

# 为什么要对 URI 进行编码

作者: [lucianolixin](#)

原文链接: <https://ld246.com/article/1569166121649>

来源网站: [链滴](#)

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



1. URL 统一资源定位符 (union resource locator) 表示资源的位置，期望提供查找资源的方法。
2. URI 统一资源标识符 (union resource identifier) 用以区分资源，是URL的超集。

## 为什么要进行URI的编码

- 传递数据中，如果存在用作分隔符的保留字
- 对可能产生歧义性的数据编码
- 不在ASCII码范围内的字符
- ASCII码中不可显示的字符
- URI中规定的保留字符
- 主要的保留字符 ":" , "/" , "?" , "#" , "[" , "]" , "@"
- 非主要保留字符 "!" , "\$" , "&" , "'" , "(" , ")" , "\*" , "+" , "," , ";" , "="
- 不安全的字符（传输环节中可能会被不正确处理），如空格、引号、尖括号。

## URL 百分号编码

- 百分号编码的方式
- pct-encoded = "%" HEXDIG HEXDIG (百分号+两个十六进制数)
- 对于HEXDIG十六进制中的字符，大小写等价
- 非ASCII码字符（中文）：建议先UTF8编码，再US-ASCII编码
- 对URI合法字符，编码与不编码等价

不同的网络开发库的URI编码的方式并不完全相同，对不同不安全字符的处理侧率不相同。