

Tema 3 Lenguaje HTML

Introducción

CRONOLOGÍA

- **1993: Hypertext Markup Language (Primera versión), publicado por la IETF como Working-Draft (no era un estándar todavía) :** Permitía principalmente el uso de párrafos, cabeceras, saltos de línea, imágenes, enlaces y listas de elementos. Además definía la Estructura Básica de Documento y se añadieron las etiquetas de negrita, itálica y formularios.
- **1995: HTML 2.0, publicado por la IETF como documento RFC-1866.** Define los elementos de INPUT, SELECT y TEXTAREA. También se redefinen los enlaces, las listas de imágenes y las cabeceras.
- **1997: HTML 3.2, publicado como Recomendación del W3C.** Añadió las tablas, los atributos de alineación y las imágenes de fondo, e incluyó scripts y hojas de estilo (CSS). Formalizó el uso de colores de fondo, textos y enlaces, así como tamaños de imágenes y alineaciones.
- **1998: HTML 4.0, publicado como Recomendación del W3C.** Incorpora el uso de frames y la posibilidad de añadir objetos. Define las etiquetas DIV y SPAN, que usadas con CSS pueden crear nuevas etiquetas.

CRONOLOGÍA

- **1999: HTML 4.01:** publicado como Recomendación del W3C.
- **2000: ISO/IEC 15445:2000 (basado en HTML 4.01 sintaxis estricta),** publicado como ISO/IEC Estándar Internacional.
- **2000: XHTML 1.0,** publicado como Recomendación del W3C (se revisó y volvió a publicar en 2002).
- **2010:HTML5**

COMPATIBILIDAD

- A lo largo del desarrollo de esta unidad se utilizará el estándar de HTML 4.01 que es compatible con los navegadores Internet Explorer , Firefox , Netscape, Opera (programas que se utilizarán para visualizar los documentos realizados).
- Podemos encontrar el estándar "HTML 4.01 Specification"
 - <http://www.w3.org/TR/html401/> (en inglés)
 - <http://www.w3.org/MarkUp/html4-updates/translations#Spanish> (en castellano).

Entornos WEB

- Se basan en la explotación de los recursos existentes el WWW (World Wide Web).
- Esta red es un conjunto de recursos de información, estructurada en forma de hipertexto y que se encuentra operativa mediante la red internet.

ENTORNOS WEB

Para ello, se emplean:

- Lenguaje de marcas que permite manejar la información en forma de hipertexto → HTML (HyperText Markup Language)
- Protocolo HTTP (HyperText Transfer Protocol)
- Uniform Resource Locator (URL): Especifica cómo a cada página de información se le asocia un "nombre" único.

ENTORNOS WEB

La estructura de un documento:

- Texto simple. Ej. Artículo de opinión.
- Imágenes: Ej. Foto.
- Enlaces a otros elementos:
 - Locales: Enlaces a lugares del mismo documento.
 - Remotos: Enlaces a otros documentos del mismo servidor o de otro distinto.

ACCESO A INFORMACIÓN

Para acceder a la información en hipertexto:

- Navegador Web: Programa que permite el acceso a los recursos mediante protocolos y lenguajes de hipertexto.
Ej. <http://www.google.es>
<https://www.marca.es>
<ftp://www.recursos.net>
- Servidor Web: Son ordenadores que atienden a peticiones de otros equipos a través de protocolos de transferencia de información (http, https, ftp, etc) En los servidores web se almacenan las páginas web que contienen los recursos a los que se quiere acceder del servidor.

SERVIDOR WEB

- Un servidor web es un programa que sirve para atender y responder a las diferentes peticiones de los navegadores, proporcionando los recursos que soliciten usando el protocolo HTTP o el protocolo HTTPS.
- Un servidor web básico cuenta con un esquema de funcionamiento muy simple, basado en ejecutar infinitamente el siguiente bucle:

FUNCIONAMIENTO GENERAL



SERVIDOR

1. Espera peticiones en el puerto TCP
2. Recibe una petición.
3. Busca el recurso.
4. Envía el recurso utilizando la misma conexión por la que recibió la petición.
5. Vuelve al primer punto.

CLIENTE

1. Cliente realiza una petición por protocolo http.
2. Servidor recibe la petición.
3. Buscar el recurso asociado.
4. Envía el recurso al cliente en el protocolo

Características de HTML

- **HTML** está estandarizado y es multiplataforma.
- **HTML** es un estándar internacional mantenido por el W3c. (World Wide Web COnsortium)
www.w3c.es
- **No es un** lenguaje de programación en el sentido estricto de la palabra. Para añadir esta capacidad de proceso es necesario añadir algún lenguaje de script como Javascript.
- **Un** documento escrito en HTML puede ser leído e interpretado por cualquier navegador: Firefox, opera. Explorer, etc.

CARACTERÍSTICAS HTML

- Es un **lenguaje de marcas o etiquetas** muy sencillo que permite describir **hipertexto** (idea de *Ted Nelson*), es decir, texto presentado de forma estructurada y agradable, con *enlaces* (**hyperlinks**) que conducen a otros documentos o fuentes de información relacionadas, y con *inserciones* multimedia (gráficos, sonido...)
- La descripción se basa en especificar en el texto la estructura lógica del contenido (**títulos, párrafos de texto normal, enumeraciones, definiciones, citas, etc**) así como los **diferentes efectos** que se quieren dar (especificar los lugares del documento donde se debe poner cursiva, negrita, o un gráfico determinado) y dejar que luego la presentación final de dicho texto se realice por un programa **intérprete** especializado (como **Explorer, Mozilla Firefox...**).

CARACTERÍSTICAS HTML

- Los documentos **HTML son ficheros de texto (ASCII)** que **se** pueden crear con cualquier editor de texto (notepad, vi, emacs, ...). La **extensión de los ficheros** que contienen código HTML **es .html o .htm**
- No obstante, escribir y mantener un gran número de páginas a mano requiere esfuerzo y conocimiento en profundidad de HTML. Para evitar este problema existen **editores visuales** como Macromedia Dreamweaver, Microsoft FrontPage, etc., que permiten editar páginas Web de una forma similar a como se edita un documento de texto. Sin embargo, estos editores pueden llegar a generar código HTML de muy baja calidad sobre el que tarde o temprano habrá que realizar una operación de limpieza y reestructuración, utilizando un editor de texto.
- **Es texto plano interpretado por un navegador.**