

Symbols in subject lines

An in-depth look at symbols



An Experian CheetahMail white paper

Symbols in subject lines

What is the advantage of using symbols in subject lines?

The age of personal emails has changed significantly due to the social media boom, and instead, people are receiving more and more promotional emails in their inboxes. Concurrently, more and more email clients are using engagement as a way of determining whether a message is spam or not. That's why it's vital to keep your subscribers engaged, enticing them to open your emails (and hopefully click within them).

So, how can your email stand out from the pack? Using symbols in your subject lines is a great way to draw in the eye and vary the messaging you regularly deploy to your subscribers. Studies show that symbols help to keep their interest up and to spark interest in those who were previously unengaged.

What is a symbol?

Every font set (like Helvetica, Arial, and Times New Roman) is made entirely out of "symbols." Each letter, number and punctuation mark you're reading right now is a "symbol" in that font set. Many font sets include extra symbols beyond letters, numbers and punctuation. These special symbols, like hearts and stars, are based on the Unicode computer industry standard character set, and are different from images. This Unicode data is encoded a number of different ways for your computer to store and understand it, and then in turn rendered on your screen using a font to show you the stylized end product.

What is Unicode?

In the most simple terms, Unicode is a collection of the thousands upon thousands of different characters used to represent different languages in their written forms, as well as symbols and mathematical symbols.

Unicode — the long version

Unicode was an effort developed in the late 1980s (and released in the early 1990s) to codify characters beyond the original and limited characters of ASCII. Containing more than 240,000 assigned characters as of its latest iteration (6.1.0, January 2012) and organized into character "blocks," Unicode is an evolving collection of all of the characters used to write various languages around the world. It also contains other characters, such as symbols, arrows and mathematical operators.

In Unicode, each character is mapped to a "code point," which looks something like U+0041, which corresponds to the English alphabet capital letter A. The word "Hello," for example, would be represented in Unicode as the following code points:

"Hello" in Unicode:

U+0048 U+0065 U+006C U+006C U+006F

Symbols in subject lines

Since there could be multiple ways to store this data (in complete form, above, or without the “U+” or even without the “U+00” sections of the code), encodings are used by computers to store and understand Unicode characters in different ways. ASCII and UTF-8 are two methods used to encode this character data for consumption by computers. UTF-8 is one of the more common methods.

ASCII

ASCII, the American Standard Code for Information Interchange, is a computing standard that predates Unicode and was created in the 1960s as a way to standardize the encoding of text characters in computing. It is a table of 128 printable and “non-printable” characters created using 8 bits, including A through Z in both upper and lower cases as well as 0-9 and common punctuation like periods and question marks. The “non-printable” characters are control characters which instructed computers to do things like beep or make carriage returns or line breaks. One problem with ASCII is that it did not contain a uniform way to encode characters other than those used in standard American English, making the rest of the 255 total characters available in 8-bit format an international free-for-all without a uniform standard. This created a problem in that, beyond the first 128 characters, there was no one set code for what to do after the 128th character. Different countries and organizations defined these extra characters however they saw fit, making exchange between computers unstandardized.

UTF-8

UTF-8, or Universal Character Set Transformation Format—8 bit, was created in the 1980s and 90s as a more complete way to render characters beyond the 95 printable characters in ASCII. The first 128 characters in UTF-8 represent the 128 printable and non-printable ASCII characters and enable backward compatibility with ASCII. In total, UTF-8 encoding makes 2^{31} (or 2,147,483,648) unique characters possible, though not all are assigned or used.

Symbols in subject lines

The original 128-character ASCII Chart

Dec	Hx	Oct	Char / Description	Dec	Hx	Oct	Html	Char / Description
0	00	000	NUL Null char	32	20	040	 	Space
1	01	001	SOH Start of Heading	33	21	041	!	! Exclamation mark
2	02	002	STX Start of Text	34	22	042	"	" Double quotes
3	03	003	ETX End of Text	35	23	043	#	# Number
4	04	004	EOT End of Transmission	36	24	044	$	\$ Dollar
5	05	005	ENQ Enquiry	37	25	045	%	Procenttecken
6	06	006	ACK Acknowledgment	38	26	046	&	& Ampersand
7	07	007	BEL Bell	39	27	047	'	' Single quote
8	08	010	BS Back Space	40	28	050	((Open parenthesis
9	09	011	HT Horizontal Tab	41	29	051)) Close parenthesis
10	0A	012	LF Line Feed	42	2A	052	*	* Asterisk
11	0B	013	VT Vertical Tab	43	2B	053	+	+ Plus
12	0C	014	FF Form Feed	44	2C	054	,	, Comma
13	0D	015	CR Carriage Return	45	2D	055	-	- Hyphen
14	0E	016	SO Shift Out / X-On	46	2E	056	.	. Period, full stop
15	0F	017	SI Shift In / X-Off	47	2F	057	/	/ Slash or divide
16	10	020	DLE Data Line Escape	48	30	060	0	0 Zero
17	11	021	DC1 Device Control 1	49	31	061	1	1 One
18	12	022	DC2 Device Control 2	50	32	062	2	2 Two
19	13	023	DC3 Device Control 3	51	33	063	3	3 Three
20	14	024	DC4 Device Control 4	52	34	064	4	4 Four
21	15	025	NAK Negative acknowledge	53	35	065	5	5 Five
22	16	026	SYN Synchronous Idle	54	36	066	6	6 Six
23	17	027	ETB End of Trans. Block	55	37	067	7	7 Seven
24	18	030	CAN Cancel	56	38	070	8	8 Eight
25	19	031	EM End of Medium	57	39	071	9	9 Nine
26	1A	032	SUB Substitute	58	3A	072	:	: Colon
27	1B	033	ESC Escape	59	3B	073	;	; Semicolon
28	1C	034	FS File Separator	60	3C	074	<	< Less than
29	1D	035	GS Group Separator	61	3D	075	=	= Equals
30	1E	036	RS Record Separator	62	3E	076	>	> Greater than
31	1F	037	US Unit Separator	63	3F	077	?	? Question mark

(continued...)

Symbols in subject lines

Dec	Hx	Oct	Char / Description	Dec	Hx	Oct	Html	Char / Description
64	40	100	@ At symbol	96	60	140	`	´ Grave accent
65	41	101	A Uppercase A	97	61	141	a	a Lowercase a
66	42	102	B Uppercase B	98	62	142	b	b Lowercase b
67	43	103	C Uppercase C	99	63	143	c	c Lowercase c
68	44	104	D Uppercase D	100	64	144	d	d Lowercase d
69	45	105	E Uppercase E	101	65	145	e	e Lowercase e
70	46	106	F Uppercase F	102	66	146	f	f Lowercase f
71	47	107	G Uppercase G	103	67	147	g	g Lowercase g
72	48	110	H Uppercase H	104	68	150	h	h Lowercase h
73	49	111	I Uppercase I	105	69	151	i	i Lowercase i
74	4A	112	J Uppercase J	106	6A	152	j	j Lowercase j
75	4B	113	K Uppercase K	107	6B	153	k	k Lowercase k
76	4C	114	L Uppercase L	108	6C	154	l	l Lowercase l
77	4D	115	M Uppercase M	109	6D	155	m	m Lowercase m
78	4E	116	N Uppercase N	110	6E	156	n	n Lowercase n
79	4F	117	O Uppercase O	111	6F	157	o	o Lowercase o
80	50	120	P Uppercase P	112	70	160	p	p Lowercase p
81	51	121	Q Uppercase Q	113	71	161	q	q Lowercase q
82	52	122	R Uppercase R	114	72	162	r	r Lowercase r
83	53	123	S Uppercase S	115	73	163	s	s Lowercase s
84	54	124	T Uppercase T	116	74	164	t	t Lowercase t
85	55	125	U Uppercase U	117	75	165	u	u Lowercase u
86	56	126	V Uppercase V	118	76	166	v	v Lowercase v
87	57	127	W Uppercase W	119	77	167	w	w Lowercase w
88	58	130	X Uppercase X	120	78	170	x	x Lowercase x
89	59	131	Y Uppercase Y	121	79	171	y	y Lowercase y
90	5A	132	Z Uppercase Z	122	7A	172	z	z Lowercase z
91	5B	133	[Opening bracket	123	7B	173	{	{ Opening brace
92	5C	134	\ Backslash	124	7C	174	|	Vertical bar
93	5D	135] Closing bracket	125	7D	175	}	} Closing brace
94	5E	136	^ Caret - circumflex	126	7E	176	~	~ Tilde
95	5F	137	_ Underscore	127	7F	177		DEL Delete

Source: www.LookupTables.com

What does this have to do with symbols in subject lines?

Email marketers often want complete listings of symbols that can be used in subject lines. As you've read, however, there are literally hundreds of thousands of possible Unicode characters that exist in computing. Each one of these characters possibly can be used as a symbol in a subject line. Some clever marketers even use them in combination, for example ☀️ ||||| ☀️ was used in an email for Jeep®, representing the front of a Jeep with the headlights and seven-slatted grill.

On top of the thousands of possible symbols and combinations, each character may or may not exist in the font and/or encoding method used by each browser and email client. For example, while a black sun with rays might exist in the font and encoding set used by Gmail™ viewed through the Chrome™ browser, it might not exist in the font and encoding set used by a BlackBerry®. Since there are so many symbols and characters, not every font set has a character that corresponds with each Unicode character, and some Unicode characters might not have corresponding code in UTF-8 (or whatever encoding method is used).

Rendering

Unicode is only a list of list character “concepts.” That is to say, the characters **A** and **Λ** and **Ა** and **Ბ** and **Გ** all are U+0041, or the conceptual capital letter “A” in the standard English or Western European alphabet. Whether each rendition of this concept of the capital letter “A” renders properly is dependent upon the font set and encoding method used. A heart symbol for instance (♥️) might be present when you use UTF-8 and the font Arial, but it doesn't exist in a standard 128-character ASCII table and might not exist in certain font sets.

What symbols can I use in my emails?

As you've probably gathered, there are thousands of characters or symbols that possibly can be used in subject lines, which is the main reason no complete “master list of email symbols” exists. You can get a good idea of what's in common Unicode Blocks at <http://www.charbase.com/block> and use the “my browser” view to identify symbols that are likely to work in an email subject line.

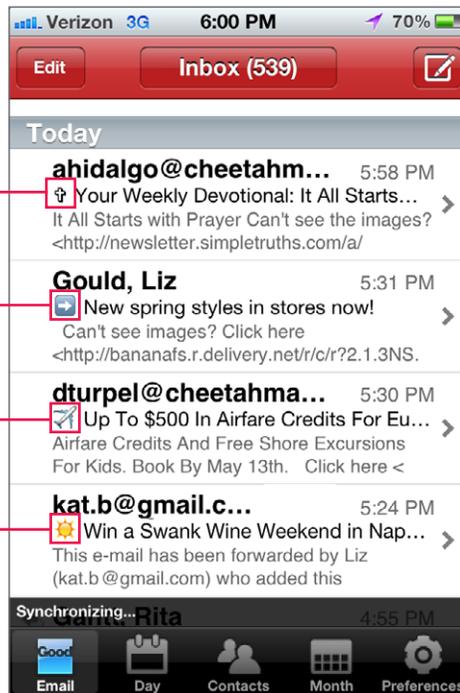
Symbols in subject lines

Tips, tricks and considerations

When is a symbol not a symbol?

There are a couple of instances in which symbols are actually converted from their character form into an image. As of the writing of this paper, both the iPhone® and Hotmail® render a few different symbols as actual images instead of characters.

Certain Unicode symbols, such as ↕ → → and ✨ pictured right, convert from character form to an image on iPhones and Hotmail.



The best way to know if a symbol will render is to test it.

Since email clients use different font sets and understand different encoding methods, there are millions of possible permutations of symbols that will render and not render. Sometimes email clients even use different fonts or encoding methods in different browsers or browser versions. Since there are so many possibilities, the best way to know if your subscribers will see the symbol properly is to send tests where your subscribers are and find out for yourself. At the bare minimum, you should be testing your most important groups of subscribers. If you know most of your subscribers read their emails in Gmail or Hotmail or in an iPhone (perhaps via Litmus reporting), at least view tests in those places to make sure most of your subscribers won't see the dreaded "empty rectangle" or question mark.

Symbols in subject lines

Symbols should make sense

A heart in a subject line about a credit card balance might not make sense. The same subject line rules still apply when using symbols — the more relevant they are, the more your subscribers will understand and enjoy them.

Keep it fresh

If you test using symbols in your subject lines and see great results at first that taper off, perhaps the “newness” has worn off. Try testing new symbols or testing symbols against no symbols.

Do symbols affect deliverability?

We’re often asked if symbols affect your chance at landing in the inbox. From what we’ve seen, the answer is no. We’ve never seen a client’s deliverability affected by symbols, which is likely due to ISPs’ current focus on using engagement data to determine whether a message is spam or not.

If symbols work well for your brand and garner higher open rates, then they could, in fact, improve your deliverability. When your customers are more engaged with your emails, many ISPs calculate a lower risk that your emails are spam.

Results

We’ve seen mixed-to-positive results when using symbols in subject lines:

- Subject lines with symbols had a higher unique open rate in 56 percent of brands we analyzed
- The black heart (♥) is the most popular symbol, but provided only a modest open rate lift
- The greatest unique open rate lifts we’ve seen were in subject lines including umbrellas (☂) and airplanes (✈)
- The most popular symbols we’ve seen (in order of popularity) are:

♥ ★ ✨ 🎵 ☀️ 🌸 ☆ 🤍 ➡️ 😊 🖤 ✈️ 🕂 → ☂

Symbols in subject lines

Resources

- A great article by Justine Jordan of Litmus on the efficacy of symbols in subject lines:
<http://litmus.com/blog/do-symbols-in-subject-lines-increase-performance>
- A compendium of Unicode Blocks, including an image representation and “your browser” view:
<http://www.charbase.com/block>

Experian CheetahMail
29 Broadway, 6th Floor
New York, NY 10006
1 855 789 8898
www.experian.com/cheetahmail



Marketing
forward 

© 2012 Experian Information Solutions, Inc. • All rights reserved

Experian and the Experian marks used herein are service marks or registered trademarks of Experian Information Solutions, Inc.

Other product and company names mentioned herein are the property of their respective owners.

08/2012