

CICLOS FORMATIVOS DE GRADO SUPERIOR:

- DESARROLLO DE APLICACIONES WEB
- DESARROLLO DE APLICACIONES MULTIPLATAFORMA
- DOBLE TITULACIÓN DAM-DAW

MODALIDAD DUAL | PRESENCIAL | PROYECTO PROPIO

MÓDULO: LENGUAJES DE MARCAS Y SISTEMAS DE GESTIÓN DE INFORMACIÓN (Código 0373)

Profesora: Elvira Girón

Curso 2019/2020

IES Clara del Rey

UNIDAD:

1. INTRODUCCIÓN A LOS LENGUAJES DE MARCAS

- 1.1. Reconocimiento de las características de lenguajes de marcas
- 1.2. Características y ventajas que proporcionan los lenguajes de marcas en el tratamiento de la información.
- 1.3. Distintas clasificaciones de los lenguajes de marcado.
- 1.4. Haciendo un poco de historia.
 - 1.4.1. SGML. Ventajas e inconvenientes. Ejemplo de SGML.
 - 1.4.2. Origen de HTML. Ventajas e inconvenientes. Ejemplo de HTML.
 - 1.4.3. XHTML.
 - 1.4.4. XML.

UNIDAD:

1. INTRODUCCIÓN A LOS LENGUAJES DE MARCAS

1.1. Reconocimiento de las características de lenguajes de marcas:

En el lenguaje escrito utilizamos tamaños de letra, negrita, numeración, viñetas, tablas, colores, etc. que acompañan a la información para que sea más fácil de entender. En definitiva estamos aplicando un determinado formato a la información que transmitimos.

En el caso de los documentos que intercambiamos a través de internet, como las páginas web, son los lenguajes de marcas quienes nos permiten aplicar dicho formato.

Dicho formato se le aplica a través de un Lenguaje específico, como es el HTML, que va etiquetando todo el texto que queremos visualizar en la página web de acuerdo al diseño elegido.

Al texto resultante de aplicarle las marcas de formato del HTML para visualizar la información, se le llama código fuente.

Dicho código, es interpretado por un programa al que se le llama agente de usuario y que son los populares navegadores web. Ejemplos:

Firefox, Chrome, Safari, Opera, Explorer

Una “marca” o etiqueta o tag, es una señal colocada dentro de un texto, con el fin de delimitar una parte del mismo y en muchos casos, aplicarle un determinado formato (aunque existen marcas con otros propósitos).

Las marcas más comunes están formadas por una palabra que describe su función encerrada entre los símbolos menor que (<) y mayor que (>). Ejemplos:

<html>, <body>, <table>, etc.

1.2. Características y ventajas que proporcionan los lenguajes de marcas en el tratamiento de la información.

TEXTO PLANO

- Los archivos de texto plano son aquellos que están compuestos únicamente por caracteres de texto a diferencia de los archivos binarios que pueden contener imágenes, sonido, archivos comprimidos, programas compilados, etc.
- Estos caracteres se pueden codificar con distintos códigos dependiendo del idioma o alfabeto que se necesite, por ejemplo: ASCII, ISO-8859-15, UTF-8
- El texto plano puede ser interpretado directamente por un simple editor de texto (ej., notepad, nano, vi, emacs, etc) a diferencia de los binarios. Lo que les permite ser independientes del SO.

COMPACIDAD

- Las etiquetas de marcado se mezclan con el propio contenido, por ejemplo: `<h3> Texto con formato </h3>`
- INDEPENDENCIA DEL DISPOSITIVO FINAL
- El mismo contenido puede ser interpretado de diferentes formas según el dispositivo final.

ESPECIALIZACIÓN

- Inicialmente los lenguajes de marcas se idearon para visualizar documentos de texto, pero progresivamente se han empezado a utilizar en otras áreas, como interfaces de usuario, síntesis de voz, sindicación de contenidos, etc.

FLEXIBILIDAD

- Los lenguajes de marcas se pueden combinar en el mismo archivo con otros lenguajes como JavaScript, PHP, con etiquetas especiales para ellos como `<script>`.
- XML ha permitido que se puedan combinar varios lenguajes de marcas diferentes en un mismo archivo, como en el caso de XHTML con MathML y SVG.

1.3. Distintas clasificaciones de los lenguajes de marcado.

Atendiendo al tipo de marcas que utilizan:

- De presentación:

Indican el formato del texto o tipografía, sin especificar su estructura, por ejemplo aumentar el tamaño de la fuente, centrar o cambiar a negrita.

El software que representa el documento debe interpretar el código en el mismo orden en que aparece.

Los procesadores de texto y en general las aplicaciones de edición profesional utilizan este tipo de marcado.

Ejemplos:

nroff, troff, RTF

TeX, Docbook (derivados de SGML)

- Descriptivo, estructural o semántico.

Indican las diferentes partes en las que se estructura el documento, pero sin especificar cómo deben representarse ni en qué orden.

XML es un metalenguaje expresamente diseñado para generar marcado descriptivo y los lenguajes derivados de XML con este propósito son: EBML, RDF, XFML, OWL y XTM.

Estos lenguajes crean documentos con estructura en árbol que almacenan información, por eso son bases de datos (BD), sin embargo no usan tablas ni respetan las reglas de integridad propias de las BD Relacionales, por ello se les llama bases de datos semiestructuradas.

Ejemplos:

ASN.1, YAML

EBML, RDF, XFML, OWL, STM (derivados de XML)

- Híbrido

Lenguajes que contienen marcas de los dos tipos anteriores indistintamente.

Ejemplos:

HTML (derivado de SGML)

XHTML, WML (derivados de XML)

Atendiendo a su funcionalidad

- Para crear documentación electrónica:
 - RTF, TeX, troff, nroff
 - ASN.1, EBML, YAML
 - Wikitexto, DocBook, LinuxDoc
- Tecnologías de Internet
 - HTML, XHTML, WML (páginas web)
 - GladeXML, XForms, XAML (formularios/interfaces de usuario)
- RSS , Atom (sindicación de contenidos)
- WSDL, SOAP, UDDI (servicios web)
- XMPP (mensajería instantánea)
- De propósito específico
 - MathML, CML (fórmulas matemáticas)
 - SSML, SRGS, VoiceXML (síntesis de voz)
 - MusicXML (partituras de música)
 - SVG, VML, X3D (gráficos vectoriales)
 - SMIL (archivos multimedia)
 - XLL (enlaces):

- XLINK (asociación de recursos)
- XML Base (URI básico)
- XPOINTER (localización de recursos)
- XSLT (transformación de documentos)
- XTM (mapas conceptuales)
- RDF, XFML, OWL, XMP (catalogación y clasificación de documentos, metadatos)
- GML (información geográfica)
- OFX (intercambio de información financiera)
- ebXML (comercio electrónico)
- XML Dsig, XML Enc, SAML, SACML, SKMS, XrML (seguridad)
 - XInclude (inclusión de archivos)

1.4. Haciendo un poco de historia.

1.4.1. SGML. Ventajas e inconvenientes. Ejemplo de SGML.

Las posibles anotaciones o indicaciones incluidos en los documentos de texto han dado lugar a lenguajes (entendiendo que en realidad son formatos de documento y no lenguajes en el sentido de los lenguajes de programación de aplicaciones) llamados lenguajes de marcas, lenguajes de marcado o lenguajes de etiquetas.

Charles Goldfarb era un investigador de IBM que propuso ideas para que los documentos de texto incluyeran la posibilidad de marcar el formato del mismo. Al final ayudó a realizar el lenguaje GML (Generalized Markup Language) de IBM el cual puso los cimientos del futuro SGML (padre de HTML y XML) ideado por el propio Goldfarb y padre de la mayoría de lenguajes de marcas actuales.

RTF es el acrónimo de Rich Text Format (Formato de Texto Enriquecido) un lenguaje creado para producir documentos de texto que incluyan anotaciones del formato. Es un formato propiedad de Microsoft, pero reconocido por la mayoría de aplicaciones de proceso de texto.

SGML es una mejora muy notable del lenguaje de GML que estandarizaba el lenguaje de marcado y que fue definida finalmente por ISO como estándar mundial en documentos de texto con etiquetas de marcado. Su responsable fue Charles Goldfarb.

Su importancia radica en que es el padre del lenguaje XML y la base sobre la que se sostiene el lenguaje HTML, dos de los lenguajes de marcas más populares de la historia.

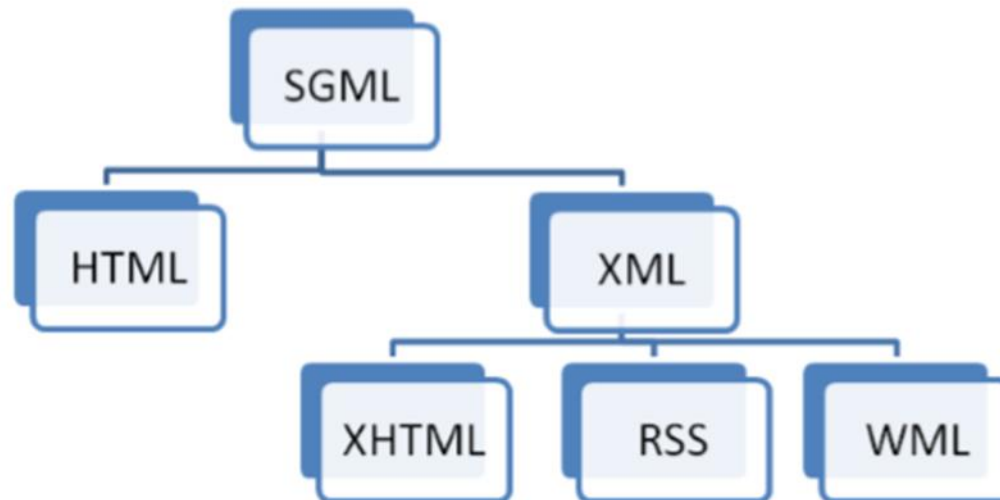
En realidad (como XML) no es un lenguaje con unas etiquetas concretas, sino que se trata de un lenguaje que sirve para definir lenguajes. Entre los lenguajes definidos mediante SGML, sin duda HTML es el más popular.

En SGML los elementos que contienen indicaciones para el texto se colocan entre símbolos < y >. Las etiquetas se cierran con el signo /. Es decir las reglas fundamentales de los lenguajes de etiquetas actuales ya las había definido SGML.

SGML aportó las etiquetas tal cual las conocemos actualmente gracias al éxito de HTML.

1.4.2. Origen de HTML. Ventajas e inconvenientes. Ejemplo de HTML

El origen de los lenguajes de marcas es SGML, del cual se derivan directamente algunos lenguajes como HTML. También se creó a partir de él por simplificación XML. Y a partir de XML se han creado muchos lenguajes como XHTML, RSS, WML, etc.



Inicialmente no podía dar soporte a todas las funciones que posteriormente tendría que realizar. Estas carencias se han ido resolviendo mediante la incorporación sucesiva de modificaciones y nuevos elementos, son las diferentes versiones de HTML.

Para superar los inconvenientes y deficiencias de la evolución poco planificada y a veces anárquica de HTML se han introducido otras tecnologías capaces de mejorar el funcionamiento y la organización de los sitios WEB; ejemplos CSS, JavaScript y los complementos del navegador.

Otro de los problemas que tiene HTML, especialmente cuando se combina con CSS, ha sido la diferente interpretación que hacen los navegadores de la misma página WEB debido al diferente ritmo en que los navegadores incorporan las nuevas propiedades CSS. Esto ha provocado la aparición de los “hacks”, pequeños trozos de código para conseguir un comportamiento homogéneo en los navegadores.

Ejemplo de HTML sencillo:

```
<!DOCTYPE html>
<html lang="es">
<head>
  <meta charset="UTF-8">
  <title>IDIOMAS</title>
</head>
<body>
  <h1>Listado de Cursos</h1>
  <h2>2019</h2>
  <p>INGLES</p>
  <p>FRANCES</p>
  <p>ALEMAN</p>
  <p>ITALIANO</p>
  <h2>Niveles</h2>
  <p>B1</p>
  <p>B2</p>
  <p>C1</p>
  <p>C2</p>
</body>
</html>
```

1.4.3. XHTML.

XHTML procede de la expresión Extensible Hypertext Markup Language.

El XHTML es un lenguaje que adapta el HTML al XML. Dicho de otro modo: se trata de una adaptación del lenguaje HTML para que sea compatible con el lenguaje XML.

Para saber qué es el XHTML, es necesario entender a qué se refieren los conceptos de HTML y XML.

El HTML es el lenguaje que, derivado del SGML, se emplea para la creación de páginas web. A través de este lenguaje de marcado, los diseñadores indican cómo se desarrollan las páginas. Lo que hacen los navegadores que utilizan los usuarios, como Mozilla Firefox, Google Chrome o Internet Explorer, es interpretar este contenido desarrollado en HTML y exhibirlo de manera tal que las personas puedan interpretarlo (como texto, imágenes, etc.).

El XML también deriva del SGML y se usa para el almacenamiento de datos. Este lenguaje (o metalenguaje) posibilita la definición de la gramática de diversos lenguajes, contribuyendo a la estructuración de la información y brindando soporte a las bases de datos.

Lo que hace el XHTML es volver compatibles ambos lenguajes (el HTML y el XML). De esta forma, mantiene la mayoría de las características del HTML, aunque con elementos del XML. El XHTML, por lo tanto, otorga mayor robustez y capacidad de adaptación, resultando importante para el desarrollo de la Web 3.0 gracias a su tipo de codificación

XHTML

- Tiene la particularidad de que puede incluir otros lenguajes también muy relevantes como pueden ser el SMIL, el MathML o incluso el que responde al nombre de SVG.

- Para poder funcionar de manera óptima, hay que destacar que se hace necesario que estén presentes y se utilicen del modo adecuado elementos tales como etiquetas en minúsculas, lo que son atributos de valores que aparezcan entrecomillados, elementos bien anidados...

- Las etiquetas que se emplean en este lenguaje que nos ocupa vienen a indicar la manera en la que tienen que ir apareciendo los elementos que dan forma a una página en concreto.

- Los atributos que también son básicos dentro del XHTML, por otro lado, se encargan de determinar lo que es el aspecto de la página en cuestión.

Numerosas son las ventajas que se considera que tiene usar el lenguaje que estamos abordando, en detrimento de otros. En concreto, entre las más significativas se hallan las siguientes:

- Los documentos creados con el mismo resulta que ofrecen un estupendo rendimiento.

- Se mantiene de manera muy sencilla.

- No menos notable es el hecho de que permite una gran facilidad a la hora de acometer lo que es la edición directa del código en cuestión.

- Es un lenguaje compatible con distintos estándares.

- Hacer uso del XHTML permite emplear herramientas actuales que ofrecen un mejor rendimiento que las de otros lenguajes.

1.4.4. XML.

Se trata de un subconjunto de SGML ideado para mejorar el propio SGML y con él definir lenguajes de marcado con sintaxis más estricta, pero más entendible.

Ha sido enormemente popular desde finales de los 90 y ha conseguido incorporar numerosos lenguajes a su alrededor para conseguir documentos muy dinámicos y con gran capacidad de formato. Es uno de los formatos de documentos más populares para exportación e importación de datos.

Actualmente está siendo sobrepasado en la mayoría de sus usos por JSON

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

```
<nombre>Jorge</nombre>
```

```
<apellido1>Sánchez</apellido1>
```

```
<dirección>
```

```
<calle>C/ Falsa nº 0</calle>
```

```
<localidad>Palencia</localidad>
```

```
<código_Postal>34001</código_Postal>
```

```
<pais>España</pais>
```

```
</direccion>
```

```
<teléfonos>
```

```
<teléfono tipo="fijo">999 999 999</teléfono>
```

```
<teléfono tipo="móvil">666 666 666</teléfono>
```

```
</telefono>
```

JSON

Abreviatura de JavaScript Object Notation, Se trata de una notación de datos procedente del lenguaje JavaScript estándar (concretamente en la versión ECMAScript de 1999). En el año 2002 se le daba soporte desde muchos de los navegadores y su fama ha sido tal que ahora se ha convertido en una notación independiente de JavaScript que compite claramente con XML en funcionalidad.

Las razones de su éxito se deben a su versatilidad, ya que permiten definir datos complejos, como arrays o código de funciones, elementos pertenecientes al mundo de la programación de aplicaciones. El éxito de JavaScript junto a la versatilidad comentada, le han convertido en el lenguaje de marcado más popular para almacenar datos.

En JSON, el texto se divide en datos y metadatos. De modo que el símbolo de los dos puntos separa el metadato del dato. Por otro lado, los símbolos de llave y corchete permiten agrupar de diversas formas los datos.

Ejemplo de código JSON:

```
{
  "nombre": "Laura",
  "apellido1": "Sánchez",
  "dirección": {
    "calle": "C/ Ebro nº 0",
    "localidad": "Valencia",
    "código Postal": 46006,
    "país": "España"
  },
  "teléfonos": [
    {
      "tipo": "fijo",
      "número": "999 999 999"
    },
    {
      "tipo": "móvil",
      "number": "666 666 666"
    }
  ]
}
```

...

ACTIVIDADES

A1.1_LM

Realiza un resumen – esquema del contenido de la unidad

A1.2_LM

Busca información sobre las distintas versiones de HTML y realiza un resumen

A1.3_LM

Busca información e indica las características más importantes de RSS y WML

BIBLIOGRAFÍA:

- Lenguajes de Marcas y Sistemas de Gestión de Información. IBERGARCETA PUBLICACIONES, S.L.
ISBN: 9788415452171

REFERENCIAS:

- <https://jorgesanchez.net/manuales/html/introduccion-html.html>