

■ 文字コードとは？

文字コードとは、「文字」と「コンピュータが扱う数値」とを対応づける**ルール（規格）**のことである。

コンピュータは文字をそのまま扱えないため、**数値（バイナリ）**に変換して記録・処理する。

たとえば、**A**という文字は内部的に **65** という数値で表される。このような対応関係を定めたものが文字コードである。

■ 主な文字コードの種類

文字コード	特徴
ASCII	アメリカ発の基本的な文字コード。英数字と記号のみ（0～127）対応。
Shift_JIS	日本語（漢字）に対応した文字コード。Windowsでよく使われた。
EUC-JP	UNIX系OSで使われた日本語対応の文字コード。
UTF-8	世界中の文字を扱える Unicode の一種。Webで広く使われている。

■ Unicodeとは？

- 世界中の文字を一つのルールで表せるようにした**統一文字コード規格**。
 - 文字に一意的番号（コードポイント）を割り当てている。
 - 実際の保存や通信には**UTF-8, UTF-16, UTF-32**などの形式で使われる。
-

■ なぜ文字コードが大事ななの？

- 適切な文字コードを使わないと**文字化け**が起こる。
 - データの保存・送信・読み込み時に**文字コードを統一**することが重要。
 - たとえば、UTF-8で保存されたファイルをShift_JISで読み込むと、「**◆**」のように表示されることがある。
-

ちなみに

絵文字 (Emoji) 🍌 は、Unicode で規格化されている。2025 年 6 月現在のバージョンは version 16.0 だ。バージョンがあがると絵柄が追加される。iPhone や Android のアプリで絵文字が見えたり見えなかったりするの、どの version の規格を採用したかで変わってくる。古いアプリは最近の絵文字がわからないのだ。

また、「この絵文字にはこの意味」という定義はするが、絵柄について絶対こうでなければいけない！という取り決めはない。(世の中にはまだ白黒画面のコンピューターもあるのだ！) 同じ絵文字でも見た目が異なることがある。

<https://www.unicode.org/reports/tr51/>