

Format Background Document: HTML 4.01

Date: July 18, 2005

Author: Grace Carpenter

GDFR data

Canonical identifier info:gdf/fred/f/html

Description HyperText Markup Language (HTML)

Alias MIME type [text/html](#)

Version 4.01

Legal or recognized owner

Name World Wide Web Consortium

Organization type Non-profit entity

Web site <http://www.w3.org/>

Specification

Document title HTML 4.01 Specification

Document type Article

Publication date December 24, 1999

Access regime Unrestricted access

Identifier <http://www.w3.org/TR/html4/>

Signature

Signature obligation Optional

External signature type File extension

Signature value .htm

Signature

Signature obligation Optional

External signature type File extension

Signature value .html

Non-GDFR data

1. General

1.1 Description (long): HTML (HyperText Markup Language) is the formatting language “widely regarded as the standard publishing language for the World Wide Web [W3C2001].” It was developed originally by Tim Berners-Lee, and then its development and specification was taken over by the W3C (World Wide Web Consortium). HTML is defined in SGML (Standard Generalized Markup Language), which is “a language for describing markup languages, particularly those used for electronic document exchange, document management, and document publishing.” [W3C 2001].

The recommendation for HTML 4.0, published in December 1997, followed by the revised recommendation, HTML 4.01, which was published in December 1999 and contains minor changes, is intended by the W3C to be the last version of HTML (Note that the W3C states that the recommendation for HTML 4.01 “obsoletes” the 4.0 recommendation. However, 4.01 is commonly referred to as “HTML 4”, a convention I’ll follow here). In January 2000 the W3C published the XHTML 1.0 recommendation, which it describes as “a reformulation of HTML 4 as an XML 1.0 application” [W3C 2000]. XHTML 1.0 is distinguished from HTML 4 primarily by its requirement of conformance to specification (put another way, documents must be well-formed in order to be considered XHTML), and by its emphasis on separation of style and content.

HTML 4 documents

- must have a doctype declaration (with one of three HTML 4.01 DTD's specified)
- must conform to the DTD

1.2 **Content type:** text

2. Category-specific

General Technical

3.1 Encoding: There is no mandate as to what character encoding an HTML document must use, nor what a browser must be able to interpret. However, UTF-8 is emerging as a standard. HTML 4 documents may specify a character encoding scheme in one of the following ways:

- an HTTP "charset" parameter in a "Content-Type" field.
- a META declaration with "http-equiv" set to "Content-Type" and a value set for "charset".
- the charset set on an element that designates an external resource.

3.2 Byte Order: Platform-dependant

3.3 Encryption: .

3.4 Human readable: yes

4. Sustainability

4.1 Proprietary: No. HTML is an open and fully documented standard.

4.2 Owner documentation: In addition to the spec, the W3C provides one or two tutorials, but not much additional documentation for users/web designers.

4.3 Other documentation: There are a huge number of HTML resources on the Web.

4.4 Adoption: HTML is the lingua franca of web browsers. Although there are great variations in the html "extensions" they support or don't, all browsers support HTML 4.

4.5 Competition: Although there are other markup languages (MathML, for instance), there is no alternative to HTML/XHTML for web page mark-up.

4.6 Licensing and patent claims:

4.7 Other preservation issues: A big issue in the preservation of HTML is the existence of external references. See [Goethals] for a discussion of this issue.

5. Lifecycle

5.1 Version Duration: HTML 4.01, which came out in 1999, is the last version of HTML. It was succeeded by XHTML 1.0, which came out in January 2000.

5.2 Version History:

HTML 2.0	Sep. 1995
HTML 3.0	Mar 1995
HTML 3.02	Jan 1997
HTML 4.0	Dec. 1997
HTML 4.01	Dec 1999
HTML 2.0	Sep. 1995
HTML 3.0	Mar 1995
HTML 3.02	Jan 1997
HTML 4.0	Dec. 1997
HTML 4.01	Dec 1999

5.3 Expected Newer Versions:

No new versions of HTML are expected.

6. Local Use (in Dspace at MIT)

7. References

[Goethals] Goethals, Andrea, “Action Plan Background: XML 1.0”, FCLA; 9 June 2003
http://www.fcla.edu/digitalArchive/pdfs/action_plan_bgrounds/xml_1_0.pdf

[W3C 1999] W3C, “HTML 4.01 Specification” W3C Recommendation; 24 December 1999
<http://www.w3.org/TR/REC-html40/>

[W3C 2000] W3C, “XHTML TM 1.0 The Extensible Hypertext Markup Language (Second Edition), A Reformulation of HTML 4 in XML 1.0” W3C Recommendation; 26 January 2000, revised 1 August 2002
<http://www.w3.org/TR/xhtml1/>

[W3C 2001] W3C, “XHTML TM 1.1 – Module-based XHTML” W3C Recommendation; 31 May 2001”
<http://www.w3.org/TR/xhtml11/>