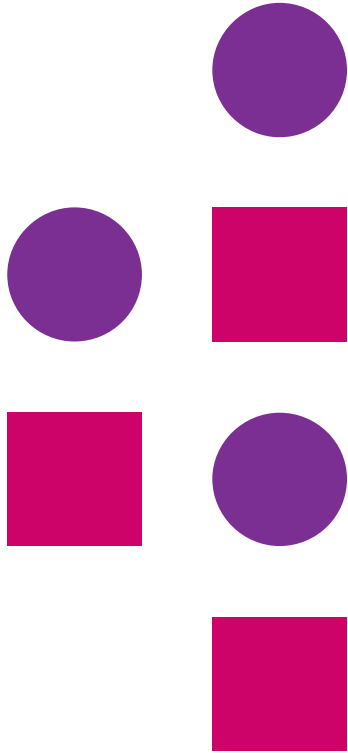
WSA | Forward looking IT finishing school

Entities, Symbols, Charset

Hypertext Markup Language 5 (HTML5)

Table of Content

- Entities
- Symbols
- Charset



Entities

(Hypertext Markup Language 5)

Entities

- Reserved characters in HTML must be replaced with character entities.
- Characters that are not present in keyboard can also be replaced by entities.
- Some characters are reserved in HTML.

  For example less than(<) or greater than(>) signs in the text browser might mix them with tags.

Character Entity

A character entity can be represented as

&entity_name;

or

&#entity_number;

so to display a less than sign(<) we must write < or <

Entities

Advantages

- An entity name is easy to remember.

Disadvantages

- Browser may not support all entity names but the support for number is good.

Javascript

Javascript makes HTML more dynamic and interactive.

Javascript is lightweight and cross platform.

The <script> tag is used to define a client-side script.

The <script> element either contains scripting statements or it points to an external script file.

JavaScript
Example

```
!DOCTYPE html>
<html>
<body>
<h1>JavaScript Example</h1>
<p>JavaScript can change the content of an HTML element:</p>
<button type="button" onclick="myFunction()">Click Me!</button>
<p id="demo">Welcome to Javascript Page</p>
<script>
function myFunction() {
    document.getElementById("demo").innerHTML = "Hello Friends!";
}
</script>
</body>
</html>
```

Symbols

(Hypertext Markup Language 5)

HTML Symbols

Many Mathematical ,technical and currency symbols ,are not present on a normal keyboard.

To add such symbols to an HTML page you can use an HTML Entity name.

If no entity name exists , we can use an entity number , a decimal , or hexadecimal reference.

HTML Symbols

- **Example**

   Output will be : €

  Output will be : €

- **Output**

  Output will be €

  Output will be €

Charset

(Hypertext Markup Language 5)

HTML Charset

To display an HTML page correctly , a web browser must know the character set used in the page.

Character Encoding

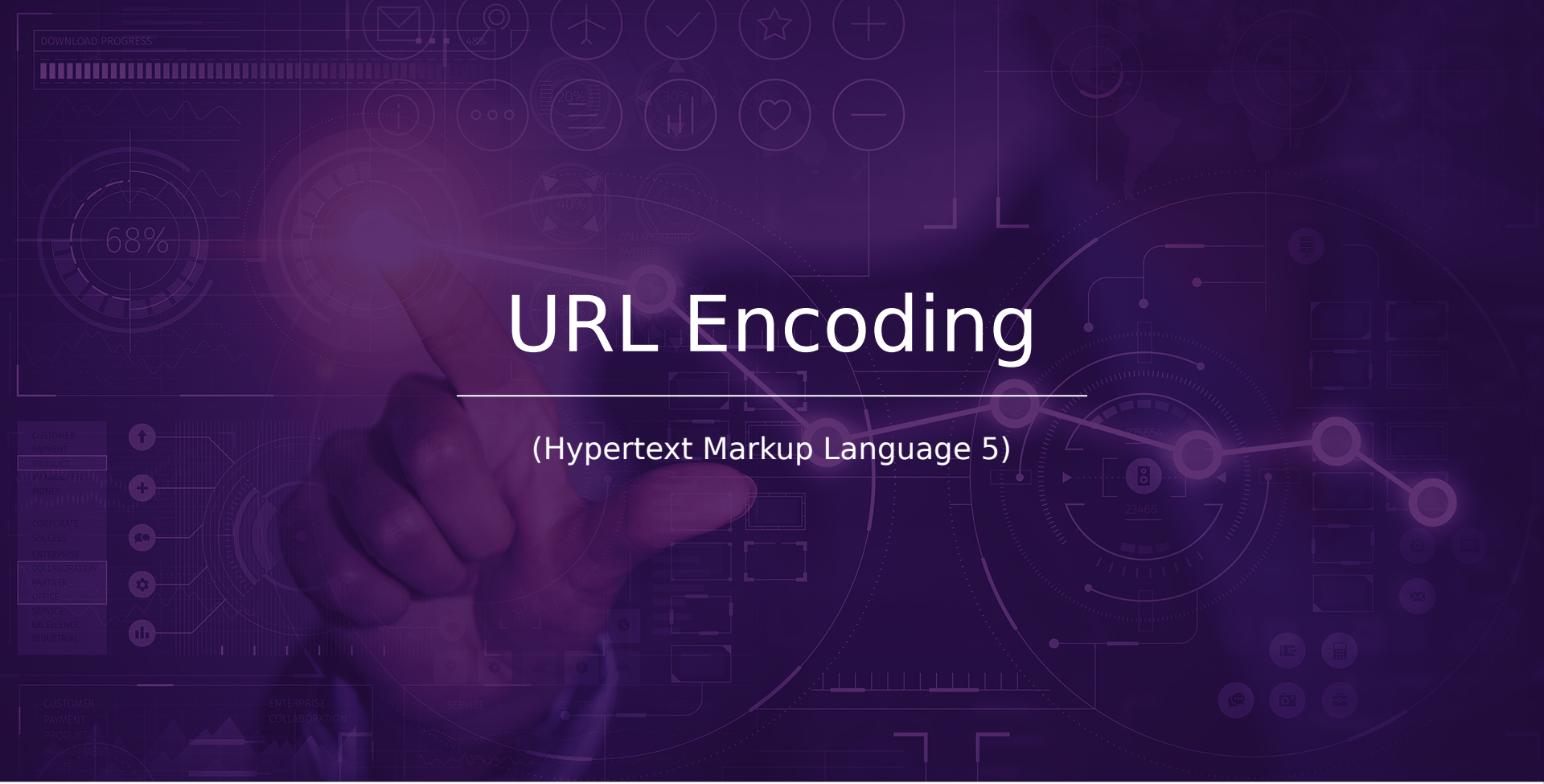
A character encoding is method of converting bytes into characters.

- ASCII was the first character encoding standard. ASCII defined 127 different alphanumeric characters.
- ANSI was the original Windows character set , with support for 256 different character code.
- ISO-8859-1 was the default character set for HTML 4. ANSI & ISO-8859-1 were so limited so default character encoding was changed to UTF-8 in HTML5.

The Charset Attribute

Charset is specified in the <meta> tag.

<meta charset=”UTF-8”>



URL Encoding

(Hypertext Markup Language 5)

URL(Uniform Resource Locator)

- A Uniform Resource Locator (URL) is used to address a document on the web
- A web address follows these rules
  - Scheme://prefix.domain:port/path/filename

URL

- scheme – defines the type of Internet service(http or https)
- prefix – defines a domain prefix(www)
- domain - defines the Internet domain
- port - defines the port number at the host (default for http is 80)
- path - defines a path at the server
- filename - defines the name of a document or resource

URL Encoding

- URL encoding converts characters into a format that can be transmitted over Internet
- Any attribute with a value that is a URL must be URL-encoded
- Example
  - <a href>
  - <img src>
  - <iframe src>

URL Encoding

- URLs can only be sent over the internet using the ASCII character-set
- URL encoding replaces unsafe ASCII characters with a “%” followed by two hexadecimal digits
- URLs cannot contain spaces. URL encoding normally replaces a space with a plus(+) sign or with %20

WSA | Forward looking IT finishing school

Web Stack Academy (P) Ltd

#83, Farah Towers,

1st floor,MG Road,

Bangalore – 560001

M: +91-80-4128 9576

T: +91-98862 69112

E: info@www.webstackacademy.com