

Unidad 1. Introducción a los Lenguajes de Marcas

Los **lenguajes de marca** son aquellos que combinan la información que contiene un documento con marcas o anotaciones. Son las reglas de codificación que indican como se ha ordenado la información. Estos designan las **etiquetas** (anotaciones con un significado especial). El organismo que regula los estándares es el W3C. Estas etiquetas facilitan la organización de la información y por lo tanto, su consulta.

Evolución de los lenguajes de marcas:

(*Examen)

- Años 60: IBM** crea el **GML** (Lenguaje de marcado general) que asentó las bases pero era complejo de usar. Poco después apareció **SGML** (Estándar generalizado de lenguaje de marca) para el software que lo usaba era externo y complejo.
- Años 80: El CERN** (Organización europea de investigación nuclear) crea **HTML** (lenguaje de marca con hipertexto)
- Años 90: W3C** hace público un estándar llamado **XML***(Es un tipo de fichero. Las siglas significan Extended Markup Language (Lenguaje de Marca Extensible). Es un núcleo estándar que hizo W3C, más sencillo que SGML y más potente que HTML.)*

Características del lenguaje de marcas:

1º **Texto plano:** Como usan texto, una de sus grandes ventajas es la facilidad de interpretación. Al ser todo texto, los documentos son independientes de la plataforma y, además, no necesitan de un programa intermediario, lo que les da una ventaja sobre los sistemas de archivos binarios.

2º **Facilidad de procesamiento:** Al ser fáciles de procesar ha permitido que se desarrollen lenguajes especializados como **CALS**, usado por las fuerzas armadas de EEUU.

3º **Flexibilidad:** Su sencillez y portabilidad los ha llevado a usarse en webs, interfaces, etc. Además se pueden combinar en función del medio multimedia a utilizar, como **SVG** (para imágenes)

4º **Compacidad:** El espacio se aprovecha al máximo, ya que el texto entre etiquetas es el contenido del documento

Tipos de lenguajes de marcas:

Puede agruparse en 3 grandes grupos:

- Lenguajes orientados a objetos:** El usado por procesadores de texto que codifican como se debe presentar al usuario (marcando negrita, cursiva...). Las marcas se suelen ocultar al usuario. El más usado es **HTML**, fundamento del **WWW***(World Wide Web)
- Lenguajes procedurales:** Aunque las etiquetas son orientadas a presentación, se integran dentro de un marco procedural que permita definir subrutinas. Los más conocidos son Tex, LaTeX y PostScript.
- Lenguajes descriptivos:** No definen qué hacer con la información sino que simplemente indican que se está representando.

Etiquetas, elementos, atributos:

-**Etiquetas:** Texto que va entre < y > que dan un nombre a la información que contienen. Puede ser de apertura o de cierre </ejemplo>

-**Elementos:** Unión de las etiquetas y su contenido, aunque hay elementos vacíos que no tienen que llevar etiquetas de fin.

-**Atributos:** Par nombre-valor encontrados dentro de una etiqueta de inicio que describe las propiedades de los elementos.

Gramática:

Los lenguajes de marcas necesitan una gramática que designa cómo se representan los elementos.

- DTD:** Definición de tipo de documento, describe la combinación de símbolos que se usan en el documento.
- XMLSchema:** Combina DTD con XML que permite especificar con más detalles diferentes tipos de datos.
- Relax NG:** Es más fácil de entender que XMLSchema ya que permite utilizar elementos intercalados.

Principales lenguajes de marcas:

- HTML:** Pudiendo mezclar texto con contenido multimedia, se enfoca en páginas web siendo el W3C el encargado de su mantenimiento.
- XML:** Se rige por reglas y por etiquetas, permite definir lenguajes específicos, según las necesidades del metalenguaje. Algunos lenguajes basados en XML son:
 - XHTML:** Usado para presentar información en la web.
 - MathML:** Intercambian información en ámbito matemático

- SVG**: Para gráficos vectoriales
- RRS y ATOM**: Agregadores de noticias para suscritos a la web
- OpenDocument**: Almacena y presenta información en paquetes de ofimática
- Open eBook**
- JSON**: Alternativa a XML para almacenar e intercambiar datos entre sistemas