

NOMBRE

dos2unix – Convertidor de archivos de texto de formato DOS/Mac a Unix y viceversa

SINOPSIS

```
dos2unix [parámetros] [ARCHIVO ...] [-n ARCH_DE_ENTRADA ARCH_DE_SALIDA ...]
unix2dos [parámetros] [ARCHIVO ...] [-n ARCH_DE_ENTRADA ARCH_DE_SALIDA ...]
```

DESCRIPCIÓN

El paquete Dos2unix incluye las utilerías dos2unix y unix2dos para convertir archivos de texto plano en formato DOS o Mac a formato Unix y viceversa.

En archivos de texto DOS/Windows, un salto de línea, también conocido como nueva línea, es una combinación de dos caracteres: un retorno de carro (CR) seguido por un salto de línea (LF). En archivos de texto Unix, un salto de línea es solamente un carácter: el salto de línea (LF). En archivos de texto Mac, antes de Mac OS X, un salto de línea era sólo un carácter retorno de carro (CR). Actualmente, Mac OS usa el estilo Unix de saltos de línea (LF).

Los archivos binarios son ignorados automáticamente, a menos que se fuerce su conversión.

Los archivos no regulares, tales como directorios y FIFO, son ignorados automáticamente.

Los enlaces simbólicos y sus destinos no son modificados por defecto. Los enlaces simbólicos pueden opcionalmente ser reemplazados, o la salida puede ser escrita al archivo destino del enlace. Los enlaces simbólicos en Windows no están soportados. Estos son siempre reemplazados, sin que se alteren sus destinos.

Dos2unix fue moldeado a partir del dos2unix que existe en SunOS/Solaris y tiene modos de conversión similares.

PARÁMETROS

- Todos los parámetros siguientes son tratados como nombres de archivo. Use este parámetro si desea convertir archivos cuyos nombres inician con un guión. Por ejemplo para convertir un archivo llamado “-foo”, use este comando:

```
dos2unix -- -foo
```

O en modo de archivo nuevo:

```
dos2unix -n -- -foo out.txt
```

-ascii

Sólo convierte los salto de línea. Éste es el modo de conversión por defecto.

-iso

Conversión entre el conjunto de caracteres DOS e ISO-8859-1. Véase también la sección MODOS DE CONVERSIÓN.

-1252

Usa la página de códigos Windows 1252 (Europa Occidental).

-437

Usa la página de códigos DOS 437 (EE. UU.). Está es la página de códigos usada por defecto para conversión ISO.

-850

Usa la página de códigos DOS 850 (Europa Occidental).

-860

Usa la página de códigos DOS 860 (Portugués).

-863

Usa la página de códigos DOS 863 (Francocanadiense).

-865

Usa la página de códigos DOS 865 (Nórdico).

- 7** Convierte caracteres de 8 bits al espacio de 7 bits.
- c, --convmode CONVMODE**
Establece el modo de conversión, Donde CONVMODE puede ser: *ascii*, *7bit*, *iso*, *mac* siendo *ascii* el valor por defecto.
- f, --force**
Fuerza la conversión de archivos binarios.
- h, --help**
Despiega la ayuda y termina el programa.
- k, --keepdate**
Mantiene la fecha del archivo de salida igual a la del archivo de entrada.
- L, --license**
Muestra la licencia del programa.
- l, --newline**
Añade salto de línea adicional.

dos2unix: Sólo los saltos de línea DOS son cambiados por dos saltos de línea Unix. En modo Mac sólo los saltos de línea Mac son cambiados por dos saltos de línea Unix.

unix2dos: Sólo los saltos de línea Unix son cambiados por dos saltos de línea DOS. En modo Mac los saltos de línea Unix son cambiados por dos saltos de línea Mac.
- m, --add-bom**
Escribe una marca de orden de bytes para UTF-8 en el archivo de salida. Nunca use esta opción cuando la codificación de salida sea distinta de UTF-8. Véase también la sección UNICODE.
- n, --newfile ARCH_DE_ENTRADA ARCH_DE_SALIDA ...**
Modo de archivo nuevo. Convierte el archivo ARCH_DE_ENTRADA y escribe la salida al archivo ARCH_DE_SALIDA. Los nombres de archivo deben ser dados en pares y los comodines *no* deben ser usados o *perderá* sus archivos.

La persona que inicia la conversión en el modo de archivo nuevo (pareado) será el propietario del archivo convertido. Los permisos de lectura/escritura del archivo nuevo serán los permisos del archivo original menos la *umask* (1) de la persona que ejecute la conversión.
- o, --oldfile ARCHIVO ...**
Modo de archivo viejo. Convierte el archivo ARCHIVO y lo sobrescribe con la salida. El programa por defecto se ejecuta en este modo. Sí se pueden emplear comodines.

En modo de archivo antiguo (in situ), el archivo convertido obtiene el mismo propietario, grupo, y permisos de lectura/escritura que el archivo original. Lo mismo aplica cuando el archivo es convertido por otro usuario quien tiene permiso de lectura en el archivo (p.e. usuario root). La conversión será abortada cuando no sea posible preservar los valores originales. Cambiar el propietario implicaría que el propietario original ya no podrá leer el archivo. Cambiar el grupo podría ser un riesgo de seguridad, ya que el archivo podría ser accesible a personas inadecuadas. La preservación del propietario, grupo, y permisos de lectura/escritura sólo está soportada bajo Unix.
- q, --quiet**
Modo silencioso. Suprime todas las advertencias y mensajes. El valor retornado es cero. Excepto cuando se emplean parámetros incorrectos.
- s, --safe**
Ignora los archivos binarios (por defecto).
- F, --follow-symlink**
Sigue los enlaces simbólicos y convierte los destinos.

-R, --replace-symlink

Reemplaza los enlaces simbólicos con los archivos convertidos (los archivos destino originales no se alteran).

-S, --skip-symlink

No altera los enlaces simbólicos ni sus destinos (por defecto).

-V, --version

Despiega la información de la versión y termina el programa.

MODOS MAC

En modo normal los saltos de línea son convertidos de DOS a Unix y viceversa. Los saltos de línea Mac no son convertidos.

En modo Mac los saltos de línea son convertidos de Mac a Unix y viceversa. Los saltos de línea DOS no son modificados.

Para ejecutar en modo Mac use el modificador `-c mac` o use los comandos `mac2unix` o `unix2mac`.

MODOS DE CONVERSIÓN

Los modos de conversión *ascii*, *7bit*, e *iso* son similares a los de los comandos `dos2unix/unix2dos` de SunOS/Solaris.

ascii

En modo *ascii* sólo los saltos de línea son convertidos. Éste es el modo de conversión por defecto.

Aunque el nombre de este modo es ASCII, el cual es un estándar de 7 bits, éste emplea 8 bits. Siempre use este modo cuando convierta archivos en Unicode UTF-8.

7bit

En este modo todos los caracteres no ASCII de 8 bits (con valores de 128 a 255) son convertidos al espacio de 7 bits.

iso Los caracteres son convertidos entre un conjunto de caracteres DOS (página de códigos) y el conjunto de caracteres ISO-8859-1 (Latín-1) de Unix. Los caracteres DOS sin equivalente ISO-8859-1, para los cuales la conversión es imposible, son convertidos en un punto. Lo mismo se aplica para caracteres ISO-8859-1 sin contraparte DOS.

Cuando sólo se emplea el parámetro `-iso`, `dos2unix` intentará determinar la página de códigos activa. Cuando esto no sea posible, `dos2unix` utilizará la página de códigos 437 por defecto, la cual es empleada principalmente en EE. UU. Para forzar una página de códigos específica emplee los parámetros `-437` (EE. UU.), `-850` (Europa Occidental), `-860` (Portugués), `-863` (Francocanadiense), o `-865` (Nórdico). La página de códigos Windows 1252 (Europa Occidental) también está soportada con el parámetro `-1252`. Para acceder a otras páginas de códigos use `dos2unix` en combinación con `iconv(1)`. `Iconv` puede convertir entre una larga lista de codificaciones de caracteres.

Nunca emplee la conversión ISO en archivos de texto Unicode. Esto corromperá los archivos codificados como UTF-8.

Algunos ejemplos:

Convierte de la página de códigos por defecto de DOS a Latín-1 de Unix.

```
dos2unix -iso -n in.txt out.txt
```

Convierte de DOS 850 a Unix Latín-1.

```
dos2unix -850 -n in.txt out.txt
```

Convierte de Windows 1252 a Unix Latín-1.

```
dos2unix -1252 -n in.txt out.txt
```

Convierte de Windows 1252 a Unix UTF-8 (Unicode).

```
iconv -f CP1252 -t UTF-8 in.txt | dos2unix > out.txt
```

Convierte de Unix Latín-1 a la página de códigos por defecto de DOS.

```
unix2dos -iso -n in.txt out.txt
```

Convierte de Unix Latín-1 a DOS 850.

```
unix2dos -850 -n in.txt out.txt
```

Convierte de Unix Latín-1 a Windows 1252.

```
unix2dos -1252 -n in.txt out.txt
```

Convierte de Unix UTF-8 (Unicode) a Windows 1252.

```
unix2dos < in.txt | iconv -f UTF-8 -t CP1252 > out.txt
```

Véase también <http://czyborra.com/charsets/codepages.html> y <http://czyborra.com/charsets/iso8859.html>.

UNICODE

Codificaciones

Existen diferentes codificaciones Unicode. En Unix y Linux los archivos Unicode son codificados comúnmente como UTF-8. En Windows los archivos de texto Unicode pueden estar codificados en UTF-8, UTF-16, o UTF-16 big endian, pero con más frecuencia son codificados en formato UTF-16.

Conversion

Los archivos de texto Unicode pueden tener saltos de línea DOS, Unix o Mac, como cualquier archivo de texto.

Todas las versiones de dos2unix y unix2dos pueden convertir archivos codificados como UTF-8, debido a que UTF-8 fue diseñado para retro-compatibilidad con ASCII.

Dos2unix y unix2dos con soporte Unicode UTF-16, pueden leer archivos de texto codificados como UTF-16 little y big endian. Para ver si dos2unix fue compilado con soporte UTF-16 escriba dos2unix -V.

Las versiones Windows de dos2unix y unix2dos siempre convierten archivos Codificados como UTF-16 a UTF-8. Las versiones Unix de dos2unix/unix2dos convierten archivos UTF-16 a la codificación de caracteres local cuando es configurado a UTF-8. Emplee el comando *locale*(1) para determinar cual es la codificación de caracteres local.

Dado que los archivos de texto formateados UTF-8 son bien soportados tanto en Windows como en Unix, dos2unix y unix2dos no tienen opción para escribir archivos UTF-16. Todos los caracteres UTF-16 pueden ser codificados en UTF-8. La conversión de UTF-16 a UTF-8 ocurre sin pérdida. Los archivos UTF-16 serán ignorados en Unix cuando la codificación de caracteres local no sea UTF-8, para evitar la pérdida accidental de texto. Cuando ocurre un error de conversión de UTF-16 a UTF-8, por ejemplo cuando el archivo de entrada UTF-16 contiene un error, el archivo será ignorado.

La conversión en modos ISO y 7-bit no funciona en archivos UTF-16.

Marca de orden de bytes

En Windows los archivos de texto Unicode típicamente tienen una marca de orden de bytes (BOM), debido a que muchos programas de Windows (incluyendo el Bloc de notas) añaden una BOM por defecto. Véase también http://es.wikipedia.org/wiki/Marca_de_orden_de_bytes_%28BOM%29.

En Unix los archivos Unicode típicamente no tienen una BOM. Se supone que los archivos de texto son codificados en la codificación de caracteres local.

Dos2unix sólo puede detectar si un archivo está en formato UTF-16 si el archivo tiene una BOM. Cuando un archivo UTF-16 no tiene una BOM, dos2unix tratará el archivo como un archivo binario.

Emplee dos2unix en combinación con *iconv*(1) para convertir un archivo UTF-16 sin una BOM.

Dos2unix nunca escribe una BOM en el archivo de salida, a menos que emplee la opción -m.

Unix2dos escribe una BOM en el archivo de salida cuando el archivo de entrada tiene una BOM, o cuando se emplea la opción `-m`.

Ejemplos Unicode

Convertir de Windows UTF-16 (con una BOM) a Unix UTF-8

```
dos2unix -n in.txt out.txt
```

Convertir de Windows UTF-16 (sin una BOM) a Unix UTF-8

```
iconv -f UTF-16 -t UTF-8 in.txt | dos2unix > out.txt
```

Convertir de Unix UTF-8 a Windows UTF-8 sin una BOM

```
unix2dos -m -n in.txt out.txt
```

Convertir de Unix UTF-8 a Windows UTF-16

```
unix2dos < in.txt | iconv -f UTF-8 -t UTF-16 > out.txt
```

EJEMPLOS

Lee la entrada desde 'stdin' y escribe la salida a 'stdout'.

```
dos2unix
dos2unix -l -c mac
```

Convierte y reemplaza a.txt. Convierte y reemplaza b.txt.

```
dos2unix a.txt b.txt
dos2unix -o a.txt b.txt
```

Convierte y reemplaza a.txt empleando modo de conversión ascii.

```
dos2unix a.txt
```

Convierte y reemplaza a.txt empleando modo de conversión ascii. Convierte y reemplaza b.txt empleando modo de conversión de 7bits.

```
dos2unix a.txt -c 7bit b.txt
dos2unix -c ascii a.txt -c 7bit b.txt
dos2unix -ascii a.txt -7 b.txt
```

Convierte a.txt del formato de Mac a Unix.

```
dos2unix -c mac a.txt
mac2unix a.txt
```

Convierte a.txt del formato de Unix a Mac.

```
unix2dos -c mac a.txt
unix2mac a.txt
```

Convierte y reemplaza a.txt manteniendo la fecha del archivo original.

```
dos2unix -k a.txt
dos2unix -k -o a.txt
```

Convierte a.txt y escribe la salida a e.txt.

```
dos2unix -n a.txt e.txt
```

Convierte a.txt y escribe la salida a e.txt, manteniendo la fecha de e.txt igual a la de a.txt.

```
dos2unix -k -n a.txt e.txt
```

Convierte y reemplaza a.txt. Convierte b.txt y escribe a e.txt.

```
dos2unix a.txt -n b.txt e.txt
dos2unix -o a.txt -n b.txt e.txt
```

Convierte c.txt y escribe a e.txt. Convierte y reemplaza a.txt. Convierte y reemplaza b.txt. Convierte d.txt y escribe a f.txt.

```
dos2unix -n c.txt e.txt -o a.txt b.txt -n d.txt f.txt
```

CONVERSIÓN RECURSIVA

Emplee dos2unix en combinación con los comandos *find*(1) y *xargs*(1) para convertir recursivamente archivos de texto contenidos en un árbol de directorios. Por ejemplo para convertir todos los archivos .txt en el árbol de directorios debajo del directorio actual escriba:

```
find . -name *.txt |xargs dos2unix
```

INTERNACIONALIZACIÓN

LANG

El idioma principal se selecciona con la variable de entorno LANG. La variable LANG consiste de varias partes. La primer parte es el código del idioma en minúsculas. La segunda es opcional y es el código del país en mayúsculas, precedido por un guión bajo. Existe también una tercera parte opcional: la codificación de caracteres, precedida por un punto. Unos cuantos ejemplos para intérpretes de comandos POSIX estándar:

export LANG=nl	Neerlandés
export LANG=nl_NL	Neerlandés, Países Bajos
export LANG=nl_BE	Neerlandés, Bélgica
export LANG=es_ES	Español, España
export LANG=es_MX	Español, México
export LANG=en_US.iso88591	Ingles, EE. UU., codificación Latín-1
export LANG=en_GB.UTF-8	Ingles, Reino Unido, codificación UTF-8

Para una lista completa de códigos de idioma y país véase el manual de gettext:
<http://www.gnu.org/software/gettext/manual/gettext.html#Language-Codes>
 <<http://www.gnu.org/software/gettext/manual/gettext.html#Language-Codes>>

En sistemas Unix puede emplear el comando *locale*(1) para obtener información específica del locale.

LANGUAGE

Con la variable de entorno LANGUAGE puede especificar una lista de prioridad de los idiomas, separados por dos puntos. Dos2unix da preferencia a LANGUAGE por encima de LANG. Por ejemplo, primero neerlandés y entonces alemán: LANGUAGE=nl:de. Antes de que pueda usar una lista de prioridad de idiomas a través de la variable LANGUAGE, primero tiene que habilitar la internacionalización, mediante asignar un valor distinto de "C" a LANG (o LC_ALL). Véase también el manual de gettext:
<http://www.gnu.org/software/gettext/manual/gettext.html#The-LANGUAGE-variable>
 <<http://www.gnu.org/software/gettext/manual/gettext.html#The-LANGUAGE-variable>>

Si selecciona un idioma que no está disponible el programa funcionará en ingles.

DOS2UNIX_LOCALEDIR

Con la variable de entorno DOS2UNIX_LOCALEDIR el LOCALEDIR asignado durante la compilación puede ser modificado. LOCALEDIR es usado para encontrar los archivos de idioma. El valor por defecto de GNU es /usr/local/share/locale. El parámetro **—version** desplegará el LOCALEDIR en uso.

Ejemplo (intérprete de comandos POSIX):

```
export DOS2UNIX_LOCALEDIR=$HOME/share/locale
```

VALOR DE RETORNO

Se regresa cero cuando el programa termina exitosamente. Cuando ocurre un error del sistema se regresará el último número de error del sistema. Para otros errores se regresa 1.

El valor de retorno es siempre cero en modo silencioso, excepto cuando se emplean parámetros incorrectos.

ESTÁNDARES

<http://es.wikipedia.org/wiki/Documento_de_texto>

<http://es.wikipedia.org/wiki/Retorno_de_carro>

<http://es.wikipedia.org/wiki/Nueva_l%C3%ADnea>

<<http://es.wikipedia.org/wiki/Unicode>>

AUTORES

Benjamin Lin – <blin@socs.uts.edu.au> Bernd Johannes Wuebben (modo mac2unix) – <wuebben@kde.org>, Christian Wurl (añadió el salto de línea extra) – <wurl@ira.uka.de>, Erwin Waterlander – <waterlan@xs4all.nl> (Mantenimiento)

Página del proyecto: <<http://waterlan.home.xs4all.nl/dos2unix.html>>

Página de SourceForge: <<http://sourceforge.net/projects/dos2unix/>>

Freecode: <<http://freecode.com/projects/dos2unix>>

VÉASE TAMBIÉN

file (1) *find* (1) *iconv* (1) *locale* (1) *xargs* (1)