

1.2. Nociones básicas sobre lenguajes de marcado

La definición más común establece que un lenguaje de marcado o lenguaje de marcas es una forma de codificar un documento que, junto con el texto, incorpora etiquetas o marcas que contienen información acerca de la estructura del texto o de su formato de presentación. Estos lenguajes pueden hacer explícita la estructura del documento que se trate, pueden indicar el contenido semántico, o pueden señalar e indicar cualquier otro tipo de información que pueda ser relevante para un uso dado.

Los lenguajes de marcas se dividen en tres grandes grupos:

- **Lenguajes de presentación:** son aquellos orientados a definir el formato o la capa de presentación del texto. Suelen ocultar las etiquetas y mostrar al usuario solamente el texto con su formato. El conocido RTF para ficheros de texto es un marcado de este tipo.
- **Lenguajes de procedimientos:** orientados también a la presentación, pero además incorporan elementos que la aplicación o programa que representa el documento debe interpretar para ejecutar acciones en función de éstos. El HTML de las páginas web es un ejemplo.
- **Lenguajes descriptivos o semánticos:** son lenguajes diseñados para representar las diferentes partes en las que se estructura un documento, y para definir su contenido. Sin embargo, y a diferencia de los anteriores, no especifican cómo deben representarse los documentos en su capa visual. Son los utilizados para facilitar el intercambio de información y datos entre aplicaciones. XML es el estándar actual para ello.

El funcionamiento de los lenguajes de marcado es simple: un elemento se destaca del resto de información mediante una marca o etiqueta:

El documento marcado se somete a un procesador, que interpreta las marcas y genera un documento final, o bien ejecuta una serie de acciones sobre el contenido etiquetado (presentación en pantalla, incorporación a una base de datos, relación con otros elementos etiquetados...). Es importante señalar que **un mismo documento puede contener al mismo tiempo diferentes lenguajes de marcado**, y que será el procesador de documentos o la aplicación que lo trate, en cada caso, la que decida que debe hacer con el contenido marcado.

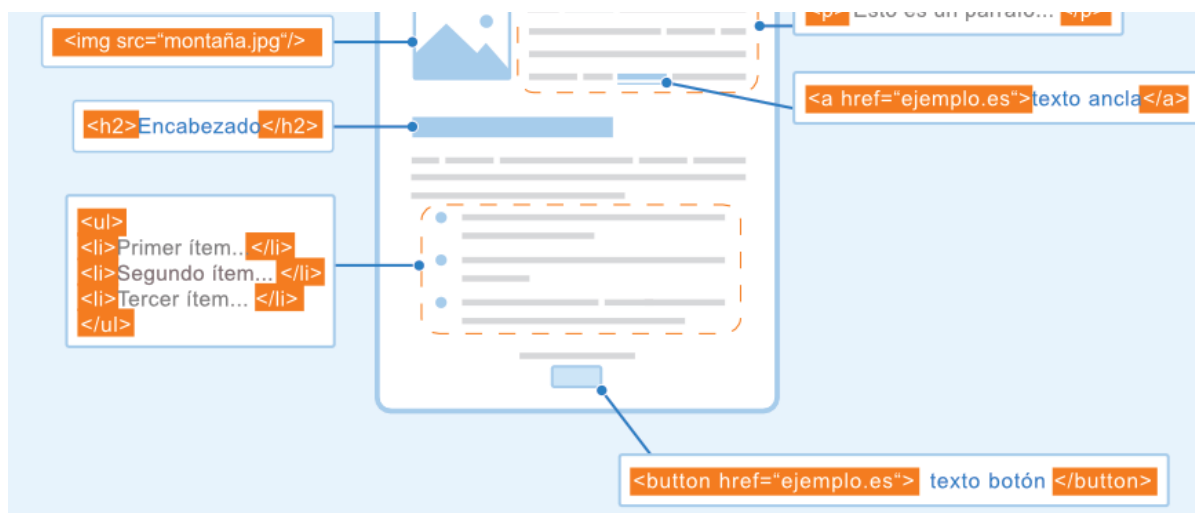


Fig. 1.

Ejemplo de marcado en HTML y su representación visual ([fuente original](#)).

Los lenguajes de marcas se usan para etiquetar, para marcar, elementos dentro de un documento. En el párrafo anterior se ha indicado que este documento se somete a un procesador. Ahora bien ¿cómo sabe el procesador cómo procesar, valga la redundancia, las marcas? Esto es posible porque todos y cada uno de **los lenguajes de marcado tienen un documento de referencia en el que se explicitan las reglas sobre cómo se estructuran los documentos, que marcas y etiquetas se usan, lo que significan, y cómo se aplican y relacionan.** Estos documentos de referencia se identifican con las siglas *DTD (Document Type Definition)*, o *XML Schema*. Por ejemplo, hay un DTD/XML Schema para HTML, otro para XML, etc. Estos documentos de referencia se publican en internet de manera abierta, de forma que cada procesador o aplicación pueda acudir a la url correspondiente, y cargarlo para saber cómo actuar ante cada documento.

El **flujo de trabajo** que se establece es el siguiente:

1. Una persona o una máquina crea un **documento de texto, sobre cuyo contenido aplica un lenguaje de marcado.**
2. Una aplicación o un **procesador accede al contenido** de ese documento.
3. **Identifica**, generalmente en la cabecera o primeras líneas de texto, los **DTD o XML Schema que debe usar para procesar el contenido.**
4. **Utiliza los url de los DTD/XML Schema para ir a la localización original, y cargar su contenido** como parámetros de trabajo.
5. Una vez cargados, **procesa el documento y su contenido marcado, de acuerdo con las reglas obtenidas** del DTD/XML Schema.

El resultado de este procesamiento puede ser la visualización de una página web, la creación de un nuevo documento, la incorporación de datos a una base de datos, u otros que puedan haber sido programados en la aplicación correspondiente. Por ejemplo, la posibilidad de crear un documento HTML, o en formato [EPUB](#), desde un documento generado por un procesador de textos, es un ejemplo del uso de lenguajes de marcado y del procesamiento que crea diferentes tipos de resultados en virtud de diferentes reglas de procesamiento, todo ello de forma transparente para el

usuario final.

Revision #18

Created 26 January 2023 11:05:39 by Jesús Tramullas

Updated 8 February 2023 17:13:28 by Jesús Tramullas