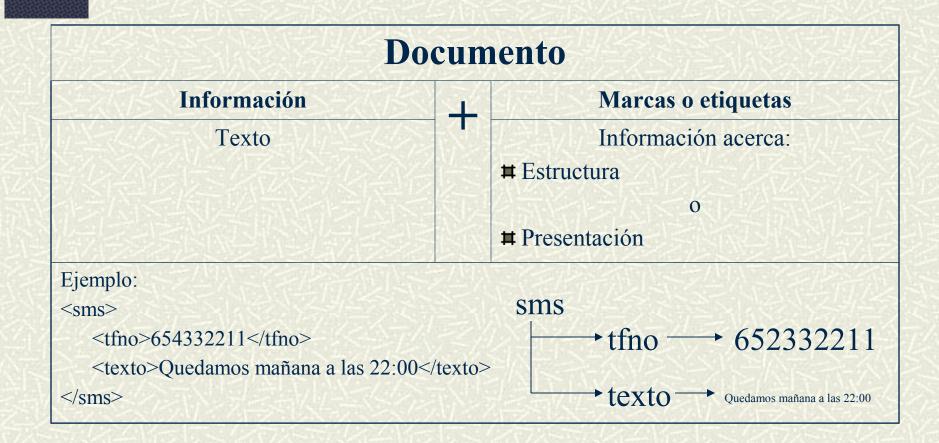
Tema 1: Lenguajes de marcas. Concepto y características

Curso 2012-2013 Sesión 1

Concepto de Lenguaje 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 de su presentación.



Tipos de lenguajes de marcas

- ★ Marcado de presentación es aquel que indica el formato del texto. Es útil para maquetar un documento para su lectura, pero insuficiente para el procesamiento automático de la información.
- ★ Marcado de procedimientos o procedural, enfocado a la presentación del texto, y visible para el usuario que edita el texto. El programa que lo edita interpreta el código en el mismo orden en que aparece. TeX, LaTeX, Postscript
- ★ Marcado descriptivo o semántico utiliza etiquetas para describir los trozos de texto, pero sin especificar cómo se representan, o en que orden. Los lenguajes expresamente diseñados para generar marcado descriptivo son el SGML y el XML.

Evolución de los lenguajes de marcas

- ➡ Surgen como solución a los formatos propios que utilizan las distintas aplicaciones.
- # El objetivo era hacer perdurar en el tiempo los documentos, principalmente documentos oficiales.

Historia

- Años 70: GML, Charles Goldfarb
- 1986: SGML, Standard Internacional ISO 8879, no almacena el diseño sino la estructura de los documentos. No fue muy difundido por su complejidad.
- 1989/90. Tim Berners-Lee creó el World Wide Web, que se basaba en el lenguaje de descripción de documentos HTML. Las páginas web son documentos con texto, vínculos y gráficos. Se basó en estándares ya existentes (ASCII + SGML).

GML

SGML

HTML

Historia

- Las desventajas de HTML, plantean la creación de otro lenguaje
 - Pensado para mostrarlo en la web pero no para la impresión o tareas de diseño.
 - Etiquetas limitadas, no es flexible.
 - No puede mostrar contenido dinámico.
 - La estructura y el diseño están mezclados.
- 1996: XML, como lenguaje estructural, para la descripción únicamente de la estructura del documento, no indicado para la representación.

■ 2000. XHTML/CSS. Es una reconversión de HTML según las reglas de XML. La apariencia del documento se describe mediante hojas de estilo

XML

XHTML

Etiquetas, elementos y atributos

- **♯** Existen 3 términos para describir las partes de un documento XML:
 - Etiqueta: texto enmarcado entre los símbolos "<" y ">". Existen etiquetas de inicio y de fin.
 - Elementos: estructuras que sirven para organizar el contenido: Constan de etiqueta de inicio, etiqueta de fin, y el contenido entre ellas. Pueden ser de contenido vacío, en cuyo caso tendrán una única etiqueta ciega: <nombre/>
 - Atributos: es una pareja etiqueta-valor que se encuentra dentro de la etiqueta de inicio e indican propiedades de los elementos.

¿Qué es XML?

... Un formato de texto?

¿Qué es XML?

```
<?xml version="1.0" encoding="IS0-8859-1"?>
<!-- catalogo de libros -->
-<catalogo>
    libro>
        <titulo>VisualBasic 2008</titulo>
        <autor>Edgar d'Andrea</autor>
        cio>35,94</precio>
        <editorial>Inforbooks</editorial>
    libro>
        <titulo>XML práctico</titulo>
        <autor>Sebastien Lecomte/Tieeri Boulanger</autor>
        cio>27</precio>
        <editorial>eni ediciones</editorial>
     </catalogo>
```

¿Qué es XML?

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!-- menu -->
<menu>
    <pla><plato>
        <nombre>ensalada mediterranea
        cio>6,5</precio>
        <descripcion>ensalalada de lechuga ,cebolla, tomate, atun y huevo duro</descripción>
    </plato>
    <pla><plato>
        <nombre>albondigas con tomate</nombre>
        <precio>5</precio>
        <descripcion>plato casero hecho con albondigas de carne</descripción>
    </plato>
    <pla><plato>
        <nombre>tiramisu</nombre>
        <precio>4.5</precio>
        <descripcion>sabroso postre elaborado con queso mascarpone</descripción>
    </plato>
</menu>
```

- **XML es...** un estándar que permite diseñar y desarrollar lenguajes de marcas.
- **XML es...** un subconjunto simplificado de SGML, que incluye las tres partes más importantes: extensibilidad, estructura y validación.
- **XML es...** un formato de texto estandarizado que sirve para representar y transportar información estructurada.

- ♯ XML es un metalenguaje.
- Los programadores web pueden ver XML como una barrera en el diseño gráfico de la web. Pero tienen herramientas para el diseño:
 - XSLT o XSL-FO.
 - CSS.
- ➡ Pero pese a que la presentación es importante, la mayoría de aplicaciones hoy en día están orientadas a datos.

Características XML

- ➡ La estructura y el diseño del documento están totalmente separadas
- ★ La estructura se puede describir mediante una DTD o un XML-esquema
- ➡ Posibilidad de definir etiquetas propias
- Asignación de atributos propios a las etiquetas

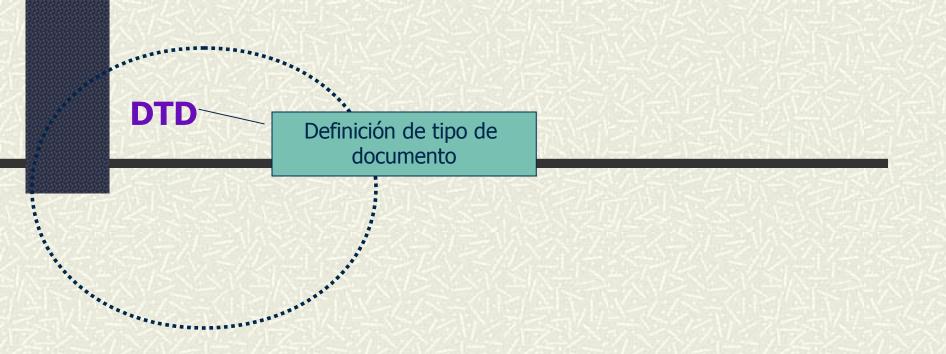
Otras características

- ♯ XML es un estándar abierto no sujeto a licencias
- **XML** es internacionalizable
- ♯ XML está basado en texto puro
- Ideal para conservar a largo plazo
- **■** XML es modular, ampliable

Inconvenientes

- **■** Requiere una transformación...
 - Una estructura, muchos diseños. El código XML es interpretado por las aplicaciones.
 - XML puro. La forma de representar los datos, dependerá del destino de los mismos.
 - Mediante hojas de estilo, CSS o XLS para mostrarlos en la web
 - Mediante un lenguaje de descripción de impresión como Postscript o PDF
 - No siempre está a la vista.

Lenguajes y tecnologías asociadas a XML



XML Schema

Describe la estructura y las restricciones de forma muy precisa

XML Schema

DOM

A través de él, los programas pueden acceder y modificar el contenido, estructura y estilo de los documentos HTML y XML

XSL/XSLT

XML Schema

DOM

"lenguaje extensible de hojas de estilo": que describen cómo se transforma o formatea la información para su presentación en un medio



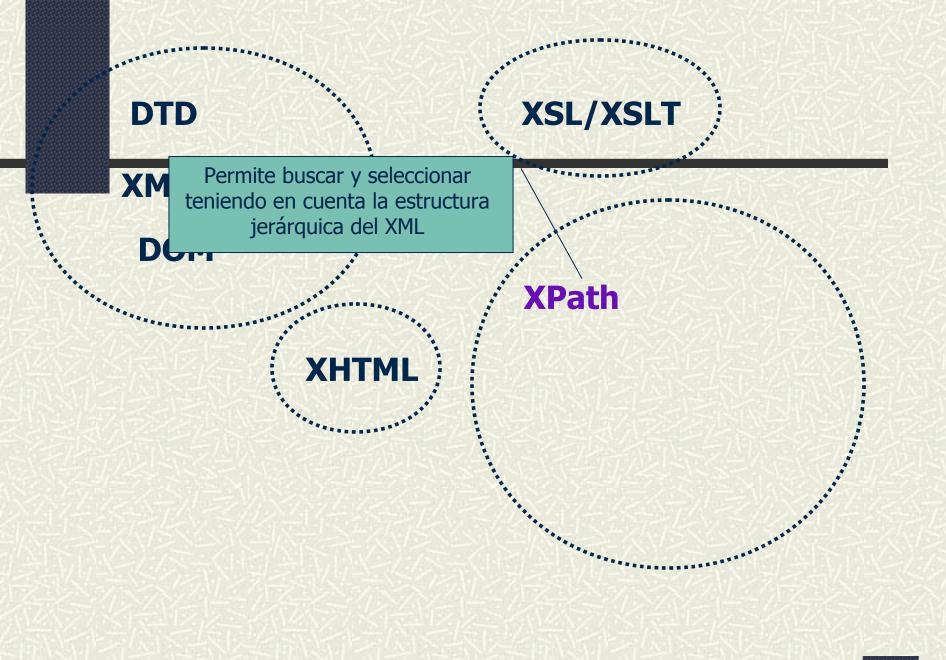
XSL/XSLT

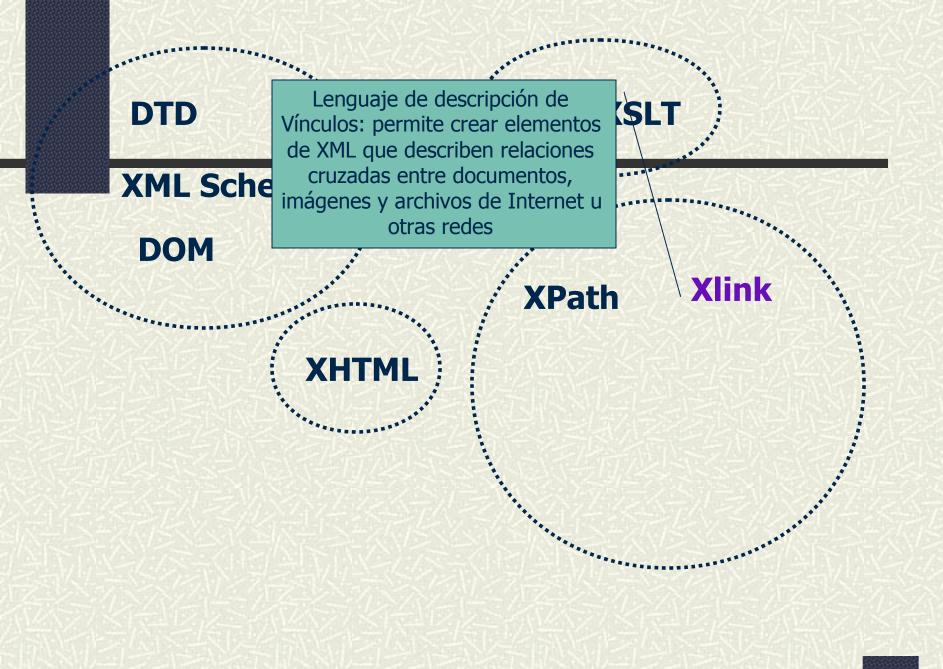
XML Schema

DOM

versión XML de HTML, con las mismas funcionalidades, pero cumple las especificaciones más estrictas de XML

XHTMI





XSL/XSLT

XML Schema

DOM

Lenguaje de punteros XML que proporciona una forma de identificar de forma única fragmentos de un documento XML con el objetivo de realizar vínculos

XPath Xlink

Xpointer



XSL/XSLT

XML Schema

DOM

Define el mecanismo básico para las referencias desde un documento XML

XPath Xlink

Xpointer

XBase

XSL/XSLT

XML Schema

DOM

XHTML

Define el método de inclusión para los documentos XML, funcionalidad básica para el ensamblaje de varios documentos. **XPath Xlink**

Xpointer

XBase

XInclude

XSL/XSLT

XML Schema

DOM

XHTML

es un lenguaje de consulta diseñado para consultar colecciones de datos XML. Similar a SQL, pero con algunas capacidades de programación **XPath** Xlink

Xpointer

XBase

XInclude

XQuery

