2.2. XML

Los primeros años del web demostraron que **el marcado con HTML no era adecuado para estructurar datos**, ni facilitaba el intercambio de información entre máquinas y aplicaciones, un intercambio de datos que dada vez era más necesario. Para dar una solución a estos problemas, el W3C (*World Wide Web Consortium*) desarrolla y publica en 1998 el estándar internacional **XML (*Extensible Markup Language)***, que es independiente de cualquier empresa o administración. Se trataba de un lenguaje de marcado puramente estructural, que no incluía ninguna información relativa a la capa de presentación de la información. S**u objetivo principal es facilitar el intercambio de información y datos estructurados entre aplicaciones**.

Esto quiere decir que se trata de un **metalenguaje**, es decir, un lenguaje de marcado pensado y diseñado para poder ser utilizado en cualquier contexto, siempre y cuando se respetan las lógica y reglas del mismo. Está basado en SGML, del que se ha dicho que es un subconjunto, pero con una reglas menos complicadas y rígidas, buscando favorecer su uso. Al ser un metalenguaje, **permite crear nuevos lenguajes de marcado**, los cuales pueden **definir nuevos tipos de documentos**, así como las **etiquetas o marcas que se usan para identificar y describir los elementos de esos tipos de documentos**. Al definir las etiquetas o marcas, también se puede establece la **sintaxis** de los elementos, en qué orden y de que manera se puede utilizar y combinar, o las dependencias o jerarquías entre ellos, si es necesario.

Las características fundamentales de XML son:

- La **estructura y el diseño son independientes**.
- Utiliza un **esquema** para definir de forma exacta las etiquetas o marcas y los atributos.
- Permite **definir** y nombrar **etiquetas** propias.
- Permite **asignar atributos** a las etiquetas.

https://www.youtube.com/embed/AZihBEg8VBk

Videotutorial: OpenWebinars, *¿Qué es XML y para qué se usa?* (2014).

Un documento XML es un fichero de texto plano, que está formado por el conjunto de información y el marcado correspondiente de sus elementos. Este fichero incluirá en su **cabecera** la **declaración** de que se trata de un **documento XML**, y las indicaciones de los url necesarios para que las aplicaciones que los usen sepan dónde encontrar l**os DTD o XML Schema**

Copyright 2026

pertinentes para su procesamiento (véase apartado 1.2). Después de esto, el documento mostrará los **datos estructurados y marcados con las etiquetas** correspondientes, de acuerdo a cómo hayan sido definidas en el DTD o XML Sxchema de referencia. **Si todo está formulado y marcado correctamente, se dice que el documento XML está bien formado**, y entonces es cuando las aplicaciones proceden a procesar su contenido.

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<!DOCTYPE recipe PUBLIC "-//Happy-Monkey//DTD RecipeBook//EN"
"http://www.happy-monkey.net/recipebook/recipebook.dtd">

<recipe>

<title>Peanutbutter On A Spoon</title>

<ingredientlist>
  <ingredient>Peanutbutter</ingredient>
</ingredientlist>

<preparation>Stick a spoon in a jar of peanutbutter, scoop
and pull out a big glob of peanutbutter.</preparation>

</recipe>
```

fig. 3. Una receta estructurada y marcada en XML, con su declaración, DTD y elementos (fuente original).

En XML las etiquetas o marcas indican el significado de los datos; no hacen referencia a su presentación o formato final. Esto es así porque **en el diseño del XML se ha separado el contenido de la presentación**. Si se quiere presentar información tomada de un documento marcado en XML, es necesario **procesar el contenido del documento con hojas de estilo externas**. Esto permite usar los mismos datos y prepararlos para diferentes medios y/o presentaciones, así como usarlos en múltiples aplicaciones. Los datos se mantienen y fluyen marcados en XML, lo que hace posible que este lenguaje de marcado se convierta en un **estándar para el intercambio de datos en la Web** (aunque no el único, véase JSON en el apartado 2.4).

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:dc="http://purl.org/dc/elements/1.1/">
  <rdf:Description>
    <dc:title>Homage to Catalonia,</dc:title>
    <dc:creator>Orwell, George, 1903-1950.</dc:creator>
    <dc:type>text</dc:type>
    <dc:publisher>London, Secker and Warburg</dc:publisher>
    <dc:date>[1938]</dc:date>
    <dc:language>eng</dc:language>
  </rdf:Description>
</rdf:RDF>
```

Fig. 4. Combinación de marcados: documento en XML con su cabecera de declaraciones, que contiene datos en RDF del estándar de metadatos Dublin Core (fuente original).

En el uso diario, en muchas páginas web pueden encontrarse combinaciones de HTML con XML y lenguajes derivados del mismo. Como estas combinaciones vienen expresadas en el marcado de

las páginas web, no se aprecian en la capa de presentación, por lo que pasan desapercibidas para el usuario final. Si se piensa sólo en XML, son numerosos los gestores de bases de datos que ya son capaces de importar o exportar datos desde ficheros en XML, y, por ejemplo, los ficheros de procesadores de textos almacenan la información en documentos XML dentro de ficheros compactados. El formato EPUB para libros electrónicos está formado por conjuntos de documentos en XML y derivados como XHTML y SVG.

Material complementario

- Morales, R. Qué es XML. *ticARTE*, 2014.
- Sánchez. J. *Fundamentos de XML*. 2016.
- desarrolloweb.com *Introducción a XML*. sin fecha.

Financiado por el Ministerio de Educación y Formación Profesional y por la Unión Europea - NextGenerationEU



Revision #21
Created 2022-11-17 23:17:03 CET by Jesús Tramullas
Updated 2023-02-01 21:03:52 CET by Jesús Tramullas

Copyright 2026