

**CURSO 2022-2023**  
CICLO SUPERIOR DE DESARROLLO DE APLICACIONES WEB  
IES AGUADULCE

---

# Lenguajes de Marcas y Sistemas de Gestión de la Información

---

Francisco Javier Sueza Rodríguez

3 de octubre de 2022

# Índice general

|          |  |          |
|----------|--|----------|
| <b>1</b> | <b>Lenguajes de Marcas y Sistemas de Gestión de la Información</b> | <b>4</b> |
| 1.1      | Definición y Clasificación de los Lenguajes de Marcas . . . . .    | 4        |
| 1.2      | Evolución de los Lenguajes de Marcas . . . . .                     | 5        |
| 1.2.1    | El origen: GML y SGML . . . . .                                    | 5        |
| 1.2.2    | La Popularización: HTML . . . . .                                  | 6        |
| 1.2.3    | La Madurez: XML . . . . .  | 6        |
|          | <b>Glosario</b>  | <b>8</b> |
|          | <b>Bibliografía</b>  | <b>9</b> |

# Índice de figuras

|  |   |
|--|---|
| 1.2.1 Ejemplo de XML en el navegador . . . . . | 7 |
|--|---|

# Tema 1

## Lenguajes de Marcas y Sistemas de Gestión de la Información

En esta unidad vamos a estudiar los aspectos básicos de los lenguajes de marcas y los sistemas de gestión de la información. Por un lado, veremos la evolución de los **lenguajes de marcas**, desde GML hasta HTML, así como sus elementos y atributos, haciendo especial énfasis en XML. A continuación, veremos en que consisten los **sistemas de gestión de la información**, en concreto los **ERP**, sus características, configuración básica, personalización,..etc.

### 1.1 Definición y Clasificación de los Lenguajes de Marcas

Los «lenguajes de marcas» sirven para **codificar un documentos**. Estos incorporan **etiquetas** o marcas con **información adicional** sobre como se estructura el texto o como se presenta. El lenguaje de marcas será el que defina que etiquetas se permiten, donde deben colocarse y que significado tienen.

Todo lenguaje de marcas esta definido en un documento denominado **DTD**, donde se establecen las marcas, los elementos utilizados por dicho lenguaje y sus correspondientes etiquetas y atributos, así como su sintaxis.

Los lenguajes de marcas se pueden clasificar, principalmente, en tres grupos:

- **Orientados a la presentación:** son los utilizados generalmente por los procesadores de texto y definen como debe presentarse el documento, es decir, el formato que tiene.
- **De procedimientos:** orientados también a la presentación, pero en este caso, dentro de un **marco procedural** que permite la definición de macros, es decir, el programa que representa el documento debe interpretar el código en el mismo orden que aparece. Algunos ejemplos son **TeX**, **LaTeX** y **Postscript**
- **Descriptivos o semánticos:** estos lenguajes no describen la presentación del documento, sino que **describen la información**, que es lo que se esta representando sin especificar como debe presentarse.

Algunos ejemplos de lenguajes de marcado agrupados por su ámbito de uso son los siguientes:

- **Documentación Electrónica:**
  - **RTF (Rich Text Format):** fue desarrollado por Microsoft en 1987 y permite el intercambio de documentos entre los diferentes procesador de texto.

- **TeX**: creado por **Donald Knuth**, este lenguaje está especialmente enfocado en la creación de textos científicos. Es considerado la mejor forma de componer fórmulas matemáticas complejas. [1]
  - **Wikitexto**: permite la creación de páginas wiki en servidores preparados para soportar este lenguaje.
  - **DocBook**: permite generar documentos separando la estructura lógica del documento de su formato, permitiendo que estos documentos puedan publicarse en diferentes formatos sin tener que modificar el documento original.
- **Tecnologías de Internet:**
- **HTML, XHTML** (Hypertext Markup Language, eXtensible Hypertext Markup Language): estos lenguajes están orientados a la creación de páginas web.
  - **RSS** (Really Simple Syndication): permite la difusión de contenido web mediante la sindicación de contenidos.
- Otros lenguajes especializados:
- **MathML** (Mathematica Markup Language): especializado en expresar los formalismos matemáticos de forma que puedan ser entendidos por diferentes aplicaciones.
  - **VoiceXML** (Voice eXtended Markup Language): permite el intercambio de información entre usuarios y una aplicación con capacidad de reconocer el habla.
  - **MusicXML**: permite el intercambio de partituras entre diferentes editores de partituras.

## 1.2 Evolución de los Lenguajes de Marcas

A finales de los **años 60** surgen unos lenguajes informáticos, diferentes de los lenguajes de programación, orientados a la gestión de la información. Con el desarrollo de los editores y procesadores de texto surgen los primeros lenguajes informáticos orientados a la descripción y estructuración de la información: **los lenguajes de marcas**. Paralelamente también surgen otros lenguajes orientados a la representación, almacenamiento y consulta de grandes cantidades de datos: lenguajes y sistemas de bases de datos.

Los lenguajes de marcas surgieron inicialmente como lenguajes formados por un conjunto de códigos que los procesadores de textos insertaban en los documentos para dirigir el proceso de presentación (impresión) mediante una impresora. Al igual que los lenguajes de programación, estos estaban **ligados** a las características de los **procesadores de texto** y las **impresoras** en los que se usaban y no permitían a los programadores abstraerse de dichas características.

Posteriormente se añadió como medio de presentación a la pantalla y se automatizó el proceso, teniendo ya solo que pulsar una combinación de teclas para lograr los resultados deseados en vez de hacerlo a mano. Este marcado estaba orientado exclusivamente a la presentación de la información, aunque posteriormente se le dieron nuevos usos surgiendo con ello el **formato generalizado**.

### 1.2.1 El origen: GML y SGML

Uno de los problemas que ha tenido la informática ha sido la **falta de estandarización** en los formatos de información usados por los diferentes programas.

En los años 60, **IBM** encargó a **Charles F. Goldfarb** la construcción de un sistema de edición, almacenamiento y búsqueda de documentos legales. Después de analizar el funcionamiento de la empresa se

llegó a la conclusión de que necesitaban un formato estándar a todos los departamentos para gestionar la documentación.

Así fue como se creó **GML**, un formato que permitía describir los documentos de tal forma que el resultado fuese independiente de la plataforma o la aplicación utilizada. Este formato evolucionó hasta que en 1986 se creó el estándar **ISO 8879** donde se especificaba el formato **SGML**, un lenguaje muy complejo y que requería de unas herramientas de software caras, por lo que su uso quedó relegado a grandes aplicaciones industriales.

### 1.2.2 La Popularización: HTML

En 1990, **Tim Berners-Lee** creó el World Wide Web y conociendo SGML, se encontró con la necesidad de compatibilizar, enlazar y organizar gran cantidad de documentos procedentes de diversos sistemas. Como solución, a partir de la sintaxis de SGML, creó un lenguaje de descripción de documentos llamado **HTML**, combinando dos estándares ya existentes:

- **ASCII**: código basado en el alfabeto latino, tal como se usa en inglés moderno [2]. Cualquier procesador de textos simple puede reconocer y almacenar este formato, permitiendo la transferencia de datos entre dos ordenadores.
- **SGML**: lenguaje que permite dar estructura al texto aplicando diferentes formatos.

**HTML** es una **versión simplificada de SGML**, ya que solo utiliza las instrucciones absolutamente necesarias. Gracias a su simplicidad, tuvo un éxito rotundo en la World Wide Web, convirtiéndose rápidamente en el **estándar general** para la **creación de páginas web**.

Actualmente, HTML es el tipo de documento más utilizado en el mundo, gracias en parte a su sencillez, ya cualquier persona puede escribir documentos en este lenguaje sin tener prácticamente conocimientos de informática.

### 1.2.3 La Madurez: XML

Uno de los problemas que surgió con HTML es que la cantidad de documentos escritos en este lenguaje creció exponencialmente, muchos de los cuales no se ceñían a ningún estándar generando bastante caos. Como respuesta a ese problema, el **W3C** estableció en 1998 el estándar internacional **XML**, un lenguaje de marcas puramente estructural, que **no incluye información sobre el diseño**, y permite la creación de etiquetas adaptadas a las necesidades, convirtiéndose con rapidez en el estándar para intercambio de datos en la web.

**XML** es un **metalenguaje** con las siguientes características:

- Permitir definir etiquetas propias.
- Permitir asignar atributos a las etiquetas.
- Utilizar un esquema para definir de forma exacta las etiquetas y sus atributos.
- La estructura y el diseño son independientes.

En realidad XML es un **conjunto de estándares** relacionados entre sí y que comprende los siguientes:

- **XLS** (eXtensible Style Language): permite definir hojas de estilo para XML e incluye capacidad de transformación de documentos.

- **XML Linking Language:** determina aspectos sobre los enlaces entre documentos XML e incluye **Xpath**, **Xlink** y **Xpointer**.
- **XML Namespaces:** proveen de un contexto donde se aplican las marcas del documento XML y que se diferencian de otras con el mismo nombre válidas en otros contextos.
- **XML Schemas:** permiten definir restricciones que se aplicarán a un documento XML. Actualmente las mas utilizadas son **DTD**.

A continuación se muestra un ejemplo de un documentos XML.

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

<biblioteca>
  <ejemplar tipo="libro" isbn="978-2-7460-4958-1" edicion="1">
    <titulo>XML practico</titulo>
    <editorial>Ediciones Eni</editorial>
    <autor>Sebastien Lecomte</autor>
    <autor>Thierry Boulanger</autor>
    <autor funcion="traductor">Ángel Belinchon Calleja</autor>
    <prestamos>
      <lector inicio="13/05/2014" devolucion="15/05/2014">Pedro López</lector>
      <lector inicio="13/07/2015" devolucion="15/07/2015">Ali Méndez</lector>
    </prestamos>
  </ejemplar>
</biblioteca>
```

Si publicáramos este código en un navegador, el resultado sería el que se muestra en la siguiente figura, capturado en un navegador Firefox.

```
-<biblioteca>
  -<ejemplar tipo="libro" isbn="978-2-7460-4958-1" edicion="1">
    <titulo>XML practico</titulo>
    <editorial>Ediciones Eni</editorial>
    <autor>Sebastien Lecomte</autor>
    <autor>Thierry Boulanger</autor>
    <autor funcion="traductor">Ángel Belinchon Calleja</autor>
  -<prestamos>
    <lector inicio="13/05/2014" devolucion="15/05/2014">Pedro López</lector>
    <lector inicio="13/07/2015" devolucion="15/07/2015">Ali Méndez</lector>
  </prestamos>
  </ejemplar>
</biblioteca>
```

Figura 1.2.1: Ejemplo de XML en el navegador

# Glosario

**ASCII** American Standard Code for Information Interchange. 5

**DTD** Document Type Definition. 3, 5

**GML** Generalized Markup Language. 5

**HTML** Hypertext Markup Language. 5

**metalenguaje** Lenguaje que permite la definición de otros lenguajes. 5

**SGML** Standard Generalized Markup Language. 5

**W3C** Word Wide Web Consortium. 5

**XML** eXtensible Markup Language. 5



# Bibliografía

- [1] Wikipedia - Tex. <https://es.wikipedia.org/wiki/TeX>.
- [2] Wikipedia - ASCII. <https://es.wikipedia.org/wiki/ASCII>.