

HTML RFC Format

Abstract

This document defines an HTML format that can be used for the production of Internet-Drafts and RFCs. [#abstract-p-1]

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on January 16, 2013.

Copyright Notice

Copyright © 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to [IETF Documents](#) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Editorial Notes

Distribution of this document is unlimited. Please send comments to [RFC Interest](mailto:rfc-interest@rfc-editor.org). [#editorial-notes-p-1]

If you are viewing a version of this document other than the HTML generated by the editor, your a missing vital information. Download a canonical version from <http://cursive.net/draft-hildebrand-html-rfc.html>.

Disclaimer: This is just a first pass at the syntax. The author is willing to add, remove, or change any of the markup used expressly or implicitly in this document. Furthermore, the author does not claim that the current CSS used for this document is useful, interesting, or anything like what the RFC editor may eventually pick. [#editorial-notes-p-3]

Working Group: Network Working Group
Series: Internet-Draft
Intended Status: Standards Track
Published: 2012-07-16
Expires: 2013-01-16
draft-hildebrand-html-rfc
Version: 01
Author: J. Hildebrand (Cisco Systems, Inc.)

The RFC Editor
will remove
this section

Table of Contents

1. [Introduction](#)
 - 1.1. [Background](#)
 - 1.2. [Overview](#)
 - 1.3. [Terminology](#)
 2. [Accessibility](#)
 3. [HTML Format](#)
 - 3.1. [Syntax](#)
 - 3.2. [Basic Structure](#)
 - 3.2.1. [HTML5](#)
 - 3.2.2. [DOCTYPE](#)
 - 3.2.3. [Root Element](#)
 - 3.2.4. [Charset Declaration](#)
 - 3.2.5. [Style](#)
 - 3.2.6. [Emphasis](#)
 - 3.2.7. [Comments](#)
 - 3.2.8. [Sections](#)
 - 3.2.9. [Appendices](#)
 - 3.2.10. [Paragraphs](#)
 - 3.2.11. [Lists](#)
 - 3.2.12. [References](#)
 - 3.2.12.1. [Internal References](#)
 - 3.2.12.2. [References to Standards](#)
 - 3.2.12.3. [Other External References](#)
 - 3.3. [More Elaborate Information](#)
 - 3.3.1. [Requirement Keywords](#)
 - 3.3.2. [Sections to be Removed by the RFC Editor](#)
 - 3.3.3. [Formatting the Table of Contents](#)
 - 3.3.4. [Images](#)
 - 3.3.5. [SVG](#)
 - 3.3.6. [Inline Code](#)
 - 3.3.7. [Blocks of Code](#)
 - 3.3.8. [ASCII Art](#)
 - 3.3.9. [Packet Formats](#)
 4. [Document Metadata](#)
 - 4.1. [Document Information](#)
 - 4.2. [Title](#)
 - 4.3. [Abstract](#)
 - 4.4. [IPR Statements](#)
 - 4.5. [Author](#)
 - 4.6. [Bibliographical Information](#)
 5. [Examples](#)
 - 5.1. [Self](#)
 - 5.2. [Code Sample](#)
 - 5.3. [Sequence Diagrams](#)
 - 5.4. [ABNF](#)
 - 5.5. [Mathematical Formulae](#)
 6. [Security Considerations](#)
 7. [IANA Considerations](#)
 8. [References](#)
 - 8.1. [Normative References](#)
 - 8.2. [Informative References](#)
 9. [Authors' Addresses](#)
- Appendix A. [Allowable Subset of HTML](#)
Appendix B. [CSS Classes with Special Meaning](#)
Appendix C. [Element IDs with Special Meaning](#)
Appendix D. [Acknowledgments](#)

1. Introduction

1.1. Background

The RFC Series has been in existence for over 40 years. During much of that time, the limitations of character set, line and page length, and graphics restrictions of RFC documents met the most immediate needs of the majority of authors and readers. As technology changed, new formats that allowed for a richer set of edit, search and display features came in to use, and tools were created to convert the plain ASCII documents to other desired formats such as HTML, PDF, and Microsoft Word. While the converted versions of the RFCs are widely available, the canonical display format remains the plain text, ASCII, line-printer structured one. The canonical source format is nroff. [#background-p-1]

Canonical source and display versions of an RFC exists for several reasons: [#background-p-2]

- to provide verification of the content of an RFC in case inconsistencies are created when a document is converted to another format or mirrored to another location
- to verify the final content of a document in cases of legal dispute

- to aid in the conversion of the RFC in to formats requested by the community

The current basic format of RFC source and display documents have two characteristics that are considered by the RFC Series Editor to be critical to the RFC Series, including: [#background-p-4]

- persistence (tools to read, edit, and print the documents remain easily accessible to everyone)
- convertibility (the plain text version is simple to convert to other formats)

That said, the very simple nature of the current display format in particular introduces a variety of limitations, the list of which has grown as changes in technology create new expectations: [#background-p-6]

- ASCII art is considered by some to be a major limitation in expressing visually-oriented information
- the internationalization of the authorship and the Internet is introducing [Unicode](#) codepoints not expressible in 7-bit ASCII
- the more common forms of display (web pages, smaller devices) make the limitations of page and line length a hindrance to the reading of an RFC
- tools for people with visual impairments may stumble over the page-oriented structure of the current format; large fonts on a screen that is not large enough to show an entire line, for example, can make the current format difficult to read, since lines do not re-wrap automatically

1.2. Overview

This memo describes a format that can be used both as the canonical input format to the RFC Series Editor (RSE), as well as an archival format. Some document authors will write documents directly in this format (perhaps with tooling to generate the more repetitive tasks), and some authors will prefer other formats as their original source, all of which **MUST** be able to generate the format described in this memo. [#overview-p-1]

This memo has the following goals: [#overview-p-2]

1. Define a strict subset of HTML appropriate for Internet-Draft and RFC Series documents
2. Serve as a comprehensive example of all of the HTML elements that are permissible

1.3. Terminology

The key words "**MUST**", "**MUST NOT**", "**REQUIRED**", "**SHALL**", "**SHALL NOT**", "**SHOULD**", "**SHOULD NOT**", "**RECOMMENDED**", "**NOT RECOMMENDED**", "**MAY**", and "**OPTIONAL**" in this document are to be interpreted as described in [RFC 2119](#)[#rfc:2119]. [#mustard-p-1]

2. Accessibility

One of the major goals of the HTML format is to ensure accessibility for the following consumers of documents written in the format: [#accessibility-p-1]

- People with impaired vision, including those that use large fonts and those that use screen readers
- People with difficulty distinguishing between colors
- People who use devices with small screens, such as cell phones
- Other groups **TBD**

Specific instances where these goals are important in the design choices of the format have been called out in the text. [#accessibility-p-3]

NOTE: designing for these consumers does not preclude the use of features they cannot use, but does require that key semantic data is not lost when read using the tools and settings that are required by a given constituency. [#accessibility-p-4]

3. HTML Format

The format specified here is a subset of HTML, deemed to be widely-implemented by common browsers at the time that the specification was created, likely to continue to be widely-implemented in the future, and unlikely to cause security issues. [#format-p-1]

3.1. Syntax

The following rules **SHALL** be enforced before submittal. [#syntax-p-1]

- The HTML source **MUST** be encoded as UTF-8, as specified in [RFC3629](#).
- The HTML source **MUST** be formatted in the manner of well-formed [XML](#), with all element start tags having matching end tags (or `<element />` for empty elements), and all elements properly nested. HTML "boolean" attributes **MUST** be formatted in the `attr='attr'` style.
- Single quotes (U+0027 APOSTROPHE: ') **MUST** be used to quote attribute values. Unquoted attribute values **MUST NOT** be used.
- HTML **SHALL** be indented using spaces (not tabs).
- Each child element **SHALL** be indented two spaces more than its parent element, unless the child element is mixed with non-whitespace-only text children of the same parent element.

- Each logical line **MUST** be terminated solely with a `\n` (U+000A: LINE FEED), otherwise known as "Unix-style" line endings.
- Other than `\n` (U+000A: LINE FEED), code points less than " " (U+0020: SPACE) (otherwise known as "control characters") **MUST NOT** be used. Any character references that would generate these code points (e.g.) **MUST NOT** be used. NOTE: this rule explicitly forbids `\t` (U+0009: CHARACTER TABULATION), `\f` (U+000C: FORM FEED), and `\r` (U+000D: CARRIAGE RETURN) from appearing in the source.
- Unicode codepoints that are unassigned at the time of publication **MUST** not be used.
- Any Unicode codepoint higher than `~` (U+007E: TILDE) **MUST** serve an explicit purpose that enhances the understanding of the document. Author names and examples are two known cases. The intent is that the document **MUST** be understandable by a reader with the ability to read technical English.
- Each text-containing element such as headings (`<h1>-<h6>`), paragraphs (`<p>`), or list items (``), **MUST** be serialized as a single line without wrapping.

NOTE: none of these rules affect the rendered output of the HTML, but are intended to increase the chance that multiple tools that process the format will generate identical syntax. In turn, this will make difference tools that operate on the HTML source easier to write. [#syntax-p-3]

3.2. Basic Structure

3.2.1. HTML5

The HTML comprising the document **MUST** be valid according to the latest version of the HTML specification at the publishing, starting with the version commonly known as [HTML5](#) [#w3c:WD-html5-20120329]. Although the HTML specification mandates several of syntax and structure rules in this document, they are called out here for emphasis. [#html5-p-1]

3.2.2. DOCTYPE

The DOCTYPE of the document **MUST** be "html", which declares that the document is compliant with [HTML5](#) [#w3c:WD-html5-20120329]. For example, the document will start with exactly this string: [#doctype-p-1]

```
<!DOCTYPE html>
```

3.2.3. Root Element

The root element of the document **MUST** be `<html>`. This element **SHOULD** include a `lang` attribute, whose value is a [RFC5646](#) [#rfc:5646] language tag describing the natural language of the document. For documents submitted to the RFC Series or Internet-Draft Series, the language tag **MUST** be "en", meaning "English". If the `lang` attribute is not present, its value should be taken to be "en". [#root-p-1]

3.2.4. Charset Declaration

In order to be correctly processed by browsers that load the HTML using a mechanism that does not provide a valid MIME content-type or charset, the HTML `<head>` element **MUST** contain a `<meta>` element, with the attributes `http-equiv='Content-Type'` and `content='text/html; charset=utf-8'`. This will look like: [#charset-p-1]

```
<meta http-equiv='Content-Type' content='text/html; charset=utf-8' />
```

3.2.5. Style

The `<head>` **SHOULD** contain an embedded [CSS](#) [#w3c:REC-CSS2-20110607] stylesheet in a `<style>` element. The styles in the stylesheet are to be set consistently between documents by the RFC Editor, according to the best practices of the day. The RFC Editor **SHALL** choose a stylesheet that does not modify the meaning of the normative text of the document. The RFC Editor **SHALL** make the stylesheet available via a standard protocol (e.g. HTTP or HTTPS) for ease of authorship. However, when a document is submitted, external stylesheets (other than "local.css" as specified below) are NOT ALLOWED. The stylesheet itself **MUST NOT** be considered as normative information. [#style-p-1]

To ensure consistent formatting, individual `style` attributes **SHOULD NOT** be used in the main portion of the document source except in highly exceptional circumstances; each use **MUST** be individually justified. [#style-p-2]

Different readers of a specification will desire different tweaks to the stylesheet. To facilitate this, the `<head>` **SHOULD** include a `<link>` to a stylesheet in the same directory as the HTML file, named "local.css", **after** the embedded stylesheet. Note that this "local.css" file will not exist for most users; browsers will correspondingly ignore this `<link>`. When the document is used canonically, these local style overrides **MUST NOT** be in effect. [#style-p-3]

For example: [#style-p-4]

```
<head>
  <style type='text/css'>
<!--
/* RFC-editor styles */
-->
  </style>
  <link rel='stylesheet' type='text/css' href='local.css' />
</head>
```

3.2.6. Emphasis

Words or phrases may be emphasized using the `` element for "**bold**", and the `` element for "*italics*". Underlining **MUST NOT** be used except for links, to avoid visual confusion. Text-only emphasis **MUST NOT** be used. [#emphasis-p-1]

The RFC Editor will set a policy that reflects the current feelings of the community as to whether this emphasis markup is allowed in documents that are submitted for publication in the RFC series. [#emphasis-p-2]

3.2.7. Comments

HTML comments **MAY** be used, but **MUST NOT** contain normative information. One example is to clarify particular choices in the HTML format. Example: [#comments-p-1]

```
<!-- Automatically generated: do not modify -->
```

3.2.8. Sections

Each section of the document **SHALL** be formatted as a `<div>` tag, with a class attribute with value "section". A document-unique, id attribute **SHOULD** be assigned to each section `<div>`. The id **MAY** be human-readable or generated. [#sections-p-1]

NOTE: [XML](#) [#w3c:REC-xml-20081126] requires id attributes to be unique across an entire document: [#sections-p-2]

Values of type ID must match the Name production. A name must not appear more than once in an XML document as a value of this type; i.e., ID values must uniquely identify the elements which bear them. [#sections-p-0]

Each section `<div>` **MUST** contain a header tag (`<h2>`-`<h6>`) of the appropriate depth, with top-level sections getting an `<h2>` tag, and each nested section getting the next higher header level. If more than five levels of headers are required, `<h6>` **MUST** be used for each deeper-nested section. However, nesting sections more than five levels deep is **NOT RECOMMENDED**. [#sections-p-4]

The text in each header tag **MUST** begin with the section number. Section numbers **MUST** begin at "1.", and **MUST** increment by one for each successive section at the same level. Subsections **MUST** be numbered by appending the subsection number to the parent section number. [#sections-p-5]

It is **RECOMMENDED** that the section number be wrapped in an `<a>` element, whose href attribute points to the corresponding section div with a local relative reference. This `<a>` element **SHOULD** have the CSS class `self-ref`. [#sections-p-6]

Within a section, each "normal" paragraph **MUST** be surrounded by a `<p>` element. [#sections-p-7]

For example: [#sections-p-8]

```
<div class='section' id='example'>
  <h2><a class='self-ref' href='#example'>1.</a> Example Section</h2>
  <p>This is a description of the example</p>
  <div class='section' id='nested'>
    <h3><a class='self-ref' href='#nested'>1.1.</a> Nested Section</h3>
    <p>This is a description of the nested section.</p>
    <p>This is the second description paragraph.</p>
  </div>
</div>
```

3.2.9. Appendices

Appendices are special cases of top-level sections. Each appendix of the document **SHALL** be formatted as a `<div>` tag, with a class attribute with value "appendix". A document-unique, id attribute **SHOULD** be assigned to each section `<div>`. The id **MAY** be human-readable or generated. Each appendix `<div>` **MUST** contain an `<h2>` element containing text that describes the purpose of the appendix. Appendices are identified to the reader with Latin capital letters A-Z, in order. It is **NOT RECOMMENDED** to have more than 26 appendices, but if required, appendices "AA.", "AB.", etc. follow Appendix Z. [#appendices-p-1]

Inside the appendix, subsections **MUST** be formatted per [Sections](#) [#sections], numbered sequentially. For example, the first subsection of "Appendix A." is "Appendix A.1.". [#appendices-p-2]

For example: [#appendices-p-3]

```
<div class='appendix' id='acknowledgements'>
  <h2>Appendix A. Acknowledgements</h2>
  <p>The author gratefully acknowledges the contributions of...</p>
  <div class='section' id='contributors'>
    <h3>Appendix A.1. Contributors</h3>
    <p>These people contributed text...</p>
  </div>
</div>
```

3.2.10. Paragraphs

Paragraphs **MUST** be contained in a [section](#)[#sections] <div> or an [appendix](#)[#appendices] <div>. A document-unique, id attribute **SHOULD** be assigned to each <p>. The id will usually be machine-generated, but **MAY** be human-readable if desired. [#paragraphs-p-1]

It is **RECOMMENDED** that each paragraph be kept relatively small compared to a "page" in previous RFC formats, so that references to each paragraph are at least as valuable as page references have been in previous formats. [#paragraphs-p-2]

3.2.11. Lists

Lists may be used inside a section <div>, and may nest in other lists as needed. However, lists **MUST NOT** be nested inside a <p> element. Unordered lists () and ordered lists () may both be used. For example: [#lists-p-1]

```
<div class='section' id='lists'>
  <h4>Unordered list</h4>
  <p id='lists-p-1'>An explanation:</p>
  <ul>
    <li>One</li>
    <li>Two</li>
    <ol>
      <li>Two.1: (this one is numbered)</li>
    </ol>
  </ul>
</div>
```

3.2.12. References

3.2.12.1. Internal References

References to other paragraphs or sections in the same document **MUST** use an <a> element with an href attribute with a fragment that points at the id attribute of the target (i.e. the id prefixed with a #). The target element **MUST** have a human-readable id attribute, which **MUST** be stable even when tooling generates new id attributes. For example: [#internal-references-p-1]

```
See <a href='#example'>Example Section</a> for more details
```

3.2.12.2. References to Standards

References to standards are special, in that they generate formal bibliographical metadata. All links to standards in the main body of the text **MUST** jump to the [bibliographical](#)[#bibliographical] entry; the href **MUST** be of the form #[series]:[number]. For example: [#standards-references-p-1]

```
href='#rfc:2119'
```

Valid series identifiers include: [#standards-references-p-3]

- 3gpp: [The 3rd Generation Partnership Project](#)
- ansi: [American National Standards Institute](#)
- ccitt: [ITU Telecommunication Standardization Sector](#)
- fips: [Federal Information Processing Standard](#)
- id: [Internet-Drafts](#)
- ieee: [Institute of Electrical and Electronics Engineers](#)
- iso: [International Organization for Standardization](#)
- itu: [International Telecommunication Union](#)
- nist: [National Institute of Standards and Technology](#)
- oasis: [Organization for the Advancement of Structured Information Standards](#)
- pkcs: [Public-Key Cryptography Standards](#)
- rfc: [Request For Comments](#)

- w3c: [The World Wide Web Consortium](#)
- xep: [XMPP Standards Foundation](#)

The text inside the link **SHOULD** be a human-readable colloquial representation of the standard name and/or number. [#standards-references-p-5]

Normative references **MUST** use <a> elements with class ref. Example: [#standards-references-p-6]

```
See: <a class='ref' href='#rfc:2119'>RFC2119</a>
```

Informative references **MUST** use <a> elements with class inforef. Example: [#standards-references-p-8]

```
See: <a class='inforef' href='#rfc:2119'>RFC2119</a>
```

3.2.12.3. Other External References

References to other documents that are not standards **SHOULD** be linked using the http: or https: URI scheme, and **MUST** be linked using a URI scheme that is widely-deployed at the time that the document is published, and which does not raise any security or stability issues. In particular, javascript: references **MUST NOT** be used. Links using the mailto: scheme **SHOULD** be limited to the author's address information. [#external-references-p-1]

[#external-references-p-2]

For example: [#external-references-p-3]

```
<a href='http://example.com/'>Example</a>
```

3.3. More Elaborate Information

This section describes how to format several types of information that occur regularly in documents for the Internet-Draft and RFC Series which are not descriptive text. [#elaborate-p-1]

3.3.1. Requirement Keywords

The [RFC2119](#)[#rfc:2119] keywords in the document **MAY** be set off with special markup. If so, they **MUST** be surrounded with a element containing the CSS class rfc2119. For example: [#2119-format-p-1]

```
If so, they <span class='rfc2119'>MUST</span> be surrounded
```

3.3.2. Sections to be Removed by the RFC Editor

The author may want to inject notes to the reader that are not to be a part of the final document that is published by the RFC editor. These notes **MAY** use any format desired by the author that would otherwise be legal in the document, but the outermost element of the note **MUST** have a CSS class with value rfceditor-remove. [#rfceditor-remove-format-p-1]

```
<div class='rfceditor-remove'>
  <h2>Editorial Notes</h2>
  <p>...</p>
</div>
```

3.3.3. Formatting the Table of Contents

The table of contents for the document **MUST** appear in a <div> element, which **SHOULD** precede any of the [sections](#)[#sections] of proper document content. The <div> element **MUST** have an id attribute with value toc. The <div> element **SHOULD** contain an <h2> element containing the string Table of Contents, followed by nested and elements describing the structure of the document, with links to each of the [sections](#)[#sections] mentioned. For example: [#toc-format-p-1]

```

<div id='toc'>
  <h2>Table of Contents</h2>
  <ul>
    <li>
      <div>1. <a href='#introduction'>Introduction</a></div>
      <ul>
        <li>
          <div>1.1. <a href='#background'>Background</a></div>
          </li>
          ...
        </li>
      </ul>
    </li>
  </ul>

```

NOTE: the Table of Contents **SHOULD NOT** be considered meta-data for the document. The [sections](#)^[#sections] themselves **SHOULD** contain all of the data that is required. ^[#toc-format-p-3]

3.3.4. Images

Include an image using the `` tag with a `src` attribute. During the editing process, it may be useful to keep the value of the `src` attribute as a file name relative to the document, or a http(s) URL. However, upon submission, the final version of a document **SHALL** include a data: URI as specified in [RFC2397](#)^[#rfc:2397]. ^[#images-p-1]

The image MIME type of the image **SHALL** be `image/png`, as specified in [RFC2083](#)^[#rfc:2083]. The RFC Editor can allow other image types in the future, at the Editor's discretion, as the state of the art and common implementation patterns change. ^[#images-p-2]

Consider how images will look when printed. Consider how your images will be used by vision-impaired readers, including those readers with color vision deficiency. For example, images **SHOULD NOT** have meaningful distinctions conveyed only by color differences, and images **SHOULD** be available in high enough resolution that readers with other vision deficiencies can zoom in to see detail. Images **MUST NOT** be animated. ^[#images-p-3]

The `alt` attribute is **REQUIRED**, and **MUST** be a complete, accessible description of the image. ^[#images-p-4]

The `height` and `width` attributes **SHOULD** be used to specify the size of the image in pixels. ^[#images-p-5]

Images **MUST** be wrapped in a `<figure>` element. The `<figure>` element **SHOULD** contain a `<figcaption>` element after the `` element, which **SHOULD** contain text that describes the image. For example: ^[#images-p-6]

```

<figure>
  <img src='data:image/png;base64,[BASE64-encoded PNG]''
      alt='A description of the diagram'
      width='236'
      height='176' />
  <figcaption>Possible workflow for processing HTML RFCs</figcaption>
</figure>

```

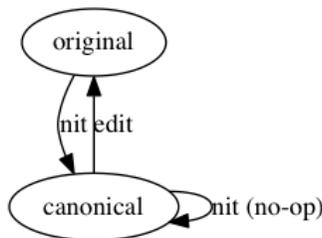


Figure 1. Possible workflow for processing HTML RFCs

Images **SHOULD NOT** be normative. Instead, the information contained in the image **SHOULD** be adequately conveyed in the textual description that accompanies the image. ^[#images-p-9]

3.3.5. SVG

[SVG](#)^[#w3c:REC-SVG11-20110816] can be included directly in the HTML source, surrounded by a `<figure>` element and succeeded by a `<figcaption>` element, as described in [Section 3.3.4](#)^[#images]. The root `<svg>` element **MUST** contain a `<title>` or `<desc>` element that fully describes the diagram for accessibility to screen readers; this is similar to the `alt` attribute on images. For example: ^[#svg-p-1]

```

<figure id='fig-2'>
  <svg xmlns='http://www.w3.org/2000/svg'>
    <title>A sample SVG</title>
    <desc>This is a sample image, with a title and description</desc>
    ...
  </svg>
  <figcaption>Sample SVG</figcaption>
</figure>

```

Might render as: `[#svg-p-3]`

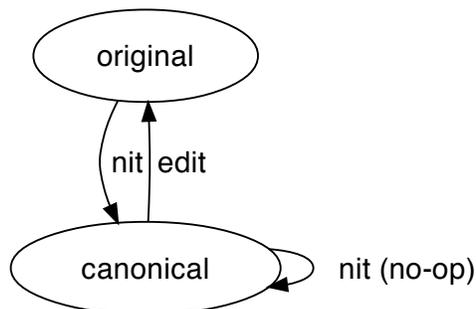


Figure 2. Sample SVG

Note that there are currently more browsers that can deal with `` elements (or their `alt` text) than are able to generate any sensible fallback rendering from SVG. Until this changes, authors might consider replacing their SVG with a rendered image. `[#svg-p-5]`

3.3.6. Inline Code

Use the `<code>` element to set aside literal references to code or protocol elements in the middle of a paragraph. If desired, the language of the code or protocol can be declared using a `class` attribute starting with `language-`. For example: `[#inline-code-p-1]`

```
Use the <code class='language-html'>&lt;code&gt;</code> element
```

3.3.7. Blocks of Code

Larger sections of code or protocol can be included using a `<pre>` element with a `class` attribute of code. If desired, the language of the code or protocol can be declared using a further `class` value starting with `language-` (multiple `class` values are separated by spaces in HTML). The text inside the `<pre>` element will be rendered in a monospace font, with whitespace maintained. For example: `[#block-code-p-1]`

```
<pre class='code language-html'>
&lt;html&gt;
  &lt;body /&gt;
&lt;/html&gt;
</pre>
```

Will be rendered as: `[#block-code-p-3]`

```
<html>
  <body />
</html>
```

Depending on author style, blocks of code **MAY** be enclosed in a `<figure>` element, with a `<figcaption>` element that describes the block. For example, see [Figure 3](#) `[#blockfigure]`. `[#block-code-p-5]`

```
<figure id='blockfigure'>
  <pre class='code language-html'></pre>
  <figcaption>A code block wrapped in a figure.</figcaption>
</figure>
```

Figure 3. A code block wrapped in a figure.

3.3.8. ASCII Art

ASCII art is still preferred by some authors in preference to an image or SVG. The RFC Editor may decide to prefer [images](#) `[#images]` or [SVG](#) `[#w3c:REC-SVG11-20110816]`, or may decide to prohibit ASCII art in the future, depending on the needs of the community at the time of publishing. Until that time, to include ASCII art, wrap a `<pre>` element with `class='ascii'` in a `<figure>` along with a `<figcaption>`, as if the `<pre>` element were an image, as specified in the [Section 3.3.4](#) `[#images]`. For example: `[#ascii-art-p-1]`

```

<figure>
  <pre class='ascii'>
      +-----+
      | original | &lt;+
      +-----+
      |   nit   |   edit
      |   v     |   |
nit (no-op) +-----+
+-----+ | canonical |
+-----+ &gt; |         |   -+
          +-----+
          |         |
          +-----+
  </pre>
  <figcaption>Sample ASCII art</figcaption>
</figure>

```

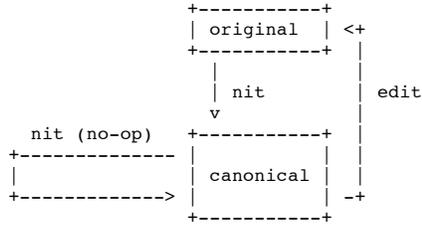


Figure 4. Sample ASCII art

3.3.9. Packet Formats

Packet format descriptions can be encoded as a `<table>` element wrapped in a `<figure>` along with a `<figcaption>`, as if the `<pre>` element were an image, as specified in [Section 3.3.4](#) [#images]. For consistent formatting, the `<table>` element should have class `pdu`. For example: [#pdu-p-1]

```

<figure>
  <figcaption>Sample packet format</figcaption>
  <table class='pdu'> [table describing the packet] </table>
</figure>

```

Would be rendered as: [#pdu-p-3]

Figure 5. Sample packet format

Offsets	Octet	0							1							2							3										
Octet	Bit	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
0	0	version				header len				differentiated services				ECN				length															
4	32	identification														R	S	D	M	fragment offset													
8	64	source IP																															
12	96	destination IP																															
16	128	options																															
...																															

4. Document Metadata

Metadata for the document **SHOULD** be easily extractable from the document by tools that ordinarily process HTML. Typically, the `class` and `id` attributes can be used to query the document using [CSS](#) [#w3c:REC-CSS2-20110607]-style selectors. The metadata scheme **SHOULD** be designed such that the element name is not required in order to select a given piece of data. Instead, any element that can contain text can be used for a given `class` or `id` to be selected. The value of the data contained by the selected element(s) consists of the concatenation of all of the text from all of the child nodes of the selected element or elements, with each run of consecutive whitespace Unicode codepoints [codepoints with the `White_Space` property, such as U+0020 (SPACE), U+0009 (CHARACTER TABULATION), U+000A (LINE FEED), U+000C (FORM FEED), U+000D (CARRIAGE RETURN), U+00A0 (NON-BREAKING SPACE), and U+2029 (PARAGRAPH SEPARATOR)] compressed to a single U+0020 (SPACE). The metadata scheme **MUST** allow unambiguous selection. [#metadata-p-1]

The `id` attribute is used to identify pieces of data that are guaranteed to be unique across the document. Any element with an `id` attribute can also be used as a fragment target in a URI by starting with the base URI of the document, then appending `"#"` (U+0023: NUMBER SIGN) and the value of the `id` attribute. In CSS, the element with a given `id` attribute value is selected by prepending the value with `"#"` (U+0023: NUMBER SIGN). For example, the following HTML in a document with the URI `http://example.com/index.html`: [#metadata-p-2]

```
<div id='example'>Important Text</div>
```

Can be targeted directly with the URL `http://example.com/index.html#example`, and the CSS selector `#example`. [#metadata-p-4]

The `class` attribute is a catch-all tagging mechanism for everything in the document that might not be unique. Multiple classes may be defined on a single element by setting the `class` attribute to a space-separated list of classes. All of the elements with a given class name can be selected in CSS by prepending the class name with "." (U+002E: FULL STOP). [#metadata-p-5]

4.1. Document Information

Information about the document as a whole. The `<div>` element with `id='document'` **SHOULD** be the first child element of the HTML body. For example: [#md-document-p-1]

```
<div id='document'>
  <div class='identifiers'>
    <div class='workgroup'>Network Working Group</div>
    <div class='series'>Internet-Draft</div>
    <div class='status'>Standards Track</div>
    <div class='published'>2012-07-07</div>
    <div class='expires'>2013-01-07</div>
    <div class='version'>00</div>
  </div>
  <div class='authors'>
    <div class='author'>
      <span class='initial'>J.</span>
      <span class='surname'>Hildebrand</span>
      <span class='company'>Cisco Systems, Inc.</span>
    </div>
  </div>
</div>
```

More details for this format will be included in future drafts of this document. [#md-document-p-3]

4.2. Title

The title of the document **MUST** appear in an `<h1>` element, which **SHOULD** follow directly after the [Document Information](#) [#md-document]. The `<h1>` element **MUST** have an `id` attribute with value `title`. For example: [#md-title-p-1]

```
<h1 id='title'>HTML RFC Format</h1>
```

4.3. Abstract

The abstract for the document **MUST** appear in a `<div>` element, which **SHOULD** follow directly after the [Title](#) [#md-title]. The `<div>` element **MUST** have an `id` attribute with value `abstract`. The `<div>` element **SHOULD** contain an `<h2>` element containing the word `Abstract`, and **MUST** contain one or more `<p>` elements containing text that describes the document succinctly. For example: [#md-abstract-p-1]

```
<div id='abstract'>
  <h2>Abstract</h2>
  <p>This document defines an HTML format...</p>
</div>
```

4.4. IPR Statements

The IPR boilerplate for the document **MUST** appear in a `<div>` element, which **SHOULD** follow directly after the [Abstract](#) [#md-abstract]. The `<div>` element **MUST** have an `id` attribute with value `ipr` and a CSS `class` of the name of the relevant IPR ruleset. The only valid values for the IPR ruleset class are `trust200902`, `noModificationTrust200902`, and `noDerivativesTrust200902` at this time. The contents of the `<div>` element are to be set correctly for the given ruleset, based on guidance from the IETF trust. For example: [#md-ipr-p-1]

```
<div id='ipr' class='trust200902'>
  <h2>Status of this Memo</h2>
  <p>...</p>
  <h2>Copyright Notice</h2>
  <p>...</p>
</div>
```

Question: should the valid IPR classes be put in an IANA registry along with their boilerplate expansions? [#md-ipr-p-3]

The RFC Editor will remove this section

4.5. Author

NOTE: this document currently uses the approach specified by "[hCard](http://microformats.org/wiki/hcard)". The author The RFC Editor recommends that the [vcarddav](http://datatracker.ietf.org/wg/vcarddav/charter/) Working Group of the IETF be tasked to will remove propose an approach for HTML embedding of vCard that is aligned with [RFC 6350](http://rfc.ietf.org/public/rfc/html/rfc6350.html). In particular, the language and this section alt id mechanisms of [RFC 6350](http://rfc.ietf.org/public/rfc/html/rfc6350.html) are not explicitly mentioned in hCard, and are required in order to fit the desire for authors' names to be representable both by English readers as well as the native language of the author. [#md-author-p-1]

This section will be augmented with normative text when an approach is decided upon. A quick example (as an existence proof) can be found in [Figure 6](#). The rendered version can be found in [Section 9](#).

```
<address class='vcard'>
  <span class='n hidden'>
    <span class='family-name'>Hildebrand</span>
    <span class='given-name'>Joe</span>
  </span>
  <span class='nickname hidden'>hildjj</span>
  <span class='fn'>Joe Hildebrand</span>
  <span class='org'>Cisco Systems, Inc.</span>
  <a class='email' href='mailto:jhildebr@cisco.com'>jhildebr@cisco.com</a>
  <div class='adr'>
    <div class='street-address'>1899 wynkoop St, Suite 600</div>
    <div><span class='locality'>Denver</span>, <span class='region'>CO</span> <span class='postal-code'>80202</span></div>
    <div class='country-name'>United States</div>
  </div>
</address>
```

Figure 6. Sample (temporary) author information

4.6. Bibliographical Information

TBD: define microformat for bibliographical data, perhaps based on the [citation](http://microformats.org/wiki/citation) work at microformats.org. [#bibliographical-p-1]

5. Examples

5.1. Self

This draft itself is a good example of how to use the format. Please view-source. [#self-example-p-1]

5.2. Code Sample

```
#include <stdio.h>
int main(int argc, char **argv)
{
    printf("Hello, IETF\n");
    return 0;
}
```

5.3. Sequence Diagrams

Include an image tag with class='sequence', where the alt text is the [WebSequenceDiagrams.com](http://www.websequencediagrams.com/) source for the diagram. [#websequencediagram-sample-p-1]

Before publication, this approach will be replaced by something more well-specified and not requiring third-party software. The RFC Editor will remove this section

```
<figure>
  <img class='sequence' alt='
title Authentication Sequence
Alice->Bob: Authentication Request
note right of Bob: Bob thinks about it
Bob->Alice: Authentication Response' />
  <figcaption>A sample sequence diagram</figcaption>
</figure>
```

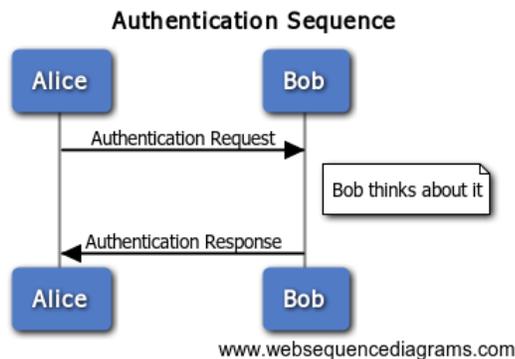


Figure 7. A sample sequence diagram

5.4. ABNF

Augmented Backus-Naur Form is a way of describing formal syntax, described in [RFC5234](#)^[#rfc:5234]. Include ABNF (without extra indentation) in a `<pre>` element, with CSS class "code" and "language-abnf". For example: ^[#abnf-p-1]

```

<pre class='code language-abnf'>
label      = top-level *4section-num
top-level  = section-num / appendix-let
section-num = 1*DIGIT "."
appendix-let = 1*CAP "."
CAP        = %x41-5A ; A-Z
DIGIT      = %x30-39 ; 0-9
</pre>

```

Is rendered as: ^[#abnf-p-3]

```

label      = top-level *4section-num
top-level  = section-num / appendix-let
section-num = 1*DIGIT "."
appendix-let = 1*CAP "."
CAP        = %x41-5A ; A-Z
DIGIT      = %x30-39 ; 0-9

```

5.5. Mathematical Formulae

For now, just use an image (as specified in [Section 3.3.4](#)^[#images]), with the alt text being a [LaTeX](http://www.latex-project.org/) formula that would produce the image. For example: ^[#math-p-1]

$$\Re z = \frac{n\pi \frac{\theta+\psi}{2}}{\left(\frac{\theta+\psi}{2}\right)^2 + \left(\frac{1}{2} \log \frac{B}{A}\right)^2}$$

Figure 8. A sample formula

Future versions of this document will likely favor [SVG](#)^[#w3c:REC-SVG11-20110816] or [MathML](#)^[#w3c:REC-MathML3-20101021] representations of formulae, if browser support and accessibility concerns are addressed. ^[#math-p-3]

6. Security Considerations

Since RFCs are sometimes exchanged outside the normal Web sandboxing mechanism (e.g. rsync to a mirror) then loaded from a local file, more care must be taken with the HTML than is ordinary on the Web. In particular, the intent with the format is to forbid any embedded code such as JavaScript as well as all mechanisms that could be used to execute code outside of the browser such as plugins or non-static media (such as video). ^[#security-p-1]

7. IANA Considerations

TBD ^[#iana-p-1]

8. References

8.1. Normative References

- [RFC2083] [Boutell](#), T., "PNG (Portable Network Graphics) Specification Version 1.0", [TXT](#), RFC 2083, March 1997.

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [TXT] [HTML] [XML], BCP 14, RFC 2119, March 1997.
- [RFC2397] Masinter, L., "The "data" URL scheme", [TXT] [HTML] [XML], RFC 2397, August 1998.
- [RFC3629] Yergeau, F., "UTF-8, a transformation format of ISO 10646", [TXT], STD 63, RFC 3629, November 2003.
- [RFC3979] Bradner, S., "Intellectual Property Rights in IETF Technology", [TXT], BCP 79, RFC 3979, March 2005.
- [RFC5378] Bradner, S. and J. Contreras, "Rights Contributors Provide to the IETF Trust", [TXT], BCP 78, RFC 5378, November 2008.
- [RFC5646] Phillips, A. and M. Davis, "Tags for Identifying Languages", [TXT], BCP 47, RFC 5646, September 2009.
- [Unicode61] The Unicode Consortium, "The Unicode Standard, Version 6.1.0", Mountain View, CA: The Unicode Consortium, 2012. ISBN 978-1-936213-02-3, [HTML], 2012.
- [W3C.REC-CSS2-20110607] Çelik, T., and H. Lie, I. Hickson, B. Bos, "Cascading Style Sheets Level 2 Revision 1 (CSS 2.1) Specification", [HTML], World Wide Web Consortium Recommendation REC-CSS2-20110607, June 2011.
- [W3C.REC-SVG11-20110816] Watt, J., and E. Dahlström, C. McCormack, D. Schepers, F. Jun, J. Ferraiolo, P. Dengler, A. Grasso, C. Lilley, D. Jackson, "Scalable Vector Graphics (SVG) 1.1 (Second Edition)", [HTML], World Wide Web Consortium Recommendation REC-SVG11-20110816, August 2011.
- [W3C.REC-xml-20081126] Sperberg-McQueen, C., and F. Yergeau, T. Bray, E. Maler, J. Paoli, "Extensible Markup Language (XML) 1.0 (Fifth Edition)", [HTML], World Wide Web Consortium Recommendation REC-xml-20081126, November 2008.
- [W3C.WD-html5-20120329] Hickson, I., "HTML5", [HTML], World Wide Web Consortium WD WD-html5-20120329, March 2012.

8.2. Informative References

- [RFC5234] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", [TXT], STD 68, RFC 5234, January 2008.
- [RFC6350] Perreault, S., "vCard Format Specification", [TXT], RFC 6350, August 2011.
- [W3C.REC-MathML3-20101021] Ion, P., and D. Carlisle, R. Miner, "Mathematical Markup Language (MathML) Version 3.0", [HTML], World Wide Web Consortium Recommendation REC-MathML3-20101021, October 2010.

9. Authors' Addresses

Joe Hildebrand

Cisco Systems, Inc.
jhildebr@cisco.com
 1899 Wynkoop St, Suite 600
 Denver, CO 80202
 United States

Appendix A. Allowable Subset of HTML

This section collects all of the elements that are allowed in the HTML RFC format. Each element is listed with a set of allowed attributes, and a list of the parent elements in which the element may be placed. The attributes `class`, `id`, and `lang` are allowed on all elements. All other elements, attributes, and nesting approaches **MUST NOT** be used. [#html-subset-p-1]

Element	Attributes	Parents
a	href, title	address, div, figcaption, h2, h3, h4, h5, li, p, span, td
address		div
blockquote		div
body		html
br		td, th
code		li, p, td
div		address, body, div, li
em		p
figcaption		figure
figure		div
h1		body
h2		div
h3		div
h4		div
h5		div
head		html
html		
img	alt, height, src, width	figure
li		ol, ul
link	href, rel, type	head
meta	content, http-equiv, name	head
ol		div
p		blockquote, div, td
pre		div, figure
span		address, div, li, p, span
strong		p, pre
svg	height, viewBox, width	figure
table		div, figure
tbody		table

td	colspan	tr
th	colspan	tr
thead		table
title		head
tr		tbody, thead
ul		div, li, td

Appendix B. CSS Classes with Special Meaning

Although the author can add class information to any element, the following class names have special meaning in an HTML RFC: [#html-classes-p-1]

Class	Meaning
adr	
appendix	
ascii	
author	
authors	
code	
company	
country-name	
date	
edge	
email	
expires	
family-name	
figref	
fn	
formula	
given-name	
graph	
hidden	
identifiers	
initial	
language-abnf	
language-c	
language-html	
locality	
n	
nickname	
node	
note	
org	
pdu	
postal-code	
published	
ref	
reflinks	
region	
rfc2119	
rfceditor-remove	
section	
sectref	
self-ref	
sequence	
series	
series-info	
status	
street-address	
surname	
title	
toc	
trust200902	
vcard	
version	

workgroup	
-----------	--

Appendix C. Element IDs with Special Meaning

Although the author can add an id attribute to any element, the following id values **SHOULD NOT** be used except for the role defined for each below: [#html-ids-p-1]

ID	Meaning
document	Data about the document, including dates, name, version, etc.
title	The title of the document, usually applied to a <h1> element.
abstract	The abstract for the document, usually applied to a <div> element that contains a heading and paragraphs of text.
ipr	<p>The Intellectual Property Rights associated with the document. The class attribute of the same element will contain a machine-readable IPR statement name from this list: [#html-ids-p-0]</p> <ul style="list-style-type: none"> trust200902: This is appropriate for most drafts, where the entire content of the draft is written by the draft's authors, or all authors of other material have given explicit permission to use their work. noModificationTrust200902: This is appropriate for drafts where the authors wish to place the additional condition that if the draft is published as an RFC, it must have no changes other than formatting. An example might be a document published by another organization that permits copying but not modification. noDerivativesTrust200902: This is appropriate for drafts not intended to be published as RFCs. pre5378Trust200902: This is appropriate for drafts that include material submitted to the IETF prior to RFC 5378 (10 Nov 2008), where the authors of that material have not given explicit permission to use their work in this draft. An example might be a draft using material from an RFC whose author has died or cannot be located, or who thinks your draft is stupid. <p>The element with this id will contain all of the IPR and status boilerplate text [#html-ids-p-2]</p> <p>Note: an IANA registry may be required for this attribute in the future. [#html-ids-p-3]</p>
venue	The venue for discussion. Inside the element tagged with this id will be one or more <a> elements that describe the discussion venue for Internet-Drafts.
toc	The Table of Contents
references	The section containing bibliographical data, including sections for normative and informative references.
normative	The section containing normative document references.
informative	The section containing informative document references.
authors	The section containing data about the authors of the document.
security	The section containing the Security Considerations for the document.
iana	The section containing the IANA Considerations for the document.
acknowledgments	The section containing the author's acknowledgments.

Appendix D. Acknowledgments

The author gratefully acknowledges the contributions of: Heather Flanagan and Patrick Linskey [#acknowledgments-p-1]