

Fundamentos de Aplicaciones y Servicios Telemáticos

2º Grado en Ingeniería de Tecnologías de Telecomunicación

Departamento de Ingeniería Telemática



Tema 01

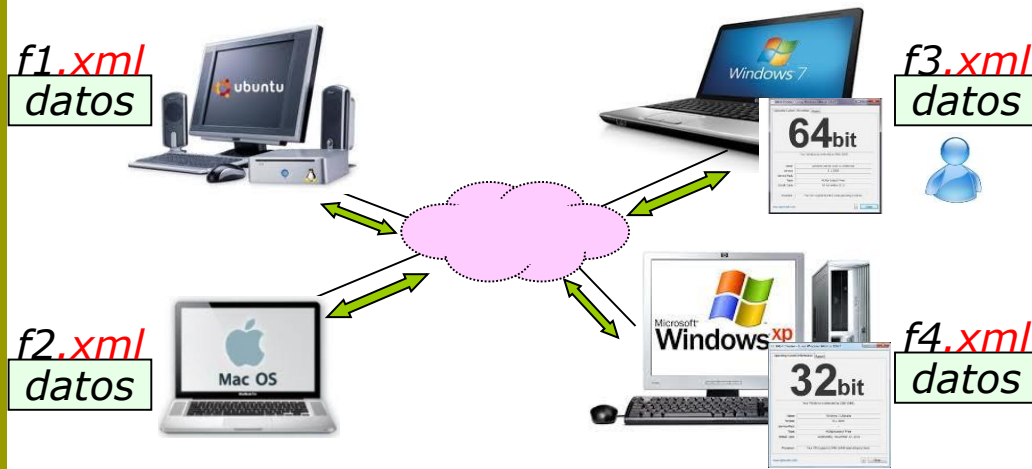
Programación Web Estática XML

XML: Contenido

- ❑ Introducción. Objetivo
- ❑ Etiquetas: Sintaxis
- ❑ Estructura de un fichero XML
- ❑ Espacios de Nombres
- ❑ Uso. Ejemplo: xHTML
- ❑ Comprobación del código

XML: Introducción

- ❑ Objetivo: Formato de datos (datos y formato separados):
 - “universal” (cualquier equipo/aplicación)
 - legible (jerárquico, etiquetas personalizadas).
- ❑ Multitud de usos:
 - Web estáticas (.xhtml) y dinámicas (.jspx), Javascript (leyendo XML del servidor), WAP/WML, SMIL (MM), BD XML, aplicaciones de usuario (MS Word .docx = .zip+.xml, ...)...
- ❑ XML es un lenguaje de etiquetas (o elementos).
- ❑ Versiones de XML normalizadas W3C: 1.0, 1.1 (amplía).



fx.xml

```
... <book category="A">
    <title>L1</title>
    <author>Nombre</author>
</book>
<book category="B">
    ...
</book>...
```

XML: Etiquetas, Sintaxis

□ Sintaxis de las etiquetas XML (case **Sensitive**):

Opcionales/Obligatorios según cada etiqueta concreta

`<etiqueta atributo1="valor" ...> Contenido </etiqueta>`

Comillas obligatorias

Obligatorios

□ Elementos SIN contenido también:

`<etiqueta atributo1="valor" ... />`

□ Además, XML soporta comentarios (se ignoran):

`<!-- Comentario -->`

□ Ejemplo: *fichero.xml*

```
... <libro category="A">
    <titulo>L1</titulo>
    <marcado />
    <!-- Texto -->
</libro>
```

□ Se exige anidación correcta

`<e1><e2>T</e1></e2>` — Mal

`<e1><e2>T</e2></e1>` — OK

XML: Etiquetas, Sintaxis (2)

- ❑ Contenido de etiquetas / Valor de propiedades:
 - Espaciados: NO se colapsan.
- ❑ Contenido de etiquetas: Caracteres prohibidos:
 - ❑ "<" (comienzo etiqueta)
 - ❑ "&" (carácter especial)
 - ❑ Soluciones:

XML admite caracteres especiales en los contenidos: > (>),...

| | |
|----------------------------------------------|---------------------------------------------|
| Caracter especial | < → &lt; & → &amp; |
| Sección CDATA (su contenido es texto) | <![CDATA[texto]]> |

```

...
<texto at="v" <& 1">Dos      palabras</texto>

<texto>&amp;puntero</texto>
<valor> <![CDATA[ &puntero5 ]]> </valor>

<valor>&lt;5</valor>
<valor> <![CDATA[ <5 ]]> </valor>...
  
```

fichero.xml

XML: Estructura de un fichero XML

□ Estructura:

fichero.xml ← Extensión usualmente reconocida

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

← [Recomendado] Versión de XML: 1.0 o 1.1 (opcionalmente codificación caracteres)

```
<!DOCTYPE raiz SYSTEM "def.dtd">
```

← [Opcional, para Validar] Definición de los Tipos de datos a permitir (nombres etiquetas/atributos, valores, anidaciones, ...): DTD o XSD (XML Schema)

```
<?xml-stylesheet type="text/css" href="estilo.css" ?>
```

← [Opcional] Estilos de presentación a aplicar (si no predeterminado, usualmente árbol): CSS o XSLT

```
<raiz>
```

...

```
</raiz>
```

← [Obligatorio] Elemento raíz (único)

} **Etiquetas XML** (las de arriba NO lo son)

XML: Múltiples DTD/XSD: Espacios de nombres

- Cada DTD/XSD un juego de etiquetas. Mismo XML puede varios DTD/XSD:
 - Pueden tener etiquetas con igual nombre.
 - Procesamiento deseado puede variar según sea de un DTD/XSD u otro
- Espacio de nombres: identificar el DTD/XSD de cada etiqueta

Identificador único (no necesita existir como URL)

| | |
|---------------------------------------|---------------------------------------|
| Definición (atributo de un elemento): | <code>xmlns:prefijo="URI"</code> |
| Uso (en elementos hijos): | <code><prefijo:etiqueta ...</code> |

- "1" juego de etiquetas puede usarse sin prefijo (sin ambigüedad):
 - Espacio de nombres por defecto: `xmlns="URI"`

f1.xml

```
<?xml version="1.0">
<root>
  <libro>
    <tit>L1</tit>
    <marcado />
  </libro>
</root>
```

f2.xml

```
<?xml version="1.0">
<raiz>
  <pele>
    <tit>L1</tit>
    <vista />
  </pele>
</raiz>
```



f.xml

Espacio por defecto

```
<?xml version="1.0">
<raiz xmlns="http://f1/x/"
      xmlns:f="http://f2/y">
  <libro>
    <tit>L1</tit>
  </libro>
  <f:pele>
    <f:tit>L1</tit>
  </f:pele>
</raiz>
```

XML: Uso. Ejemplo: XHTML

□ *Uso de ficheros XML:*

1º Fichero XML: con las etiquetas deseadas (metalenguaje).

- *Bajo la sintaxis de etiquetas XML*

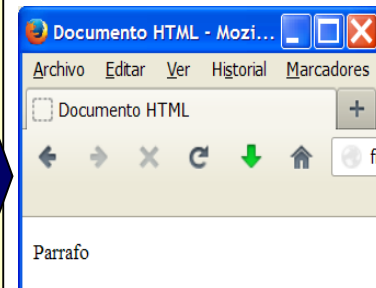
2º Ficheros externos (opcionales):

- *DTD o XSD (etiquetas/atributos, contenidos... permitidos): para poder comprobar si un XML está bien (validar)*
- *CSS o XSLT: Estilos de presentación personalizados.*
- *Interpretación de las etiquetas (qué hacer ante cada etiqueta): el que le quiera dar el programa que lea el XML.*
 - *Ej: Un navegador interpreta ficheros XML "XHTML".*

```
<?xml version="1.0">
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
    "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
  <head>
    <title>Documento HTML</title>
  </head>
  <body>
    <p>Parrafo</p>
  </body>
</html>
```



pagina.xhtml



XML: Uso. XHTML vs HTML

- Estándar XML (y así XHTML): exige que ante fallo se pare
 - Motivo: mejorar la velocidad de interpretación (garantizar fichero correcto y que la aplicación no tenga que solucionar los problemas)

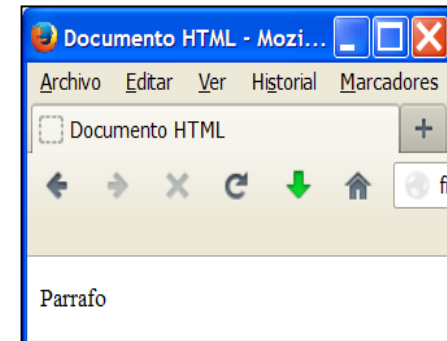
pagina.html (HTML)

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"
    "http://www.w3.org/TR/html4/strict.dtd">

<html>...
<body>
    <p>Parrafo
</body></HTML>
```

↓

El navegador "lo pasa"

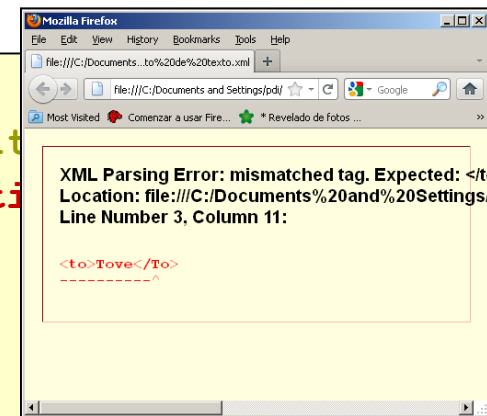


pagina.html (XHTML)

```
<?xml version="1.0">
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transiti
    "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transiti
<html xmlns="http://www.w3.org/1999/xhtml">...
<body>
    <p>Parrafo
</body></html>
```

↓

El navegador para



XHTML: diferencias con HTML

□ Estructura del documento. **Es obligatorio:**

■ **DOCTYPE** (ejemplos correctos):

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
```

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1 //EN"
"http://www.w3.org/TR/xhtml11/DTD/xhtml11.dtd">
```

■ Atributo **xmlns** en `<html>` (ejemplo correcto):

```
<html xmlns="http://www.w3.org/1999/xhtml">
```

■ `<html>` `<head>` `<title>` `<body>`

□ Elementos (ejemplos incorrectos):

■ Correctamente anidados:

```
<b><i>negrita e itálica</b></i>
```

■ Cerrados:

```
<p>párrafo y foto sin cerrar 
```

■ En minúsculas:

```
<BODY> <P>párrafo</P> </BODY>
```

□ Atributos (ejemplos incorrectos):

■ En minúsculas:

```
<img SRC="foto.gif" ALT="foto" />
```

■ Valor entrecomillado:

```

```

■ Minimización prohibida:

```
<input type="checkbox" name="vehículo" value="coche" checked />
```

XML: Comprobación del código

- ❑ Comprobaciones sobre un fichero XML: **dos tipos**
 - Bien formado ("well-formed"): sintaxis de etiquetas correcta.
 - ❑ Ante error la interpretación debe parar
 - Válido: conforme al DTD/XSD (etiquetas/atributos, contenidos, anidaciones, ... permitidos).
 - ❑ Válido implica bien formado, no al revés.
- ❑ Herramientas:
 - Entornos de desarrollo: Eclipse, Firefox plugin,...
 - Webs específicas: e.g. <http://validator.w3.org/>
 - ❑ Bien formado:
 - http://www.w3schools.com/xml/xml_validator.asp
 - ❑ Valido:
 - <http://www.validome.org/>

Referencias

□ XML:

- *Estándar XML 1.1:* <http://www.w3.org/TR/xml11/>
- *Uso práctico (1.0, 1.1):* <http://www.w3schools.com/xml/>