

ASCII (ang. American Standard Code for Information Interchange, Standardowy Amerykański Kod Wymiany Informacji.) jest to standard kodowania znaków w komputerach i w Internecie. Przypisuje kolejne liczby naturalne znakom z określonego zbioru. Podstawowy kod ASCII to 7-bitowy kod przyporządkowujący liczby z zakresu 0-127 literom (alfabetu angielskiego), cyfrom, znakom przestankowym i innym symbolom oraz poleceniom sterującym. Np. litera "a" jest zakodowana jako liczba 97, znak spacji - 32. Pierwsze 32 znaki, od 0 do 31, są znakami sterującymi, używanymi do sterowania procesem przetwarzania danych (np. znak końca wiersza); kod 127 przypisany jest klawiszowi Delete.

Litery, cyfry oraz inne znaki drukowane tworzą zbiór znaków ASCII. Jest to 95 znaków o kodach 32-126. Pozostałe 33 kody (0-31 i 127) to tzw. kody sterujące służące do sterowania urządzeniem odbierającym komunikat, np. drukarką czy terminalem.

Ponieważ kod ASCII jest 7-bitowy, a większość komputerów operuje na 8-bitowych bajtach, dodatkowy bit jest wykorzystany na powiększenie zbioru kodowanych znaków. W komputerach typu PC stosowany jest rozszerzony zestaw ASCII składający się z 256 znaków - 128 symboli tradycyjnego ASCII oraz 128 znaków dodatkowych obejmujących m.in. symbole do tworzenia ramek oraz symbole matematyczne. Wprowadzenie rozszerzonego kodu ASCII pozwoliło również na dodanie niestandardowych liter występujących w językach narodowych (np. polskie „ogonki”). Zadanie to realizowane jest przez tzw. strony kodowe, czyli rozszerzenia 128-znakowego zestawu o charakterystyczne znaki narodowe. Właśnie Rozszerzenia ASCII, wykorzystujące ósmy bit (np. norma ISO 8859, rozszerzenia firm IBM lub Microsoft), nazywane są stronami kodowymi.

**Unicode** jest to uniwersalny standard kodowania znaków mający w zamierzeniu obejmować wszystkie pisma używane na świecie. Definiują go dwa standardy - Unicode oraz ISO 10646. Znaki obu standardów są identyczne. Standardy te różnią się w drobnych kwestiach, m.in. Unicode określa sposób składu.

Unicode rozwijany jest przez konsorcjum, w którego skład wchodzi ważne firmy komputerowe, producenci oprogramowania, instytuty naukowe, agencje międzynarodowe oraz grupy zainteresowanych użytkowników. Konsorcjum współpracuje z organizacją ISO.

Standard Unicode obejmuje przydział przestrzeni numeracyjnej poszczególnym grupom znaków, nie obejmuje zaś sposobów bajtowego kodowania znaków.

Dzięki Unicode można wyświetlać znaki charakterystyczne dla różnych języków. Obejmuje języki Ameryki, Europy, Środkowego wschodu, Afryki, Indii, Azji i języki pacyficzne (ponadregionalne), ale również inne symbole historyczne czy techniczne. Unicode zamienia, konwertuje i wyświetla

wielojęzyczne teksty, symbole techniczne i matematyczne a dzięki temu rozwiązuje wiele międzynarodowych problemów językowych w systemach obliczeniowych związanych ze standardami charakterystycznymi dla konkretnych narodów. Nie wszystkie nowoczesne lub archaiczne języki są obecnie uwzględnione w standardzie

Zestaw znaków Unicode może być używany praktycznie we wszystkich znanych typach kodowania. Unicode jest modelowany podobnie jak w zestawach znaków ASCII (Amerykański standardowy kod wymiany informacji). Używa opisu literowego i liczbowej wartości dla każdego kodowanego znaku. 16 bitowa wartość jest definiowana jako liczba w systemie szesnastkowym wraz z przedrostkiem **U** na przykład , **U+0041** przedstawia **A**, a unikalna nazwa dla tej wartości to **LATIN CAPITAL LETTER A**.

### **Unicode zgodny z ASCII i ISO**

Unicode jest zgodny z kodami ASCII i jest wspierany przez wiele programów. Pierwsze 128 znaków Unicode odpowiadają kodom ASCII i mają tę samą wartość bajtu. W Unicode, przykładowo, znaki *U+0020* i *U+007E* są równoważne w ASCII znakom *0x20* i *0x7E* . Nicco inaczej jest w ASCII, który wspiera łaciński alfabet i używa siedmiobitowego zestawu znaków. Unicode używa szesnastobitowych wartości dla każdego znaku.

Cytowane za: [developer.mozilla.org](http://developer.mozilla.org). digipedia, wikipedia