

Representación de la información alfanumérica (texto)

Contenidos:

- 1) ASCII
- 2) Unicode: sistemas de codificación UTF-8, UTF-16, UTF-32

Para representar texto o caracteres alfanuméricos se han propuesto muchos sistemas de codificación distintos. Desde el punto de vista de nuestro propósito, los que más nos interesan son: ASCII y UTF-8/UTF-16/UTF-32

No se trata de adquirir un conocimiento a fondo de la historia, etcétera de estos sistemas de codificación, si no más bien:

- conocer sus nombres y saber un poco como funcionan
- ser consciente de que cuando trabajamos con texto (bases de datos, programación, etcétera) es importante comprender que no todos los archivos tienen porque estar en el mismo formato y que pueden existir incompatibilidades, que te irás encontrando.....

1. ASCII

Estudia el artículo de la Wikipedia: [ASCII](https://es.wikipedia.org/wiki/ASCII) (<https://es.wikipedia.org/wiki/ASCII>)

Ideas importantes a sacar:

- **caracteres de control** y **caracteres imprimibles**
- diferencia entre ASCII y ASCII extendido
- concepto de páginas de códigos (code pages): ver
 - CP850: [Página de códigos 850](#)
 - CP-1252 ([Win-1252](#)),

Otro juego de caracteres muy importante en desarrollo web es: [ISO-8859-1](https://es.wikipedia.org/wiki/ISO/IEC_8859-1) (https://es.wikipedia.org/wiki/ISO/IEC_8859-1): échale un vistazo

2. Unicode: sistemas de codificación UTF-8, UTF-16, UTF-32

Hoy en día se tiende a utilizar Unicode para todo. Lee los artículos siguientes de la wikipedia:

- Unicode
 - Conceptos importantes a entender:
 - Diferencia entre **Unicode** y los **sistemas de codificación** de Unicode (UTF-8/16/32)
 - **BOM**
 - **little-endian** y **big-endian**
 - ver el ejemplo de "Codificación de caracteres"
- [UTF-8](#)
- [UTF-16](#)