

Title: Codificación UTF16 y UTF8

### Keyword

UTF-16  
Unicode  
Codificación  
Pares  
Multilingüe  
Caracter

Topic: Codificación UTF16

Notes: UTF16 es un formato de codificación de caracteres que permite representar todos los puntos de código válidos de Unicode, que son más de un millón. Utiliza una longitud variable, ya que los caracteres se codifican con una o dos unidades de 16 bits. Los caracteres más comunes, como los del Plano Multilingüe Básico (BMP), se representan con una sola unidad de 16 bits, mientras que los caracteres menos comunes, requieren pares sustitutos.

### Questions

¿Qué es UTF-16?

¿Cuáles ventajas y desventajas presenta?

¿En qué entornos de programación se utiliza comúnmente?

Una de las ventajas de UTF-16 es su capacidad para manejar una amplia gama de caracteres con un compromiso razonable entre espacio y rendimiento. Sin embargo, su naturaleza de longitud variable puede introducir complejidad en el manejo de cadenas y errores en software que no está diseñado para manejar adecuadamente los pares sustitutos.

En comparación con UTF-8, es menos eficiente para textos que contienen principalmente caracteres ASCII, ya que utiliza al menos dos bytes por carácter. Sin embargo, es más eficiente para textos que incluyen caracteres de idiomas asiáticos o símbolos que requieren más de un byte en UTF-8.

Summary: UTF-16 es un formato de codificación Unicode de longitud variable que utiliza una o dos unidades de 16 bits para representar caracteres. Es ideal para manejar una amplia gama de caracteres, aunque su uso es menos común en la web debido a la popularidad de UTF-8. Es ampliamente utilizado en sistemas operativos y entornos de programación.



# Title: Codificación UTF16 y UTF8

## Keyword

UTF-8  
Unicode  
Codificación  
ASCII  
Web  
Bytes

## Topic: Codificación UTF8

**Notes:** UTF-8 es un estándar de codificación de caracteres que permite representar cualquier carácter Unicode. Esto significa que puede codificar texto en cualquier idioma del mundo, así como una amplia gama de símbolos. UTF-8 es una codificación de longitud variable, lo que significa que utiliza entre 1 y 4 bytes para representar un carácter. Los caracteres ASCII se representan con 1 byte, lo que hace que UTF-8 sea compatible con ASCII.

## Questions

¿Qué es UTF-8 y por qué es importante?

Una de las principales ventajas de UTF-8 es su compatibilidad con versiones anteriores. Debido a que los primeros 128 caracteres de UTF-8 son idénticos a ASCII, el texto ASCII simple también es válido UTF-8, facilitando la transición de sistemas antiguos a Unicode. Eficiencia en almacenamiento.

¿Cuáles son sus ventajas sobre otras codificaciones?

UTF-8 se ha convertido en la codificación dominante en la 'WWW'. Según W3Techs, más del 98% de los sitios web utilizan UTF-8, por su capacidad de compatibilidad con la mayoría de los sistemas operativos y aplicaciones.

## Summary:

UTF-8 es un estándar de codificación de caracteres que permite representar cualquier carácter Unicode. Es una codificación de longitud variable que utiliza entre 1 y 4 bytes para representar un carácter. UTF-8 es compatible con ASCII facilitando la transición de sistemas antiguos a Unicode. Es la codificación dominante en la "WWW" por su eficiencia en el uso del espacio.