

American Standard Code for Information Interchange

El código **ASCII** ([acrónimo inglés](#) de **American Standard Code for Information Interchange** — *Código Estadounidense Estándar para el Intercambio de Información*), pronunciado generalmente [áski], es un [código de caracteres](#) basado en el [alfabeto latino](#) tal como se usa en inglés moderno y en otras lenguas occidentales. Fue creado en [1963](#) por el Comité Estadounidense de Estándares (ASA, conocido desde 1969 como el Instituto Estadounidense de Estándares Nacionales, o [ANSI](#)) como una refundición o evolución de los conjuntos de códigos utilizados entonces en [telegrafía](#). Más tarde, en [1967](#), se incluyeron las minúsculas, y se redefinieron algunos códigos de control para formar el código conocido como **US-ASCII**.

El código ASCII utiliza 7 bits para representar los caracteres, aunque inicialmente empleaba un bit adicional ([bit de paridad](#)) que se usaba para detectar errores en la transmisión. A menudo se llama incorrectamente ASCII a otros [códigos de caracteres de 8 bits](#), como el estándar [ISO-8859-1](#) que es una extensión que utiliza 8 bits para proporcionar caracteres adicionales usados en idiomas distintos al inglés, como el español.

ASCII fue publicado como estándar por primera vez en 1967 y fue actualizado por última vez en 1986. En la actualidad define códigos para 33 caracteres no imprimibles, de los cuales la mayoría son [caracteres de control](#) obsoletos que tienen efecto sobre como se procesa el texto, más otros 95 caracteres imprimibles que les siguen en la numeración (empezando por el carácter espacio).

Casi todos los sistemas informáticos actuales utilizan el código ASCII o una extensión compatible para representar textos y para el control de dispositivos que manejan texto.

Como otros códigos de representación de caracteres, el ASCII especifica una correspondencia entre cadenas de bits y una serie de símbolos (alfanuméricos y otros), permitiendo de esta forma la comunicación entre dispositivos digitales así como su procesamiento y almacenamiento. El código de caracteres ASCII — o una extensión compatible (ver más abajo) — se usa casi en todos los ordenadores, especialmente con [ordenadores personales](#) y [estaciones de trabajo](#). El nombre más apropiado para este código de caracteres es "US-ASCII".

ASCII es, en sentido estricto, un código de siete bits, lo que significa que usa

!	"	#	\$	%	&	'	()	*	+	,	-	.	/	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?	
@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	[\]	^	_
`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	{		}	~	~

cadenas de bits representables con siete dígitos binarios (que van de 0 a 127 en base decimal) para representar información de caracteres. En el momento en el que se introdujo el código ASCII muchos ordenadores trabajaban con grupos de ocho bits ([bytes](#) u [octetos](#)), como la unidad mínima de información; donde el octavo bit se usaba habitualmente como [bit de paridad](#) con funciones de control de errores en líneas de comunicación u otras funciones específicas del dispositivo. Las máquinas que no usaban la comprobación de paridad asignaban al octavo bit el valor cero en la mayoría de los casos, aunque otros sistemas como las computadoras [Prime](#), que ejecutaban [PRIMOS](#) ponían el octavo bit del código ASCII a uno.

El código ASCII define una relación entre caracteres específicos y secuencias de bits; además de reservar unos cuantos códigos de control para el procesador de textos, y no define ningún mecanismo para describir la estructura o la apariencia del texto en un documento; estos asuntos están especificados por otros lenguajes como los [lenguajes de etiquetas](#).

Historia

El código ASCII se desarrolló en el ámbito de la [telegrafía](#), y se usó por primera vez comercialmente como un código de teleimpresión impulsado por los servicios de datos de [Bell](#). Bell había planeado usar un código de seis bits, derivado de [Fieldata](#), que añadía puntuación y letras

minúsculas al más antiguo código de teleimpresión [Baudot](#), pero se les convenció para que se unieran al subcomité de la Agencia de Estándares Estadounidense (ASA), que habían empezado a desarrollar el código ASCII. Baudot ayudó en la automatización del envío y recepción de mensajes telegráficos, y tomó muchas características del [código Morse](#); sin embargo, a diferencia del código Morse, Baudot usó códigos de longitud constante. Comparado con los primeros códigos telegráficos, el código propuesto por Bell y ASA resultó en una reorganización más conveniente para ordenar listas (especialmente porque estaba ordenado alfabéticamente) y añadió características como la '[secuencia de escape](#)'.

La Agencia de Estándares Estadounidense (ASA), que se convertiría más tarde en el Instituto Nacional Estadounidense de Estándares ([ANSI](#)), publicó por primera vez el código ASCII en 1963. El ASCII publicado en 1963 tenía una flecha apuntando hacia arriba (↑) en lugar del circunflejo (^) y una flecha apuntando hacia la izquierda en lugar del guión bajo (_). La versión de 1967 añadió las letras minúsculas, cambió los nombres de algunos códigos de control y cambió de lugar los dos códigos de control ACK y ESC de la zona de letras minúsculas a la zona de códigos de control.

ASCII fue actualizado en consecuencia y publicado como ANSI X3.4-1968, ANSI X3.4-1977, y finalmente ANSI X3.4-1986.

Otros órganos de estandarización han publicado códigos de caracteres que son idénticos a ASCII. Estos códigos de caracteres reciben a menudo el nombre de ASCII, a pesar de que ASCII se define estrictamente solamente por los estándares ASA/ANSI:

- La Asociación Europea de Fabricantes de Ordenadores ([ECMA](#)) publicó ediciones de su clon de ASCII, ECMA-6 en 1965, 1967, 1970, 1973, 1983, y 1991. La edición de 1991 es idéntica a ANSI X3.4-1986.
- La Organización Internacional de Estandarización ([ISO](#)) publicó su versión, ISO 646 (más tarde ISO/IEC 646) en 1967, 1972, 1983 y 1991. En particular, ISO 646:1972 estableció un conjunto de versiones específicas para cada país donde los caracteres de puntuación fueron reemplazados con caracteres no ingleses. ISO/IEC 646:1991 La International Reference Version es la misma que en el ANSI X3.4-1986.
- La Unión Internacional de Telecomunicaciones ([ITU](#)) publicó su versión de ANSI X3.4-1986, Recomendación ITU T.50, en 1992. A principios de la [década de 1970](#) publicó una versión como Recomendación CCITT V.3.
- [DIN](#) publicó una versión de ASCII como el estándar DIN 66003 en 1974.
- El Grupo de Trabajo en Ingeniería de Internet ([IETF](#)) publicó una versión en 1969 como [RFC 20](#), y estableció la versión estándar para Internet, basada en ANSI X3.4-1986, con la publicación de [RFC 1345](#) en 1992.
- La versión de [IBM](#) de ANSI X3.4-1986 se publicó en la literatura técnica de IBM como [página de códigos 367](#).

El código ASCII también está incluido en su probable relevo, [Unicode](#), constituyendo los primeros 128 caracteres (o los 'más bajos'). Algunos observadores consideran el código ASCII el estándar de software más exitoso que nunca se haya promulgado.

Los caracteres de control ASCII

El código ASCII reserva los primeros 32 códigos (numerados del 0 al 31 en decimal) para [caracteres de control](#): códigos no pensados originalmente para representar información imprimible, sino para controlar dispositivos (como [impresoras](#)) que usaban ASCII. Por ejemplo, el carácter 10 representa la función "nueva línea" (line feed), que hace que una impresora avance el papel, y el carácter 27 representa la tecla "escape" que a menudo se encuentra en la esquina superior izquierda de los [teclados](#) comunes.

El código 127 (los siete bits a uno), otro carácter especial, equivale a "suprimir" ("delete"). Aunque esta función se asemeja a otros caracteres de control, los diseñadores de ASCII idearon este código para poder "borrar" una sección de [papel perforado](#) (un medio de almacenamiento popular hasta la década de 1980) mediante la perforación de todos los agujeros posibles de una posición de carácter concreta, reemplazando cualquier información previa. Dado que el código 0 era ignorado, fue posible dejar huecos (regiones de agujeros) y más tarde hacer correcciones.

Muchos de los caracteres de control ASCII servían para marcar paquetes de datos, o para controlar

protocolos de transmisión de datos (por ejemplo ENquiry, con el significado: ¿hay alguna estación por ahí?, ACKnowledge: recibido o "acuse de recibo", Negative Acknowledge: No recibido, Start Of Header: inicio de cabecera, Start of TeXt: inicio de texto, End of TeXt: final de texto, etc.). ESCape y SUBstitute permitían a un protocolo de comunicaciones, por ejemplo, marcar datos binarios para que contuviesen códigos con el mismo código que el carácter de protocolo, y que el receptor pudiese interpretarlos como datos en lugar de como caracteres propios del protocolo.

Los diseñadores del código ASCII idearon los caracteres de separación para su uso en sistemas de cintas magnéticas.

Dos de los caracteres de control de dispositivos, comúnmente llamados [XON](#) y [XOFF](#) generalmente ejercían funciones de caracteres de [control de flujo](#) para controlar el flujo a hacia un dispositivo lento (como una impresora) desde un dispositivo rápido (como un ordenador), de forma que los datos no saturasen la capacidad de recepción del dispositivo lento y se perdiesen.

Los primeros usuarios de ASCII adoptaron algunos de los códigos de control para representar "metainformación" como final-de-línea, principio/final de un elemento de datos, etc. Estas asignaciones a menudo entraban en conflicto, así que parte del esfuerzo de convertir datos de un formato a otro comporta hacer las conversiones correctas de metainformación. Por ejemplo, el carácter que representa el final-de-línea en ficheros de texto varía con el [sistema operativo](#). Cuando se copian archivos de un sistema a otro, el sistema de conversión debe reconocer estos caracteres como marcas de final-de-línea y actuar en consecuencia.

Actualmente los usuarios de ASCII usan menos los caracteres de control, (con algunas excepciones como "retorno de carro" o "nueva línea"). Los lenguajes modernos de etiquetas, los protocolos modernos de comunicación, el paso de dispositivos basados en texto a basados en gráficos, el declive de las teleimpresoras, las tarjetas perforadas y los papeles continuos han dejado obsoleta la mayoría de caracteres de control.

Caracteres imprimibles ASCII

El código el carácter espacio, designa al espacio entre palabras, y se produce normalmente por la barra espaciadora de un teclado. Los códigos del 33 al 126 se conocen como caracteres imprimibles, y representan letras, dígitos, signos de puntuación y varios símbolos.

El ASCII de siete bits proporciona siete caracteres "nacionales" y, si la combinación concreta de hardware y software lo permite, puede utilizar combinaciones de teclas para simular otros caracteres internacionales: en estos casos un backspace puede preceder a un acento abierto o grave (en los estándares británico y americano, pero sólo en estos estándares, se llama también "opening single quotation mark"), una [tilde](#) o una "marca de respiración".

Rasgos estructurales

- Los dígitos del 0 al 9 se representan con sus valores prefijados con el valor 0011 en binario (esto significa que la conversión [BCD-ASCII](#) es una simple cuestión de tomar cada unidad bcd y prefijarla con 0011).
- Las cadenas de bits de las letras minúsculas y mayúsculas sólo difieren en un bit, simplificando de esta forma la conversión de uno a otro grupo.

La [RFC 1345](#) (publicada en Junio de 1992) y el [registro IANA de códigos de caracteres](#), reconocen los siguientes nombres alternativos para ASCII para su uso en Internet.

- ANSI_X3.4-1968 (nombre canónico)
- ANSI_X3.4-1986
- ASCII
- US-ASCII (nombre MIME recomendado)
- us
- ISO646-US
- ISO_646.irv:1991
- iso-ir-6

- IBM367
- cp367
- csASCII

De estos, sólo los nombres "US-ASCII" y "ASCII" se usan ampliamente. A menudo se encuentran en el parámetro de "código de caracteres" opcional en la cabecera Content-Type de algunos mensajes [MIME](#), en el elemento equivalente "meta" de algunos documentos [HTML](#), y en la parte de declaración de codificación de carácter de la cabecera de algunos documentos [XML](#).@

Variantes de ASCII

A medida que la tecnología informática se difundió a lo largo del mundo, se desarrollaron diferentes estándares y las empresas desarrollaron muchas variaciones del código ASCII para facilitar la escritura de lenguas diferentes al inglés que usaran alfabetos latinos. Se pueden encontrar algunas de esas variaciones clasificadas como "[ASCII Extendido](#)", aunque en ocasiones el término se aplica erróneamente para cubrir todas las variantes, incluso las que no preservan el conjunto de códigos de caracteres original ASCII de siete bits.

La [ISO 646](#) (1972), el primer intento de remediar el sesgo pro-inglés de la codificación de caracteres, creó problemas de compatibilidad, pues también era un código de caracteres de 7 bits. No especificó códigos adicionales, así que reasignó algunos específicamente para los nuevos lenguajes. De esta forma se volvió imposible saber en qué variante se encontraba codificado el texto, y, consecuentemente, los procesadores de texto podían tratar una sola variante.

La tecnología mejoró y aportó medios para representar la información codificada en el octavo bit de cada byte, liberando este bit, lo que añadió otros 128 códigos de carácter adicionales que quedaron disponibles para nuevas asignaciones. Por ejemplo, [IBM](#) desarrolló páginas de código de 8 bits, como la [página de códigos 437](#), que reemplazaba los caracteres de control con símbolos gráficos como sonrisas, y asignó otros caracteres gráficos adicionales a los 128 bytes superiores de la página de códigos. Algunos sistemas operativos como [DOS](#), podían trabajar con esas páginas de código, y los fabricantes de [ordenadores personales](#) incluyeron soporte para dichas páginas en su hardware.

Los estándares de ocho bits como [ISO 8859](#) y [Mac OS Roman](#) fueron desarrollados como verdaderas extensiones de ASCII, dejando los primeros 127 caracteres intactos y añadiendo únicamente valores adicionales por encima de los 7-bits. Esto permitió la representación de un abanico mayor de lenguajes, pero estos estándares continuaron sufriendo incompatibilidades y limitaciones. Todavía hoy, [ISO-8859-1](#) y su variante [Windows-1252](#) (a veces llamada erróneamente ISO-8859-1) y el código ASCII original de 7 bits son los códigos de carácter más comúnmente utilizados.

[Unicode](#) y [Conjunto de Caracteres Universal](#) (UCS) [ISO/IEC 10646](#) definen un conjunto de caracteres mucho mayor, y sus diferentes formas de codificación han empezado a reemplazar ISO 8859 y ASCII rápidamente en muchos entornos. Mientras que ASCII básicamente usa códigos de 7-bits, Unicode y UCS usan "code points" o apuntadores relativamente abstractos: números positivos (incluyendo el cero) que asignan secuencias de 8 o más bits a caracteres. Para permitir la compatibilidad, Unicode y UCS asignan los primeros 128 apuntadores a los mismos caracteres que el código ASCII. De esta forma se puede pensar en ASCII como un subconjunto muy pequeño de Unicode y UCS. La popular codificación [UTF-8](#) recomienda el uso de uno a cuatro valores de 8 bits para cada apuntador, donde los primeros 128 valores apuntan a los mismos caracteres que ASCII. Otras codificaciones de caracteres como [UTF-16](#) se parece a ASCII en cómo representan los primeros 128 caracteres de Unicode, pero tienden a usar 16 a 32 bits por carácter, así que requieren de una conversión adecuada para que haya compatibilidad entre ambos códigos de carácter.

La palabra *ASCIIbético* (o, más habitualmente, la palabra "inglesa" *ASCIIbetical*) describe la ordenación según el orden de los códigos ASCII en lugar del orden alfabético.

La abreviatura ASCIIZ o ASCIZ se refiere a una cadena de caracteres terminada en cero (del inglés "zero").