

Tema 1 Introduccion a los lenguajes de marcas

CODIFICACIÓN DE LA INFORMACIÓN

- Los pc usan binario
- Codificamos la informacion para q los pc la puedan procesar
- Todos los datos del PC son binarios pero a nivel formal se considera binario lo que no es texto
- Nombre.extension

Tipos de archivos:

Planos

Ficheros de texto
ASCII, UNICODE

Configuracion.ini .inf
Codigo fuente:.sql .c .java
Pagina web: .html .php .xml
Enriquecidos:.tex

Binarios.

- No son de texto.
- Necesitan un formato para interpretarse.
- Imagen: .jpg .gif .bmp
- Vídeo: .mpg .mov .avi
- Comprimidos: .zip .gz .rar .tar
- Ejectables: .exe .com .cgi
- Procesadores de texto: .doc .odt

¿Por qué los lenguajes de marcas?

Antiguamente los ficheros planos solo transmitían texto .txt

Los binarios, que sí permitían formatos, requerían de software específico .exe

La solución fue darle formato mediante etiquetas a los ficheros planos

Siguen siendo ficheros planos, pero poseen etiquetas que el ordenador puede entender

Historia

1º Procesadores de texto

2º Marking up. Anotaciones que se hacían a los documentos han dado lugar a los lenguajes de marcas

3º Charles Goldfarb. Padre de los lenguajes de marca

4º TeX y LaTeX. En los 70 Donald Knuth crea TeX

5º RTF acrónimo de Rich Text Format

6º SGML. Versión de GML Padre del XML y base del HTML

7º PostScript

8º HTML. Tim Bernes Lee

9º XML mas sencillo de sgml mas potente que html

Tipos de Lenguajes de marcas:

Orientados a la presentación. Al texto se añaden marcas que indican cómo presentarlo. Procesadores de Texto y HTML.

Orientados a la descripción. Las marcas dan significado al texto pero no indican como presentarlo. Ej.- Indican si el texto es un nombre , una matrícula de coche. Ideal para almacenar datos. SGML, XML

- Orientados a procedimientos. Las marcas se interpretan como órdenes. El archivo contiene instrucciones a realizar con el texto. LaTeX o PostScript

HTML JSON XML LaTeX PostScript

Organizaciones Desarrolladoras de Lenguajes de Marcas:

ISO 1947

WORLD WIDE WEB CONSORTIUM W3C 1994

XML

metalenguaje que permite definir lenguajes de marcas
se puede utilizar cualquier editor de texto

Características destacables:

Proviene de SGML y es compatible con él
Representa datos de forma estructurada
Gramática de OBLIGADO CUMPLIMIENTO.

Usos

Contenido web. Podría reemplazar a HTML (XML tiene una sintaxis rígida y HTML no).
Intercambio de información entre aplicaciones.
Computación distribuida. Intercambio de información a través de redes.

Lenguajes basados en XML:

XHTML. Adaptación de HTML a las normas de XML.

RDF Sirve para desarrollar documentos que describan recursos

RSS

SMIL

SVG

MATHML

VOICEXML

Otras tecnologías relacionadas que utilizan XML:

- DTD. Nos permite componer nuestras propias estructuras y nombres de elementos para nuestros documentos.
- XML Schema. Misma función que DTD pero es un lenguaje en XML.
- Namespacing, espacios de nombres. Nos permite crear documentos más consistentes mediante el uso de múltiples vocabularios en un mismo documento.
- XPath. Lenguaje de consulta que permite seleccionar o acceder a partes de un documento XML.
- CSS. (Cascading Style Sheets) Permiten dar formato al documento.
- XSL. (extensible Style Sheet Language) Como CSS con más posibilidades.
- XSLT. Para transformar nuestros documentos de un tipo a otro.
- XQuery. Permite consultar datos de los documentos XML.

- DOM. Document Object Model, permite acceder a la estructura jerárquica del documento normalmente para utilizarla dentro de un lenguaje de programación
- SAX. Simple API for XML, permite el uso de herramientas para acceder a la estructura jerárquica del documento XML, se usa mucho en Java