



XML

Héctor Baiges, Daniel Esquina, Jorge Matas, Alejandra Volkova

Contenidos

- Introducción
- Versiones
- Características
- DTD
- Ejemplo de código
- Extra: XHTML
- Puntos fuertes vs. puntos débiles
- Conclusiones



Introducción

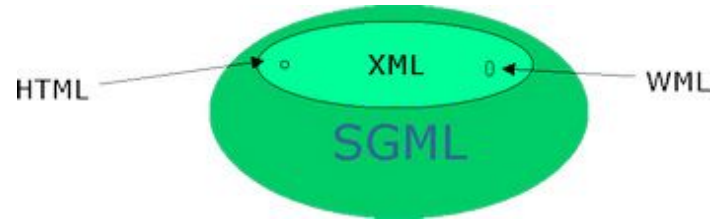
E^XTENSIBLE M^ARKUP L^ANGUAGE

Contenedor estructurado de información

Estándar, subconjunto de SGML (Standard Generalised Markup Language)

Fácilmente manipulable mediante un lenguaje de programación (JavaScript, PHP...)

Uso práctico: RSS (Really Simple Syndication)





Versiones

11 Abril 2013: Canonical XML Version 2.0

16 Agosto 2006: Extensible Markup Language (XML) 1.1 (Second Edition)

26 Noviembre 2008: Extensible Markup Language (XML) 1.0 (Fifth Edition)



Características: Inicio del documento

```
<?xml version="1.0" encoding="utf-8"? standalone="yes">
```

Siempre al principio del documento.

Única vez que se usa una etiqueta que contiene “?” en su nombre.



Características: Etiquetas y atributos

CaseSensitive

```
<nombre atributo="valor"> contenido </nombre>
```

```
<tag></tag>equivale a <tag/>
```

Caracteres reservados:

(> , < , & , ' , ")

Comentarios:

```
<!-- Esto es un comentario. -->
```



Características: Estructura de árbol

```
<raíz>
  <hijo>
    <nieto> ... </nieto>
    <nieto> ... </nieto>
  </hijo>
</raíz>
```

Solo una única raíz en el documento.

Igual que HTML, genera un árbol DOM (Document Object Model) accesible desde JavaScript.



Características: Entidades

<	<
>	>
&	&
'	'
"	"

Declaración:

```
<!ENTITY nombre "Cristina">
```

Referencia:

```
<texto>
```

```
    <autor>La escritora de la novela es &nombre;</autor>  
</texto>
```




Características: CDATA

Construcción xml que permite poner cualquier carácter y evita que sea interpretado por el parseador:

```
<![CDATA[  
<html>  
  <head>  
    <title> ... </title>  
  </head>  
</html>  
]]>
```



Document Type Definition

```
<?xml version="1.0" encoding="utf-8"? standalone="no">
```

Declarado dentro del propio documento XML:

```
<!DOCTYPE etiqueta[  
<!ELEMENT etiqueta (nombre)>  
<!ELEMENT nombre (#PCDATA)>  

```

Declarado en un fichero externo:

```
<!DOCTYPE coche SYSTEM "http://www.sitio.com/dtd/coche.dtd">
```



Document Type Definition

DTD:

```
<!ELEMENT mensaje (de, a, texto)>
<!ATTLIST mensaje prioridad (normal | urgente) normal>
<!ELEMENT texto(#PCDATA)>
<!ATTLIST texto idioma CDATA #REQUIRED>
<!ATTLIST texto traduccion CDATA #IMPLIED>
```

XML:

```
<mensaje prioridad="urgente">
  <de>Alfredo Reino</de>
  <a>Hans van Parijs</a>
  <texto idioma="holandés"> Hallo Hans, hoe gaat het?...</texto>
</mensaje>
```



Ejemplo

```
<?xml version="1.0"?>
<!DOCTYPE catalog SYSTEM "catalog.dtd">
<Catalog>
  <Book id="bk101">
    <Author>Garghentini, Davide</Author>
    <Title>XML Developer's Guide</Title>
    <Genre>Computer</Genre>
    <Genre>IT</Genre>
    <Price>44.95</Price>
    <PublishDate>2000-10-01</PublishDate>
    <Description>An in-depth look at creating applications
      with XML.</Description>
  </Book>
</Catalog>
```



Extra: XHTML

- Versión extendida de HTML con características de XML.
- 1998-2000. Entre HTML 4.01 y HTML 5. En desuso.
- Muy estricto: menos errores.
- Última versión: XHTML 1.1
- Inicio: `<?xml>` y `<!DOCTYPE html>`

En comparación con HTML:

- Los atributos y etiquetas van en minúsculas
- Elimina varias etiquetas y atributos marcados como “deprecated” en HTML4 (``, `<center>`). Tampoco se admiten `<frameset>`. Hay dos DTDs de retrocompatibilidad para esto.
- Algunos atributos son obligatorios (``, ``)
- No se admiten etiquetas sin cerrar o mal cerradas (`<i>txt</i>`)



Puntos fuertes

- Fácilmente procesable
- Separa radicalmente contenido y formato
- Diseñado para cualquier lenguaje
- Es extensible
- Relativamente fácil de leer para las personas

Puntos débiles

- Es muy estricto
- Tiene mucho “envoltorio” (*verbose*)
- Genera un árbol de etiquetas difícil de manejar
- Otros lenguajes de datos estructurados como JSON son más simples y flexibles



Referencias

<http://www.mundolinux.info/que-es-xml.htm>

<https://www.xml.com/axml/axml.html>

<https://support.office.com/es-es/article/xml-para-principiantes-a87d234d-4c2e-4409-9cbc-45e4eb857d44>

Albeza, B. (2006). XHTML + CSS ;de una maldita vez!

Bartolomé Sintés, M. (2015) XML: Lenguaje de marcas extensible. Accesible desde
<http://www.mclibre.org/consultar/xml/lecciones/xml-dtd.html>

Reino Romero, A.(2000) Introducción a XML en castellano. Accesible desde
<http://deim.urv.cat/~pedro.garcia/SOB/IntroXML.pdf>



Créditos

- **Héctor:** XHTML, puntos fuertes y débiles
- **Daniel:** características, DTD
- **Jorge:** características, XHTML, puntos fuertes y débiles
- **Alexandra:** introducción, versiones, DTD