XHTML (Extensible HyperText Markup Language) is a family of XML markup languages that mirror or extend versions of the widely used Hypertext Markup Language (HTML), the language in which web pages are written.

Filename extension .xhtml, .xht, .xml, .html, .htm

Internet media type application/xhtml+xml

Developed by World Wide Web Consortium

Type of format Markup Language

Extended from XML, HTML

Stands for Extensible HyperText Markup Language

Application Application of XML

Function Extended versions of Html that is stricter and Xml-based

Nature Restrictive subset of XML and needs to be parsed with standard XML parsers

Origin World wide Web Consortium Recommendation in 2000.

Versions Xhtml 1, Xhtml 1.1; xhtml2; xhtml5

**Overview of HTML and XHTML**

**HTML** is the predominant mark up language for web pages. HTML creates structured documents by denoting structural semantics for text like headings, lists, links, quotes etc. It allows [images](/difference/JPEG_vs_PNG) and objects to be embedded to create interactive forms. It is written as tags surrounded by angle brackets - for example, <html>. Scripts in languages like [JavaScript](/difference/Java_%28programming_language%29_vs_JavaScript) can also be loaded.

**XHTML** is a family of XML languages which extend or mirror versions of HTML. It does not allow omission of any tags or use of attribute minimization. XHTML requires that there be an end tag to every start tag and all nested tags must be closed in the right order. For example, while <br> is valid in HTML, it would be required to write <br /> in XHTML.

**Features of HTML vs XHTML documents**

HTML documents are composed of elements that have three components- a pair of element tags – start tag, end tag; element attributes given within tags and actual, textual and graphic content. HTML element is everything that lies between and including tags. (Tag is a keyword which is enclosed within angle brackets).

XHTML documents has only one root element. All elements including variables must be in lower case, and values assigned must be surrounded by [quotation marks](/difference/Double_Quotes_vs_Single_Quotes), closed and nested for being recognized. This is a mandatory requirement in XHTML unlike HTML where it is optional. The declaration of DOCTYPE would determine rules for documents to follow.

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 or 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: <br/>. 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: <br />.

## XHTML vs HTML Specification

HTML and XHTML are closely related and therefore can be documented together. Both HTML 4.01 and XHTML 1.0 have three sub specifications – strict, loose and frameset. The difference opening declarations for a document distinguishes HTML and XHTML. Other differences are syntactic. HTML allows shortcuts like elements with optional tags, empty elements without end tags. XHTML is very strict about opening and closing tags. XHTML uses built in language defining functionality attribute. All syntax requirements of XML are included in a well formed XHTML document.

Note, though, that these differences apply only when an XHTML document is served as an application of XML; that is, with a MIME type of application/xhtml+xml, application/xml, or text/xml. An XHTML document served with a MIME type of text/html must be parsed and interpreted as HTML, so the HTML rules apply in this case. A style sheet written for an XHTML document being served with a MIME type of text/html may not work as intended if the document is then served with a MIME type of application/xhtml+xml. For more information about MIME types, make sure to read MIME Types.

This can be especially important when you’re serving XHTML documents as text/html. Unless you’re aware of the differences, you may create style sheets that won’t work as intended if the document’s served as real XHTML.

Where the terms “XHTML” and “XHTML document” appear in the remainder of this section, they refer to XHTML markup served with an XML MIME type. XHTML markup served as text/html is an HTML document as far as browsers are concerned.

**How to migrate from HTML to XHTML**

As recommended by W3C following steps can be followed for migration of HTML to XHTML (XHTML 1.0 documents):

* Include xml:lang and lang attributes on elements assigning language.
* Use empty-element syntax on elements specified as empty in HTML.
* Include an extra space in empty-element tags: <html />
* Include close tags for elements that can have content but are empty: <html></html>
* Do not include XML declaration.

Carefully following W3C’s guidelines on compatibility, a user agent (web browser) should be able to interpret documents with equal ease as HTML or XHTML.

**How to migrate from XHTML to HTML**

To understand the subtle **differences between HTML and XHTML**, consider the transformation of a valid and well-formed XHTML 1.0 document into a valid HTML 4.01 document. To make this translation requires the following steps:

* 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.
* Remove the XML namespace (xmlns=URI). HTML has no facilities for namespaces.
* Change the document type declaration from XHTML 1.0 to HTML 4.01.
* If present, remove the XML declaration. (Typically this is: <?xml version="1.0" encoding="utf-8"?>).
* 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.
* Change the XML empty-element syntax to an HTML style empty element (<br/> to <br>).