

Lenguajes de marcas

Introducción a los lenguajes de marcas

Objetivos

- Conocer el mundo de los lenguajes de marcas, identificando sus características más generales, reconociendo sus ventajas, estudiando su clasificación y estructura e identificando lo más relevante.
- Conocer el ámbito de aplicación de estos lenguajes, con ejemplos reales de su uso.

```
<?xml version="1.0" encoding="iso-8859-8" standalone="yes" ?>
<CURRENCIES>
  <LAST_UPDATE>2004-07-29</LAST_UPDATE>
  <CURRENCY>
    <NAME>dollar</NAME>
    <UNIT>1</UNIT>
    <CURRENCYCODE>USD</CURRENCYCODE>
    <COUNTRY>USA</COUNTRY>
    <RATE>4.527</RATE>
    <CHANGE>0.044</CHANGE>
  </CURRENCY>
  <CURRENCY>
    <NAME>euro</NAME>
    <UNIT>1</UNIT>
    <CURRENCYCODE>EUR</CURRENCYCODE>
    <COUNTRY>European Monetary Union</COUNTRY>
    <RATE>5.4417</RATE>
    <CHANGE>-0.013</CHANGE>
  </CURRENCY>
</CURRENCIES>
```

Definición

Según Wikipedia "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 adicional acerca de la estructura del texto o su presentación."

Tipos de lenguajes de marcas: de presentación.

- Define el formato del texto.
- Este tipo de marcado es útil para maquetar la presentación de un documento para su lectura, pero resulta insuficiente para el procesamiento automático de la información.
- El marcado de presentación resulta fácil de elaborar. Sin embargo resulta complicado de mantener o modificar, por lo que su uso se ha ido reduciendo en proyectos grandes en favor de otros tipos de marcado más estructurados.
- Se puede tratar de averiguar la estructura de un documento de esta clase buscando pistas en el texto. Por ejemplo, el título puede ir precedido de varios saltos de línea (o renglón), y estar ubicado centrado en la página web. Varios programas pueden deducir la estructura del texto basándose en esta clase de datos, aunque el resultado suele ser bastante imperfecto.
- Un ejemplo de marcado de presentación, puede ser RTF.

Tipos de lenguajes de marcas. Marcado de procedimientos.

- Está enfocado hacia la presentación del texto, sin embargo, también es visible para el usuario que edita el texto.
- El programa que representa el documento debe interpretar el código en el mismo orden en que aparece. Por ejemplo, para formatear un título, debe haber una serie de directivas inmediatamente antes del texto en cuestión, indicándole al software instrucciones tales como centrar, aumentar el tamaño de la fuente, o cambiar a negrita. Inmediatamente después del título deberá haber etiquetas inversas que reviertan estos efectos.
- En sistemas más avanzados se utilizan macros o pilas que facilitan el trabajo.
- Algunos ejemplos de marcado de procedimientos son nroff, troff, TeX. Este tipo de marcado se ha usado extensivamente en aplicaciones de edición profesional, manipulados por tipógrafos calificados, ya que puede llegar a ser extremadamente complejo.

Tipos de lenguajes de marcas. Marcado descriptivo o semántico.

- Utiliza etiquetas para describir los fragmentos de texto, pero sin especificar cómo deben ser representados, o en qué orden.
- Los lenguajes expresamente diseñados para generar marcado descriptivo son el SGML y el XML.
- Las etiquetas pueden utilizarse para añadir al contenido cualquier clase de metadatos. Por ejemplo, el estándar Atom, un lenguaje de sindicación, proporciona un método para marcar la hora «actualizada», que es el dato facilitado por el editor de cuándo ha sido modificada por última vez cierta información.
- Una de las virtudes del marcado descriptivo es su flexibilidad: los fragmentos de texto se etiquetan tal como son, y no tal como deben aparecer. Estos fragmentos pueden utilizarse para más usos de los previstos inicialmente. Por ejemplo, los hiperenlaces fueron diseñados en un principio para que un usuario que lee el texto los pulse. Sin embargo, los buscadores los emplean para localizar nuevas páginas con información relacionada, o para evaluar la popularidad de determinado sitio web.

Tipos de lenguajes de marcas. Marcado descriptivo o semántico.

- El marcado descriptivo también simplifica la tarea de reformatear un texto, debido a que la información del formato está separada del propio contenido.
- El marcado descriptivo está evolucionando hacia el marcado genérico. Los nuevos sistemas de marcado descriptivo estructuran los documentos en árbol, con la posibilidad de añadir referencias cruzadas. Esto permite tratarlos como bases de datos.

Uso de los lenguajes de marcas

- Mensajería instantánea (XMPP)
- Servicios Web (WSDL y SOAP)
- Páginas Web (HTML)
- Sindicación de contenidos (RSS)
- Gráficos vectoriales X3D
- Documentación electrónica (RTF)

Características comunes a los lenguajes de marcas

- Son en texto plano
- Interoperabilidad o independencia: al estar en texto plano pueden abrirse y editarse desde cualquier máquina, con independencia de la arquitectura de la misma. Solo hay que tener en cuenta la codificación.
- Flexible y fáciles de crear: basta con un editor de texto plano para crearlos.