

Allegato 4

Set di caratteri ammessi per la costituzione di un alias

Premessa

Nonostante i caratteri ammessi da AGCOM per la costituzione dell'Alias siano quelli sotto elencati, è bene sottolineare come non tutti siano invece supportati dalla rete GSM e dunque da operatori di telefonia mobile e device di destinazione (telefoni cellulari). In alcuni casi è infatti possibile che i caratteri meno comuni (simboli internet, simboli di valuta, simboli matematici, etc) vengano sostituiti o troncati dagli operatori. Consigliamo pertanto di utilizzare per il proprio Alias:

- numeri dallo 0 al 9
- lettere non accentate dalla a alla z, sia maiuscole che minuscole
- spazio

Indicazioni AGCOM per la costituzione di un Alias

Con riferimento al paragrafo 6.2.1 “GSM 7 bit Default Alphabet” dello standard tecnico “Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Alphabets and language-specific information”, 3GPP TS 23.038 version 11.0.0 (2012-10) Release 11, per la costituzione degli alias è ammissibile solo quanto segue.

1. le lettere dell'alfabeto internazionale minuscole e maiuscole:
 - [ABCDEFGHIJKLMNOPQRSTUVWXYZ (codici HEX rispettivamente da 41 a 5A)
 - [abcdefghijklmnopqrstuvwxyz (codici HEX rispettivamente da 61 a 6A)
2. Le lettere minuscole accentate presenti nella tastiera italiana:
 - [èèùìò (codici HEX rispettivamente da 04 a 08)
 - [à (codice HEX 7F)
3. Le cifre da 0 a 9
 - [0123456789 (codici HEX rispettivamente da 30 a 39)
4. Comuni segni di punteggiatura:
 - [SP (spazio: codice HEX 20)
 - [! (punto esclamativo: codice HEX 21)

- [‘ (apostrofo: codice HEX 27)
- [, (virgola: codice HEX 2C)
- [. (punto: codice HEX 2E)
- [: (due punti: codice HEX 3A)
- [; (punto e virgola: codice HEX 3B)
- [? (punto interrogativo: codice HEX 3F)

I precedenti caratteri non possono essere preceduti dal carattere spazio.

Non è consentito l’uso consecutivo di spazi

- [“ (virgolette, codice HEX 22)

In un Alias, possono essere presenti esclusivamente due virgolette: una come apertura ed una come chiusura. La prima non può precedere un spazio e la seconda non può seguire uno spazio.

5. Comuni simboli di valuta

- [€ (euro: codice di due caratteri HEX 1B 65)
- [£ (lira: codice HEX 01)
- [\$ (dollaro: codice HEX 02)

6. Comuni simboli matematici

- [% (percentuale: codice HEX 25)
- [((parentesi tonda aperta: codice HEX 28)
- [) (parentesi tonda chiusa: codice HEX 29)
- [+ (più: codice HEX 2B)
- [- (meno o anche trattino: codice HEX 2D)
- [= (uguale: codice HEX 3D)

7. Simboli utilizzati in internet:

- [@ (chiocciolina o “at” : codice HEX 00)
- [_ (sottolineato o “underscore” : codice HEX 11)
- [# (cancelletto o “hash” : codice HEX 23)
- [& (and: codice HEX 26)
- [* (asterisco o “star” : codice HEX 2A)

In definitiva, la lista dei caratteri ammessi e le relative codifiche in esadecimale ETSI da utilizzare nella trasmissione degli SMS/MMS nonché le relative codifiche in esadecimale UTF-8 da utilizzare nella comunicazione verso la banca dati dell’Autorità sono:

Carattere	Codifica ETSI	Codifica UTF-8	Carattere	Codifica ETSI	Codifica UTF-8	Carattere	Codifica ETSI	Codifica UTF-8
@	00	40	8	38	38	Z	5A	5A
£	01	C2 A3	9	39	39	a	61	61

\$	02	24	:	3A	3A	b	62	62
è	04	C3 A8	;	3B	3B	c	63	63
é	05	C3 A9	=	3D	3D	d	64	64
ù	06	C3 B9	?	3F	3F	e	65	65
ì	07	C3 AC	A	41	41	f	66	66
ò	08	C3 B2	B	42	42	g	67	67
_	11	5F	C	43	43	h	68	68
SP	20	20	D	44	44	i	69	69
!	21	21	E	45	45	j	6A	6A
"	22	22	F	46	46	k	6B	6B
#	23	23	G	47	47	l	6C	6C
%	25	25	H	48	48	m	6D	6D
&	26	26	I	49	49	n	6E	6E
'	27	27	J	4A	4A	o	6F	6F
(28	28	K	4B	4B	p	70	70
)	29	29	L	4C	4C	q	71	71
*	2A	2A	M	4D	4D	r	72	72
+	2B	2B	N	4E	4E	s	73	73
,	2C	2C	O	4F	4F	t	74	74
-	2D	2D	P	50	50	u	75	75
.	2E	2E	Q	51	51	v	76	76
0	30	30	R	52	52	w	77	77
1	31	31	S	53	53	x	78	78
2	32	32	T	54	54	y	79	79
3	33	33	U	55	55	z	7A	7A
4	34	34	V	56	56	à	7F	C3 A0
5	35	35	W	57	57	€	1B 65	E2 82 AC
6	36	36	X	58	58			
7	37	37	Y	59	59			

Di conseguenza, la tabella dei caratteri utilizzabili organizzata secondo la codifica ETSI è la seguente.

				b7	0	0	0	0	1	1	1	1
				b6	0	0	1	1	0	0	1	1
				b5	0	1	0	1	0	1	0	1
b4	b3	b2	b1	HE X	0	1	2	3	4	5	6	7

0	0	0	0	0	@ 40		SP 20	0 30		P 50		p 70
0	0	0	1	1	£ C2 A3	- 5F	! 21	1 31	A 41	Q 51	a 61	q 71
0	0	1	0	2	\$ 24		" 22	2 32	B 42	R 52	b 62	r 72
0	0	1	1	3			# 23	3 33	C 43	S 53	c 63	s 73
0	1	0	0	4	è C3 A8			4 34	D 44	T 54	d 64	t 74
0	1	0	1	5	è C3 A9		% 25	5 35	E 45	U 55	e 65	u 75
0	1	1	0	6	ù C3 B9		& 26	6 36	F 46	V 56	f 66	v 76
0	1	1	1	7	ì C3 AC		' 27	7 37	G 47	W 57	g 67	w 77
1	0	0	0	8	ò C3 B2		(28	8 38	H 48	X 58	h 68	x 78
1	0	0	1	9) 29	9 39	I 49	Y 59	i 69	y 79
1	0	1	0	A			* 2A	: 3A	J 4A	Z 5A	j 6A	z 7A
1	0	1	1	B		1)	+ 2B	; 3B	K 4B		k 6B	
1	1	0	0	C			, 2C		L 4C		l 6C	
1	1	0	1	D			- 2D	= 3D	M 4D		m 6D	
1	1	1	0	E			. 2E		N 4E		n 6E	
1	1	1	1	F				? 3F	O 4F		o 6F	à C3 A0

Note:

- SP corrisponde al carattere spazio;

- 1) non è un carattere ma indica il codice (HEX 1B) da anteporre per indicare i caratteri presenti nella Extension table. In particolare alla codifica HEX 1B 65 corrisponde il carattere €, la cui codifica UTF-8 è E2 82 AC. Il carattere € è l'unico carattere utilizzabile della Extension table;

- In ciascuna cella è riportata in prima riga il carattere relativo alla codifica secondo lo standard 3GPP TS 23.038 version 11.0.0 (2012-10);

- In seconda riga è riportata la relativa codifica UTF-8 da utilizzare nelle comunicazioni verso il DB dell'AGCOM;