



# UD1

## Introducción a lenguajes de marcas

CFGs DESARROLLO DE  
APLICACIONES WEB

# UD1

## Introducción a lenguajes de marcas



**Criterios de evaluación**



**Contenidos**

# CRITERIOS DE EVALUACIÓN

CE1.a Se han identificado las características generales de los lenguajes de marcas.

CE1.b Se han reconocido las ventajas que proporcionan en el tratamiento de la información.

CE1.c Se han clasificado los lenguajes de marcas e identificado los más relevantes.

CE1.d Se han diferenciado sus ámbitos de aplicación.

CE1.e Se han reconocido la necesidad y los ámbitos específicos de aplicación de un lenguaje de marcas de propósito general.

CE1.f Se han analizado las características propias de diferentes lenguajes de marcas.

CE1.g Se ha identificado la estructura de un documento y sus reglas sintácticas.

CE1.h Se ha contrastado la necesidad de crear documentos bien formados y la influencia en su procesamiento.

CE1.i Se han identificado las ventajas que aportan los espacios de nombres.

## CONTENIDOS

---

Clasificación

---

Características y ámbitos de aplicación

---

Estructura y sintaxis.

---

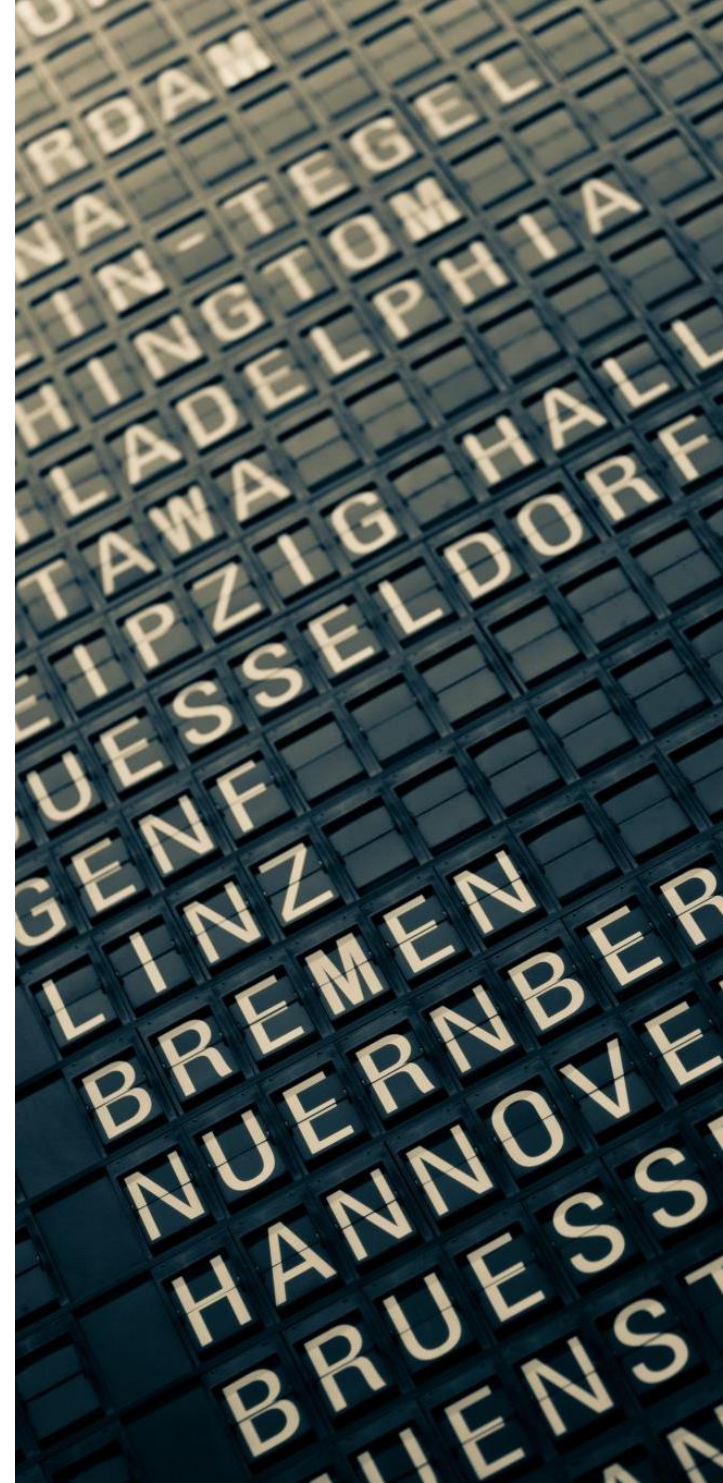
Elaboración de documentos bien formados

---

Utilización de espacios de nombres.

# Concepto de lenguajes 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 su presentación.
- El lenguaje de marcas más extendido es HTML (HyperText Markup Language, Language de marcado de hipertexto).
- Una **marca** es una señal colocada dentro de un texto, con el fin de delimitar una parte de este.
- Las marcas más comunes están formadas por una palabra que describe su función encerrada entre los símbolos menor que (<) y mayor que (>).
- Es muy habitual que aparezcan por parejas, una de comienzo y otra de fin.





# Concepto de lenguajes de marcas

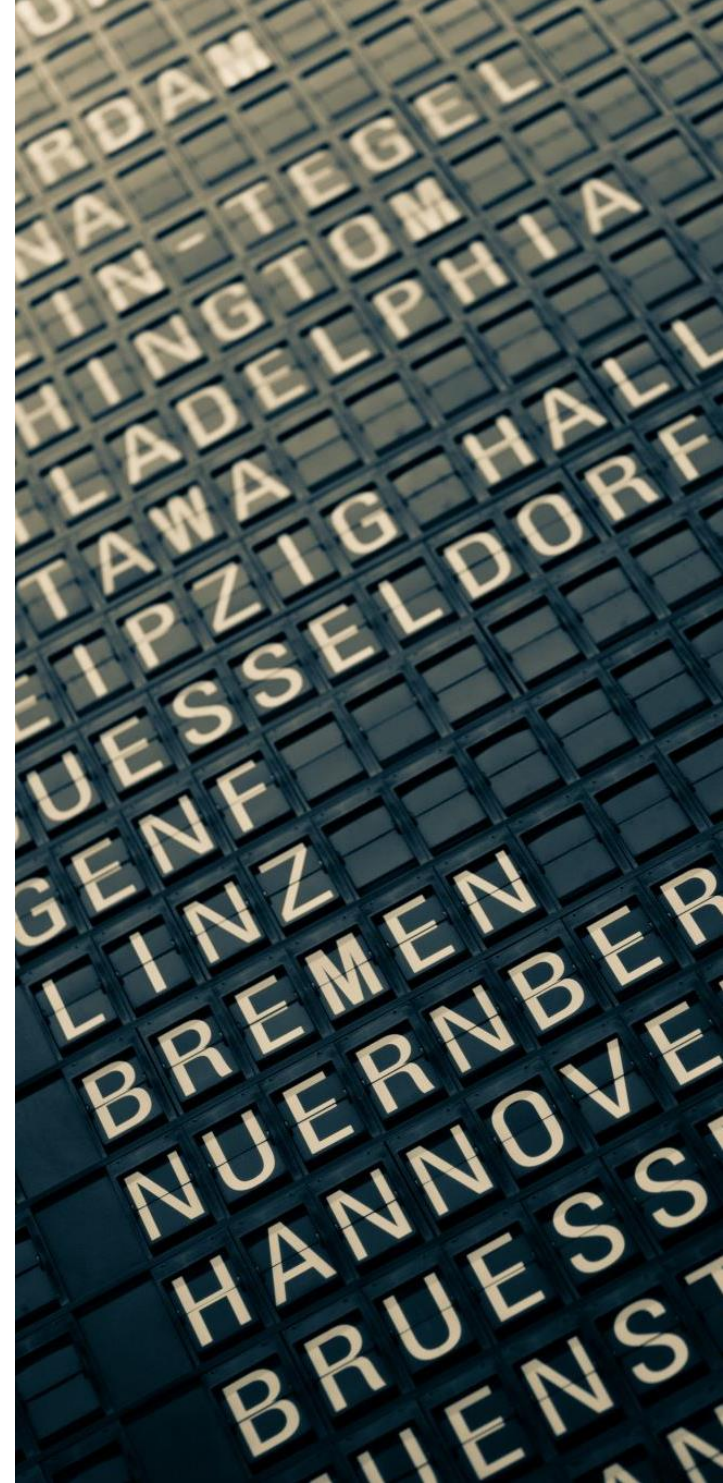
- Ejemplos:

## HTML:

```
<h1>Lenguaje de marcas</h1>
```

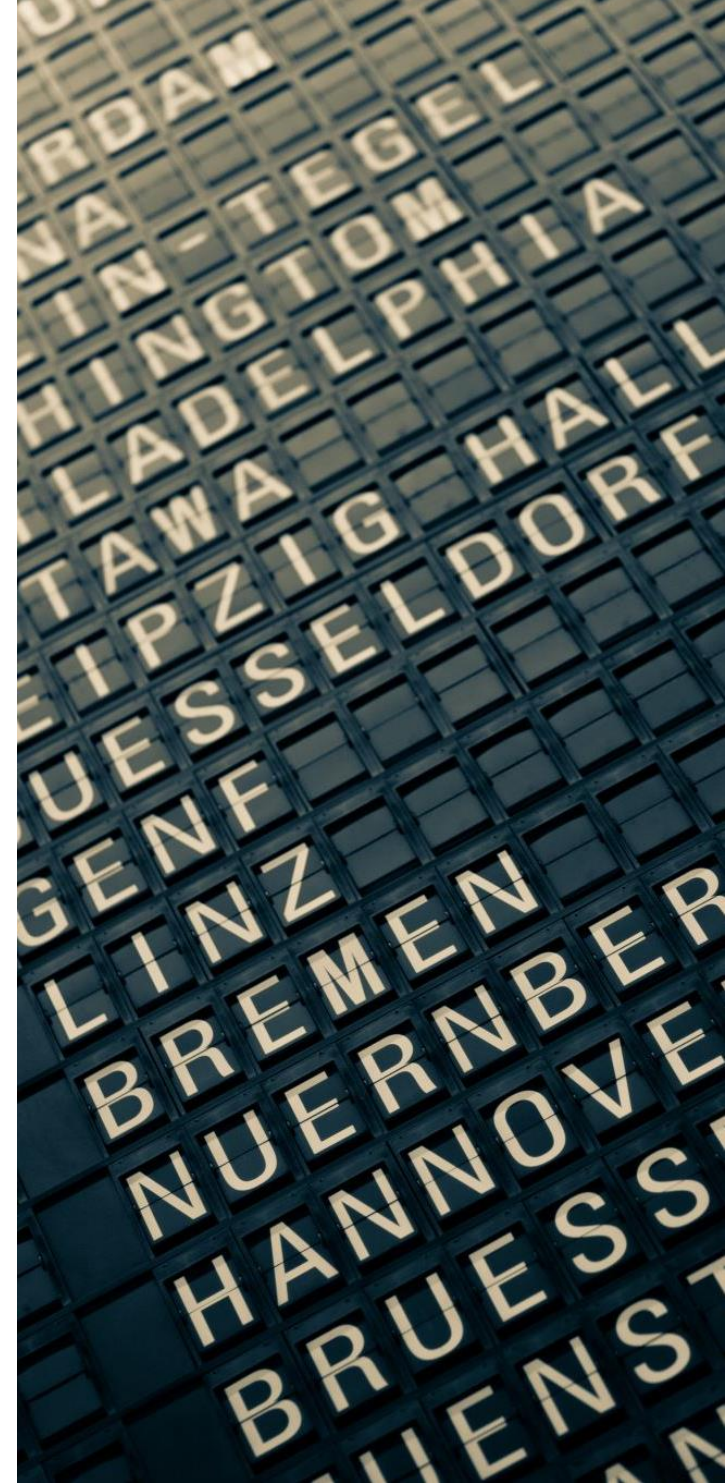
## XML:

```
<noticia>  
    <fecha>25/09/2015</fecha>  
    <lugar>Almería</lugar>  
    <titular>Inicio del curso</titular>  
    <desarrollo>IES los  
Sauces</desarrollo>  
</noticia>
```



# Características comunes de los lenguajes de marcas

- **Texto plano.** Compuestos únicamente por caracteres de texto.
- **Compacidad.** Las instrucciones de marcado se mezclan con el propio contenido.
- **Independencia del dispositivo final.** El mismo documento puede ser interpretado de diferentes formas.
- **Especialización.** Se idearon para visualizar documentos de texto, pero se han empezado a utilizar en muchas otras áreas: gráficos vectoriales, sindicación de contenidos, interfaces de usuario, etc...
- **Flexibilidad.** Se pueden combinar en el mismo archivo con otros lenguajes.



# Clasificación de los lenguajes de marcas

Los lenguajes de marcas de pueden clasificar, según su uso, finalidad y características.



# Clasificación de los lenguajes de marcas

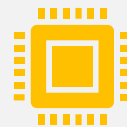


**Atendiendo a su funcionalidad:**



**Para crear documentación electrónica.**

RTF, TeX, troff, Wikitexto, DocBook, ...



**Tecnologías de Internet.**

HTML, XHTML, WML, RSS, Atom, ...



**De propósito específico.**

MathML, VoiceXML, MusicXML, SVG, XLINK, XSLT, ...

# Clasificación según su propósito: Lenguajes de marcas de presentación

**Lenguajes de marcas de presentación.** Su objetivo principal es como estructurar documentos y facilitar la presentación visual.

- **HTML:** Usado para contenido web, permitiendo a los navegadores renderizar texto, imágenes y multimedia.
- **XHTML:** Variante de HTML que sigue una sintaxis más estricta. Basada en XML
- **LaTeX:** Usa para publicación científica, este permite crear documentos muy estructurados y muy detallados, especialmente matemáticas y ciencias.
- **Markdown:** Lenguaje ligero que permite escribir texto plano con marcas simples para convertirlo en HTML u otros formatos.

# Clasificación según su propósito: Lenguaje de marcas de datos


**Lenguajes de marcas de datos.** Estos lenguajes están orientados a describir y estructurar datos, sin centrarse en cómo se verán, sino en cómo se organizan y se pueden intercambiar entre sistemas.

- XML: Diseñado para describir y transportar datos, permitiendo la creación de lenguajes personalizados para definir tipos de información.
- JSON: No es un lenguaje de marcado en el sentido tradicional, usado para el intercambio de datos de manera legible entre humanos y las máquinas.
- YAML: Un formato legible para almacenar datos, similar a JSON, usado para archivos de configuración y almacenamiento de datos.

# Clasificación según su propósito: Lenguaje de marcas semánticos

- Los **lenguajes de marcas semánticos** están diseñados para añadir significado al contenido, lo que permite que las máquinas comprendan mejor la información presentada.
- RDF (Resource Description Framework) : Un modelo estándar para descripción de información de una web.
- Microdata y RDFa (Resource Description Framework in Attribute) : Extensiones de HTML que permiten incrustar metadatos en las páginas web para mejorar la interpretación del contenido por motores de búsqueda y otros servicios.

# Clasificación según su complejidad



- Lenguajes de marcas ligeros
  - Markdown
  - YAML
- Lenguajes de marcas pesados o complejos
  - XML
  - HTML



# Clasificación según su adaptabilidad

- **Lenguajes de marcas fijas**
  - Markdown
  - HTML
- **Lenguajes de marcas extensibles:** Estos lenguajes permite crear nuevas etiquetas y la personalización para adaptarse a diferentes necesidades de código.
  - XML Los usuarios pueden crear sus propias etiquetas dependiendo de la aplicación o el dominio.
  - SGML Es el predecesor de XML, permite la creación de nuevos lenguajes adaptados a diferentes contextos.

# Clasificación según su uso

- **Lenguajes de marcas web**
  - HTML y XHTML: son la base para la presentación del contenido en los navegadores web.
  - CSS (aunque no es un lenguaje de marcas) complementa a HTML para la presentación
- **Lenguajes de marcas para la publicación de documentos:** Para la creación de documentos bien estructurados en diferentes medios
  - LaTeX: Muy utilizado en la academia y la publicación científica..
  - DocBook: Usado par documentación técnica compleja.
- **Lenguajes para el intercambio de datos**
  - XML. Intercambio de información entre aplicaciones web y otros sistemas.
  - JSON y YAML. Muy usados en servidores web y aplicaciones de programación para transmitir datos entre servidores y clientes.

# Clasificación según su uso semántico

- **Lenguajes de marcas generales:** Utilizados para estructurar contenido web, pero con capacidades limitadas para representar datos semánticos.
  - HTML (con algunas etiquetas semánticas).
- **Lenguajes de marcas para la web semánticas:** Ayudan a los sistemas a entender el significado de los datos de la web
  - RDF: Esencial para la creación de ontologías (vocabulario que se está utilizando) y la estructuración de datos para la web semántica.
  - RDFa y Microdata: Extensiones de HTML que permiten añadir metadatos semánticos.

## Otros lenguajes especializados



MathML representación del formalismo matemático.



VoiceXML para el intercambio de información entre el usuario y una aplicación.



MusicXML Para el intercambio de partituras entre distintos editores de partituras.

# Metalinguaje

## Ejemplo

Ejemplo de fórmula

$$k = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

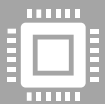
```
1 <?xml version="1.0"?>
2 <math xmlns="http://www.w3.org/1998/Math/MathML">
3
4 <mrow>
5 <mi>x</mi>
6 <mo>=</mo>
7 <mfrac>
8 <mrow>
9 <mrow>
10 <mo>-</mo>
11 <mi>b</mi>
12 </mrow>
13 <mo>&#xB1;<!--PLUS-MINUS SIGN--></mo>
14 <msqrt>
15 <mrow>
16 <msup>
17 <mi>b</mi>
18 <mn>2</mn>
19 </msup>
20 <mo>-</mo>
21 <mrow>
22 <mn>4</mn>
23 <mo>&#x2062;<!--INVISIBLE TIMES--></mo>
24 <mi>a</mi>
25 <mo>&#x2062;<!--INVISIBLE TIMES--></mo>
26 <mi>c</mi>
27 </mrow>
28 </mrow>
29 </msqrt>
30 </mrow>
31 <mrow>
32 <mn>2</mn>
33 <mo>&#x2062;<!--INVISIBLE TIMES--></mo>
34 <mi>a</mi>
35 </mrow>
36 </mfrac>
37 </mrow>
38 </math>
```



# Origen: SGML



SGML son las siglas de Standard Generalized Markup Language o "Estándar de Lenguaje de Marcado Generalizado" para documentos.



Debido a los problemas de compatibilidad entre dispositivos para intercambio de documentos dado que cada aplicación utilizaba sus propias marcas, surge en 1986 SGML (ISO 8879).

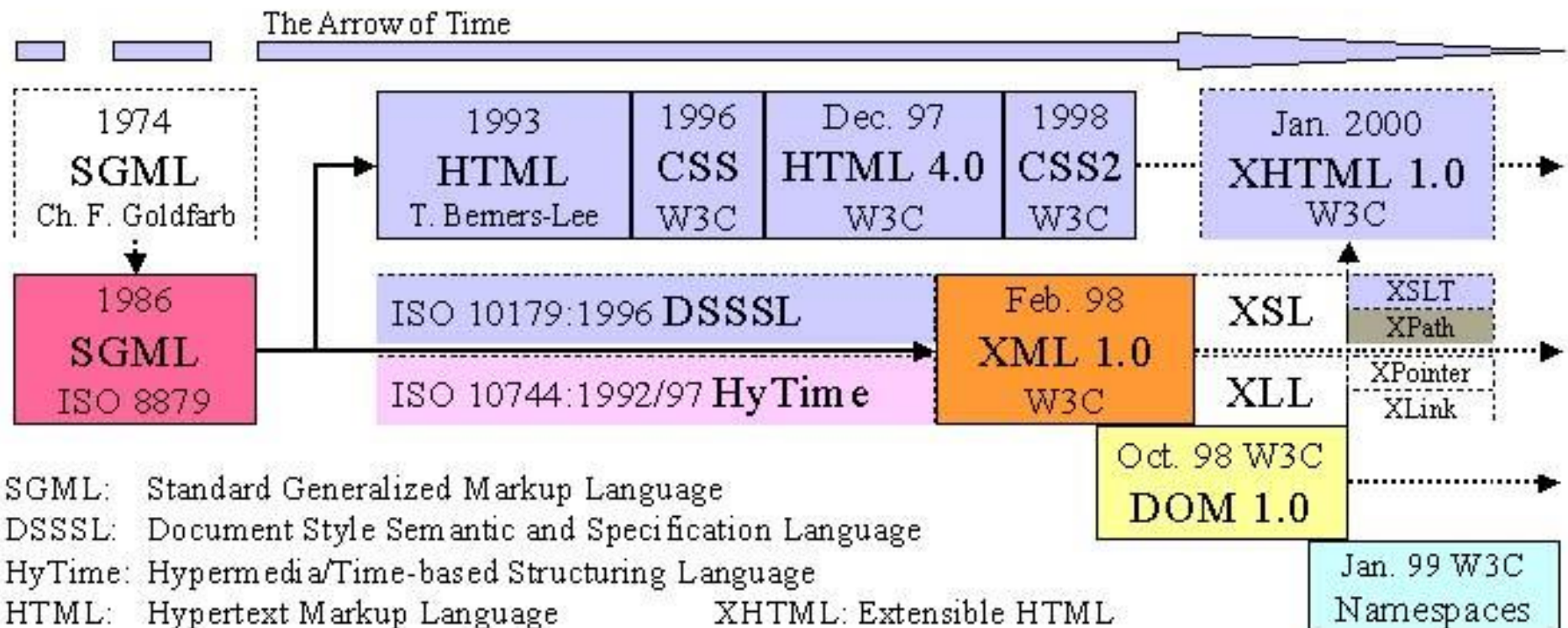


SGML es un metalenguaje, es decir, un conjunto de normas que permiten crear otros lenguajes de marcas.



Éste era un lenguaje muy complejo y requería unas herramientas software muy caras.

# SGML/XML - History



SGML: Standard Generalized Markup Language

DSSSL: Document Style Semantic and Specification Language

HyTime: Hypermedia/Time-based Structuring Language

HTML: Hypertext Markup Language

XHTML: Extensible HTML

CSS: Cascading Style Sheets

DOM: Document Object Model

XML: Extensible Markup Language

XSL: XML Stylesheet Language


XLL: XML Linking Language

XQL: XML Query Language

ISO: Int. Organization for Standardization

W3C: World Wide Web Consortium

# Comparativa XML -SGML



XML	SGML
Su uso es sencillo	Su uso es muy complejo
Trabaja con documentos bien formados, no exige que estén validados.	Sólo trabaja con documentos válidos
Facilita el desarrollo de aplicaciones de bajo coste	Su complejidad hace que las aplicaciones para procesar SGML sean muy costosas.
Compatibilidad e integración con HTML.	No hay una compatibilidad con HTML
Formateo y estilos fáciles de aplicar	Formateo y estilos complejos