 2007.
18. [^] "HTML 4 – 4 Conformance: requirements and recommendations" [↗](#). Retrieved December 30, 2009.
19. [^] "HTML 4.0 Specification" [↗](#). World Wide Web Consortium. April 24, 1998. Retrieved November 16, 2008.
20. [^] "HTML 4.01 Specification" [↗](#). World Wide Web Consortium. December 24, 1999. Retrieved November 16, 2008.
21. [^] ^a ^b ISO (2000). "ISO/IEC 15445:2000 – Information technology – Document description and processing languages – HyperText Markup Language (HTML)" [↗](#). Retrieved December 26, 2009.
22. [^] Cs.Tcd.le [↗](#). Cs.Tcd.le (2000-05-15). Retrieved on 2012-02-16.
23. [^] Connolly, Daniel (6 June 1992). "MIME as a hypertext architecture" [↗](#). CERN. Retrieved 24 October 2010.
24. [^] Connolly, Daniel (15 July 1992). "HTML DTD enclosed" [↗](#). CERN. Retrieved 24 October 2010.
25. [^] Connolly, Daniel (18 August 1992). "document type declaration subset for HyperText Markup Language as defined by the World Wide Web project" [↗](#). CERN. Retrieved 24 October 2010.
26. [^] ^a ^b Connolly, Daniel (24 November 1992). "Document Type Definition for the HyperText Markup Language as used by the World Wide Web application" [↗](#). CERN. Retrieved 24 October 2010. See section "Revision History"
27. [^] Berners-Lee, Tim; Connolly, Daniel (June 1993). "Hypertext Markup Language (HTML) Internet Draft version 1.1" [↗](#). IETF IIR Working Group. Retrieved 18 September 2010.
28. [^] Berners-Lee, Tim; Connolly, Daniel (June 1993). "Hypertext Markup Language (HTML) Internet Draft version 1.2" [↗](#). IETF IIR Working Group. Retrieved 18 September 2010.
29. [^] Berners-Lee, Tim; Connolly, Daniel (28 November 1994). "HyperText Markup Language Specification – 2.0 INTERNET DRAFT" [↗](#). IETF. Retrieved 24 October 2010.
30. [^] "HTML 3.0 Draft (Expired!) Materials" [↗](#). World Wide Web Consortium. December 21, 1995. Retrieved November 16, 2008.
31. [^] "HyperText Markup Language Specification Version 3.0" [↗](#). Retrieved June 16, 2007.

32. [^] [Raggett, Dave \(28 March 1995\). "HyperText Markup Language Specification Version 3.0" ↗. *HTML 3.0 Internet Draft Expires in six months*. World Wide Web Consortium. Retrieved 17 June 2010.](#)
33. [^] [Bowers, Neil. *Web lint: Just Another Perl Hack* ↗.](#)
34. [^] [Lie, Håkon Wium; Bos, Bert \(April 1997\). *Cascading style sheets: designing for the Web* ↗. Addison Wesley Longman. p. 263. Retrieved 9 June 2010.](#)
35. [^] ["HTML5" ↗. World Wide Web Consortium. June 10, 2008. Retrieved November 16, 2008.](#)
36. [^] ["HTML5, one vocabulary, two serializations" ↗. Retrieved February 25, 2009.](#)
37. [^] ["W3C Confirms May 2011 for HTML5 Last Call, Targets 2014 for HTML5 Standard" ↗. World Wide Web Consortium. 14 February 2011. Retrieved 18 February 2011.](#)
38. [^] ["XHTML 1.0: The Extensible HyperText Markup Language \(Second Edition\)" ↗. World Wide Web Consortium. January 26, 2000. Retrieved November 16, 2008.](#)
39. [^] ["XHTML 1.1 – Module-based XHTML — Second Edition" ↗. World Wide Web Consortium. February 16, 2007. Retrieved November 16, 2008.](#)
40. [^] ["XHTML 2.0" ↗. World Wide Web Consortium. July 26, 2006. Retrieved November 16, 2008.](#)
41. [^] ["XHTML 2 Working Group Expected to Stop Work End of 2009, W3C to Increase Resources on HTML5" ↗. World Wide Web Consortium. July 17, 2009. Retrieved November 16, 2008.](#)
42. [^] ["W3C XHTML FAQ" ↗.](#)
43. [^] ["HTML5" ↗. World Wide Web Consortium. October 24, 2008. Retrieved November 16, 2008.](#)
44. [^] ["Activating Browser Modes with Doctype" ↗. Hsivonen.iki.fi. Retrieved on 2012-02-16.](#)
45. [^] ["On SGML and HTML" ↗. World Wide Web Consortium. Retrieved November 16, 2008.](#)
46. [^] ["XHTML 1.0 – Differences with HTML 4" ↗. World Wide Web Consortium. Retrieved November 16, 2008.](#)
47. [^] ["Korpela, Jukka \(July 6, 1998\). "Why attribute values should always be quoted in HTML" ↗. Cs.tut.fi. Retrieved November 16, 2008.](#)
48. [^] ["Objects, Images, and Applets in HTML documents" ↗. World Wide Web Consortium. December 24, 1999. Retrieved November 16, 2008.](#)
49. [^] ["H56: Using the dir attribute on an inline element to resolve problems with nested directional runs" ↗. *Techniques for WCAG 2.0*. W3C. Retrieved 18 September 2010.](#)
50. [^] ["The Named Character Reference &apost;" ↗. World Wide Web Consortium. January 26, 2000.](#)
51. [^] ["The Unicode Standard: A Technical Introduction" ↗. Retrieved 2010-03-16.](#)
52. [^] [Berners-Lee, Tim; Fischetti, Mark \(2000\). *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by Its Inventor*. San Francisco: Harper. ISBN \[\[Special:BookSources/978-0-06-251587-X|978-0-06-251587-X\]\].](#)
53. [^] [Raggett, Dave \(2002\). "Adding a touch of style" ↗. W3C. Retrieved October 2, 2009. This article notes that presentational HTML markup may be useful when targeting browsers "before Netscape 4.0 and Internet Explorer 4.0". See the \[list of web browsers\]\(#\) to confirm that these were both released in 1997.](#)
54. [^] [Tim Berners-Lee, James Hendler and Ora Lassila \(2001\). "The Semantic Web" ↗. Scientific American. Retrieved October 2, 2009.](#)
55. [^] [Nigel Shadbolt, Wendy Hall and Tim Berners-Lee \(2006\). "The Semantic Web Revisited" ↗. IEEE Intelligent Systems. Retrieved October 2, 2009.](#)
56. [^] ["XHTML 1.0 The Extensible HyperText Markup Language \(Second Edition\)" ↗. World Wide Web Consortium. 2000, revised 2002. Retrieved December 7, 2008. "XHTML Documents which follow the guidelines set forth in Appendix C, "HTML Compatibility Guidelines" may be labeled with the Internet Media Type "text/html" \[RFC2854\], as they are compatible with most HTML browsers. Those documents, and any other document conforming to this specification, may also be labeled with the Internet Media Type "application/xhtml+xml" as defined in \[RFC3236\]."](#)
57. [^] ["RFC 2119: Key words for use in RFCs to Indicate Requirement Levels" ↗. Harvard University. 1997. Retrieved December 7, 2008. "3. SHOULD This word, or the adjective "RECOMMENDED", mean that there may exist valid reasons in particular circumstances to ignore a particular item, but the full implications must be understood and carefully weighed before choosing a different course."](#)
58. [^] ["XHTML 1.1 – Module-based XHTML — Second Edition" ↗. World Wide Web Consortium. 2007. Retrieved December 7, 2008. "XHTML 1.1 documents SHOULD be labeled with the Internet Media Type text/html as defined in \[RFC2854\] or application/xhtml+xml as defined in \[RFC3236\]."](#)
59. [^] ["Ted Nelson \(29 January 1999\). "Ted Nelson's Computer Paradigm, Expressed as One-Liners" ↗. Xanadu. Retrieved 26 July 2010.](#)
60. [^] ["HTML Design Constraints" ↗, W3C Archives](#)
61. [^] ["WWW:BTB – HTML" ↗, Pris Sears](#)
62. [^] [Freeman, E \(2005\). *Head First HTML*. O'Reilly.](#)
63. [^] ["Jakob Nielsen \(January 3, 2005\). "Reviving Advanced Hypertext" ↗. Retrieved June 16, 2007.](#)
64. [^] ["Sauer, C.: WYSIWIKI – Questioning WYSIWYG in the Internet Age. In: Wikimania \(2006\)](#)
65. [^] ["Spiesser, J., Kitchen, L.: Optimization of HTML automatically generated by WYSIWYG programs. In: 13th International Conference on World Wide Web, pp. 355–364. WWW '04. ACM, New York, NY \(New York, NY, U.S., May 17–20, 2004\)](#)
66. [^] ["XHTML Reference: blockquote" ↗. Xhtml.com. Retrieved on 2012-02-16.](#)
67. [^] ["Doug Engelbart's INVISIBLE REVOLUTION" ↗. Invisiblerevolution.net. Retrieved on 2012-02-16.](#)

- [HTML 4.01](#), the most recent finished specification (1999) ↗
- [HTML5](#), the upcoming version of HTML ↗
- [Dave Raggett's Introduction to HTML](#) ↗
- [Empty elements in SGML, HTML, XML and XHTML](#) ↗
- [Tim Berners-Lee Gives the Web a New Definition](#) ↗
- [Comparison Matrix of HTML5-oriented Software](#) ↗

Tutorials

- [HTML Dog](#) ↗
- [HTML.net](#) ↗
- [pguides HTML tutorial](#) ↗


Wikimedia Commons has media related to: [HTML](#)


Wikibooks has a book on the topic of [HyperText Markup Language](#)


Wikiversity has learning materials about [HTML](#)


Wikiversity has learning materials about [HTML Challenges](#)


Look up [Appendix:Hyper Text Markup Language](#) in Wiktionary, the free dictionary.

Document markup languages (list) [show]	
World Wide Web Consortium [hide]	
Products and standards	Recommendations
	Notes
	Working Drafts
	Guidelines
	Initiative
	Deprecated
Organizations	World Wide Web Foundation · SVG Working Group · WebOnt · W3C Device Description Working Group · WHATWG ·
Software	Agora · Argo · Arena · Amaya · CERN httpd · Libwww · Line Mode Browser ·
Conference-related	IW3C2 · World Wide Web Conference · WWW1 ·
ISO standards [show]	

Rate this page

[View page ratings](#)

[What's this?](#)

Trustworthy Objective Complete

Well-written

I am highly knowledgeable about this topic (optional)

Submit ratings

Categories: [HTML](#) | [Markup languages](#) | [Technical communication](#) | [World Wide Web Consortium standards](#)

This page was last modified on 10 August 2012 at 11:22.

Text is available under the [Creative Commons Attribution-ShareAlike License](#); additional terms may apply. See [Terms of use](#) for details.

Wikipedia® is a registered trademark of the [Wikimedia Foundation, Inc.](#), a non-profit organization.

[Contact us](#)

[Privacy policy](#) | [About Wikipedia](#) | [Disclaimers](#) | [Mobile view](#)

