



# Codificaciones

Vazquez Rocha Jorge Ivan

3 de febrero de 2020

## Resumen

El presente documento contiene una breve descripción de las codificaciones *ASCII*, *ISO 8859*, *Unicode* y *UTF-8*, así como una tabla con diferentes codificaciones para algunos caracteres.

## 1. ASCII

El *ASCII* es un código numérico que representa los caracteres, usando una escala decimal del 0 al 127. Esos números decimales son convertidos por la computadora en números binarios para ser posteriormente procesados. Este código nos permite codificar combinaciones de caracteres y símbolos que usan las computadoras internamente para gestionar texto en palabras, instrucciones y letras<sup>1</sup>.

## 2. ISO 8859-1

*ISO 8859-1* es una norma de la ISO que define la codificación del alfabeto latino, incluyendo los diacríticos (como letras acentuadas, ñ), y letras especiales, necesarios para la escritura de algunas lenguas originarias de Europa occidental. Esta norma pertenece al grupo de juegos de caracteres de la ISO que se caracterizan por poseer la codificación ASCII en su rango inicial (128 caracteres) y otros 128 caracteres para cada codificación, con lo que en total utilizan 8 bits<sup>2</sup>.

---

<sup>1</sup><https://tecnologia-informatica.com/que-es-codigo-ascii/>  
<http://informatica.dgenp.unam.mx/recomendaciones/codigo-ascii>

<sup>2</sup>[https://es.wikipedia.org/wiki/ISO/IEC\\_8859-](https://es.wikipedia.org/wiki/ISO/IEC_8859-1)

### 3. Unicode

*Unicode* es un set de caracteres universal, es decir, un estándar en el que se definen todos los caracteres necesarios para la escritura de la mayoría de los idiomas hablados en la actualidad que se usan en la computadora. Además, en *Unicode* existen distintas formas de codificar el mismo carácter. Por ejemplo, la letra á se puede representar mediante dos bytes en una codificación y con cuatro bytes, en otra<sup>3</sup>.

### 4. UTF-8

*UTF-8* es un formato de codificación que utiliza 1 byte para representar caracteres en el set ASCII, dos bytes para caracteres en otros bloques alfabéticos y tres bytes para el resto del BMP. Para los caracteres complementarios se utilizan 4 bytes<sup>4</sup>.

### 5. Tabla comparativa

Carácter	Código ASCII		ISO 8859-1	
	Decimal	Hexadecimal	Decimal	Hexadecimal
a	97	61	97	61
b	98	62	98	62
c	99	63	99	63
A	65	41	65	41
B	66	42	66	42
C	67	43	67	43
á	-	-	225	E1
é	-	-	233	E9
í	-	-	237	ED
ó	-	-	243	F3
ú	-	-	250	FA
ñ	-	-	241	F1
ç	-	-	191	BF
ï	-	-	161	A1

---

<sup>3</sup><https://www.w3.org/International/articles/definitions-characters/index.es>

<sup>4</sup><https://www.w3.org/International/articles/definitions-characters/index.es>