

# INTRODUCCIÓN A LOS LENGUAJES DE MARCAS

JOSÉ DOMINGO MUÑOZ

IES GONZALO NAZARENO

ENERO 2020



- Se utilizan para cualquier tipo de documento electrónico (textos, presentaciones, gráficos, tecnologías de Internet, matemáticas, música, multimedia, ...)
- Incluyen anotaciones (marcas o etiquetas) con información adicional.
- Utilizado desde los años 60: Scribe, roff, TEX, SGML, HTML, ...
- Desde los años 80 ha predominado el uso de formatos binarios privativos y cerrados.
- A finales de los 90 surgió XML (eXtensible Markup Language), que se ha convertido en un estándar.
- En los últimos años han nacido nuevos lenguajes de marcas que actualmente se usan en muchas ocasiones: json, yaml, markdown, ...



# ¿CÓMO PUEDO ESTRUCTURAR LA INFORMACIÓN?

La forma más simple de estructurar la información en un fichero de texto, es separandola por un signo: CSV (Comma Separated Values):

```
LANG,TITLE,AUTHOR,YEAR,PRICE,CATEGORY
en,Everyday Italian,Giada De Laurentiis,2005,30.00,COOKING
en,Harry Potter,J K. Rowling,2005,29.99,CHILDREN
en,XQuery Kick Start,(James McGovern;Per Bothner;Kurt Cagle;James Linn;Vaidyanathan Nagarajan),2003,49.99,WEB
en,Learning XML,Erik T. Ray,2003,39.95,WEB
```

- No puedo representar estructuras complejas.
- La posición del dato es importante
- Todas las líneas tienen que tener el mismo número de elementos
- Si tengo conjuntos de elementos, tengo que usar otro signo
- Son difíciles de tratar por un ordenador (programa)



# ¿CÓMO PUEDO ESTRUCTURAR LA INFORMACIÓN?

Para estructurar la información en ficheros de textos usamos los lenguajes de marcas:

- XML
- JSON
- YAML
- HTML
- Markdown
- ...



# JSON

```
"bookstore": {  
  "book": [  
    {  
      "title": {  
        "lang": "en",  
        "text": "Everyday Italian"  
      },  
      "author": "Giada De Laurentiis",  
      "year": "2005",  
      "price": "30.00",  
      "category": "COOKING"  
    },  
    {  
      "title": {  
        "lang": "en",  
        "text": "Harry Potter"  
      },  
      "author": "J K. Rowling",  
      "year": "2005",  
      "price": "29.99",  
      "category": "CHILDREN"  
    },  
    ...  
  ]  
}
```

```
{  
  "title": {  
    "lang": "en",  
    "text": "XQuery Kick Start"  
  },  
  "author": [  
    "James McGovern",  
    "Per Bothner",  
    "Kurt Cagle",  
    "James Linn",  
    "Vaidyanathan Nagarajan"  
  ],  
  "year": "2003",  
  "price": "49.99",  
  "_category": "WEB"  
},  
{  
  "title": {  
    "lang": "en",  
    "text": "Learning XML"  
  },  
  "author": "Erik T. Ray",  
  "year": "2003",  
  "price": "39.95",  
  "category": "WEB"  
}  
}  
}  
]
```

# XML

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<bookstore>
  <book category="COOKING">
    <title lang="en">Everyday Italian</title>
    <author>Giada De Laurentiis</author>
    <year>2005</year>
    <price>30.00</price>
  </book>
  <book category="CHILDREN">
    <title lang="en">Harry Potter</title>
    <author>J K. Rowling</author>
    <year>2005</year>
    <price>29.99</price>
  </book>
  <book category="WEB">
    <title lang="en">XQuery Kick Start</title>
    <author>James McGovern</author>
    <author>Per Bothner</author>
    <author>Kurt Cagle</author>
    <author>James Linn</author>
    <author>Vaidyanathan Nagarajan</author>
    <year>2003</year>
    <price>49.99</price>
  </book>
</bookstore>
```

# YAML

```
bookstore:
  book:
    - title:
        lang: "en"
        text: "Everyday Italian"
        author: "Giada De Laurentiis"
        year: "2005"
        price: "30.00"
        category: "COOKING"
    - title:
        lang: "en"
        text: "Harry Potter"
        author: "J K. Rowling"
        year: "2005"
        price: "29.99"
        category: "CHILDREN"
    - title:
        lang: "en"
        text: "XQuery Kick Start"
        author:
          - "James McGovern"
          - "James Linn"
          - "Vaidyanathan Nagarajan"
        year: "2003"
        price: "49.99"
        category: "WEB"
```



```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//E
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
<html>
  <head>
    <title>Biblioteca</title>
  </head>
  <body>
    <h1>bookstore</h1>
    <table>
      <tr>
        <td>Category</td>
        <td>Lenguaje</td>
        <td>Title</td>
        <td>Author</td>
        <td>Year</td>
        <td>Price</td>
      </tr>
      <tr>
        <td>COOKING</td>
        <td>en</td>
        <td>Everyday Italian</td>
        <td>Giada De Laurentiis</td>
        <td>2005</td>
        <td>30.00</td>
      </tr>
```

```
      <tr>
        <td>CHILDREN</td>
        <td>en</td>
        <td>Harry Potter</td>
        <td>J K. Rowling</td>
        <td>2003</td>
        <td>49.99</td>
      </tr>
      <tr>
        <td>WEB</td>
        <td>en</td>
        <td>XQuery Kick Start</td>
        <td>James McGovern<br/>Per Bothner<br/>Kurt Cag
        <td>2003</td>
        <td>49.99</td>
      </tr>
    </table>
  </body>
</html>
```





# CLASES DE LENGUAJES DE MARCAS

Se suelen distinguir tres clases de lenguajes de marcas:

- **Marcado de Presentación:** Donde sólo se muestra la presentación, pero es difícil extraer información. (**HTML, Markdown**)
- **Marcado de Procedimientos:** Donde se incluyen instrucciones de cómo hay que procesar el texto.
- **Marcado Descriptivo o semántico:** Donde no se especifica cómo hay que procesar el texto sino que se utilizan etiquetas que describen el texto. (**XML, Json, Yaml**)

Aunque en general la mayoría de lenguajes de marcas tienen características de más de una clase.

[codebeautify.org](https://codebeautify.org)

