

クリエイタースクールで学ぶ「JSP/ JAVAサーブレット」

NO.01 : サーブレット (日本語処理)

- 1-1 文字コード -

文字コードとは、文字を数字で表したものです。
コンピュータは文字を扱う際、この文字コードを使用します。
日本語処理で用いられる文字コードは、主に JIS SJIS EUCです。
以下に各文字コードについて解説します。

- JIS -

日本工業規格で定められています。
英数字と半角カナを 1バイト、漢字を 2バイトで表します。



バイトとは、情報量の単位のことです。

- SJIS -

Microsoft社が策定した文字コードです。
WindowsやMacintoshでデフォルトで使用されています。

- EUC -

AT&T社がUNIXのために策定した文字コードです。
UNIXとは、AT&T社で開発されたOSのことで、
インターネット上の多くのWWWサーバはUNIXを採用しています。

- Unicode -

Unicodeは、すべての文字が 2バイトで表され、
全世界で使用される主要な文字を集めた文字集合です。
Javaはプログラム内での内部処理を Unicodeベースで行ないます。