9. Language Attribute in Web Pages

In a multilingual web environment, the information about natural language of a web page could not be inferred from the character encoding as the character encoding does not enable unambiguous identification of a natural language. There *must* be a 1:1 mapping between encoding and language for this inference to work. For example, a single character encoding could be used for many languages, eg, Devanagari (Unicode U+900-97F) could encode both Hindi and Sanskrit, as well as many other languages.

The language attribute unambiguously specifies the natural language of web page content. It should always be used to indicate the primary language of the web page. If the language changes within the main page container element this should also be reflected in a sub container element.

The content language information can be useful for many applications. Some of these work at the level of the document as a whole, some work on appropriately labeled document fragments. A few possible applications, which should have the language tags information are presented here:

Authoring tools

Authoring tools can supply appropriate spelling and grammar checking based on the language of a segment.

Translation tools

Translation tools can use the tags to help recognize sections of text in a particular language.

Accessibility

The language attribute assists speech synthesizers and Braille translators in W3C-Web Accessibility Initiative (WAI).

Font selection

User-agents can (and do) use the content language to select language-appropriate fonts, which improves the overall user experience of the page.

Page rendering

CSS2 uses the language attribute powerfully as a pseudo class. For example, it may be required to use different font size for different font sizes.

Search

Search engines can group or filter results based on the user s linguistic preferences.

It is also common to use meta tags to specify keywords that a search engine may use to improve the quality of search results. When several meta elements provide language-dependent information about a document, search engines may filter on the meta elements, using associated language attributes, and display search results according to the language preferences of the user.

Parsing

The language information may be used to extract or identify specific text, while processing the file.

References:

www.w3.org