

# HTML, XHTML, XML y códigos de control

audiencia de destino: XHTML/HTML coders (using editors or scripting), script developers (PHP, JSP, etc.)

## pregunta

¿Cómo manejo los códigos de control (es decir, la extensión 'C0' U+0000-U+001F y 'C1' U+007F-U+009F) en XML, XHTML y HTML?

[\[Pasar directamente a la respuesta\]](#)

## información preliminar

A veces, las aplicaciones preexistentes crean datos e incorporan códigos de control. Por lo tanto, a veces es importante comprender de qué manera son compatibles los controles en lenguajes de etiquetas, al migrar estas aplicaciones o sus datos a la Web.

Existen dos extensiones de juego de caracteres Unicode que están asignadas como códigos de control. El estándar Unicode no utiliza estos controles en particular y hace que su definición dependa de la aplicación. Si la aplicación no especifica su utilización, deberán ser interpretados conforme a la semántica de ISO/IEC 6429. La mayoría reconocerá muchos de los 6429 controles: ACK, NAK, BEL, LF, FF, VT, CR y otros. La familia ISO 8859 y otros estándares de caracteres basan sus códigos de control en el estándar ISO 6429.

Los códigos de control de la extensión U+0000-U+001F se denominan **extensión "C0"**. Esta extensión comienza con el control NUL (Nulo) U+0000. Los códigos de control de la extensión U+0080-U+009F se denominan **extensión "C1"**. DEL (Eliminar) U+007F también es un control y se encuentra junto al comienzo de la Extensión C1.

## respuesta



[Deutsch](#) [English](#) [Polski](#)  
[Português-BR](#) [Română](#)  
[Русский](#) [Українська](#)

Este documento es una traducción. En caso de discrepancias o errores, la única versión normativa es el último original en inglés. Los derechos de autor originales corresponden al W3C, como puede verse al final de la página.

Traductor: [Spanish Translation Team, Spanish Translation US.](#)



[Inicio](#) > [Recursos](#) > [Artículos](#)

## en esta página

[Pregunta](#)

[Información preliminar](#)

[Respuesta](#)

[Manejo de códigos de control](#)

[Compatibilidad con códigos de control](#)

[A propósito](#)

[Lecturas complementarias](#)

## Manejo de códigos de control

Los códigos de control se deben reemplazar con el [etiquetado apropiado](#). Debido a que XML proporciona una manera estándar de codificar los datos estructurados, la representación de los códigos de control en lugar del etiquetado además revertiría las ventajas reales derivadas de la utilización de XML. La utilización de los códigos de control en HTML y XHTML nunca es apropiada, debido a que estos idiomas de etiquetado se utilizan para representar texto y no datos. La única vez que se podría necesitar la siguiente información sería en el hipotético caso en el que no se pudieran limpiar los datos preexistentes que contengan códigos de control.

Si los datos no son realmente textuales y son binarios, sería más práctico codificarlos, por ejemplo, mediante base64 o como valores hexadecimales, a fin de garantizar que sólo los caracteres compatibles se utilizan en el texto del idioma de etiquetado. (Y, por supuesto, mediante la decodificación del texto cuando se leen los archivos). Se debe tener en cuenta que el esquema XML proporciona tipos de datos para estas codificaciones.

Otra alternativa es almacenar los datos en un documento externo y hacer referencia a ellos desde el documento XML.

En XML 1.1, si es necesario representar el código de control de manera explícita, la alternativa más sencilla es utilizar una NCR (referencia de caracteres numéricos). Por ejemplo, el código de control ESC (Escape) U+001B estaría representado por las referencias de caracteres numéricos `&#x1B;` (hexadecimal) o `&#27;` (decimal).

## Compatibilidad con códigos de control

La siguiente tabla resume los idiomas de etiquetado que son compatibles con los códigos de control:

| Controles              | Extensión              | HTML 4     | XHTML 1.0  | XML 1.0     | XML 1.1    |
|------------------------|------------------------|------------|------------|-------------|------------|
| C0, excepto HT, LF, CR | U+0000 (NUL)           | Ilegal     | Ilegal     | Ilegal      | Ilegal     |
|                        | U+0001-U+001F          | Ilegal     | Ilegal     | Ilegal      | NCR        |
| HT, LF, CR             | U+0009, U+000A, U+000D | Compatible | Compatible | Compatible  | Compatible |
| DEL + C1               | U+007F-U+009F          | Ilegal     | Ilegal     | Compatible  | NCR        |
| NEL                    | U+0085                 | Ilegal     | Ilegal     | (permitido) | Compatible |

- El control NUL (Nulo) es ilegal y, por lo tanto, no puede estar representado por NCR ni codificado directamente en idiomas de etiquetado.
- HTML, XHTML y XML 1.0 no son compatibles con la extensión C0, excepto el caso de HT (Tabulación horizontal)

U+0009, LF (Salto de línea) U+000A y CR (Retorno de carro) U+000D. Es compatible con la extensión C1; es decir, los controles se pueden codificar directamente o se pueden representar como NCR (referencias de caracteres numéricos).

- XML 1.1 restringe la extensión C1, excepto en el caso de NEL U+0085 (la nueva línea de EBCDIC), además de la extensión C0. Sin embargo, XML 1.1 permite que los controles estén representados por NCR (referencias de caracteres numéricos).

## a propósito

Mientras que la familia ISO 8859 reserva la extensión C1 para los controles, los juegos de caracteres de Microsoft (por ejemplo, 1250-1258) colocan los caracteres en esta extensión. A veces, los autores de contenidos utilizan los puntos de codificación de caracteres de Microsoft por error al crear NCR en lugar de utilizar valores Unicode. Debido a la prevalencia de este error, existen diversos exploradores que muestran los caracteres de Microsoft en esta extensión. Es un comportamiento incorrecto y engaña aún más al desarrollador de contenidos al confirmar de manera incorrecta el valor equivocado. Finalmente, el problema podrá descubrirse cuando alguna aplicación utilice los datos o cuando el explorador que cumple con los estándares no muestre el carácter que se desee.

## lecturas complementarias

- Para obtener más información acerca de la extensión C0, consultar la [Tabla de códigos Unicode: Controles C0 y latín básico](#)
- Para obtener más información acerca de la extensión C1, consultar la [Tabla de códigos Unicode: Controles C1 y suplemento de latin-1](#)
- El documento denominado [Unicode en XML y otros idiomas de etiquetas](#) contiene pautas relacionadas con la utilización del estándar Unicode en conjunto con los idiomas de etiquetado tales como XML.

Dinos qué piensas (en Inglés).

 [Envíanos un comentario](#)

Suscripción a feed RSS.

 [Nuevos recursos](#)

 [Noticias de la página de inicio](#)

Twitter (Noticias de la página de inicio)



[@webi18n](#)

---

De: Tex Texin, XenCraft, François Yergeau, Alis Technologies. Traductor: [Spanish Translation Team](#), Spanish Translation US..

---

Traducido del inglés con fecha 2003-09-06. Traducción modificada por última vez el 2010-01-05 10:56 GMT.

Para ver el historial de cambios del documento, busque [qa-controls](#) en la bitácora de internacionalización.

Copyright © 2003-2010 W3C® (MIT, ERCIM, Keio, Beihang), All Rights Reserved. W3C [liability](#), [trademark](#), [document use](#) and [software licensing](#) rules apply. Your interactions with this site are in accordance with our [public](#) and [Member](#) privacy statements.

