

概要

文字名とは、ISO/IEC 10646 ならびに Unicode において各文字 (符号位置) に与えられた名前です。ラテン大文字で付けられています。

例えば、ラテン文字の大文字 A (U+0041) に対しては、「LATIN CAPITAL LETTER A」という名前が付けられています。平仮名の「あ」(U+3042) は「HIRAGANA LETTER A」です。

記号類や空白、制御コードの類にも与えられています。「。」(U+0046) の文字名は「FULL STOP」。ASCII のハイフンにもマイナスにも使われる「-」(U+002D) は「HYPHEN-MINUS」(ハイフンマイナス) です。

漢字については、U+4E00 にある漢字「一」の文字名が「CJK UNIFIED IDEOGRAPH-4E00」となっているように、符号位置によって識別するようになっています。

規格間の対応付け

文字名は単に Unicode の符号位置の別名であるだけでなく、コード変換表の役割も果たしています。

1990 年代以降、ISO/IEC 10646 以外の文字コード規格も、各文字について上記のような文字名を与えるようになっています。2000 年に制定された JIS X 0213 はもちろんの事、10646 以前に制定された文字コード規格、例えば JIS X 0201 や JIS X 0208 は、1990 年代の改正時に文字名を定義しています。国際標準の ISO/IEC 646 や ISO/IEC 8859-1 も各文字に 10646 に対応する文字名を定義しています。

例えば、JIS X 0201 はラテン文字集合の符号位置 4/1 のラテン大文字 A について文字名「LATIN CAPITAL LETTER A」であると規定しています。これは Unicode の符号位置 U+0041 にあたります。また、JIS X 0208 は区点位置 1-29 にあるダッシュ記号について、文字名「EM DASH」と規定しています。Unicode の符号位置 U+2014 にあたります。

これにより、既存の文字コード規格と Unicode ならびに ISO/IEC 10646 との間で文字の対応付けがなされています。

関連項目

- ・ ISO/IEC 10646
- ・ Unicode
- ・ 代替名称

参考

- ・ JIS X 0213 のコード対応表 - 当サイト。JIS と Unicode の対応する符号位置に、その文字名も記されています。