

11. インターネットを流れる情報 (インターネットと文字コード)

[11.1] 電子メール

<メール使用文字の大原則>  
7ビットのASCIIコードを使用する

<文字コードの種類>  
1バイトコード(半角文字)  
・JISコード  
・ASCIIコード  
・EBCDICコード

2バイト(漢字)コード  
・JIS漢字コード  
・シフトJISコード  
・EUCコード(拡張UNIXコード)  
・Unicode

	JIS	シフトJIS	EUC	Unicode
漢	3441	8ABF	B4C1	6F22
字	3B7A	8E9A	BBFA	5B57

漢字(2バイトコード)と半角英数字(1バイトコード)が混在する例

3441 「漢」 34 「4」 41 「A」

漢 字 と A B C  
3441 3B7A 2448 41 42 43 JIS  
8ABF 8E9A 82C6 41 42 43 シフトJIS

日本語を利用するためのしくみ

文字の符号化方式「ISO-2022-JP」 JIS漢字コードを使用する

JIS漢字コードの開始点 1B 24 42 \$B

終了点 1B 28 42 (B

(例)

こ ん に ち は \$B\$3\$s\$K\$A\$0 (B  
2433 2473 244B 2441 244F

添付ファイル送受信のしくみ base64

メールのヘッダ Content-Transfer-Encoding: base64  
(変換符号化方式)

文字情報(テキスト形式)でないデータをテキスト形式に変換する方法

8ビット×3=24ビット      6ビット×4に分割      頭に00をつけて8ビットにする

例      1111 1111 1111 1111 1111 1111      (真っ白な写真)

111111 111111 111111 111111

00111111 00111111 00111111 00111111  
(63)      (63)      (63)      (63)

6ビットでは0~63の数字に変換できる      base64

<base64 対応表>

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
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
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
26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51
0	1	2	3	4	5	6	7	8	9	+	/														
52	53	54	55	56	57	58	59	60	61	62	63														

## [11.2] ウェブの文字コード

JIS 漢字、シフト JIS、EUC、Unicode いずれも OK

<ルール>

ウェブページの先頭で文字コードを指定する

<meta http-equiv="Content-Type" content="text/html; charset=Shift\_JIS">

ブラウザによる自動判別

Internet Explorer      「表示」 - 「エンコード」  
Netscape Navigator      「表示」 - 「文字コードセット」

- ・日本語(自動選択)
- ・日本語(ShiftJIS)
- ・日本語(EUC)
- ・Unicode . . .

Unicode への対応