



# Convertir HTML a .po y a la inversa

De GALPon WiKi

Este como está dispoñíbel en  galego no wiki de **Trasno** ([http://wiki.trasno.net/Converter\\_HTML\\_a\\_.po\\_e\\_%C3%A1\\_inversa](http://wiki.trasno.net/Converter_HTML_a_.po_e_%C3%A1_inversa))  
This HowTo is available in  english in the **Tupi** wiki ([http://www.maefloresta.com/wiki/index.php?title=Main\\_Page](http://www.maefloresta.com/wiki/index.php?title=Main_Page)) .

## Introducción

Este **como** nace de la necesidad de disponer de una forma sencilla de traducción, de archivos de ayuda y de páginas web, hechas con HTML (.htm, .html, .xhtml).

Pensado, inicialmente, para que los desarrolladores faciliten un método estándar para la traducción, a los traductores sin perfil técnico, sirve también para aquellos traductores más avanzados que quieren simplificar procesos, o facilitar métodos más sencillos a compañeros de equipo.

En este documento vamos a ver los siguientes pasos para completar el proceso:

### **Trabajo a cargo del(os) desarrollador(es) o del coordinador de traducciones**

- 1 Preparación de las herramientas y archivos
- 2 Conversión de HTML a .po

### **Trabajo a cargo del(os) traductor(es)**

- 3 Proceso de traducción

### **Trabajo a cargo del(os) desarrollador(es) o del coordinador de traducciones**

- 4 Conversión de .po a HTML
- 5 Traducción de .xml

### **Trabajo que debe definirse si lo hace el traductor o el desarrollador/coordinador**

- 6 Traducción (?) de imágenes

Documentación de base

- <http://translate.sourceforge.net/wiki/toolkit/html2po>
- <http://opensource.bureau-cornavin.com/html2pot-pot2html/index.html>

## Preparación de las herramientas y archivos

Necesitamos tener instalado:

- el paquete de herramientas de traducción «*translate-toolkit*» y las de desarrollo «*python*» y «*python-dev*»

- En Debian: `apt-get install translate-toolkit python python-dev`

Empleo como ejemplo la aplicación **Tupí** en la que estuve trabajando estos días pasados y que fue lo que me dio la idea de hacer este manual. El objetivo es convertir a archivos **.po** los archivos de la ayuda integrada en la aplicación, que están contruidos con HTML.

Examinando el archivo de fuentes *tupi\_0.1-git12.tar.gz* (una vez descomprido se llama *tupi-0.1*) encontramos la carpeta **help**

```
/home/usuario/Tupi/tupi-0.1/src/components/help/help$
```

haciendo un «ls» vemos que contiene las siguientes carpetas

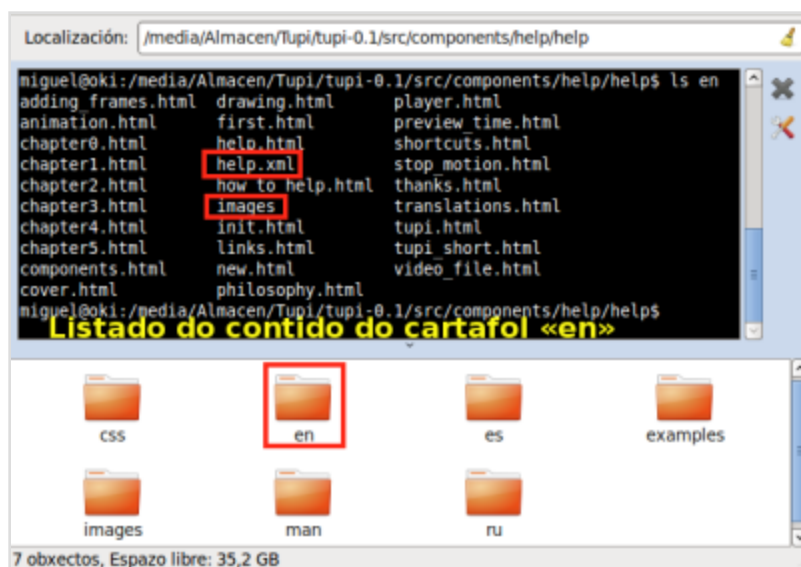
```
css en es examples images man ru
```

destaco en `<color font color="#FF0000">rojo</font>` los paquetes de idioma veamos el contenido de la carpeta «en»

```
ls en
```

que nos da la siguiente salida:

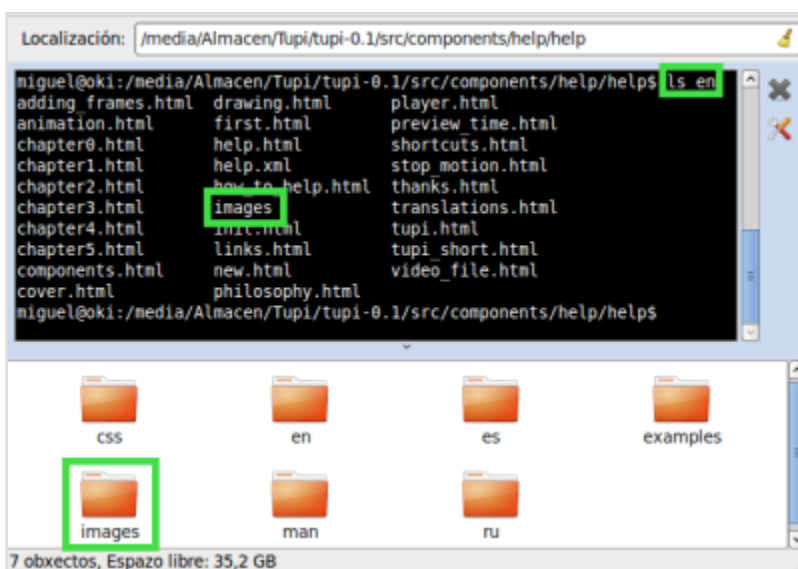
```
adding_frames.html  drawing.html      player.html
animation.html      first.html        preview_time.html
chapter0.html       help.html         shortcuts.html
chapter1.html       help.xml          stop_motion.html
chapter2.html       how_to_help.html  thanks.html
chapter3.html       images            translations.html
chapter4.html       init.html         tupi.html
chapter5.html       links.html        tupi_short.html
components.html     new.html          video_file.html
cover.html          philosophy.html
```



**Fig. num. 1** – Vista de la carpeta «help» y el destaque de la carpeta «en» en un explorador gráfico

Fijémonos en un detalle, tenemos dos carpetas con el nombre «images»

- una en la ruta → /home/usuario/Tupi/tupi-0.1/src/components/help/help
- y otra en la ruta → /home/usuario/Tupi/tupi-0.1/src/components/help/help/en



**Fig. num. 2** – Observad lo marcado en verde

Normalmente, en la primera, se encuentran las imágenes inamovibles o genéricas, y en la segunda las imágenes que cambian en función de la traducción, como son las capturas de interfaz, las de mensajes, etc. De estas segundas hablaremos más adelante.

Volver al inicio

## Conversión de HTML a .po

Crear los archivos de traducción «.po» a partir de los «.html» originales

```
html2po en l10n
```

```

↓   ↓   ↓
↓   ↓   nombre que le damos a la carpeta en el que vamos a guardar los «.po»
↓   nombre de la carpeta en la que están los «.html» originales
orden de conversión html a(to) po
```

hecho esto facemo otra vez un «ls»

css en es examples images **l10n** man ru

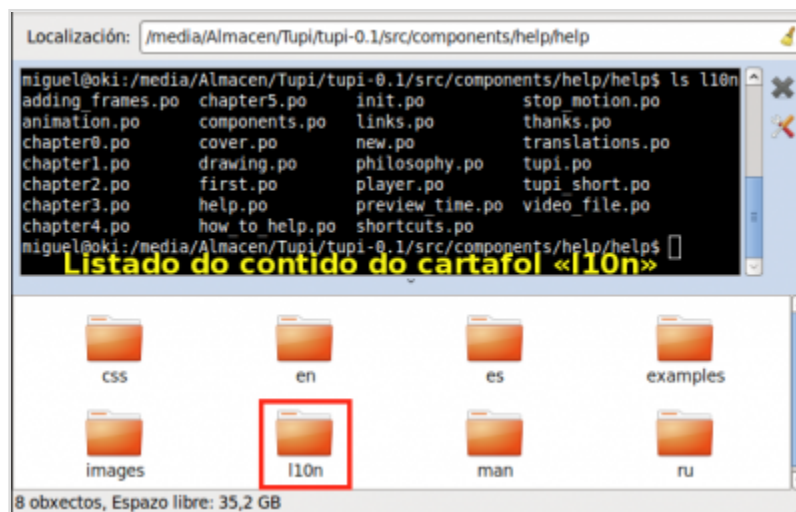
y comprobamos que ahora tenemos una nueva carpeta «l10n»

veamos su contenido con un «ls l10n»

```

adding_frames.po  chapter5.po  init.po        stop_motion.po
animation.po      components.po links.po       thanks.po
chapter0.po      cover.po     new.po        translations.po
chapter1.po      drawing.po   philosophy.po  tupi.po
chapter2.po      first.po     player.po     tupi_short.po
chapter3.po      help.po      preview_time.po video_file.po
chapter4.po      how_to_help.po shortcuts.po

```



**Fig. num. 3** – Vista de la carpeta «help» y el destaque de la carpeta «l10n» en un explorador gráfico

Observemos, en diferencia con el listado anterior, que no existe el archivo/carpeta «images» ni el archivo correspondiente al «help.xml»

#### **Nota:**

Existe otra posibilidad, extraer los archivos HTML a «.pot» esta es una buena forma de generar «actualizadores». Es más sencilla para el desarrollador/coordinador, sin embargo algo más compleja de uso para los traductores con un bajo perfil técnico.

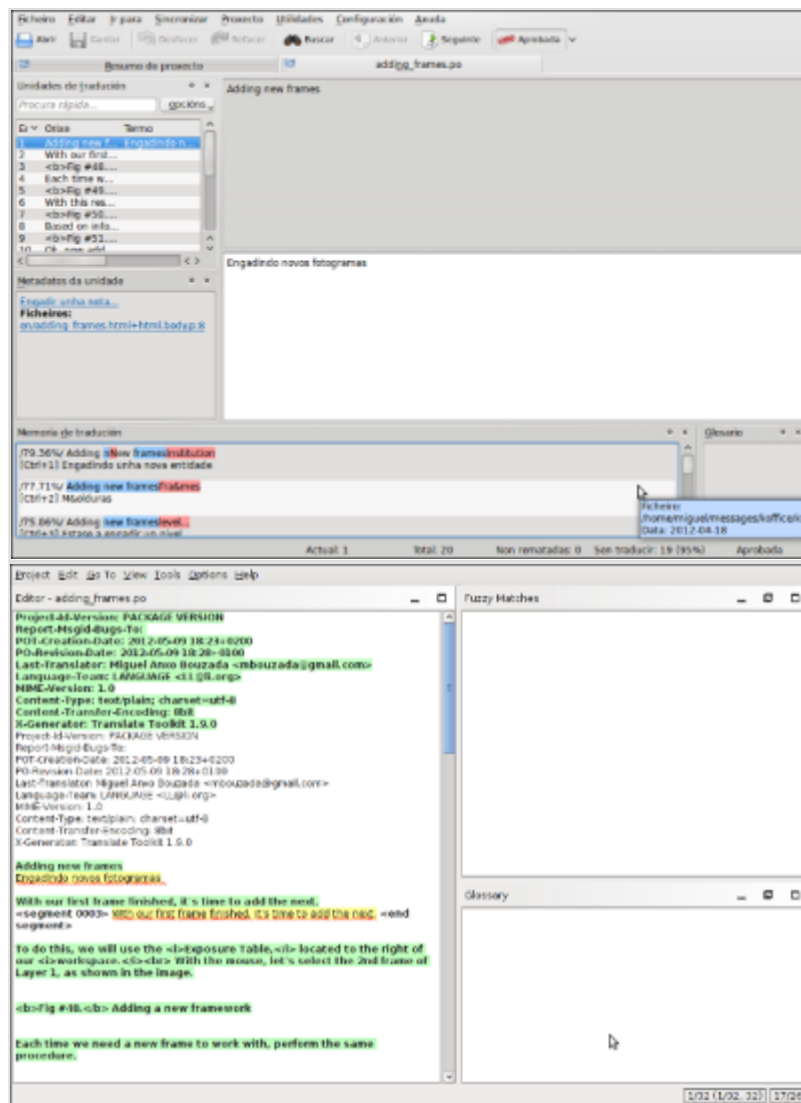
La orden sería: `html2po -P en l10n` en la que la opción `-P` indica que la extracción se haga a archivos `.pot`

[Volver al inicio](#)

## Proceso de traducción

Una vez que ya disponemos de los archivos «.po» procedemos a traducirlos como solemos hacerlo, empleando las herramientas de CAT (Traducción Asistida por Computador) que mejor se adapten a nosotros. Otra posibilidad es enviar estos archivos a algún servicio de Traducción en Web, como Pootle (<http://translate.sourceforge.net/wiki/pootle/index>) o Transifex (<https://www.transifex.net/>) .





**Fig. num. 4** – Vista de distintas interfaces de traducción: PoEdit, Virtaal, Lokalize y OmegaT

En este punto es importante destacar que **no es necesario** generar los archivos binarios «.mo».

Volver al inicio

## Conversión de .po a HTML

Crear los archivos «.html» ya traducidos a partir de los «.po» que estuvimos (o estuvieron) traduciendo

```
po2html -t en l10n gl
```

```

↓      ↓ ↓ ↓ ↓
↓      ↓ ↓ ↓ nombre que le damos a la carpeta de destino
↓      ↓ ↓ ↓ en este caso empleamos el indicador del idioma gallego «gl»
↓      ↓ ↓ carpeta en la que están los archivos «.po» ya traducidos
↓      ↓ carpeta en la que están los originales, para ser empleados como «modelos»
↓      opción de uso de modelo (t de template)
orden de conversión po a(to) html

```

Un nuevo «ls» nos dirán que fue creada la carpeta «gl»

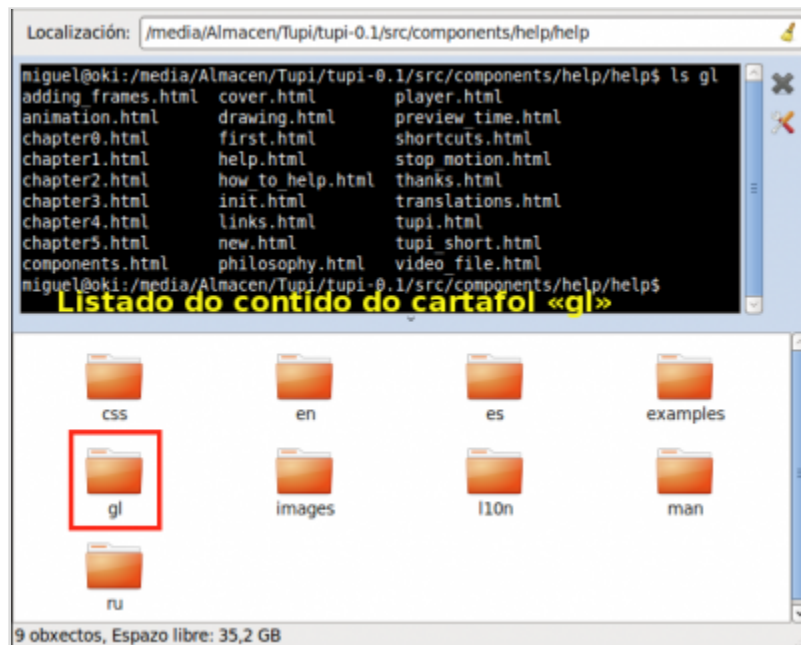
```
css en es examples gl images l10n man ru
```

veamos ahora el contenido de esta carpeta «gl» con la orden «ls gl»

```

adding_frames.html  cover.html      player.html
animation.html      drawing.html    preview_time.html
chapter0.html       first.html      shortcuts.html
chapter1.html       help.html       stop_motion.html
chapter2.html       how_to_help.html thanks.html
chapter3.html       init.html       translations.html
chapter4.html       links.html      tupi.html
chapter5.html       new.html        tupi_short.html
components.html     philosophy.html video_file.html

```



**Fig. num. 5** – Vista de la carpeta «help» y el destaque de la carpeta «gl» en un explorador gráfico

Quedamos en que en esta carpeta nos faltan la carpeta «images» y el archivo «help.xml». Para el caso del archivo «.xml» veremos a continuación como proceder, para el caso de la carpeta «images» lo veremos más adelante, finalizando este manual.

Volver al inicio



## Traducción de .xml

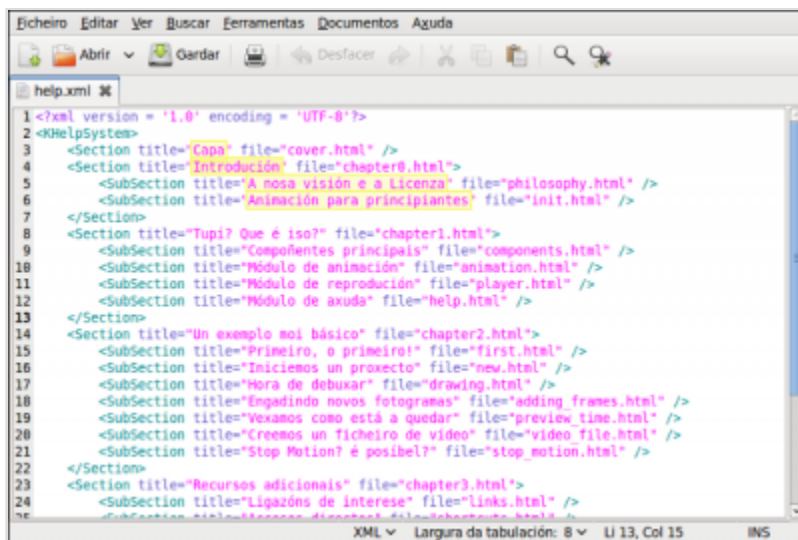
La traducción de un archivo «.xml» deberemos hacerla empleando un editor de texto sencillo, nunca editores de texto con formato, como pueden ser Abiword, LibreOffice, Apache-Openoffice, etc.

Un archivo «.xml» suele tener un esquema semejante la este:

```
<?xml version = '1.0' encoding = 'UTF-8'?>
<KHelpSystem>
  <Section title="Capa" file="cover.html" />
  <Section title="Introducción" file="chapter0.html">
    <SubSection title="Nuestra visión y la Licencia" file="philosophy.html" />
    <SubSection title="Animación para principiantes" file="init.html" />
  </Section>
  <Section title="Tupí? Que es eso?" file="chapter1.html">
    <SubSection title="Componentes principales" file="components.html" />
    <SubSection title="Módulo de animación" file="animation.html" />
    <SubSection title="Módulo de reproducción" file="player.html" />
    <SubSection title="Módulo de ayuda" file="help.html" />
  </Section>
</KHelpSystem>
```

Sólo se debe traducir **el contenido entre comillas** «"» de los títulos (*Section title*, *SubSection title*) y respetar rigurosamente las dobles comillas inglesas que delimitan el texto. **Nunca** debemos traducir los nombres de archivos (*file="cover.html"* p.ej.) ni ninguna otra etiqueta.

Aconsejamos que empleéis un editor de texto sencillo que emplee coloreado en función del código, hoy en día lo hacen la mayoría de los que acompañan a una distribución, o, en el peor de los casos, emplean un añadido para hacerlo. El ejemplo siguiente es una captura del editor *Gedit*.



**Fig. num. 6** – Vista del archivo «help.xml»



se destacan los elementos que se traducen

[Volver al inicio](#)

## Traducción (?) de imágenes

Esto es algo muy sencillo de entender, y no muy fácil de hacer, ya que requiere de bastante paciencia y, a veces, de algún conocimiento de tratamiento de imágenes.

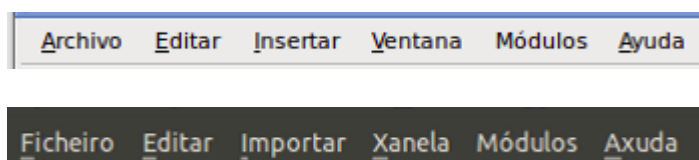
Todo comienza con la instalación de la aplicación en el idioma objeto de la traducción y su correcta ejecución. Debemos ir aplicando las acciones tal y como aparecen en las imágenes del original e iremos haciendo las oportunas capturas.

Estas capturas, una vez «procesadas» (cómo se puede ver a continuación) irán en una carpeta llamada *images*, incluida en la carpeta que creamos anteriormente. En nuestro ejemplo en:

```
/home/usuario/Tupi/tupi-0.1/src/components/help/help/gl/images
```

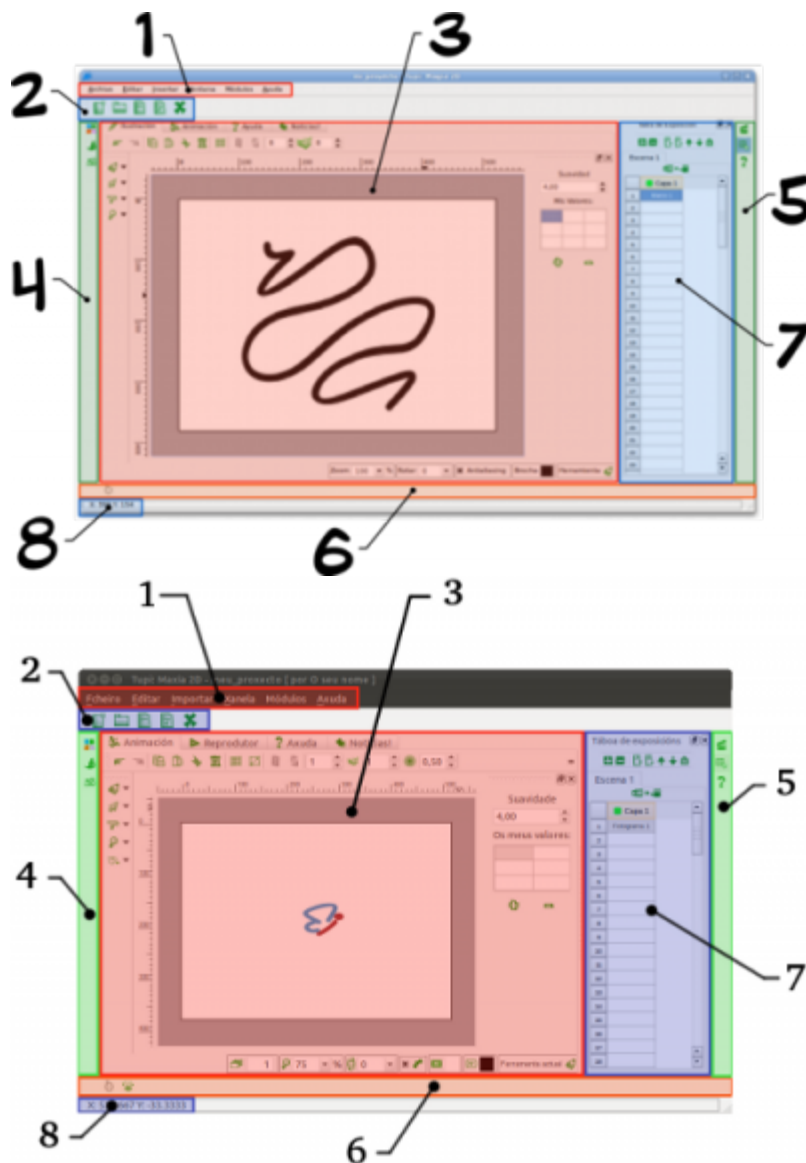
**Importante:** las imágenes deben respetar **exactamente** el mismo nombre. Cualquier cambio en él hará que no se encuentre esa imagen.

Simplemente una captura de un recorte



**Fig. num. 7** – Vista del menú superior de la aplicación en castellano y en gallego  
Nota: la diferencia que se observa en el contenido se debe a que corresponden a versiones diferentes.

O, a veces, tendremos que emplear alguna herramienta de dibujo como el *Gimp* o *Inkscape*.



**Fig. num. 5** – Vista del esquema de la aplicación en castellano y en gallego

Volver al inicio

Obtenido de «[http://www.galpon.org/wiki/index.php/Convertir\\_HTML\\_a\\_.po\\_y\\_a\\_la\\_inversa](http://www.galpon.org/wiki/index.php/Convertir_HTML_a_.po_y_a_la_inversa)»

- Esta página fue modificada por última vez el 17 jun 2012, a las 09:39.
- Esta página ha sido visitada 368 veces.
- El contenido está disponible bajo los términos de la GNU Free Documentation License 1.2.