

# Internationalization in Debian

## How to improve further?

A Maps and Sponge Bob Presentation

Nicolas François

Fosdem, February 2009



# Introduction

Different parts of Debian are translated

- Debian Installer
- Debconf Templates
- Website
- Programs
  - native packages
  - upstream translations
- Package Descriptions
- Manpages
- Documentation

This relies on

- Translators/Reviewers
- Maintainers
- Developers
- Localization Teams
- Infrastructure
  - BTS
  - Package Repositories
  - Statistics/Robots





# Translation Processes

The Debian Installer is translated using PO files. Translators usually commit directly their work in the SVN repository.

There is a total of 4837 strings (41281 words), among them 1016 strings are mandatory to enable a language in the installer (10103 words).

<http://d-i.alioth.debian.org/doc/i18n> contains information for translators (updating translations, adding new languages, ...) and maintainers of Debian Installer packages.



# Status in Lenny

There are currently 70 translations

<http://d-i.alieth.debian.org/l10n-stats>

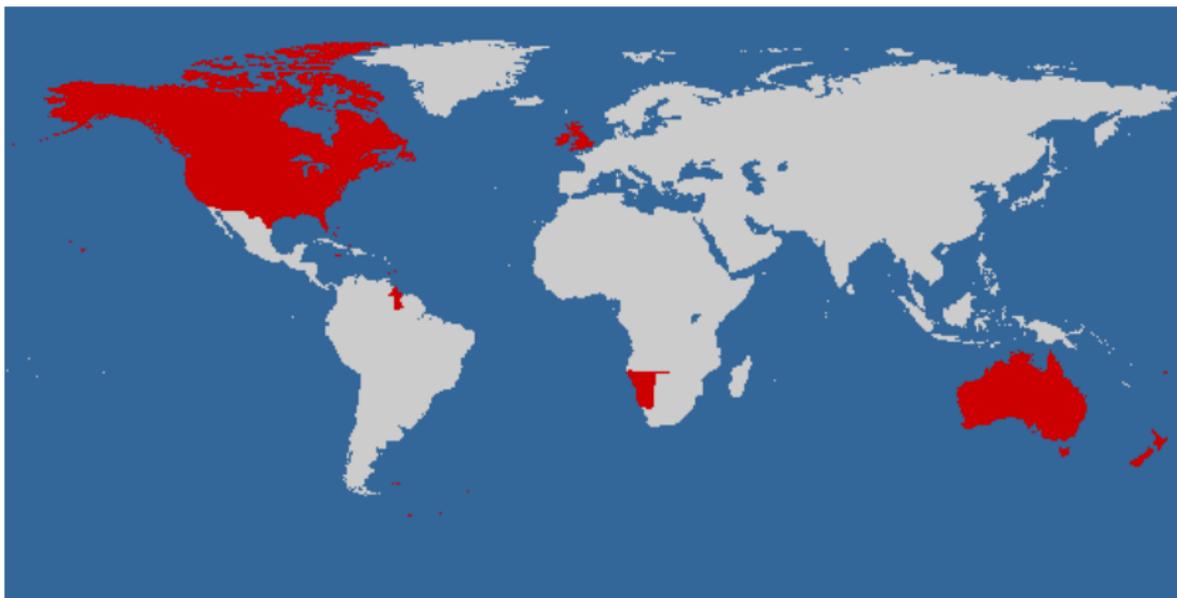
The Lenny installer supports 63 languages.

new languages		
Amharic	Ethiopia	45.0 M
Gaelic	Ireland	0.4 M
Marathi	India	90.0 M
Northern Sami	Norway	0.02M
Serbian	Serbia and Montenegro	11.1 M
Welsh	United Kingdom	0.5 M

lost languages		
Estonian	Estonia	1.0 M



# Status in Potato - 2000



Number of supported languages: 1 (5% world population)

Cacophonix



el: Κακοφωνίξ

# Status in Woody - 2002



Number of supported languages: 16 (36-38% world population)

Pegleg

SOL  
LIGET OMNIBUS  
!!



# Status in Sarge - 2005



Number of supported languages: 40 (45-49% world population)



# Status in Etch - 2007



Number of supported languages: 57 (57-61% world population)

Dogmatix



bn: gNoyartumix

# Status in Lenny - 2009?



Vitalstatistix



Number of supported languages: 63 (58-61% world population)

Detailed Status

# Work in progress



Number of languages: 80

Fulliautomatix



nl: Hoefnix

# How to improve further?

The Debian Installer has a very good I18N infrastructure

- Release schedule
- String freeze (3 weeks)
- Direct access to the repository

Some translations got completed very lately (more than 10 in the last week)

- Keep your translations up-to-date
- Commit often
- Nominate a backup coordinator / team

Announce earlier the failures to meet the release criteria to a wider audience (debian-user\*).

Merge with some Ubuntu effort



# The Debconf Templates

## Debconf Templates

Debconf Templates contain all interactions with the system administrator while installing packages (or installing the entire system).

- Templates are translated using PO files (generated with po-debconf)
- PO files are usually assigned to translators
- Coordination by the localization teams<sup>1</sup>
- Allowing the translation of all Debconf messages using a gettext-based system was a Lenny release goal.
- 85 languages, 62 with more than 10%

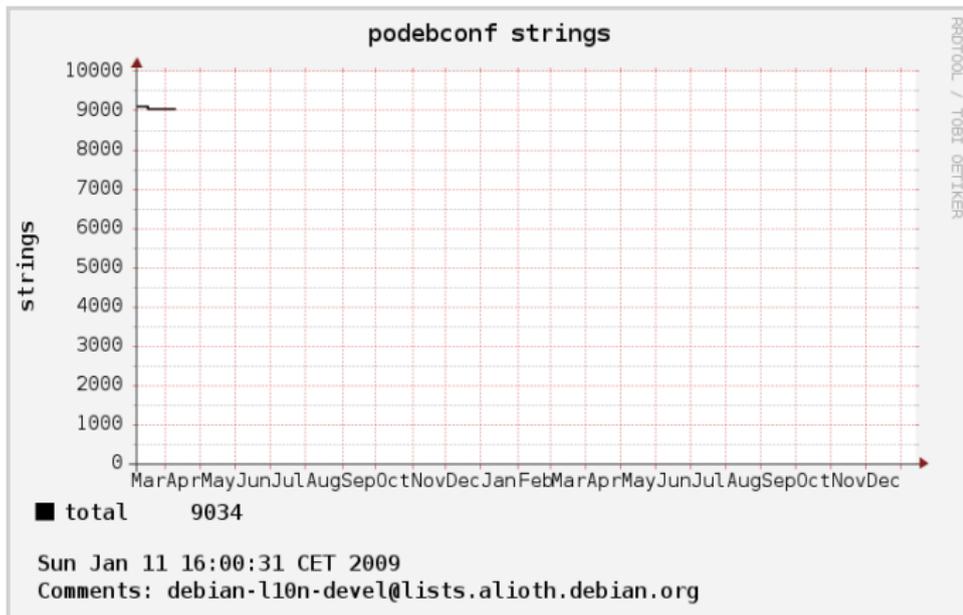
<sup>1</sup><http://www.debian.org/international/l10n/po-debconf/fr>  
[http://i18n.debian.net/debian-l10n/french/fr.by\\_type.html](http://i18n.debian.net/debian-l10n/french/fr.by_type.html)

Bacteria



da: Remouladine

# Fully translated in Etch



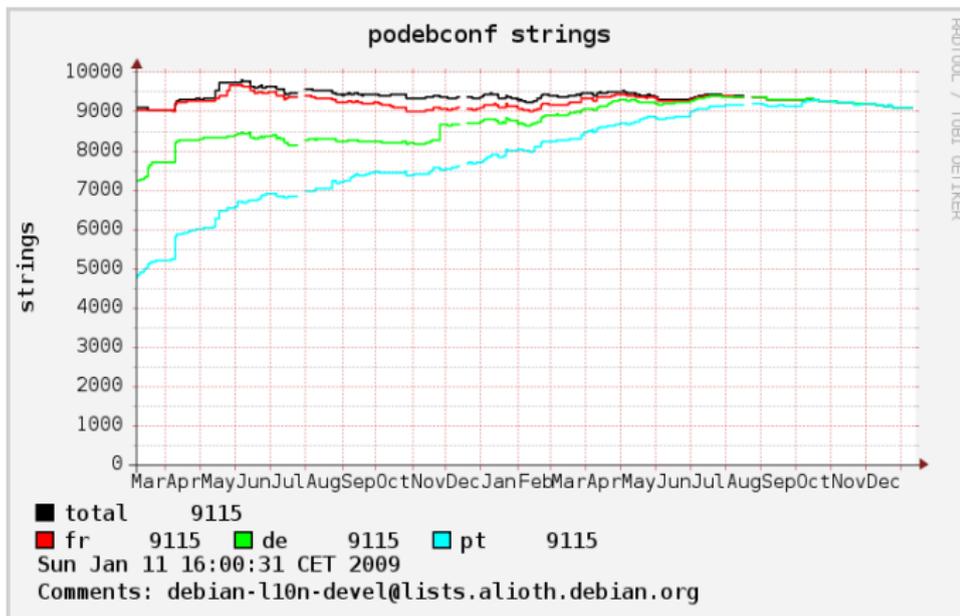
Languages fully translated in Etch: 0

Cacophonix



ga: Sgreuchanaix

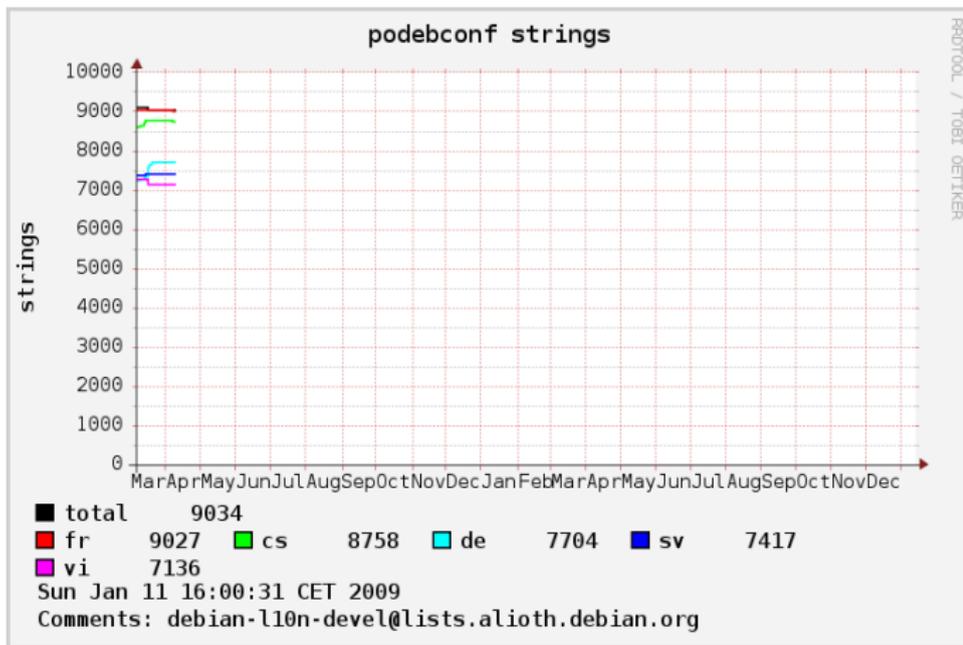
# Fully translated in Lenny



Languages fully translated in Lenny: 3



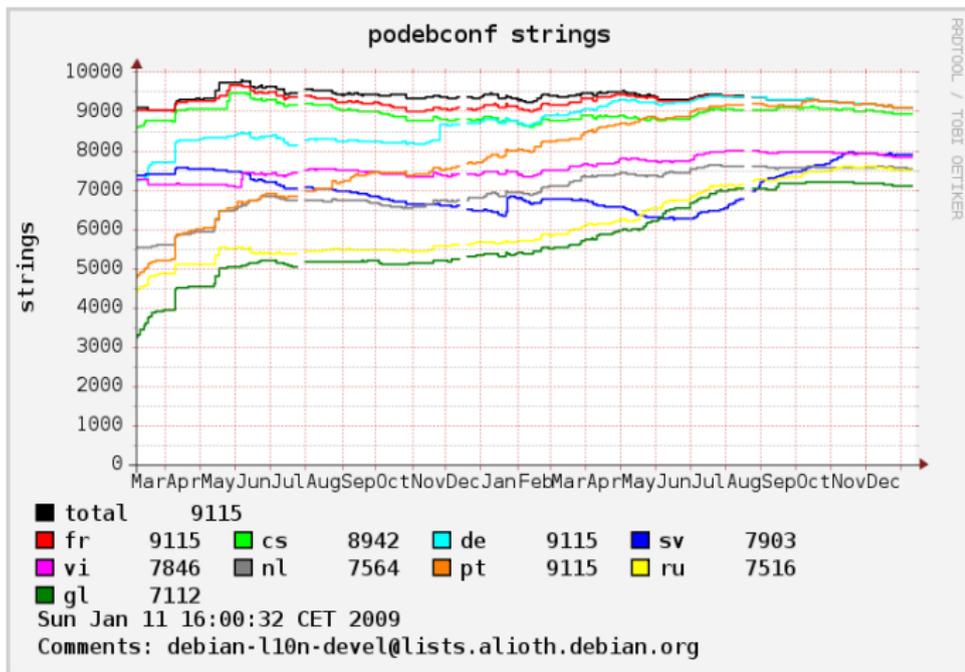
# More than 75% in Etch



Languages with more than 75% translated in Etch: 5



# More than 75% in Lenny



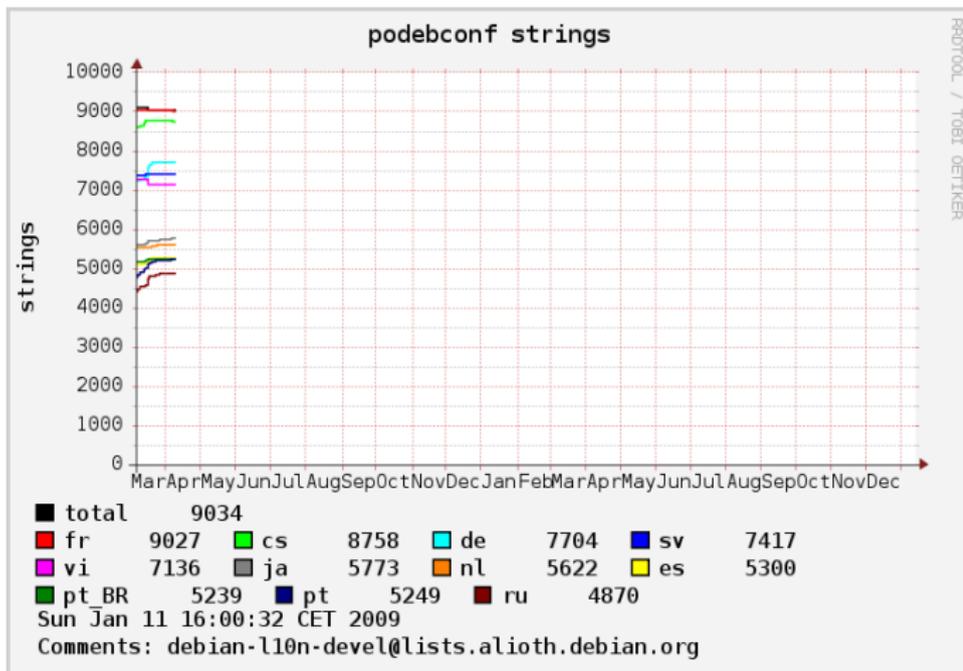
Languages with more than 75% translated in Lenny: 9

Impedimenta



pt\_BR: Naftalina

# More than 50% in Etch

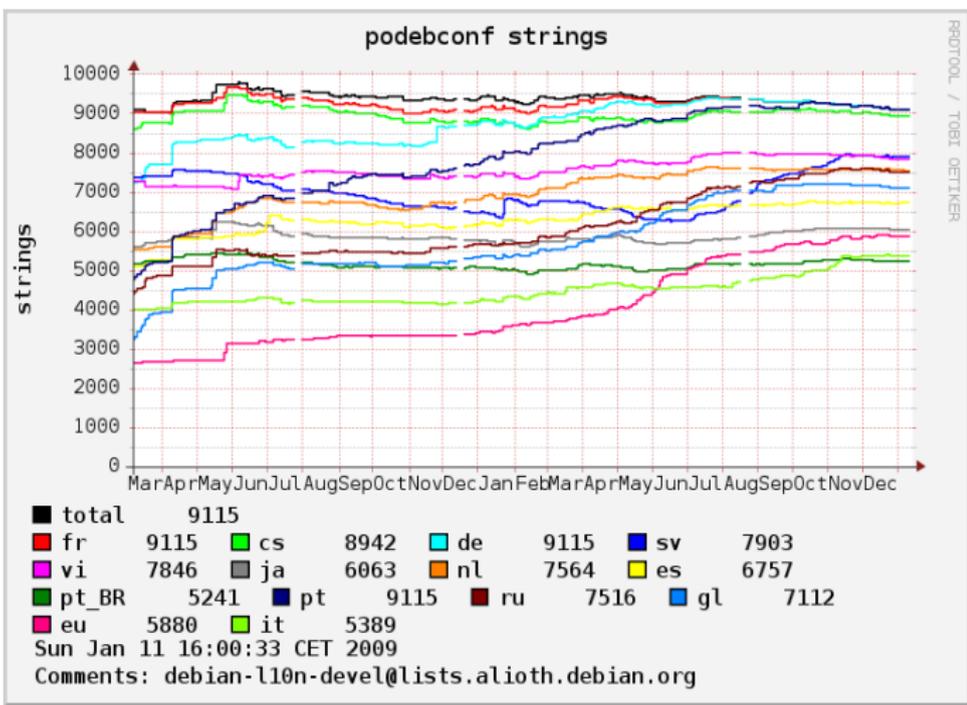


Languages with more than 50% translated in Etch: 11



Detailed Status

# More than 50% in Lenny



Languages with more than 50% translated in Lenny: 14



# What helped the Debianconf Templates translation?

- Support by the release team and maintainers
  - Unblock requests did not lower the translation statistics.
  - Easy unblock for translation only uploads
- I18N NMU campaign<sup>2</sup>
  - Reduces the number of translation lagging in the BTS
  - Provides a schedule to translators
  - Ensures a timely commit of the translation
- Call for translations
  - Provides a schedule to translators
  - Ensures a timely commit of the translation
- Long period of stability
  - French and German translations completed since July.
- Lower usage of debconf
  - Same number of messages, even with an higher number of packages.

---

<sup>2</sup>[http://i18n.debian.net/debian-110n/110n-nmu/nmu\\_bypackage.html](http://i18n.debian.net/debian-110n/110n-nmu/nmu_bypackage.html)



# How to improve further?

## ● Maintainers

- Motivate translators by integrating their translations
- Send calls for translations when you change a debconf template.
- Send calls for translations once in a while even if all your translations are up-to-date.

## ● Translators

- Try to define a goal for Squeeze (all packages with more than a given popcon score, etc.)
- Regularly check your statistics
- Improve the quality with a review process

## ● Coordinators/Infrastructure

- mid-term I18N NMU campaign
- I18N NMU campaign coordination
- Keep track of the packages for which there were no calls for translations for a long time.
- Automatic unblock of translation only uploads during (the early stages of) a freeze.



# Debian Web Site

## Debian Web Site

The Debian Web Site (<http://www.debian.org>) is one of the first contact with Debian and a primary source of information for users and non-users. Other official (Wiki) and non-official (Planet, \*.debian.net) source of information are rarely translated.

The website contains WML pages and scripts.



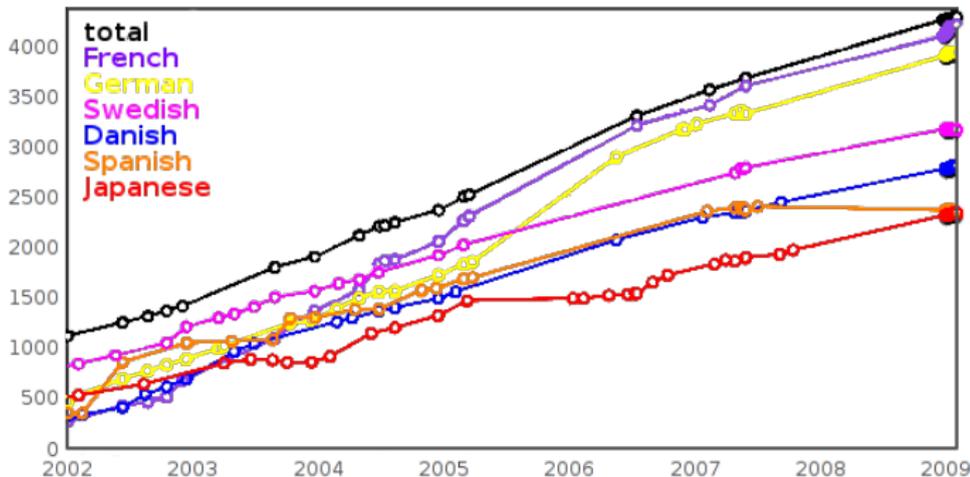
# Translation

- WML pages are translated individually and a PO file contains common strings used by scripts.
- The Web Site uses content negotiation to distribute the right translations to users.
- Markers are placed in the translation to identify the version of the English pages used for the translation.
- There are 4279 pages to translate (12368135 bytes)
- The PO file contains 887 strings
- Translators have access to the website repository (CVS).
- 2 translators are needed to maintain a translation.



# Status

- The Debian Web Site is translated in 36 languages.
- 7 languages have more than 50% of the pages translated: da, de, en, es, fr, ja, sv.
- <http://www.debian.org/devel/website/stats/>



# How to improve further?

- Notification for new pages or outdated pages
- Access to CVS
- Formatting changes, hidden changes
- DSA variability
  - tool to create DSA?
  - tool to translate DSA?
- No processes / tools to translate Wikis.

## Switching to a PO based translation

- improve the quality
- reduce the effort
- but very big effort to do the conversion

Bacteria



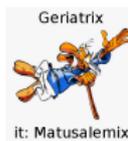
# Program Translations

- Debian localization teams do not focus on the translation of programs.
- There are other localization teams or projects for these translations:
  - Gnome
  - KDE
  - Mozilla
  - OpenOffice.org
  - Fedora
  - TP
  - decathlon<sup>3</sup>
  - translate@thon<sup>4</sup>
  - ...
- Except for native packages

---

<sup>3</sup><http://translate.sourceforge.net/wiki/decathlon/mainpage>

<sup>4</sup><http://translate.sourceforge.net/wiki/guide/translateathon>



# Status

- 11014 PO strings
  - 16 teams with more than 30%
- 
- Swedish (76%)
  - Japanese (68%)
  - French (67%)
  - Chinese, as spoken in China (63%)
  - Russian (61%)
  - Korean (60%)
  - German (54%)
  - Chinese, as spoken in Taiwan (53%)
- Portuguese (49%)
  - Vietnamese (46%)
  - Greek (45%)
  - Hungarian (38%)
  - Finish (36%)
  - Italian (36%)
  - Norwegian Bokmål (32%)
  - Spanish (32%)



# How to improve further?

- Lack of tracking

- <http://www.debian.org/intl/l10n/po/> tracks all the PO files distributed in Debian.
- An additional page is needed for Debian native packages.

- Call for translations

- Access to the repository



# Package Descriptions

- Started in 2001
- Revived in 2006
- `http://ddtp.debian.net`
  - Database
  - Mail interface
  - Web interface
- Distribution of the translations
  - now officially in the archive<sup>5</sup>
  - apt (apt-cache, aptitude, ...)
  - `http://packages.debian.org`
  - will be included in the Lenny CDs?

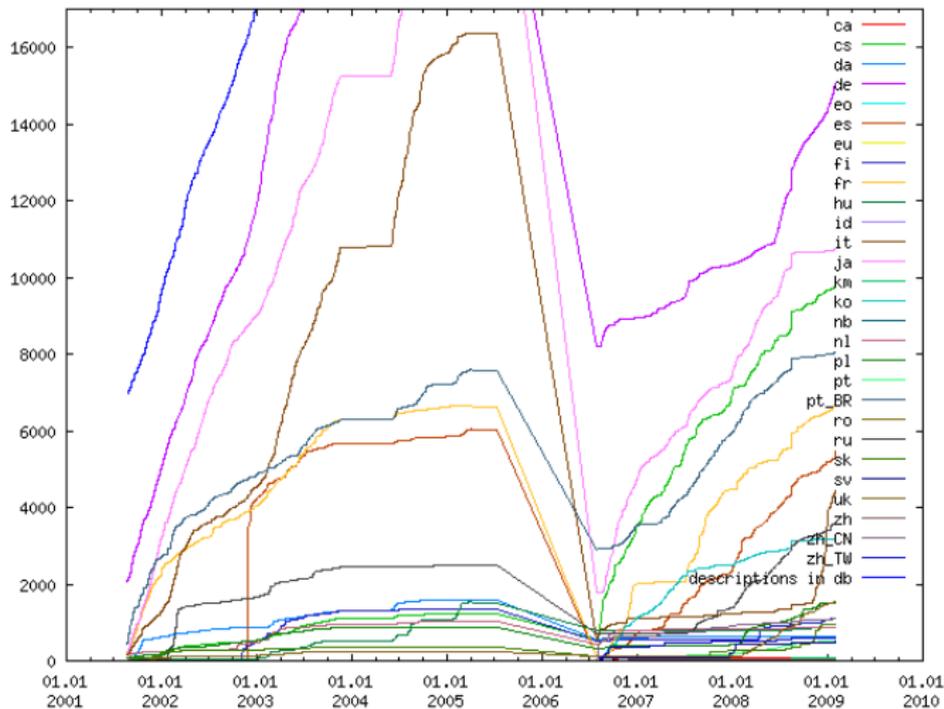
---

<sup>5</sup>`http://ftp.us.debian.org/debian/dists/unstable/main/i18n/`



# Status

55760 package descriptions (25986 actives)



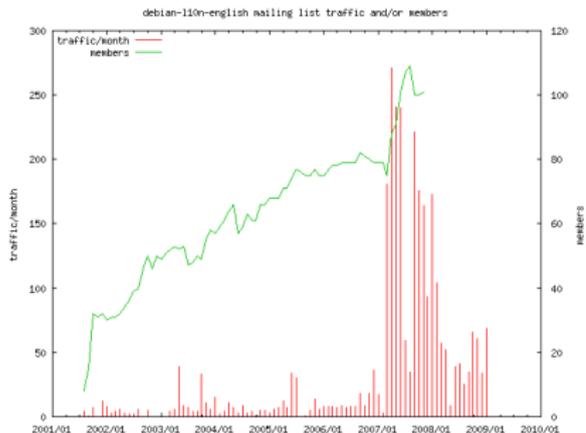
Dogmatix



eo: Idefiks

# English Translation

- debian-l10n-english exists since August 2001.
- Very few messages up to February 2007.



## Smith Review Project <sup>6</sup>

<sup>6</sup><http://wiki.debian.org/I18n/SmithReviewProject>  
[lists.debian.org/debian-l10n-english/2007/03/msg00178.html](http://lists.debian.org/debian-l10n-english/2007/03/msg00178.html)



# Review Process

- Focus on
  - debconf templates
  - package descriptions
  - review requests from Maintainers
- Based on popcon scores, for packages with debconf templates
  - improve the debconf templates
  - package descriptions
- Use the same robot process as other teams for tracking  
`http://i18n.debian.net/debian-l10n/english/en.by\_status.html`



# Status

- Long process
  - notify the package maintainer
  - Call for debconf templates review (day 7)
  - Review
  - Send the review to the BTS (day 16)
  - Call for translation updates (day 19)
  - Translation updates
  - Send the patch to the package maintainer (day 32)
- 153 packages reviewed so far
- 672 packages have debconf templates



# How to improve further?

- The team is very small
  - Volunteers?
  
- Maintainer could send more requests for reviews
  - Non-native speakers
  - Native speakers
  
- Tracking what packages/versions were reviewed or need to be reviewed.



# Support from Maintainers

- Provide a standard translation framework
- Request translations when you add or change strings
  - `podebconf-report-po --call` from the PO directory.
  - before the upload!
- Request translations from time to time when you plan an upload.
- Help translators
  - provide hints for cryptic strings and meaning of variables
  - use `--previous` when you generate PO files.

## Example

```
#, fuzzy
#| msgid "A string with a typos."
msgid "A string with no typos."
msgstr "Une chaine avec une erreur."
```



# podebconf-report-po session

```
<update your POT and PO files>
```

```
$ cd debian/po
```

```
$ podebconf-report-po --call
```

- It sends a notification to [debian-i18n@debian.org](mailto:debian-i18n@debian.org) with the list of existing translations, with the current POT file, and a 10 days deadline
  - an editor is open with the mail body and the mail headers.
  - a confirmation is asked before sending the mails.
- Then sends a request to update the translations to the translators listed in the PO files (with the language teams in copy), with the current PO file, and a 10 days deadline.
  - an editor is open with the mail body and the mail headers.
  - a confirmation is asked before sending the mails.
- You can ignore bounces caused by wrong email address / subscribers-only mailing lists.

# Translators

- Motivate other translators
- Translate and Review
- Report English bugs
- Use the compendia: <http://i18n.debian.net/compendia/po/fr/compendium-fr-LATEST.po>
  - to improve the consistency of your translations (conflicts tagged with #-#-#-#-#)
  - to start new translations
- Improve the quality of your translations
  - spellchecker (acheck, PO translation tools)
  - checker tools (msgfmt, pofilter, PO translation tools)
- Organize translate@thon



# Translation tools

- PO generator tools
  - gettext
  - translate toolkit
  - poxml
  - po4a
- Translation editors
  - virtaal
  - kbabel/lokalize
  - poedit
  - gtranslator
  - editor (vim/emacs mode)
- On-line translation editors
  - Pootle
  - Rosetta

Pegleg

SOL  
L'OGIET OMIBUS  
!!



# Questions? Ideas?



Thanks for your attention.

Debian Installer

Debconf Templates

Debian Web Site

Program Translations

Package Descriptions

English Translation

Miscellaneous

**Backup**

oooooooooooo oooooooooo

oooo

ooo

oo

oooo

oooo

# Backup

# Backup

Fichier Édition Affichage Aller à Signets Outils Onglets Aide

Précédent Suivant Arrêter Recharger Accueil Historique Signets Plus petit Plus grand

http://118n.debian.net/debian-110n/french/fr.by\_type.html#podebconf Aller à

dcB:schedule BTS : Google : debian 110n

Debian -- État d'ava... FrontPage - Debian ... Coordination of deb...

**podebconf**

apt	podebconf	fr.po	Christian Perrier	rfr	2008-11-06 18:39:44	<a href="#">[2008-11-00053]</a>
dbconfig-common	podebconf	fr.po	Christian Perrier	rfr	2009-02-06 07:13:54	<a href="#">[2009-02-00050]</a>
ddclient	podebconf	fr.po	Christian Perrier	rfr	2009-01-30 07:22:11	<a href="#">[2009-01-00444]</a>
			Christian Perrier	rfr	2009-01-31 12:38:00	<a href="#">[2009-01-00461]</a>
			Christian Perrier	lfcf	2009-02-01 13:48:46	<a href="#">[2009-02-00004]</a>
			Christian Perrier	bts	2009-02-04 17:27:59	<a href="#">[2009-02-00034]</a> #51409
debian-installer	podebconf	fr.po	Christian Perrier	rfr	2008-08-01 07:30:32	<a href="#">[2008-08-00000]</a>
			Christian Perrier	rfr	2008-08-01 07:23:42	<a href="#">[2008-08-00003]</a>
			Christian Perrier	rfr	2008-08-05 07:14:06	<a href="#">[2008-08-00062]</a>
emacspeak-ss	podebconf	fr.po	Christian Perrier	bts	2008-12-02 10:42:02	<a href="#">[2008-12-00173]</a> #50755
fglnx-driver	podebconf	fr.po	Christian Perrier	rfr	2009-01-24 14:22:13	<a href="#">[2009-01-00370]</a>
fwsnort	podebconf	fr.po	Franck Joncourt	lfcf	2008-10-23 17:29:54	<a href="#">[2008-10-00144]</a>
harden	podebconf	fr.po	Thomas	wontfix	2006-04-20	<a href="#">[2006-04-000633]</a> #36393

# Backup

Fichier Édition Affichage Aller à Signets Outils Onglets Aide

Précédent Suivant Arrêter Recharger Accueil Historique Signets Plus petit Plus grand

http://www.debian.org/international/i10n/po-debconf/fr

dcB:schedule BTS : Google : debian i10n

Debian -- État d'ava... FrontPage - Debian ... Coordination of deb...

**Section : main**

Paquet	Note	Fichier	Traducteur	État	Traducteur	Date	Bogue
<a href="#">unattended-upgrades</a>	50% (1t;1f;0u)	<a href="#">fr.po</a>	Steve Petruzzello	bts	Christian Perrier	2009-01-10 06:54:02	<a href="#">511373</a>
<a href="#">dot.l10n</a>	92% (25t;2f;0u)	<a href="#">fr.po</a>	Christian Perrier	bts	Christian Perrier	2009-02-04 17:27:59	<a href="#">514092</a>
<a href="#">zabbix</a>	70% (7t;2f;1u)	<a href="#">fr.po</a>	Yves Rütschli				
<a href="#">tpconfig</a>	66% (4t;1f;1u)	<a href="#">fr.po</a>	Michel Grentzinger	bts	Christian Perrier	2009-01-02 17:51:01	<a href="#">510512</a>
<a href="#">mserv</a>	50% (3t;3f;0u)	<a href="#">fr.po</a>	Christian Perrier				
<a href="#">mopd</a>	57% (4t;2f;1u)	<a href="#">fr.po</a>	Laurence Colombet	bts	Christian Perrier	2009-02-04 17:28:29	<a href="#">514094</a>
<a href="#">emacspeak-ss</a>	0% (0t;0f;16u)			bts	Christian Perrier	2008-12-02 10:42:02	<a href="#">507550</a>

**Section : contrib**

Paquet	Note	Fichier	Traducteur	État	Traducteur	Date	Bogue
<a href="#">l10n</a>	96% (31t;0f;1u)	<a href="#">fr.po</a>	Christian Perrier	bts	Christian Perrier	2008-12-27 07:21:01	<a href="#">509255</a>

**Section : non-free**

**Les paquets qui utilisent po-debconf pour lesquels la traduction est à jour :**

**Section : main**

# LICENCE

Slides are GPL2 or later, but some images may be subject to copyright (and subject to fair use)