

Definición de Lenguaje de Marcas (I)

Los lenguajes de marcas (LM) son aquellos que combinan la información del documento con anotaciones (marcas o etiquetas) relativas a la estructura del texto o la forma de representarlo.

Cada lenguaje concreto especificará las marcas que se pueden utilizar junto con su significado y el lugar donde se colocan.

2

Definición de Lenguaje de Marcas (II)

Ejemplo:

```
<articulo>
  <titulo>Lenguajes de Marcas</titulo>
  <autor>Oscar Hernando</autor>
  <cuerpo>
    Ejemplo de <negrita>LM</negrita>
    ...
  </cuerpo>
</articulo>
```

3

Definición de Lenguaje de Marcas (III)

El término marca proviene del inglés *mark-up* y se refiere a la técnica de marcar documentos manuscritos con lápices de colores para indicar las diferentes tipografías que se utilizarían en imprenta.



4

Definición de Lenguaje de Marcas (IV)

Los lenguajes de marcas no son lenguajes de programación ya que no disponen de los elementos esenciales de éstos: variables, funciones, estructuras de control.

No obstante, un documento puede combinar un lenguaje de marcas que establezca la estructura y formato del documento con algún lenguaje de programación (Javascript, PHP, etc) que aporte mayor funcionalidad y dinamismo.

5

Evolución de los Lenguajes de Marcas (I)

- En los años 60 con el comienzo de uso de ordenadores para la edición de documentos se traslada la práctica del marcado.
- Cada empresa y/o aplicación utiliza sus propias marcas para definir los diferentes elementos, lo que provoca que sea imposible el intercambio de documentos entre plataformas.

6

Evolución de los Lenguajes de Marcas (II)

- IBM trata de resolver el problema de falta de compatibilidad describiendo GML (*Generalized Markup Language*).
- GML utiliza marcas genéricas para definir el formato y la estructura de los documentos, de manera que el mismo documento puede utilizarse en diferentes dispositivos. Basta con establecer un perfil para cada dispositivo.

7

Evolución de los Lenguajes de Marcas (III)

- En 1986 GML pasa a manos de la ISO y se convierte en SGML (ISO 8879), *Standard Generalized Markup Language*, software libre y de código abierto.
- SGML no es un lenguaje sino un metalenguaje que permite definir lenguajes de marcas como HTML, RFT, etc.

8

Evolución de los Lenguajes de Marcas (IV)

- Ventajas e inconvenientes de SGML
 - Ventajas:
 - Reutilización de datos.
 - Integridad y control sobre los datos.
 - Portabilidad.
 - Adaptabilidad.
 - Inconvenientes:
 - Complejidad.

9

Evolución de los Lenguajes de Marcas (V)

- Estructura de SGML
 - **Prólogo:** contiene la estructura.
 - Declaración: indica que el documento es SGML y algunos parámetros.
 - Definición del Tipo de Documento (DTD): indica la sintaxis particular del lenguaje creado.
 - **Instancia del documento:** contiene los datos.

10

Evolución de los Lenguajes de Marcas (VI)

- Ejemplo de SGML: Ciclos formativos
 - Vocabulario: cicloformativo, módulo, título, contenido, unidad.
 - Reglas:
 - Un cicloformativo tiene varios módulos.
 - Un módulo se compone de título y contenido.
 - Un contenido tiene varias unidades.
 - Detrás de una unidad sólo puede ir otra unidad o fin de contenido.
 - Detrás de un módulo sólo puede ir otro módulo o fin de cicloformativo.

11

Evolución de los Lenguajes de Marcas (VII)

■ Ejemplo de SGML: Ciclos formativos

```
<cicloformativo>
  <módulo>
    <título>Lenguajes de Marcas</título>
    <contenido>
      <unidad>1.Introduccion</unidad>
      <unidad>2.Lenguajes para la web</unidad>
      ...
    </contenido>
  </módulo >
  ...
</cicloformativo>
```

12

Características de los Lenguajes de Marcas (I)

- **Texto plano:** los archivos de texto plano están compuestos únicamente por caracteres de texto, de manera que pueden ser interpretados por un simple editor de texto y son independientes del S.O. facilitando la interoperabilidad.

- **Compacidad:** las marcas se mezclan con el propio contenido del documento.

<h2>Contenido</h2>

13

Características de los Lenguajes de Marcas (II)

- **Independencia del dispositivo final:** el mismo documento puede ser interpretado de diferentes formas dependiendo del dispositivo final (móvil, ordenador, impresora, etc).

- **Especialización:** se pueden utilizar en muchas áreas: texto, gráficos vectoriales, notación científica, sindicación, etc.

- **Flexibilidad:** en un documento se pueden combinar lenguajes de marcas y lenguajes de programación.

14

Clasificación de los Lenguajes de Marcas (I)

■ **Orientados a presentación:** indican como debe presentarse la información del documento.

Es el tipo utilizado tradicionalmente por los procesadores de texto.

En la mayoría de las ocasiones las marcas se ocultan al usuario, de manera que el texto se muestra con su formato final (WYSIWYG)

Ejemplos:

- nroff, troff, RTF
- TeX, Docbook (derivados de SGML)

15

Clasificación de los Lenguajes de Marcas (II)

■ **Procedurales:** las marcas están también orientadas a presentación pero se integran dentro de un marco que permite definir secuencias de acciones (macros) y subrutinas.

Ejemplos:

- LaTeX
- Postscript

16

Clasificación de los Lenguajes de Marcas (III)

- **Descriptivos:** establecen el significado de cada elemento del documento pero no indican cómo debe representarse la información ni en qué orden.

Estos lenguajes crean documentos con estructura de árbol en los que se almacena información, de ahí que se les considere bases de datos.

Ejemplos:

- ASN.1, YAML
- EBML, RDF, XFML, OWL, XTM (derivados de XML)

17

Clasificación de los Lenguajes de Marcas (IV)

- **Funcionalidad:**

- Creación de documentación electrónica:

- RTF, TeX, troff, nroff.
- ASN.1, EBML, TAML.
- Wikitexto, DocBook, LinuxDoc.

- Tecnologías de Internet:

- Web: HTML, XHTML, WML.
- Interfaces usuario: GladeXML, Xforms, XAML.
- Sindicación: RSS, Atom.
- Servicios web: WSDL, SOAP, UDDI.
- Mensajería instantánea: XMPP.

18

Clasificación de los Lenguajes de Marcas (V)

■ **Funcionalidad:**

- Propósito específico:
 - Matemáticas: MathML, CML.
 - Gráficos vectoriales: SVG, VML, S3D.
 - Síntesis voz: SSML, SRGS, VoiceXML.
 - Partituras: MusicXML.
 - Multimedia: SMIL.
 - Transformación documentos: XSLT.
 - Mapas conceptuales: XTM.
 - Información geográfica: GML.
 - Información financiera: OFX.
 - (...)

19

Organizaciones y Estándares (I)

La normalización o estandarización es el proceso mediante el cual se definen unas normas o reglas que garanticen el correcto funcionamiento y compatibilidad de elementos contruidos de forma independiente.

En el ámbito de los lenguajes de marcas las dos organizaciones más importantes son:

- ISO (*International Organization for Standardization*)
- W3C (*World Wide Web Consortium*)

20

Organizaciones y Estándares (II)

- **ISO** (1947): es el organismo encargado de promover el desarrollo de normas internacionales de fabricación comercio y comunicación para todas las ramas industriales excepto eléctrica y electrónica.

Esta formada por una red de institutos de 163 países y aunque sus normas no son de obligado cumplimiento si son generalmente aceptadas.

Tras el éxito de GML, en 1986 publicó SGML (ISO 8879) con rango de Estándar Internacional.

21

Organizaciones y Estándares (III)

- **W3C** (1994): comunidad internacional que se encarga de desarrollar estándares que garanticen el crecimiento de la web a largo plazo.

Uno de sus primeros trabajos fue normalizar HTML. Con crecimiento de la web surgieron presiones para ampliar HTML y el W3C decidió que la solución era crear unas reglas para que cualquiera pudiera crear un lenguaje de marcas adaptado a sus necesidades pero manteniendo una estructura común, así nace XML (1998).

22