Lenguajes de Marcas

UT1 – Reconocimiento de las características de los lenguajes de marcas

- Lenguajes de marcas.
 - 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.

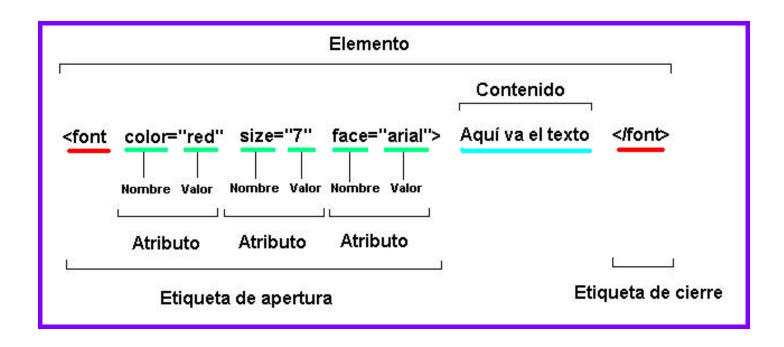
- Si en un lenguaje de marcas quisiésemos definir la frase:
 - "Comienza una nueva aventura con los lenguajes de marcas"
- usaríamos
- La etiqueta indica un nuevo párrafo y el final del mismo.
- Las etiquetas <i> y sirven para delimitar las partes del texto que van en cursiva y negrita respectivamente.

```
<!DOCTYPE html>
  <html>
     <head>
          <title>Título que aparecerá en la barra del navegador</title>
     </head>
     <body>
          <!- esto es un comentario -->
          <h1>Encabezado de nivel 1</h1>
          Esto es un parrafo con información
                    . las lineas salen pegadas aun que dejemos
                    espacios, saltos de linea, etc. <br >> Break, es un cambio
                    de línea
          <br><strong>texto resaltado </strong>
          Otro parrafo
          <h3>Encabezado de nivel 3</h3>
    </body>
</html>
```

```
<persona>
  <nombre>Ana</nombre>
  <apellidos>Ramos Ruiz</apellidos>
  <dni>123456</dni>
  <datos_familiares>
     <estado>casada</estado>
     <hijos>2</hijos>
  </datos familiares>
</persona>
```

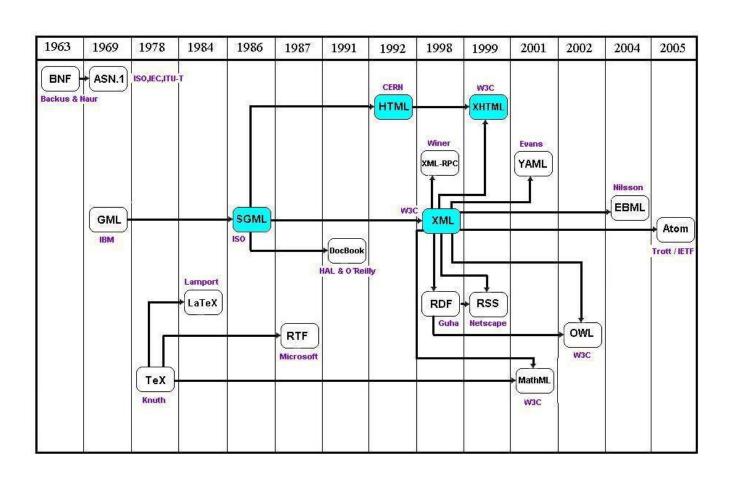
- Etiquetas.
 - Texto incluido entre los símbolos menor que < y mayor que >. El texto incluido dentro de los símbolos será explicativo de la utilidad de la etiqueta.
 - <h2> Encabezado de nivel 2 </h2>

- Atributos de las etiquetas.
 - Las etiquetas pueden presentar modificadores que llamaremos atributos, que permitirán definir diferentes posibilidades de la acción de la etiqueta.
 - Página principal del MIT



- Por ejemplo: Elementos HTML
 - Básicos. Documento, cabecera, título, cuerpo.
 - Enlaces.
 - Listas.
 - Marquesinas.
 - Imágenes.
 - Mapas.
 - Tablas.
 - Formularios.
 - •
 - Se tratarán con detalle en la siguiente unidad.

• • 3. Orígenes: SGML.



• • 3. Orígenes: SGML.

- Standard Generalized Markup Language. (Lenguaje de marcado de aplicaciones generales)
- Origen de HTML y XML.
- Lenguaje para definir lenguajes.
- SGML se basa en contenidos y no tiene en cuenta el formato.
- Permite escribir un documento origen y aplicar diferentes formatos.

• • 4. HTML

- Hypertext Markup Language (Lenguaje de Marcado de Hipertexto)
 - Describe tanto estructura como contenido.
 - Versión 5 de HTML
 - Se especifica:

<!DOCTYPE html>

• • • 4. HTML

- o Base para el crecimiento de la Web
- Ha incluido más de 100 etiquetas nuevas sin contar las específicas de los exploradores.
- Sobrecarga del servidor web.
 - Debe dar formato a la página dependiendo del navegador de destino.

• • • 4. HTML

o Diferentes definiciones de DTD.

- <!DOCTYPE html PUBLIC "-//IETF//DTD HTML 2.0//EN">
- <!DOCTYPE html PUBLIC "-//W3C//DTD HTML 3.2 Final//EN">
- <!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN" "http://www.w3.org/TR/html4/strict.dtd">
- <!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN" "http://www.w3.org/TR/html4/loose.dtd">

• • • 5. XML

o XML.

- Extensible Markup Language
- Conjunto de reglas para codificar documentos en formato legible para la máquina.
- Simplificación y adaptación de SGML para adaptarlo a las nuevas tecnologías como Internet.
 - Carece de funcionalidades avanzadas.
- Puede usarse en diferentes plataformas como bases de datos, hojas de cálculo, editores de texto, etc.

• • • 5. XML

- El servidor web envía el mismo documento XML independientemente del navegador.
- Si quiere dar formato al documento envía una hoja de estilo acompañante y el propio navegador del cliente se encarga de aplicar el estilo.
- Permite añadir nuevas etiquetas.
- o Independiente del sistema operativo.

4. Organizaciones desarrolladoras.

- ISO (International Standard Organization).
 - Nacida el 23 de Febrero de 1947
 - Promueve la creación de estándares internacionales.
 - Ejemplo: Pdf → Norma iso 32000

4. Organizaciones desarrolladoras.

- W3C (World Wide Web Consortium).
 - Creada el 1 de Octubre de 1994.
 - Organización internacional que produce recomendaciones para la World Wide Web.
 - Ejemplo: Especificación HTML5

- Tipos de lenguajes de marcas.
 - Marcas de presentación.
 - Marcas de procedimiento.
 - Marcas descriptivas.

- Marcas de presentación.
 - Las marcas solamente indican el formato del texto.
 - Se interpretan en el orden en el que aparecen.
 - Cada marca tiene su apertura y cierre.
 - Los documentos extensos son complejos de mantener.
 - Ejemplo: Lenguaje RTF

Ejemplo RTF.

{\rtf1\ansi{\fonttbl\f0\fswiss Helvetica;}\f0\pard Esto es un texto en {\b negrita}.\par }

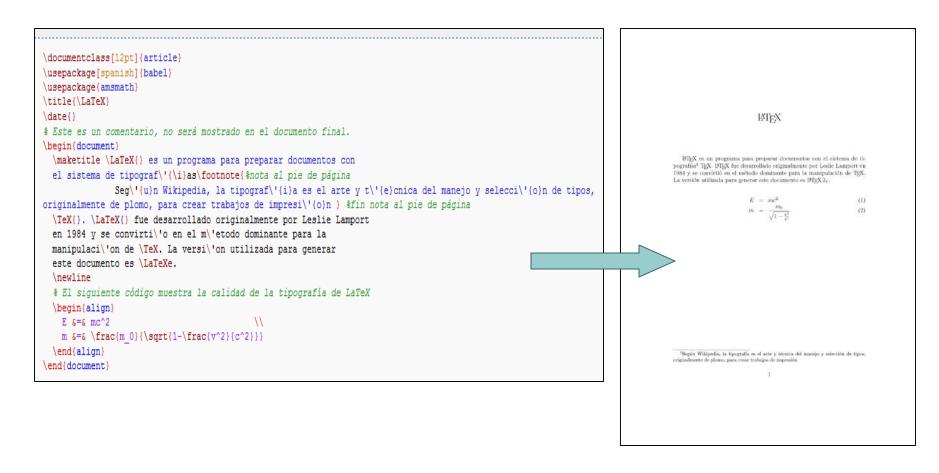
o Se representa como:

Esto es un texto en negrita

- Marcas de procedimientos.
 - Orientado a la presentación, es decir, también se indica el formato.
 - En este caso se indican los procedimientos que deberá seguir el SW encargado de la presentación.
 - Ejemplo: Tex, LaTex.

5. Clasificación

Ejemplo documento en laTeX.



- Marcas descriptivas.
 - Se utilizan las marcas para describir fragmentos de texto sin especificar cómo debe representarse.
 - El formato está separado del contenido permitiendo flexibilidad a la hora de dar un formato al texto.
 - Ejemplo: Lenguaje SGML, XML.

- Los lenguajes de marcas utilizados en la actualidad en entornos web son:
 - HTML
 - XHTML
- XHTML es una extensión de XML apropiada para web, mucho más estricta que HTML.

- Diferencias HTML vs XHTML.
 - En XHTML los elementos deben ser correctamente anidados
 - En XHTML los documentos debe ser <u>gramaticalmente correctos</u>
 - Los nombres de etiqueta deben estar en minúscula.
 - XML distingue entre minúsculas y mayúsculas.
 - Todos los elementos XHTML deben estar cerrados

- En HTML algunos elementos pueden ser incorrectamente anidados el uno dentro del otro de esta manera:
 - <i>Este texto esta en negrita y en itálica</i></i>
- En XHTML todos los elementos deben ser correctamente anidados el uno dentro del otro de esta manera:
 - <i>Este texto esta en negrita y en itálica</i>

- Todos los elementos XHTML deben estar anidados dentro de la etiqueta del elemento raíz. Todos los otros elementos pueden tener subelementos. Estos deben estar en pares y correctamente anidados dentro de su elemento paternal. La estructura de documento básica es:
 - <html>
 - <head> ... </head>
 - <body> ... </body>
 - </html>

- La gramática son el conjunto de reglas y principios que gobiernan el uso de un lenguaje.
- Los lenguajes de ordenador, como cualquier lenguaje, también poseen una gramática propia.
- A continuación veremos unos ejemplos de algunos lenguajes de marcado específicos para observar sus gramáticas.

MatML (Mathematical Markip Language)

- Lenguaje de marcas dirigido a la representación de fórmulas matemáticas.
- Codifica tanto notación como significado.
- No todos los navegadores son capaces de interpretarlo.
- Ejemplo:
- https://www.eyeasme.com/Joe/MathML/MathML_browser_test.html

MathML

```
FÓRMULA
                            (a + b)<sup>2</sup>
                                           → CONTENIDO
  PRESENTACIÓN
<math>
                                          <math>
<mrow>
                                           <apply>
 <msup>
                                            <power/>
  <mfenced>
                                             <apply>
                                              <plus/>
   <mrow>
    <mi>a</mi>
                                              <ci>a</ci>
                                              <ci>b</ci>
    <mo>+</mo>
    <mi>b</mi>
                                             </apply>
                                            <cn>2</cn>
   </mrow>
  </mfenced>
                                           </apply>
  <mn>2</mn>
                                          </msup>
</mrow>
```

SVG (Scalable Vector Graphics)

- Permite representar gráficos vectoriales en los navegadores.
- Los archivos deben llevar extensión svg.
- Veamos un ejemplo de código SVG.

```
<?xml version="1.0" standalone="no"?>
<!DOCTYPE svg PUBLIC "-//W3C//DTD SVG 20010904//EN"
"http://www.w3.org/TR/2001/REC-SVG-20010904/DTD/svg10.dtd">
<svg width="500" height="400" xmlns="http://www.w3.org/2000/svg">
<rect x="25" y="10" width="200" height="100"
     fill="orange" stroke="blue" stroke-width="3" />
<rect x="225" y="25" width="50" height="85"
    fill="blue" stroke="red" stroke-width="8" />
<circle cx="70" cy="110" r="25"</pre>
    fill="green" stroke="red"/>
<circle cx="220" cy="110" r="25"</pre>
    fill="green" stroke="red"/>
<path fill="pink" d="M 316 9"
     L 396 9 L 359 46 L 359 94 C 359 94 357 100 384 102
     L 384 103 L 331 103 L 331 102 C 356 100 354 94 354 94
     L 354 46 L 316 9 z "/>
stroke="green" stroke-width="8" x1="25" y1="150" x2="400" y2="150" />
<polyline fill="red" points="20, 250 85, 350 120, 350" />
<polygon fill="green" points="250,250 297, 284 279,340 220, 340 202, 284" />
<ellipse fill="blue" cx="400" cy="300" rx="72" ry="50"/>
</svq>
```

7. Gramáticas

- Si copiamos el código anterior en un documento de texto y lo guardamos con la extensión SVG, obtendremos al abrirlo con nuestro navegador la siguiente imagen.
- No todos los navegadores son compatibles.

