

Ficheros de Marcas

Se puede decir que los archivos de marcas son una **manera diferente de almacenar información** en ordenadores.

El **objetivo** principal de los ficheros de marcas es:

- intentar recoger las mejores características de los archivos de texto y binarios
- evitar sus problemas.

Los ficheros de marcas toman **como base los archivos de texto** para aprovecharse de las características más interesantes de este tipo de archivos:

- La facilidad de creación y lectura.
- El cumplimiento de **estándares de almacenamiento** definidos y públicos.

Los archivos de marcas también intentan conseguir las características más interesantes de los **ficheros binarios**, como:

- La incorporación de **metadatos**.
- La definición de la **estructura** de los datos.

Los lenguajes de marcas, se han convertido en una de las formas habituales de representar datos:

- Internet **-La Web-**, que está basada totalmente en los lenguajes de marcas.
- Muchos **programas**, utilizan algún lenguaje de marcas para almacenar sus **datos de configuración o de resultados**:
 - Internamente los formatos de documentos de Microsoft Office o de OpenOffice/LibreOffice, están basados en lenguajes de marcas.
 - Microsoft Visual Studio guarda su configuración utilizando lenguajes de marcas.

Las marcas son una serie de códigos que se incorporan a los documentos digitales para determinar su **formato**:

- la forma en que se han de imprimir,
- la estructura de los datos, etc.

Son **anotaciones que se incorporan a los datos** pero que no forman parte de ellos

Las marcas deben ser fácilmente distinguibles del texto normal:

- por su posición,
- porque siguen algún tipo de sintaxis, etc.

Las más usadas son las que están formadas por

- textos descriptivos
- rodeadas de los símbolos de "menor " (<) y "mayor" (>)



```
< nombre > Manel Puig Garcia </ nombre >
```

```
<persona >
  <nombre > Manel Perez Garcia </ nombre >
  <nombre > Pedro González Parada </ nombre >
  <nombre > María Pozos Calle </ nombre >
</ persona >
```

Estas marcas pueden ser ensambladas para **indicar estructuras de datos.**

```
= Personas =
* Manel Perez Garcia
* Pedro González Parada
* María Pozos Calle
```

Wiki Markup los caracteres "=" en la primera posición de una línea se usan para indicar que el texto es un título de apartado y el * para cada elemento de la lista

```
\ Section {Personas} \ Begin {itemize}
\ Item Manel Perez Garcia
\ Item Pedro González Parada
\ Item María Pozos Calle
\ End {itemize}
```

TeX utiliza las barras invertidas para indicar el inicio de las marcas

Las **características** más interesantes que ofrecen los **lenguajes de marcas** se encuentran:

- Que se basan en el **texto plano**.
- Que permiten utilizar **metadatos**.
- Que son **fáciles de interpretar y procesar**.
- Que son **fáciles de crear** y suficientemente **flexibles** para representar datos muy diversos.

Clasificación de los lenguajes de marcas

Según cuál es su objetivo básico:

- Lenguajes **procedimentales y de presentación**, orientados a especificar cómo se debe representar la información.
- Lenguajes **descriptivos o semánticos**: orientados a describir la estructura de los datos que contiene.

Esta es la clasificación **más aceptada**, pero, nos podemos encontrar lenguajes que tengan aspectos de los dos grupos y permitan tanto definir la manera de presentar la información como definir su estructura

Procedimentales y de presentación

HTML,
TeX,
Wikitext

...

Si tomamos un ejemplo sencillo utilizando el lenguaje de marcas ligero Wiki markup, que utiliza *Mediawiki* (programa con el que se ha desarrollado la Wikipedia), que nos mostrará lo que se ve en la Figura 1.7:

```
= Clase =
```

```
== Asignatura: XML ==
```

```
[[Archivo: xml.png]]
```

```
" 'Profesor' ":
```

```
* " Manel Puig "
```

```
" Alumnos " '
```

```
* Frederic Puig
```

```
* Filomeno García
```

```
* Manel Puigdevall
```

Figura 1.7 Representación del texto en formato wikicode

Classe

Asignatura: XML



Professor:

- Manel Puig

Alumnes

- Frederic Puig
- Filomeno Garcia
- Manel Puigdevall

Descriptivos o semánticos

XML,
JSON

...

```
<alumnos >  
  <persona >  
    <nombre > Pedro </ nombre >  
    <apellido > Perez </ apellido >  
  </ persona >  
  <persona >  
    <nombre > Manuel </ nombre >  
    <apellido > Garcia </ apellido >  
  </ persona >
```

Se describe qué estructura lógica tiene el documento, ignorando de qué manera será representada en los programas

Estructura de los datos y su **significado** (contenido semántico de las etiquetas)

Se deduce que Pedro es el nombre de una persona que es un alumno