Converting to HTML 5 from HTML 4 or XHTML

HTML-5.com is an HTML User's Guide and quick reference of HTML elements and attributes for web developers who code HTML web pages, not only for HTML 5 but for HTML coding in general, with demos and examples of HTML code plus a cheat sheet for web developers. TV Series & Actors and Actresses. Follow TV Series and HTML 5 on Google+.

HTML-5.com ➢ itemscopehttp://data-vocabulary.org/BreadcrumbHTML 5 ➢ itemscopehttp://data-vocabulary.org/BreadcrumbHTML Tutorials ➢ itemscopehttp://data-vocabulary.org/Breadcrumb
Converting to HTML 5 from HTML 4 or XHTML

Converting to HTML 5

Converting from 2000 W3C standard HTML

(Some links on this page take you to details in the HTML Tag Reference. Bookmark this page in your Favorites so you can come back to it later.)

Converting to HTML 5 polyglot code from a strict implementation of the previous version of HTML is as simple as 1-2-3:

1. Remove the PUBLIC FPI and SYSTEM identifiers from the DOCTYPE declaration.
2. Replace any obsolete or deprecated HTML tags or constructs with HTML code that is compliant with the HTML 5 standard. For example, make sure that any set of <col> tags for HTML table columns always have a **colgroup** element as their parent.
3. Start taking advantage of new features of HTML 5, such as converting from <div> tags to HTML 5 sectioning tags.

Converting from 20th Century versions of HTML (prior to 2000)

Converting from earlier versions of HTML or transitional HTML code requires more extensive changes, since there were significant changes to the HTML standard between 1997 and 2000 in order to support XML-based parsers, mobile devices with stricter parsing rules, cacheable client-side templates and aggregation with other types of content such as RSS feeds.

1. Make sure the document starts with an xml declaration and a DOCTYPE declaration.
2. Make sure the <html> top element tag includes the `xmlns="http://www.w3.org/1999/xhtml"` attribute. (The namespace URI for xHTML and later verions of HTML includes the year "1999" because it was assigned that year while the W3C HTML Recommendation released in 2000 was still being developed.)
3. Make sure that all tags are matched with end tags, or are self-closed with `/>`. Make sure the element names in the start and end tags match and are lower case.
4.
Make sure that all attribute values are enclosed in quotation marks. Make sure boolean attributes are coded in their full form using the attribute name in quotes as the value (`attribute="attribute"`) when the value is true and completely omitting it when the value is false. The full form will be properly understood by web browsers parsing polyglot documents with either the HTML syntax or the XML syntax of HTML 5. Avoid using a

---

~~minimized form for boolean attributes, such as~~ `selected`~~, or values with an empty string,~~ such as in `selected=""`, which XPath treats as false rather than true.

Note that the HTML 5 specifications explicitly state that:

<blockquote style="margin-top: 0"> The values "true" and "false" are not allowed on boolean attributes. </blockquote>
This is because browsers that look at the coded value for boolean attributes would treat the string `"false"` as *false* while browsers that only look for the presence or absence of the attribute would treat that code as *true*, resulting in very inconsistent behavior.

Boolean attributes that may need to be changed include:

`async`
>   change to `async="async"`

`checked`
>   change to `checked="checked"`

`compact`
>   change to `compact="compact"`

`declare`
>   change to `declare="declare"`

`defer`
>   change to `defer="defer"`

`disabled`
>   change to `disabled="disabled"`

`ismap`
>   change to `ismap="ismap"`

`multiple`
>   change to `multiple="multiple"`

`noresize`
>   change to `noresize="noresize"`

`noshade`
>   change to `noshade="noshade"`

`nowrap`
>   change to `nowrap="nowrap"`

`open`
>   change to `open="open"`

`readonly`

> change to `readonly="readonly"`

`required`

> change to `required="required"`

`reversed`

> change to `reversed="reversed"`

`scoped`

> change to `scoped="scoped"`

`selected`

> change to `selected="selected"`

Note that `"true"` and `"false"` *are* valid values for some *non-boolean attributes*, in particular enumerated attributes such as the `draggable` attribute.

Detecting which version of HTML is being used

A good way to determine which version of HTML a web site is using is to submit the URL of the web site to the W3C Markup Validation Service. Possible results include:

HTML5

> indicates that the site has already been converted to HTML 5. For example, Google's web site is using HTML 5.

XHTML 1.0 Strict

> indicates that the site is using the 2000 W3C standard version of HTML. For example, the W3C web site itself adheres to that standard.

XHTML 1.0 Transitional or HTML 4.01 Transitional

> indicates that the site is using a transitional format between the 1997 HTML 4 standard and the 2000 W3C standard version of HTML. For example, AltaVista uses the HTML 4.01 Transitional format and Microsoft's web site uses the XHTML 1.0 Transitional format. A web site that is using the 2000 XHTML 1.0 Transitional format is easier to convert to HTML 5 than one that is using the 1997 HTML 4.01 Transitional format.

HTML 4.01 Strict

> indicates that the site is using the older 1997 HTML 4 version of HTML. For example, the Yahoo! web site uses the older HTML standard.

> indicates that the site is using a transitional format between the 1997 HTML 4 standard and the 2000 W3C standard version of HTML. For example, Microsoft's web site uses this transitional format.

If the W3C HTML validator reports any errors while checking the HTML syntax of your own web site, fixing those errors will make the rest of the conversion to HTML 5 go easier.

back to top

THE END