

Traducción asistida por ordenador

*Fran Diéguez
Oficina de Software Libre da USC*



OFICINA DE
SOFTWARE LIBRE

www.usc.es/osl



OFICINA DE
SOFTWARE LIBRE



www.usc.es/osl

Sobre o relator

- *Técnico / Desenvolvedor na Oficina de Software Libre da USC*
- *Colaborador e tradutor oficial do proxecto de tradución GNOME*

Guión

- *Pequena introdución*
- *Formatos empregados*
- *Aplicativos empregados*
- *Ciclo da tradución en aplic. Software libre*

Que é o CAT

- *Supón a inclusión da tecnoloxía informática para a axuda ao tradutor profesional ou na consecución dunha tradución.*
- *En software libre existen múltiples ferramentas en continuo desenvolvemento*

Que é i18n = internacionalización?

- *Preparar o software para ser empregado*
 - *Con textos “internacionais”*
 - *En múltiples “locales”*
- *Texto traducido (dinámico e estático)*
 - *Dirección do texto*
 - *Formato de datas e números e moeda*
- *Afecta a*
 - *Deseño*
 - *Formatos*
 - *Contidos (dependentes de culturas)*

Formatos empregados

- *Tecnoloxía Gettext*
- *XLIFF*
- *Memorias de tradución TMX*
- *Glosarios TBX*

Gettext

“gettext é a librería de internacionalización de GNU (i18n). É amplamente empregada para escribir programas multinguaxe.”

Gettext

- *A máis empregada no mundo do SwL*
- *Ferramenta consistente*
- *Suporte a memorias de tradución*
- *Simple pero á vez eficaz*
- *Idónea para tradución de software, mais non para textos e documentación.*

Gettext

- *Vantaxes*
 - *extracción automática de cadeas do código fonte*
 - *xeración de ficheiros portábeis e eficientes en execución*
- *po & mo*
- *ferramentas de xestión*
 - *xgettext, msgmerge & msgfmt*
- *cobertura multiidioma real*
- *plurais, polisemias*

Gettext

- *Como traballa?*
 - *Marcado de código*
“cadea sen marcar”
_ (“Cadea marcada”)

Gettext

Extracción de cadeas

xgettext [OPCIÓN] [FICHEROENTRADA]...

*Extrae as cadeas traducibles a partires
dun grupo de ficheiros dados.*

Xera .po (pofles)

Gettext

Os .po (pofiles)

Cabeceiras

Persoas que editaron o arquivo

Licenzas

Configuración do arquivo

Idioma

URL do proxecto

Datas edición e creación

Último tradutor

Formas plurais

Outras

Corpo

Texto orixinal

Tradución final

Gettext

- *Formas especiais*
 - %d %i %s
 - <etiquetas> </etiquetas>
 - referencias a imaxes
 - formas plurais, etc...

Gettext

- *Fusión de archivos de traducción*

msgmerge [OPCIÓN] def.po ref.pot

- *Fusiona dous ficheiros .po de estilo uniforme nun só.*
- *Traballa como unha memoria de tradución.*

Gettext

- *Mofiles*
 - *Ficheiros binarios só entendíbeis pola computadora*

– *msgfmt [OPCIÓN] fcheiro.po ...*

Xera un catálogo binario de mensaxes, só entendíbel pola computadora, a partires da descrición da tradución realizada.

Gettext

- *Os .mo (mofiles)*
 - *proprios de cada arquitectura (binarios)*
 - *fonte de onde o programa en execución extrae as cadeas*

Gettext

- *Como sabe o idioma que “falamos”?*
 - *Emprega un sistema de variábeis do sistema que representan unívoca da configuración do idioma.*

LC_CTYPE: o sistema de caracteres e a conversión de maiúsculas (UTF-8, ...)

LC_COLLATE: a colación do texto (diferencias Mai/min, ...)

LC_MONETARY: formato da moeda (\$ 10.0, 10.0 €)

...

Gettext

- *Como calcula as traducións?*
 - *Segue a xerarquía se as variables de antes están configuradas e se dispón da tradución para dito idioma.*
 - *Se non ten a tradución axeitada para o idioma ou non están configuradas as variábeis recolle a cadea que está no texto orixional do programa (sempre inglés).*

XLIFF

Moito rebumbio na rede

Formato baseado en XML

Creado para estandarizar a localización

XLIFF foi creado por OASIS en 2002.

*Preparado para tradución de software,
documentación, textos, ...*

Extremadamente extensíbel

XLIFF

- *Permite traducción de moitos formatos*
 - *Páxinas web: HTML*
 - *Documentación: Docbook*
 - *Documentos de texto: ODF/DOC*
 - *XML*
 - *FrameMaker*
 - *Paquetes de recursos de software (formatos distintos coma .properties, .po, .msg, .java, etc.)*

XLIFF

- *Para o tradutor*
 - *Permítelle centrarse na tradución*
 - *Separa o texto do formato*
 - *Almacenamento de información que é útil no proceso de localización*
 - *Introduce roles dentro do ciclo de tradución*
- *Para o enxeñeiro de software*
 - *Facilítalle a creación de ferramentas por estar baseado en XML*

Memorias de tradución

- *Destacamos dous formatos*
 - *Baseadas en pofile*
 - *Pofile compendium: non son máis que fusións de ficheiros po.*
 - *Baseados en XML*
 - *TMX: estandarizado e con boas ferramentas*

Glosarios

- *Podémoslos resumir nun formato*
 - *TBX:*
 - *Baseado en XML*
 - *A maioría dos aplicativos teñen soporte.*
 - *Son pares de cadea orixe / tradución*

Ferramentas



OFICINA DE
SOFTWARE LIBRE

www.usc.es/osl

Tipos de aplicativos

- *A base sempre vai ser gettext*
- *Aplicativos de escritorio*
Lokalize, poEdit, gtranslator,
CrowdSight
- *Plataformas Web*
Pootle, entrans, rosetta

Lokalize

- *Suporte para memorias de traducción, incluso TMX*
- *Suporte para Glosarios TBX*
- *Corrector ortográfico integrado*
- *Xestión de proxectos de traducción*
- *Estatísticas*

Poedit

- *Multiplataforma (GNU/Linux, Ms. Windows, Mac OS X)*
- *Corrector ortográfico integrado*
- *Xestor de proxectos de tradución*
- *Soporte para memorias de tradución, mais non TMX*
- *Estatísticas*

OmegaT

- *Baseado integramente en XML*
 - *Memorias de traducción TMX*
 - *Glosarios TBX*
 - *Estatísticas por proxecto*
 - *Corrector ortográfico integrado*

Pootle

- *Sistema baseado en Web*
- *Soporte de xestión de proxecto*
- *Xestión de usuarios contribuidores*
- *Baseado na tecnoloxía Gettext máis,*
 - *Importa e exporta en distintos formatos na nova versión*
 - *Recomendábel para grupos de tradución sen coñecementos técnicos.*

Entrans

- *Sistema baseado en web*
- *Xestión de varios proxectos*
- *Xestión de usuarios contribuidores*
- *Empregado por Mancomun para traducir Firefox e Openoffice no seu día*
- *Extinto debido as melloras nos outros aplicativos web*

Rosseta / Launchpad

- Accesíbel dende <http://translations.ubuntu.com>
- Open Source dende fai menos de 3 meses
- Xestión de múltiples proxectos
- Xestión de usuarios contribuíntes
 - Roles: tradutor oficial, contribuidor orixinal
- Xestión de suxestións de traduzóns
- Xestión de memorias de tradución
- Tradución multilinguaxe
- Empregado por Ubuntu para localizar o seu Sistema Operante

Ciclo da tradución en SwL

- 1. recoller última versión das cadeas*
- 2. fusión cos que temos nós (msgmerge)*
- 3. tradución*
- 4. revisión (ortográfica, gramatical, estilística, msgfmt)*
- 5. envío da tradución ao desenvolvedor*
- 6. Ir a 1 : actualización e engadido de máis cadeas*
- 7. Ir aleatoriamente a 1, 2, 3, 4, 5 ou 6*

Preguntas



OFICINA DE
SOFTWARE LIBRE

www.usc.es/osl

Grazas



OFICINA DE
SOFTWARE LIBRE

www.usc.es/osl