

Lenguaje Mercado

Introducción

Introducción. Definición

- 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.
- El lenguaje de marcas más extendido es el **HTML** "HyperText Markup Language" (Lenguaje de marcado de hipertexto), fundamento del **World Wide Web**.
- ☞ Los lenguajes de marcado suelen confundirse con **lenguajes de programación**. Sin embargo, no son lo mismo, ya que el lenguaje de marcado no tiene funciones aritméticas o variables, como sí poseen los lenguajes de programación.

Características de los lenguajes de marcas

- Desde un punto de vista más técnico, un **lenguaje de marcas** (*markup language*, en inglés) es un sistema que permite incluir anotaciones, llamadas marcas o etiquetas, en un texto. Las finalidades de esto pueden ser varias.
- Por ejemplo, indicar como se va a presentar el texto cuando se visualice en un navegador web.
- Así, el lenguaje HTML (*HyperText Markup Language*) utiliza la etiqueta

** . . . **

para indicar que el texto se tiene que representar en negrita:

** Esto se vería en negrita **

Características de los lenguajes de marcas

- A los lenguajes de marcas también se les llaman **lenguajes de etiquetaje o marcado** y entre sus características principales encontramos las siguientes:
 - **Separan el contenido de su presentación.** Es decir, por un lado tenemos el contenido del documento propiamente dicho y, por el otro, cómo queremos que se muestre. Por ejemplo, la presentación de un documento no será igual si se usa un dispositivo móvil, un ordenador o una impresora.
 - **Se codifican mediante archivos de texto plano, no binarios.** Se puede, por tanto, codificar las marcas con un simple editor de texto. Con todo, existen multitud de programas de edición para los lenguajes de marcas que facilitan la codificación.
 - **Facilitan la interoperabilidad.** El uso de archivos de texto plano posibilita el intercambio entre programas distintos e incluso entre sistemas operativos diferentes.
 - **Emplean etiquetas o marcas.** Las etiquetas (también llamadas *tags*, del inglés) permiten marcar el texto y, normalmente, van mezcladas con el propio documento sobre el que se trabaja. Una marca tiene la siguiente forma:

<nombre_etiqueta>Elemento</nombre_etiqueta>

Clasificación de los lenguajes de marcas

Se acostumbra a agrupar los lenguajes de marcas en tres categorías, aunque no son excluyentes. De hecho, existen lenguajes de marcas, como HTML, **que pueden formar parte de más de una de ellas:**

- ❑ **Lenguajes de marcado de presentación:** donde sólo se muestra la presentación, pero es difícil extraer la información.
 - ❑ **Lenguajes de marcado procedimentales:** donde se incluyen instrucciones de cómo hay que procesar el texto.
 - ❑ **Lenguajes de marcado descriptivos:** donde no se especifica cómo hay que procesar el texto sino que se utilizan etiquetas que describen el texto.
-

Clasificación de los lenguajes de marcas

- ❑ **Lenguajes de marcado de presentación.** Son los que usaban hasta hace poco los procesadores de textos como el MS Word (ahora ya usa el XML como el OpenOffice Writer). Las etiquetas indican, por ejemplo, “esto está en cursiva”, “esto está justificado a la derecha”, etc. Este tipo de lenguajes son útiles para maquetar la presentación de un documento para su lectura, pero no sirven para el procesamiento automático de la información ya que son poco flexibles y difíciles de mantener. Frecuentemente, el usuario no es consciente del marcado ya que se emplean programas de alto nivel para realizarlo. El RTF (*Rich Text Format*) es un tipo de lenguaje de marcado de este tipo.

👉 En Word puedes ver las marcas pulsando el icono  de la interfaz de Microsoft Word

Clasificación de los lenguajes de marcas

- ▣ **Lenguajes de marcado procedimentales.** Como las anteriores, están enfocadas a la presentación del texto, con todo, las etiquetas pueden incluir la ejecución de algún comando o rutina que indica como se tiene que tratar el texto. Normalmente las etiquetas son visibles para el usuario que edita el texto. Ejemplos de estos lenguajes son TeX, LaTeX y Postscript.

Clasificación de los lenguajes de marcas

- **Lenguajes de marcado descriptivos.** En esta categoría se encuentra el XML (*eXtensible Markup Language*) y el SGML y sus derivados (HTML, etc). Ambos son estándares del W3C, que es el **organismo encargado de la estandarización de las tecnologías sobre las que se sustenta la Web.**

Estos lenguajes no indican qué hacer con el texto, sino qué es el texto. Como su propio nombre indica, lo describen. La idea clave es que las marcas no determinan el procesamiento del documento de manera fija (como los procedimentales), ya que dicho procesamiento se determina a partir de necesidades concretas, y se beneficia de la estructura lógica del documento caracterizada a través de sus marcas.

Clasificación de los lenguajes de marcas

Además de la clasificación anterior, conviene mencionar los principales lenguajes de marcas en el área de las tecnologías Web:

- ❑ **SGML** (*Standar Generalized Markup Language*). En realidad es un metalenguaje, es decir, se usa para crear otros lenguajes de marcas. Por ejemplo, la gramática del HTML se creó a partir de él. Es muy potente pero también compleja de utilizar. Motivo por el cual se creó el XML.
- ❑ **XML**. Es un subconjunto de SGML. Por tanto, es también un metalenguaje pero mucho más fácil de usar. Su aplicación va más allá de Internet, siendo un estándar para el intercambio de información estructurada entre diferentes plataformas.
- ❑ **HTML** (*HyperText Markup Language*). Es el lenguaje de marcas más conocido ya que se usa para convertir documentos de texto en páginas web. Para solventar sus carencias se creó XHTML.
- ❑ **XHTML**. Es una reformulación de HTML en XML. En otras palabras, emplea el mismo vocabulario de HTML pero con reglas sintácticas tomadas del XML, mucho más estrictas.

Ejemplos de Lenguajes de Marcas

```
<!DOCTYPE motd [ <!E
<motd>
<!-- created: 2003-12-12--
<sentence>Do not throw
out the <keep>baby</>
with the
<refuse>dirty</>,
<refuse>stinky</>,
<refuse>bathwater</>.
</>
<!-- finish this later-->
</motd>
```

SGML

```
<!DOCTYPE html>
<html>
<!-- created 2010-01-01 -->
<head>
<title>sample</title>
</head>
<body>
<p>Voluptatem accusantium
totam rem aperiam.</p>
</body>
</html>
```

HTML

```
<?xml version="1.0"
<!DOCTYPE html PUBLIC
<html xmlns="http://www.w3
<head><title>XYZ</title>
</head>
<body>
<p>
voluptatem accusantium d
totam rem aperiam eaque
</p>
</body>
</html>
```

XHTML

```
<?xml version="1.0"?>
<quiz>
<question>
Who was the forty-second
president of the U.S.A.?
</question>
<answer>
William Jefferson Clinton
</answer>
<!-- Note: We need to add
more questions later.-->
</quiz>
```

XML

Características de los Lenguajes de Marcas

- ❑ Uso de texto plano.
- ❑ Compacidad: las instrucciones de marcado se entremezclan con el propio contenido.
- ❑ Facilidad de procesamiento.
- ❑ Flexibilidad.

Herramientas

- Editor de textos
 - Buscar y reemplazar
 - Cortar, copiar y pegar
 - Marcar texto
 - Resaltado de sintaxis (Syntax highlighting)
 - Formateo automático de texto (sangría automática, comentarios, ...)
 - Deshacer y rehacer
 - Macros
- ¿Editor genérico o específico? ¿Emacs23 con nxml-mode o conglomerate?
- ¿IDE? •
- Parser o analizador sintáctico: Integrado o no en el editor.
- Validador: Integrado o no en el editor

Herramientas

W3C

- El World Wide Web Consortium, abreviado W3C, es un consorcio internacional que produce recomendaciones para la World Wide Web.
-

Ejemplos de Lenguajes de Marcas

roff

.TL

The Road

.AU

Yours Truly

.AI

Koi Thikana Kahin Nahi

.sp 0.3i

.LP

.QS

ROAD, n. A strip of land along which one may pass from where
it is too tiresome to be to where it is futile to go.

.rj

\[em] Ambrose Bierce, \c

SGML

```
<!doctype linuxdoc system>
```

```
<article>
```

```
<titlepag>
```

```
  <TITLE>Ejemplo de formato SGML</TITLE>
```

```
  <author>&copy; 2010 Alberto Molina Coballes,
```

```
  <tt/alberto@gonzalonazareno.org/
```

```
    <abstract>
```

```
      Resumen del documento ...
```

```
    </abstract>
```

```
</titlepag>
```

```
<toc>
```


L^AT_EX

```
\documentclass{article}

\title{Ejemplo de formato LaTeX}
\author{\copyright 2010 Alberto Molina Coballes}
\email{alberto@gonzalonazareno.org}
\begin{abstract}
  Resumen del documento ...
\end{abstract}

\tableofcontents
```
