

Treinamento essencial em HTML - referência à sintaxe

Como o HTML costuma ser a primeira experiência das pessoas com programação ou codificação, pode haver muita apreensão sobre o aprendizado. Felizmente, a sintaxe HTML é relativamente simples e fácil de aprender. A maioria das pessoas aprende o básico do HTML e começa a criá-lo no mesmo dia.

Uma linguagem de marcação

HTML é uma linguagem de marcação. Isso significa que o conteúdo da página é "marcado" por tags para identificar o conteúdo dentro das tags. Um parágrafo, por exemplo, pode ser identificado colocando a tag de abertura "<p>" antes do conteúdo dos parágrafos e a tag de fechamento "</p>" no final de um parágrafo. O parágrafo completo ficaria assim: <p> Este é um parágrafo. </p>

As tags consistem do símbolo menor-que (<) seguido dos caracteres que identificam a tag (o "p" para parágrafo) e do símbolo maior-que para o terminar (>). As tags de fechamento de um elemento são exatamente iguais a uma tag de abertura, exceto que uma barra (/) irá preceder os caracteres das tags.

Embora a maioria dos elementos exija uma tag de abertura e fechamento, a tag de fechamento é opcional para alguns elementos e opcional para outros. Embora haja exceções à regra, na maior parte, qualquer elemento que tenha conteúdo associado à tag precisa de uma tag de fechamento também.

Estrutura básica do documento

O essencial de todos os documentos HTML consiste em três tags básicas. Primeiro, a tag html (<html>) é necessária para identificar o documento como um arquivo HTML. Diretamente dentro da tag html, você encontrará o elemento head, ou cabeça (<head>). Na "head" do

documento é onde você encontrará os metadados do documento, o título do documento e links para recursos externos, como o CSS e seus JavaScripts. Pense na "head" do documento como a área que não contém nenhum conteúdo visual da página. Em vez, ela contém informações sobre o documento e os recursos que ajudam a fazer a página funcionar.

Logo após a "head" do documento, você encontrará o "body", ou corpo (<body>). O "body" é onde você encontrará todo o conteúdo efetivo da página. Cabeçalhos, parágrafos, imagens, listas, tabelas e outros conteúdos são colocados aqui.

Portanto, o documento mais básico possível HTML ficaria assim:

```
<html>

  <head>

  </head>

  <body>

  </body>

</html>
```

DOCTYPES

Se você já olhou para páginas HTML, provavelmente notou uma tag longa e um pouco intimidante pouco antes da tag HTML de abertura. Esta é uma declaração do DOCTYPE ou tipo do documento. O DOCTYPE é um componente muito importante e muitas vezes incompreendido, das páginas HTML. Essencialmente, ele informa ao browser qual versão de HTML (ou XHTML) deve ser esperada. Assim ele sabe quais regras de sintaxe usar ao renderizar sua página. O tipo de documento usado deve ser baseado na versão do HTML que você está usando para criar a página. Para obter mais informações sobre DOCTYPES e a história por trás deles, confira a excelente seção sobre

DOCTYPEs de Mark Pilgrim em seu livro Dive Into HTML5. (<http://diveintohtml5.info/semantics.html#the-doctype>)

Aqui estão alguns dos tipos de documentos HTML mais comuns:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01  
Transitional//EN" "http://www.w3.org/TR/html4/  
loose.dtd">
```

HTML 4.0 transitório

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0  
Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/  
xhtml1-transitional.dtd">
```

XHTML 1.0 transitório

```
<!DOCTYPE HTML> HTML5
```

Estrutura de código

Documentos HTML criam estrutura aninhando elementos dentro um do outro. Você pode agrupar uma seção do conteúdo da página agrupando-os em um elemento `<div>` ou `<section>`, por exemplo. Ao aninhar uma tag dentro de outra, você deve primeiro fechar qualquer elemento filho antes de fechar qualquer elemento pai.

Essa sintaxe, por exemplo, seria incorreta:

```
<p>Você deve fechar todas as tags aninhadas  
<strong>primeiro!</p></strong>
```

Em vez disso, a sintaxe correta seria:

```
<p>Você deve fechar todas as tags aninhadas  
<strong>primeiro!</strong></p>
```

O HTML também possui regras específicas sobre quais elementos podem ser aninhados em outros elementos. Um parágrafo (<p>), por exemplo, não pode ser aninhado dentro de um título (<h1>).

Principalmente, essas regras são baseadas no tipo de conteúdo que a tag representa. O HTML 4 possui dois tipos básicos de conteúdo: **block** e **inline**. Elementos **block** normalmente ocupam sua própria linha dentro de um documento, geralmente empilhando um em cima do outro conforme são encontrados. Elementos **inline** podem aparecer dentro do fluxo de elementos de nível de bloco. Elementos como forte (), negrito (), ênfase (), itálico (<i>) e span () são sempre encontrados dentro de elementos em nível de bloco.

O HTML5, por outro lado, expande o conceito de tipos de conteúdo para sete categorias diferentes de conteúdo, com alguns elementos pertencentes a várias categorias. Isso substitui o conceito de elementos **block** e **inline** e criou algumas regras adicionais de sintaxe de aninhamento. Na maioria das vezes, as regras sobre elementos de aninhamento fazem sentido e são fáceis de entender. Você deve estar preparado, no entanto, para que leve algum tempo antes de entender todas as regras que envolvem os elementos de aninhamento. Uma forma de garantir que seu código seja estruturado corretamente é validar seu código por meio de um serviço como o serviço de validação de marcação do W3C. (<http://validator.w3.org/>)

Conteúdo de seccionamento

Ao marcar documentos, é natural ter seções de conteúdo que sejam definidas separadamente umas das outras, de forma a descrever o relacionamento desse conteúdo.

Atributos do elemento

Alguns elementos podem ser aprimorados através do uso de atributos. Os atributos permitem que você forneça mais informações ou funcionalidades adicionais ao conteúdo. Atributos são adicionados à tag

de abertura de um elemento e consistem em duas partes, o nome e o valor. Embora a sintaxe varie com base na versão do HTML que você está usando, é prática comum colocar valores entre aspas.

```
<h1 class="headline">Título principal do artigo</h1>
```

Elementos substituídos

Alguns elementos HTML representam o conteúdo que é substituído por um recurso externo, como uma imagem, um controle de formulário ou um arquivo de vídeo. Esses elementos são referidos como elementos substituídos e geralmente têm uma largura ou altura predeterminada. Em alguns casos, os elementos terão atributos que informam ao navegador onde encontrar um recurso externo, como uma imagem ou um vídeo.

```

```

Comentando código

Geralmente, é útil deixar comentários para você ou para outros desenvolvedores no seu código. Talvez você queira se lembrar do que um determinado script faz, ou mostrar para um colega de trabalho como estruturar adequadamente um conteúdo específico. Para fazer isso, você usará comentários. Os comentários devem ser escritos em sua própria linha e normalmente precedem ou aparecem diretamente após o código a que se referem. Os comentários começam com um símbolo de menor-que, um ponto de exclamação e dois hífenes (<!--) e terminam com dois hífenes seguidos por um símbolo de maior-que (-->).

Sintaxe de comentário:

```
<!-- Isso é um comentário -->
```

Usando caracteres especiais

Certos caracteres são reservados para o uso do HTML, o que significa que você não deve usá-los fora de sua finalidade específica. Os símbolos

de menor-que (>) e menor-que (<), por exemplo, são reservados para tags, e usá-los em qualquer lugar fora de uma tag pode causar erros de análise nos browsers. Para usar esses caracteres reservados e outros caracteres especiais ou símbolos, você pode usar o que é conhecido como entidade de caractere nomeado.

Esses códigos especiais informam ao browser para substituir o código da entidade por um caractere específico. Essas entidades começam com um e comercial (&) seguido pelo nome da entidade e terminam com um ponto-e-vírgula (;). Para exibir um e comercial, por exemplo, você digitaria `&` no lugar do verdadeiro e comercial. Você pode encontrar listas de entidade de caractere nomeadas dentro das especificações HTML, embora o formato dessa página possa ser um pouco difícil de se ler. Você também pode encontrar uma lista útil de entidades de caracteres na página HTML Reference e uma lista mais abrangente na Wikipedia. (http://en.wikipedia.org/wiki/List_of_XML_and_HTML_character_entity_references)