


Tema 01 - Introducción a los lenguajes de marcas

Lenguaje de marcas y sistemas de gestión de la información

1. Introducción a los lenguajes de marcas

Un "lenguaje de marcas" es un modo de codificar un documento donde, junto con el texto, se incorporan etiquetas o marcas con información relativa a la estructura del texto o su formato de presentación.

Todo lenguaje de marcas está definido en un documento denominado DTD (Document Type Definition) donde se establecen:

- Las marcas.
 - Los elementos utilizados por dicho lenguaje.
 - Las etiquetas y atributos
 - Su sintaxis y normas de uso
- 

20/09/2021

Estimada presidenta:

Bla bla bla bla bla

Don Antonio

<carta>

<fecha>20/09/2021</fecha>

<presentación>Estimada
presidenta</presentación>

<contenido>Bla bla bla bla bla...</contenido>

<firma>Don Antonio</firma>

</carta>

1. Introducción a los lenguajes de marcas

En un mismo documento pueden combinarse varios lenguajes de marcas.

Los lenguajes de marcas se pueden clasificar en:


- **LM de presentación:** Define el formato del texto → **HTML**
- **LM de procedimientos:** Orientado también a la presentación pero, en este caso, el programa que representa el documento debe interpretar el código en el mismo orden en que aparece.
- **LM descriptivo o semántico:** Describen las diferentes partes en las que se estructura el documento pero sin especificar cómo deben representarse.

2. Evolución de los lenguajes de marcas

Los lenguajes de marcas surgieron, inicialmente, como lenguajes formados por un conjunto de etiquetas que los procesadores de texto introducían en los documentos para dirigir su proceso de presentación.

Como en el caso de los lenguajes de programación, inicialmente este conjunto de etiquetas estaban ligadas a las características de una máquina, programa o procesador de textos concreto.

Este marcado estaba exclusivamente orientado a la presentación de la información, aunque pronto se percataron de sus posibilidades y le dieron nuevos usos que resolvían una gran variedad de necesidades.





Radicado: IBA-02-2017-017877
Fecha: 09/03/2017 16:28
Sede: Ibagué

La EMPRESA DOCUMENTAL le hace la devolución de la Factura No. MD71562 por la siguiente causa:

- ☐ EXTEMPORÁNEA.
- ☐ ERROR EN LIQUIDACIÓN DE LA FACTURA - DOC. EQUIVALENTE
- ☐ NO CORRESPONDE A LA ORDEN DE COMPRA Y/O A LO RECIBIDO
- ☐ NO ES PARA LA UNIVERSIDAD COOPERATIVA DE COLOMBIA

- OBSERVACIONES: Estas facturas las enviaron a la Sede Ibagué junto con otras que no eran para la sede; las facturas que tenían las órdenes de compras correctas, se remitieron a las sedes que correspondían, sin embargo, las que no corresponden a la orden de compra se devuelven para que se adjunte la correcta y se remitan a las sedes que corresponde.

Respetuosamente le solicitamos hacer las correcciones del caso y la entrega a la Universidad dentro de los tiempos establecidos.

Para su información le recordamos que tiene plazo para entregar esta documentación hasta el día 25 de cada mes o antes si el día 25 no es día hábil, en los horarios de 8:00 am a 11 a.m. y 2:00 pm a 6:00 pm de lunes a jueves y 8:00 am a 11 a.m. y 2:00 pm a 5:00 pm los viernes.; el documento debe ser entregado en la Unidad de Correspondencia de la sede únicamente, adjuntando copia de la orden de pedido.

Cordialmente,

OSCAR AUGUSTO DIAZ TRIANA

Analista Centro de Administración Documental

2.1 GML (Generalized Markup Language)

Uno de los problemas que se conocen desde hace décadas en la informática es la falta de estandarización en los formatos de información usados por los distintos programas.

Para resolver este problema, en los años sesenta IBM se encargó de establecer un formato estándar para todos los documentos que se manejaban en la empresa. Con ello se lograba gestionar cualquier documento en cualquier departamento y con cualquier aplicación, sin tener en cuenta dónde ni con qué se generó el documento. Dicho formato tenía que ser válido para los distintos tipos de documentos legales que utilizaba la empresa, por tanto, debía ser flexible para que se pudiera ajustar a las distintas situaciones.

El formato de documentos que se creó como resultado de este trabajo fue GML, cuyo objetivo era describir los documentos de tal modo que el resultado fuese independiente de la plataforma y la aplicación utilizada.

2.1 GML (Generalized Markup Language)

```
<gml:Curve>
  <gml:segments>
    <gml:LineStringSegment>
      <gml:posList>187453.376 429000.792 187444.313 429004.534</gml:posList>
    </gml:LineStringSegment>
    <gml:LineStringSegment>
      <gml:posList>187444.313 429004.534 187442.564 428999.999 187441.651 428997.63</gml:posList>
    </gml:LineStringSegment>
    <gml:LineStringSegment>
      <gml:posList>187441.651 428997.63 187453.094 428994.092</gml:posList>
    </gml:LineStringSegment>
    <gml:LineStringSegment>
      <gml:posList>187453.094 428994.092 187456.661 428992.989 187458.431 428992.407</gml:posList>
    </gml:LineStringSegment>
    <gml:LineStringSegment>
      <gml:posList>187458.431 428992.407 187459.987 428998.063</gml:posList>
    </gml:LineStringSegment>
    <gml:LineStringSegment>
      <gml:posList>187459.987 428998.063 187453.376 429000.792</gml:posList>
    </gml:LineStringSegment>
  </gml:segments>
</gml:Curve>
```


2.2 SGML (Standard Generalized Markup Language)

El formato GML evolucionó hasta que dio lugar a SGML. Éste era un lenguaje muy complejo y requería de unas herramientas de software caras. Por ello su uso ha quedado relegado a grandes aplicaciones industriales.

```
<!DOCTYPE motd [ <!EL
<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

2.3 HTML

En 1989/90 Tim Berners-Lee creó el World Wide Web y se encontró con la necesidad de organizar, enlazar y compatibilizar gran cantidad de información procedente de diversos sistemas. Para resolverlo creó un lenguaje de descripción de documentos llamado HTML, que, en realidad, era una combinación de dos estándares ya existentes:

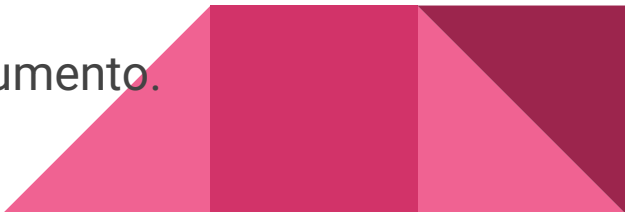
- **ASCII:** Es el formato que cualquier procesador de textos sencillo puede reconocer y almacenar. Por tanto es un formato que permite la transferencia de datos entre diferentes ordenadores.
- **SGML:** Lenguaje que permite dar estructura al texto, resaltando los títulos o aplicando diversos formatos al texto.



2.3 HTML

Ventajas: HTML es una versión simplificada de SGML, ya que sólo se utilizaban las instrucciones absolutamente imprescindibles. Era tan fácil de comprender que rápidamente tuvo gran aceptación logrando lo que no pudo SGML, HTML se convirtió en un estándar general para la creación de páginas web.

Desventajas:

- El lenguaje no es flexible, ya que las etiquetas son limitadas.
 - No permite mostrar contenido dinámico.
 - La estructura y el diseño están mezclados en el documento.
- 

2.4 XML

Para resolver estos problemas de HTML se estableció, en 1998, el estándar internacional XML, un lenguaje de marcas puramente estructural que no incluye ninguna información relativa al diseño. Está convirtiéndose con rapidez en estándar para el intercambio de datos en la Web. A diferencia de HTML las etiquetas indican el significado de los datos en lugar del formato con el que se van a visualizar los datos.

```
<?xml version="1.0"?>
<quiz>
  <qanda seq="1">
    <question>
      Who was the forty-second
      president of the U.S.A.?
    </question>
    <answer>
      William Jefferson Clinton
    </answer>
  </qanda>
  <!-- Note: We need to add
  more questions later.-->
</quiz>
```

XML

2.4 XML

XML es un metalenguaje caracterizado por:

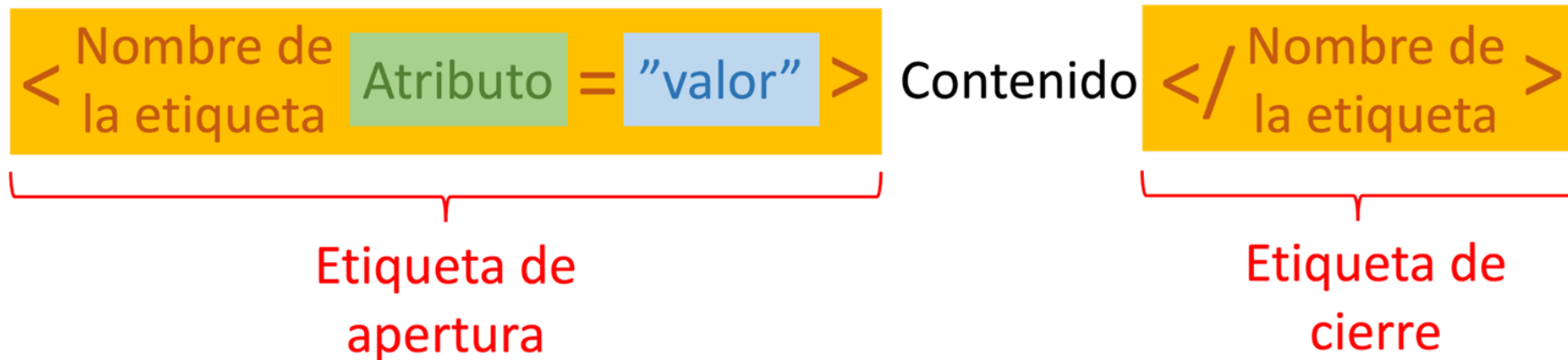
- Permitir definir etiquetas propias.
- Permitir asignar atributos a las etiquetas.
- Utilizar un esquema para definir de forma exacta las etiquetas y los atributos.
- La estructura y el diseño son independientes.



2.5 Comparación de HTML y XML

HTML	XML
<ul style="list-style-type: none">- Aplica un conjunto limitado de etiquetas sobre un único tipo de documento.- Modelo de hiperenlaces simple.- El navegador es un visor de páginas.- El problema de la "no compatibilidad" y las diferencias entre navegadores ha alcanzado un punto en el que la solución es difícil	<ul style="list-style-type: none">- Especifica cómo deben definirse conjuntos de etiquetas aplicables a un tipo de documento.- Modelo de hiperenlaces complejo.- El navegador es una plataforma para el desarrollo de aplicaciones.- Fin de la guerra de los navegadores y etiquetas propietarias.

3. Etiquetas



3. Etiquetas

Etiqueta

Se pueden distinguir dos tipos:

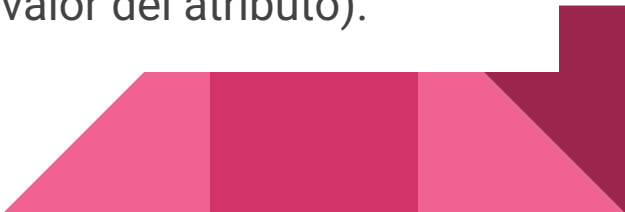
- Una **etiqueta de apertura**: es aquella que se escribe entre los caracteres "<" y ">". Dentro de la etiqueta siempre encontraremos de manera obligatoria el nombre de la etiqueta. No se puede colocar cualquier palabra como nombre de la etiqueta, sino que tienen unos nombres específicos ya que cada una de ellas se utiliza con un objetivo concreto.
- La mayor parte de las etiquetas de apertura cuentan con una **etiqueta de cierre**. Esta etiqueta de cierre indica dónde termina de actuar la etiqueta. Se caracteriza por escribirse entre los caracteres "</" y ">". En este caso solo se indica el nombre de la etiqueta que finaliza, sin los atributos.

3. Etiquetas

Atributos

Algunas etiquetas de HTML cuentan con la posibilidad de añadirles atributos (algunos opcionales y otros obligatorios). Con estos atributos se indica cierta información sobre la etiqueta y, normalmente, suelen acompañarse de un valor. Este conjunto Atributo = “valor” se coloca siempre detrás del nombre de la etiqueta y antes del caracter “>”.

El orden de los atributos no es importante en la etiqueta (siempre y cuando se coloquen detrás del nombre de la etiqueta) pero solo deben escribirse una vez (en caso contrario estaríamos reescribiendo constantemente el valor del atributo).



3. Etiquetas

Atributos

Tres tipos de atributos según los valores que acepta:

- **Valores finitos:** Son aquellos atributos que solo admiten un número limitado de valores (si introduces algún valor que no sea uno de estos entonces el valor no será válido).
- **Valores libres:** Son aquellos atributos en los que puedes introducir cualquier valor.
- **Valores booleanos:** Son aquellos atributos que solo admiten como valor "true" o "false".

3. Etiquetas

Contenido

El contenido es la información (el texto) a la que queremos aplicar la etiqueta. Este contenido tiene que colocarse entre las etiquetas de apertura y de cierre.

Comentarios

Cuando escribimos en HTML es importante ir añadiendo comentarios en los que se documenten aquellos detalles que son importantes para entender el código. Estos fragmentos de información son ignorados por el navegador, es decir, no los tiene en cuenta cuando va a visualizar la página web.

Siempre deben escribirse entre los caracteres “<!--” y “-->”.

Tu Nombre:

Tu Email:

Asunto:

Mensaje:

Envía Mensaje