



链滴

HTML 转义的替代方案 --Unicode Small Form Variants

作者: [linker](#)

原文链接: <https://ld246.com/article/1526279822893>

来源网站: [链滴](#)

许可协议: [署名-相同方式共享 4.0 国际 \(CC BY-SA 4.0\)](#)

<p>为了预防 XSS 攻击,防止用户输入中包含非法内容,URL 传递特殊符号等,人们经常需要对用户输入做转义.

但是 Unicode 提供一个更简单的方法: 一套 Small Form Variants,可以替换常的标点符号,在人类看来是一样的,但是确不同于普通的符号.

Position Decimal Name Appearance

0xFE50 65104 SMALL COMMA ,

0xFE51 65105 SMALL IDEOGRAPHIC COMMA 、

0xFE52 65106 SMALL FULL STOP .

0xFE54 65108 SMALL SEMICOLON ;

0xFE55 65109 SMALL COLON :

0xFE56 65110 SMALL QUESTION MARK ?

0xFE57 65111 SMALL EXCLAMATION MARK !

0xFE58 65112 SMALL EM DASH -

0xFE59 65113 SMALL LEFT PARENTHESIS (

0xFE5A 65114 SMALL RIGHT PARENTHESIS)

0xFE5B 65115 SMALL LEFT CURLY BRACKET {

0xFE5C 65116 SMALL RIGHT CURLY BRACKET }

0xFE5D 65117 SMALL LEFT TORTOISE SHELL BRACKET ⸮

0xFE5E 65118 SMALL RIGHT TORTOISE SHELL BRACKET ⸯ

0xFE5F 65119 SMALL NUMBER SIGN #

0xFE60 65120 SMALL AMPERSAND &

0xFE61 65121 SMALL ASTERISK *

0xFE62 65122 SMALL PLUS SIGN +

0xFE63 65123 SMALL HYPHEN-MINUS -

0xFE64 65124 SMALL LESS-THAN SIGN <

0xFE65 65125 SMALL GREATER-THAN SIGN >

0xFE66 65126 SMALL EQUALS SIGN =

0xFE68 65128 SMALL REVERSE SOLIDUS \

0xFE69 65129 SMALL DOLLAR SIGN \$

0xFE6A 65130 SMALL PERCENT SIGN %

0xFE6B 65131 SMALL COMMERCIAL AT @ </p>

<p>相比传统的方案,替换成 Unicode Small Form Variants 可以避免到处 escape/un-escape .</p>