



Lenguaje de Marcas y Sistemas de Gestión de Información

TEMA 1



¿Qué es un lenguaje de marcas

- Etiquetas
- Marcas
- Anotaciones

Permite:

- Hacer explícita la estructura de un documento
- Contenido semántico
- Cualquier otra información



¿Cómo se definen?

A partir de un documento DTD (Document Type Definition).

En él se establecen

- Marcas
- Elementos
- Etiquetas
- Atributos
- Sintaxis
- Normas de uso.

```
<carta>  
<fecha>22/11/2006</fecha>  
<presentacion>Estimado cliente:</presentacion>  
<contenido>bla bla bla bla ...</contenido>  
<firma>Don Jose Gutiérrez González</firma>  
</carta>
```



Clasificación de los LM

Presentación

Procedimientos

Descriptivo



Ejemplos

Documentación electrónica:

- RTF (Rich Text Format)
- Tex
- Wikitexto
- DocBook

Tecnologías de internet:

- HTML
- XHTML
- RSS

Lenguajes especializados:

- MathML
- VoiceXML
- MusicXML



Evolución

- Origen en los años 70 con los procesadores de texto para dirigir los procesos de impresión

<times 14><color verde><centrado> Este texto es un ejemplo para mostrar la utilización primitiva de las marcas**</centrado></color></times 14>**
<color granate><times 10><cursiva>Para realiza este ejemplo se utilizan etiquetas de nuestra invención. **</cursiva>** Las partes importantes del texto pueden resaltarse usando la **<negrita>negrita</negrita>**, o el **<subrayar>subrayado</subrayar></times 10></color>**

Este texto es un ejemplo para mostrar la utilización primitiva de las marcas
*Para realiza este ejemplo se utilizan etiquetas de nuestra invención. Las partes importantes del texto pueden resaltarse usando la **negrita**, o el subrayado*



Evolución

- Pasan a ser un medio de representación en pantalla no solo de impresión
- Se pasa de escribir la propia etiqueta a pulsar un botón o combinación de teclas
- Pero internamente se siguen utilizando las etiquetas



GML (Generalized Markup Language)

- Falta de estandarización
- IBM crea un formato estándar para todos los documentos que manejaba la empresa
- Debía ser flexible para cualquier documento de la empresa



SGML (Standard Generalized Markup Language)

- GML evoluciona en 1986 para crear el estándar ISO 8879
- Complejo y con herramientas de software caras
- Aplicaciones industriales

```
<email>
  <remitente>
    <persona>
      <nombre> Pepito </nombre>
      <apellido> Grillo </apellido>
    </persona>
  </remitente>
  <destinatario>
    <direccion> pinocho@hotmail.com </direccion>
  </destinatario>
  <asunto>¿quedamos?</asunto>
  <mensaje> Hola, he visto que ponen esta noche la película que querías ver. ¿Te apetece ir?</mensaje>
</email>
```



HTML (HyperText Markup Language)

- En 1989/90 Tim Berners-Lee creó el World Wide Web
- Organizar, enlazar y compatibilizar la información
- Se crea un lenguaje de descripción de documentos llamado HTML, que es una combinación de:
 - ASCII
 - SGML
- Versión simplificada de SGML con sólo instrucciones imprescindibles
- Sencillo y con gran aceptación



HTML (HyperText Markup Language)

- Los navegadores interpretan el lenguaje y es visualizado por la pantalla
- Desventajas:
 - No soporta tareas de impresión
 - No es flexible
 - No muestra contenido dinámico
 - Estructura y diseño mezclados

```
<html>
  <head>
    <title> Ejemplo</title>
  </head>
  <body b bgcolor="#ffffff">
    <p>
      <b>20 de octubre de 2010</b>
    </p>
    <p><b> Bienvenido al modulo de "Lenguajes de Marcas
y Sistemas de Gestión de Información" </b></p>
    <p> En este curso aprender&acute;s, entre otras
cosas:<br/>
    <ul>
      <li>Las ventajas que ofrece XML </li>
      <li>La creaci&acute;n de documentos bien
formados </li>
      <li>La creaci&acute;n de DTD</li>
    </ul>
    </p>
  </body>
</html>
```



XML (eXtensible Markup Language)

- Para resolver los problemas de HTML, W3C establece en 1998 un estándar puramente estructural sin información relativa al diseño
- Las etiquetas indican el significado en lugar del formato
- Es un metalenguaje que:
 - Permite definir etiquetas propias
 - Permite asignar atributos a las etiquetas
 - Utiliza un esquema para definir de forma exacta las etiquetas y atributos
 - La estructura y el diseño son independientes



XML (eXtensible Markup Language)

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE note SYSTEM "Note.dtd">
<note>
  <to>Tove</to>
  <from>Jani</from>
  <heading>Reminder</heading>
  <body>Don't forget me this weekend!</body>
</note>
```



Etiquetas

- Un lenguaje de marcas utiliza etiquetas intercaladas en un documento de texto
- Estas etiquetas son interpretadas posteriormente por intérpretes
 - Ej: en HTML el intérprete es un navegador
- Las etiquetas se escriben encerradas entre ángulos **<** **>** y se utilizan dos etiquetas, una de inicio y otra de fin, para saber dónde empieza y termina lo que queremos representar
- La diferencia entre la etiqueta de inicio y fin (apertura y cierre), la de cierre lleva la barra inclinada **/** antes del código

<etiqueta> texto que sufrirá las consecuencias de la etiqueta **</etiqueta>**

*Las etiquetas deben ir siempre en minúscula tal y como especifica W3C



Herramientas XML

Editores XML

Procesadores XML



XML: Estructura y Sintaxis

- Lenguaje de etiquetas que guardan información de manera **ordenada**
 - **No representa datos en sí, solo organiza la estructura**
- Forma potente de guardar la información
- Formato universal por SSOO y dispositivos móviles
- Propiedades:
 - Documento de texto
 - Extensión .xml
 - Con parejas de etiquetas
 - Estructuradas en forma de árbol
 - Organiza el documento
 - Editado por cualquier editor de texto
 - Interpretado por cualquier navegador

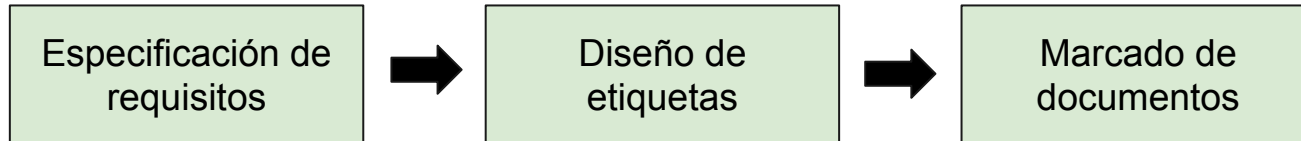


XML: Características

- Orientado a web
 - Compatible con los protocolos HTTP y URL
- Si verifica las reglas de XML es compatible con SGML
- No se requieren conocimientos de programación
- Sencillos de crear
- Difusión asegurada
 - Cualquier procesador de documentos XML puede leer un documento XML
- Legible para las personas
- Formal y conciso
- Extensible y adaptable
- Orientado a objetos



XML: Etapas de creación



*Los comentarios van entre etiquetas **<!-- -->**



XML: el prólogo

- Precede al documento:
 - Declaración de XML
 - Versión
 - Codificación de caracteres
 - Autonomía de documento

`<?xml version="1.0" encoding="UTF-8" standalone="yes"?>`

- Declaración del tipo de documento:

`<!DOCTYPE nombre_tipo>`



XML: Ejemplo

```
<?xml version="1.0" encoding="iso-8859-1"?>
```

```
<!DOCTYPE libro>
```

```
<libro>
```

```
  <titulo>XML practico </titulo>
```

```
  <autor>Sebastien Lecomte</autor>
```

```
  <autor>Thierry Boulanger</autor>
```

```
  <editorial>Ediciones Eni</editorial>
```

```
  <isbn>978-2-7460-4958-1</isbn>
```

```
  <edicion>1</edicion>
```

```
  <paginas>347</paginas>
```

```
</libro>
```



XML: Normas

Para que un documento XML este **bien formado**:

- Debe existir un solo elemento raíz
- Toda etiqueta debe tener su apertura y su cierre
 - Cuando no hay contenido se puede autocerrar
- No puede cerrarse un elemento que contenga otro elemento sin cerrar
- Caracteres reservados:
 - `< y >`
 - `&`
 - `"`
 - `'`
 - `]]>`



XML: Atributos

- No se organizan en jerarquía
- No contienen elementos
- No tiene estructura lógica

```
<?xml version="1.0" encoding="iso-8859-1"?>
<!DOCTYPE producto>
<producto>
  <codigo>G45</codigo>
  <nombre color="negro" precio="12.56">Gorro de lana</nombre>
</producto>
```



XML: Bien formados

- Debe tener un prólogo con la declaración
- Debe tener un elemento raíz del que todos los demás elementos estén anidados
- Cumplir las normas descritas previamente