i18n and l10n of i4o

Thursday October 18th, 2017

Dirk Hoffmann

Internationalisation and Localisation of Indico

Thursday October 18th, 2017

Dirk Hoffmann

Overview

- Aim and Principles of i18n in coding

- How is it done in python/Indico?

- Transifex, practical hints, interplay of coder and translator

- Development and Deployment

- Status: languages and manpower

- Wishlist / Discussion

i18n for software

- Improve user experience

- "Our" (CERN/HEP) context is exceptional.
  - Daily work with english, even in home institutes
  - High level of communication competence (invented the web) and agility (>1000 collaborators working on the same project)

- Home institute administration and other domains have different needs
  - Very often no or little competence even in reading english

Be aware of "faux amis" Language space

- What is a Conference?
  - Conference (en) = Colloque (fr)
  - Conférence (fr) = Seminar (en)

- An Event?
  - Ereignis   ? Événement   ?

- "Meaning space" of different languages is not a bidirectional matching
  - How do you say "no" in Japanese?

- Linguists know this effect (Sapir-Whorf).

Be aware of "faux amis" Language space

- What is a Conference?
  - Conference (en) = Colloque (fr)
  - Conférence (fr) = Seminar (en)
- An Event?
  - Ereignis / Veranstaltung? Événement / Manifestation ?
- "Meaning space" of different languages is not a bidirectional matching
  - How do you say "no" in Japanese?
- Linguists know this effect (Sapir-Whorf).

Be aware of "faux amis" Language space

- What does **Conference**? (partially obscured)
  - ...lloque (fr)
  - ...minar (en)
- A...
- "..." ...is ...nar...
  - ...in Japanese?
- Linguists know this effect (Sapir-Whorf).

Nom de dieu de putain de bordel de merde de saloperie de connard d'enculé de ta mère

Blocks rather than words

- Do not answer   Do not confirm
- Do not — answer — confirm

- Nicht — antworten — bestätigen
- Ne - répondez - pas — ?

- Cancel? — Annulation ? — ¿Anulación? O prefiere(s) « cancelar » ?

Technical implementation

- Python: **gettext** (standard module)
  - Load PO (portable object) file, *here french*
  - Replace `"Apple"` by `_("Apple")`
  - API returns `"Pomme"` *for french*
- PO files are universal (not depending on programming language)
- Templates (POT files) are extractable by various tools

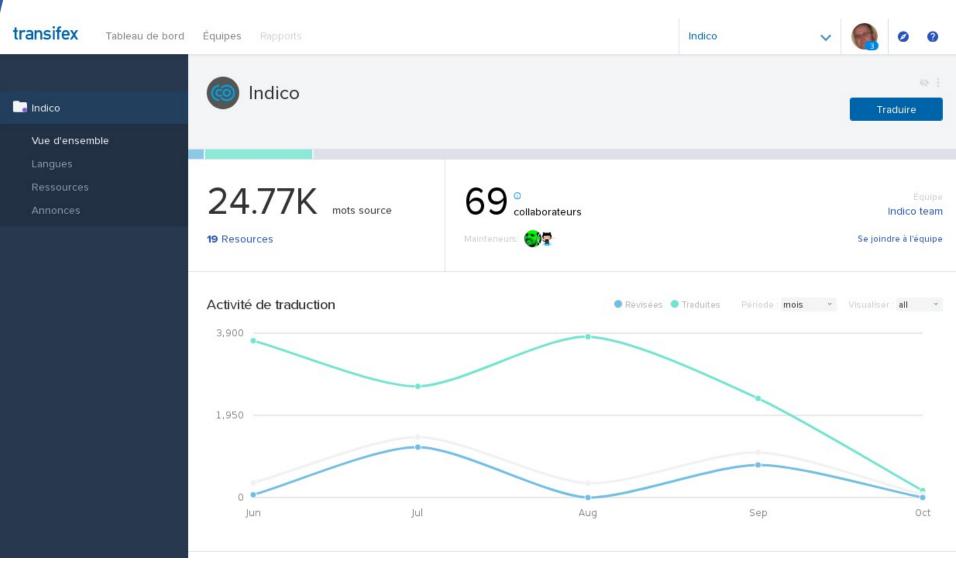- Of course, UTF-8 programming is ideal here!

Transifex

- Translators edit PO files, each for one language

- Web service transifex.com allows "easy" editing and crowd/cloud translation

Transifex

Indico – i18n

Transifex

- Translators edit PO files, each for one language

- Web service transifex.com allows "easy" editing and crowd/cloud translation

- Procedure

  - Translator gets a Transifex account

  - Requests to join a "project" (language)

  - Approval

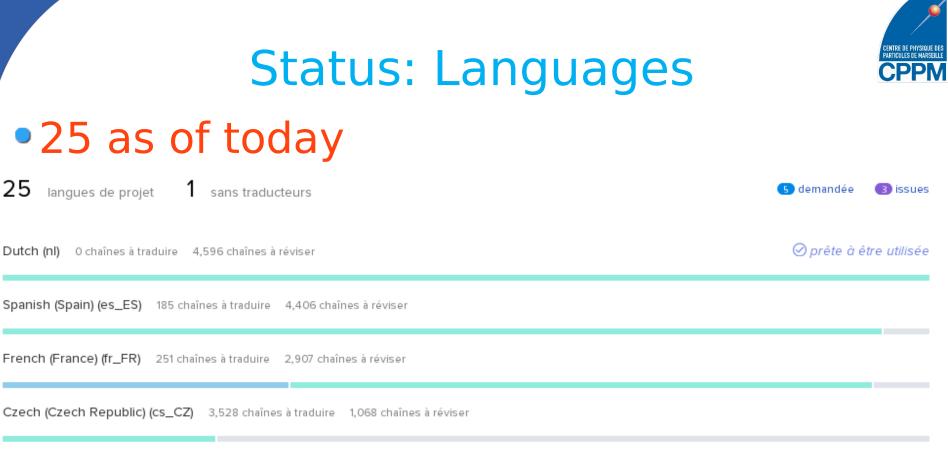 - "Manual" contact (email) necessary, no notifications!

Communication on i18n lists

- CERN e-groups exist for *some* languages
  - project-indico-i18n
  - project-indico-i18n-de
  - project-indico-i18n-es
  - project-indico-i18n-fr
- To be replaced by talk.getindico.io?
  - https://talk.getindico.io/t/portuguese-pt-translation-group/37

Status: Languages

- ## 25 as of today

25 langues de projet   **1** sans traducteurs

🔵 **5** demandée   🟣 **3** issues

Dutch (nl)   0 chaînes à traduire   4,596 chaînes à réviser   ⊘ *prête à être utilisée*

Spanish (Spain) (es_ES)   185 chaînes à traduire   4,406 chaînes à réviser

French (France) (fr_FR)   251 chaînes à traduire   2,907 chaînes à réviser

Czech (Czech Republic) (cs_CZ)   3,528 chaînes à traduire   1,068 chaînes à réviser

German (de)   3,545 chaînes à traduire   952 chaînes à réviser

Italian (it)   3,791 chaînes à traduire   805 chaînes à réviser

Russian (ru)   3,808 chaînes à traduire   22 chaînes à réviser

Turkish (tr)   4,012 chaînes à traduire   584 chaînes à réviser

afficher toutes les langues   Demander une langue

Transifex Workflow

- Extract/update of POT files
  - Indico → Transifex
- Edit/review translations
  - Translators ↔ Reviewers
- Extract/update of PO files
  - Transifex → Indico
- New Indico Release
- GOTO 0

Manpower

- Return from experience with two languages
  - L1:
    - Start with 3 volunteers, two afternoons on phone conference, then "one week of concentrated work"
    - 100% done at that time.
      Now 5%/63% to be translated/reviewed.
    - Spontaneous proposal on github ("issue"). Long reaction.
  - L2:
    - Started "alone". Punctual proposals (8) registered in Transifex.
    - One bad experience with amateur translation
    - 75% still to be translated today
    - *But "new hope" again since yesterday.*

Summary: How to start — A proposal

- Prepare yourself (and your team)

- Request Transifex membership

- Edit, discuss, review

- Test on your system (?)

- Notify Indico developers

- Enjoy

- Iterate with next release ;-)

Summary

- It *seems* easy. Transifex probably best tool on the market. (Hopefully will stay free of charge.)

- Review and collaboration/discussion is essential for quality of result
  ⇒ Additional communication channels needed

- Many good-will declarations end up with weak, incomplete or outdated result

- Complexity/importance underestimated?
  More sustained effort? Open for discussion!

- Programmers must write i18n-aware strings!

Wishlist / Discussion

- Closer contact around deployment key dates?

- Legal aspects, intellectual property?

- Get more contributions

    Productivity ↔ Motivation

- Wishes, volunteers?