

# Origen de los Lenguajes de Marcas

SGML – HTML – DTD – XMLS –  
RDF - RDFS

# SGML

- En los años 60, IBM intentó resolver los problemas asociados al tratamiento de documentos en diferentes plataformas a través de GML (Generalized markup Language).
- En 1978 (ANSI) empezó a trabajar en las especificaciones para los procesadores de textos y el resultado fue el lenguaje SGML
- 1986, pasó a manos de la ISO y se convirtió en la norma 8879, SGML (Standart Generalized Markup Language).
- Los lenguajes de marcas no son equivalentes a los lenguajes de programación.

# LM vs. SGML

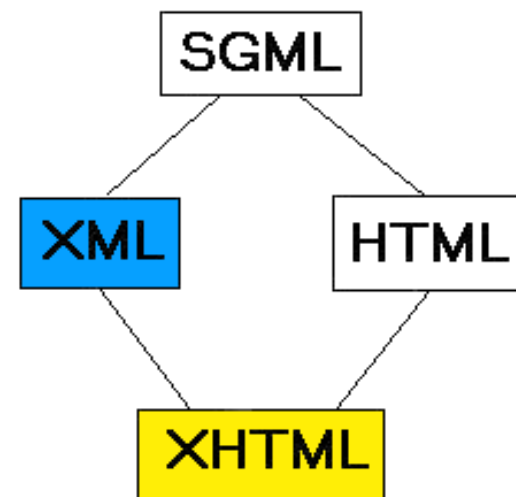
- Un lenguaje de marcas
  - se describen las reglas para el procesamiento de un texto, para describir los diferentes caracteres y sus características de impresión
- lenguaje de marcas generalizado
  - no especifica cómo deben verse las cosas o el proceso que se ha de realizar, sino que provee información sobre la estructura del documento e identifica las partes lógicas y el tipo de elementos que constituyen el documento

# Lenguajes Hipertextuales

- Los lenguajes de marcas, también denominados lenguajes de marcado o lenguajes de descripción de documentos, construyen un conjunto de reglas que definen todo aquello que es parte de un documento digital, pero que no pertenece al texto del mismo.
- Se trata de lenguajes orientados a definir la estructura y la semántica de un documento.
- En realidad, más que de lenguajes, podríamos hablar de metalenguajes o sistemas formales mediante los cuales se añade información o codificación a la forma digital de un documento bien para controlar su procesamiento, bien para representar su significado.

# Lenguajes Hipertextuales

- Niveles de información de un documento:
  - los datos que conforman el contenido (**caracteres de contenido**).
  - información superpuesta al contenido, que es lo que constituye el etiquetado (**caracteres de etiquetado**).
- Objetivos del Lenguaje de Marcado
  - Especifica las **operaciones tipográficas y las funciones** que debe ejecutar el programa navegador/visualizador sobre dichos elementos.
  - **Separa** un texto en los **elementos** de los que se **compone**, como por ejemplo un párrafo, un capítulo, etc.
- Tipos de Marcación
  - **específica**: describe cómo ha de formatearse el documento: fuente, tamaño, color, etc.
  - **estructural**: describe la estructura del documento: titular, párrafo, etc.



# Lenguajes de Marcado

- **Tipos**

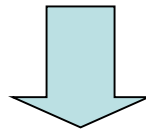
- **Procedimiento o procesado:** Describen la forma y el significado de las operaciones tipográficas que van a ser aplicadas a cada uno de los elementos del documento. Se refiere, pues, a la apariencia física o formato (fuente, estilo de letra, tamaño, etc.) tanto del documento en pantalla como del documento impreso.
- **Estructural o descriptivo:** En los lenguajes estructurales las marcas o anotaciones únicamente describen la estructura lógica del documento digital y/o la descripción del contenido, no su tipografía.

- **Usos de los lenguajes de marcas:**

- Describir contenidos (Bases de Datos).
- Definir el formato de los datos (Procesadores de texto).
- Realizar las dos funciones anteriores al tiempo (HTML).

# Ejemplo de Uso SGML

- <p>En este texto, algunas palabras aparecen en **<b>negrita</b>**, otras en **<i>cursiva</i>** y otras en **<b><i>negrita y cursiva</i><b>**



- En este texto, algunas palabras aparecen en **negrita**, otras en *cursiva* y otras en ***negrita y cursiva***

# ¿Cómo es un DTD?

Consta de una sucesión de comandos SGML enmarcados por los signos "<" y ">", estos comandos Son fundamentalmente:

- **ELEMENT**: sirve para definir una etiqueta.
- **ENTITY**: para indicar ciertos elementos del texto en forma de sucesiones de caracteres **ASCII** y para utilizar dentro del texto los caracteres reservados para la sintaxis de SGML, como los signos “menor que” o “mayor que”.
- **ATTLIST** permite establecer atributos complementarios para determinados elementos concretos. Todos los atributos pertenecientes a un elemento se resumen aquí en una sola lista en la que se establecen los nombres de los atributos, así como los valores autorizados de cada atributo .

Una sección marcada se inicia con la secuencia de caracteres "<![", seguida de una o más claves que informan sobre la manera de interpretar la información que se encuentra dentro de la sección. A continuación va el carácter "[" y detrás el contenido en sí de la sección marcada, que puede ser todo lo extenso que se desee.

Finalmente se cierra la sección mediante la secuencia de caracteres "]]". Las claves para la especificación de las secciones marcadas son **INCLUDE**, **IGNORE**, **CDATA**, **RCDATA** o **TEMP**.

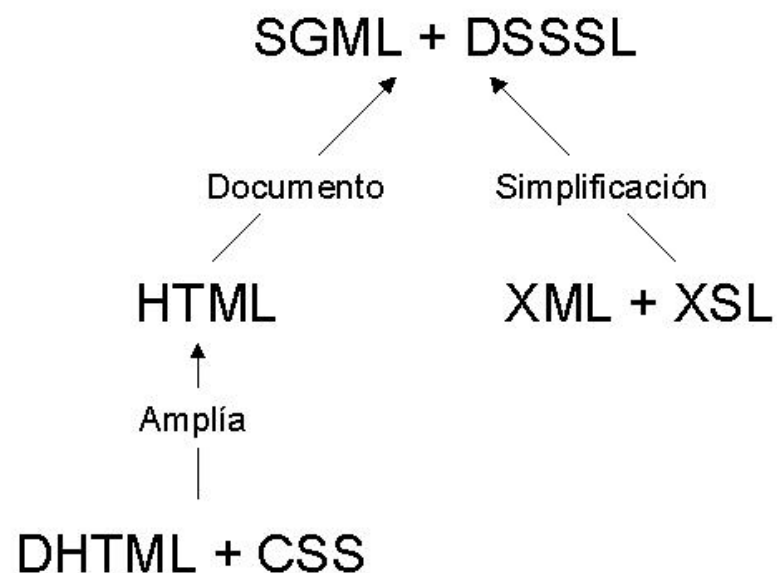


# Lenguajes SGML: HTML, XML

- **HTML: HTML (*Hipertext Markup Language*)** o lenguaje de marcas de hipertexto es el lenguaje que permite la generación de hipertextos en la [World Wide Web](#).
- **XML o *eXtensible Markup Language***, es un lenguaje de marcado derivado del lenguaje SGML que proporciona una sintaxis superficial para documentos estructurados, pero que **no impone restricciones semánticas** sobre el significado de los mismos. XML permite suministrar, recibir y procesar información en la [Web](#) y puede interoperar tanto con SGML como con HTML.
- XML no es una nueva versión de HTML y, aunque ambos proceden de un mismo metalenguaje, el SGML, el origen y enfoque que se ha seguido en ambos es muy distinto

# Relaciones Entre Lenguajes

	HTML/DHTML	XML	SGML
<b>Gramática</b>	Fija y no ampliable	Extensible	Extensible
<b>Estructura</b>	Monolítica	Jerárquica	Jerárquica
<b>Nº de marcas</b>	Fijas	Sin limite	Sin limite
<b>Complejidad</b>	Baja	Mediana	Alta
<b>Diseño de páginas</b>	Fijado por tags. Etiquetas con atributos CSS en DHTML	CSS o XSL	DSSSL
<b>Enlaces</b>	Simples enlaces	Poderosos enlaces (XLL)	HyTime
<b>Exportabilidad (formatos/aplicaciones)</b>	No	Sí	Sí
<b>Validación</b>	Sin validación	Pueden validarse	Obligatorio DTD
<b>Búsquedas</b>	Simple y a veces resuelta por <i>scripts</i> o CGI	Potente búsqueda. Con capacidad para personalizarla	Son posibles potentes búsquedas.
<b>Indización/Catalogación de páginas web</b>	Sólo lo permite los atributos de la etiqueta <META>, e implementaciones como DC.	Una descripción abierta y personalizable con el RDF.	Algún proyecto como TEI, DLI, etc.



Fuente: María Isabel García Arenas.

Curso XML 1ª Edición. Curso XML

<http://geneura.ugr.es/~maribel/xml/introduccion/index.shtml>

# Otros Lenguajes de Marcas

- Pero existen otros muchos lenguajes que abarcan diferentes aspectos:
  - Aspectos semánticos: (RDF, Topic Maps, DAML+OIL, OWL, XFML, etc. )
  - Vocabularios específicos para tratar diferentes aspectos dentro de XML (VoiceXML).
  - La posibilidad de añadir objetos no textuales, como gráficos u objetos en 3 dimensiones (VRML) o diagramas orientados a objetos (UML).

# ¿Cuándo Usar un SGML?

- Cuando exista la necesidad de intercambiar documentos entre diferentes sistemas de computación o de edición.
- Cuando los documentos tengan una larga vida de uso.
- Cuando sea fundamental la estructura de un documento.
- Cuando se utilice una base de datos para el almacenamiento y recuperación de los elementos del documento.

# Partes de un SGML

- **Declaración SGML:** Especifica el DTD, si se omite se usa una por defecto. La declaración SGML dice al usuario qué puede y qué no puede estar contenido en el documento SGML.
- **Declaración de tipo de documento (DTD):** En la DTD (*Document Type Definition* o Definición del Tipo de Documento) se identifica la estructura del documento, esto es, aquellos elementos que son necesarios en la elaboración de un documento o un grupo de documentos estructurados de manera similar.
- **Instancia de Documento:** contiene el documento en sí que incluye tanto el texto como el marcado. El contenido del documento tiene objetos SGML que siguen la estructura del árbol definido en la DTD.

# Bibliografía

- María Jesús Lamarca Lapuente.  
*Hipertexto: El nuevo concepto de documento en la cultura de la imagen:*
  - SGML.  
<http://www.hipertexto.info/documentos/sgml.htm>.
  - Lenguajes  
Hipertexto [http://www.hipertexto.info/documentos/lenguajes\\_h.htm](http://www.hipertexto.info/documentos/lenguajes_h.htm).