

HTML aula 26

Conjuntos de caracteres

Para mostrar uma página web corretamente o navegador precisa saber o conjunto de caracteres a ser usado.

O que é codificação de caracteres

O primeiro padrão de codificação de caracteres foi o ASCII.

O padrão ASCII definiu 128 caracteres alfanuméricos

Números de 0 a 9

Letras em Inglês: A-Z

Alguns caracteres especiais como ! \$ + - () ! @ < >

Conjunto ASCII

Os valores de 0 até 31 são usados para caracteres de controle.

ASCII de 32 até 126 são letras, dígitos e símbolos.

127 é um caracter de controle.

de 128 até 255 não são usados.

Padrão HTML5

O Unicode (UTF-8) cobre a maioria dos caracteres e símbolos no mundo.

O conjunto de caracteres padrão do HTML5 é o UTF-8

Conjunto UTF-8

É idêntico ao ASCII de 0 até 127.

Não usa de 128 até 159.

Idêntico ao ANSI de 160 até 255.

Contínua de 256 até mais de 10000 caracteres.

Regra @charset

Você pode especificar o conjunto de caracteres usado em código CSS usando:

```
@charset "UTF-8";
```

Ele deve ser o primeiro elemento no arquivo CSS.

Atributo charset

Para mostrar uma página corretamente, você deve indicar ao navegador qual o conjunto de caracteres usado na página:

```
<meta charset="UTF-8">
```


URL

URL é outra palavra para endereço web.

Uma URL pode ser composta de:

Palavras: www.google.com

Endereço IP: 74.125.196.105

Palavras são mais fáceis de lembrar!

Localizador de Recurso Uniforme

Os navegadores requisitam páginas dos servidores usando URL.

Uma URL é usada para endereçar um documento ou dado na web.

Sintaxe URL

Um endereço como:

`https://www.google.com/teste/oi.html`

Segue a sintaxe:

`esquema://prefixo.domínio:porta/caminho/nomearquivo`

Sintaxe URL

esquema://prefixo.domínio:porta/caminho/nomearquivo

esquema - define o tipo do serviço (comum **http** ou **https**)

prefixo - define o prefixo de um domínio (padrão **www** para http)

domínio - define o nome do domínio.

porta - define o número da porta, padrão 80 para http, não precisa digitar.

caminho - define um caminho no servidor (ou pasta raiz)

nomearquivo - define o nome do arquivo ou documento.

Esquemas URL comuns

http - protocolo de transferência de hipertexto.

https - http seguro, criptografado

ftp - protocolo de transferência de arquivos (download e upload)

file - um arquivo no computador

Codificação URL

URLs podem ser enviadas usando ASCII apenas.

Se a URL tiver código fora do ASCII ele deve ser convertido antes.

Conversão:

Caracteres não-ASCII são transmitidos em um formato ASCII.

Caracteres não-ASCII são trocados por % e dígitos hexadecimais.

URLs não podem conter espaços, espaços são trocados por + ou %20

Exemplo de troca

Á é trocado por %C3%81

Lista completa:

https://www.w3schools.com/tags/ref_urlencode.asp

Validação HTML

Você pode validar o código usando:

<https://validator.w3.org/>