Localization of Open Source Software for Pakistani Languages

Huda Sarfraz
Center for Language Engineering
Al-Khawarizmi Institute of Computer Science
University of Engineering & Technology, Lahore
Pakistan
Software Localization

- Customization of software for a specific language-region pair
- Vital role in enabling information and communication technology access where language barriers exist
- Over 60 languages spoken in Pakistan
- 10% English literacy rate
## Open Source Software

<table>
<thead>
<tr>
<th>Software</th>
<th>Number of Languages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Firefox</td>
<td>66</td>
</tr>
<tr>
<td>Thunderbird</td>
<td>49</td>
</tr>
<tr>
<td><strong>OpenOffice.org</strong></td>
<td><strong>19</strong></td>
</tr>
<tr>
<td>Pidgin</td>
<td>16</td>
</tr>
<tr>
<td>VLC Media Player</td>
<td>48</td>
</tr>
<tr>
<td>7zip</td>
<td>13</td>
</tr>
<tr>
<td>GIMP</td>
<td>13</td>
</tr>
<tr>
<td>Audacity</td>
<td>26</td>
</tr>
<tr>
<td><strong>Ubuntu</strong></td>
<td><strong>28</strong></td>
</tr>
</tbody>
</table>
Localization Terminology

- Localization
- Internationalization
- Locale
Development Prerequisites

- Language enabled
- Keyboards installed
- Fonts installed
Team Structure

- **Language experts**
  - Translate
  - Resolve localization issues (date formats, currency formats)

- **Technical experts**
  - Assist in translation
  - Integrate translations and other localization elements to produce the localized product
Localization Process

1. select OSS
2. select localization tools
3. register localization
4. compile translation resources
5. extract strings
6. translate
7. translation resources
8. review translation
9. process translation
10. test localization
11. release final localized product
12. update localization
13. release final localized product
OSS Selection Criteria

- Internationalization
- Pakistani language character set support (support for Arabic script)
- Right-to-left/bidirectional script (RTL/bidi) support
  - Numbers:
    - The conference will be held from the 12th to the 16th of January
      کانفرنس 12 سے 16 جنوری کو ہو گی۔
  - Placeholders:
    - Are you sure you want to unsubscribe from %S?
      کیا آپ واقعی %S سے رکنیت ختم کرنا چاہتے ہیں؟
- User-friendly
- Active project
- Multi-platform
Localization Tools Selection

- Translation management tools
- Localization tools
Translation Management Tools

- OmegaT
- Translate Toolkit
- Olifant
- Other collaborative web based tools
OmegaT

- Computer aided translation (CAT) tool
  - Open source
  - Cross-platform

- Features
  - Terminology glossary
  - Translation memory (TMX - translation memory eXchange)
  - Multiple file format support

- Demo

Sample OmegaT project
Localization Tools

- XUL based localization: SeaMonkey (Demo)
  - Mozilla Translator
- GNU gettext based localization: OpenOffice.org
  - Application specific build tools
- Qt based localization: Psi
  - Qt linguist
  - Demo
Localization Registration

- Coordination with software community
- Community support
- Official releases
- More accessibility
- Prevention of duplication of effort
- Processes vary with communities
Translation Resources

Compilation

- Terminology glossary by National Language Authority, Pakistan
- Previously localized software
  - OpenOffice.org 2.4.0 ur-PK
  - SeaMonkey 1.1.19 ur-PK
  - Psi 0.14 ur-PK
  - Mozilla Suite 1.5 ur-PK
  - Firefox 1.0.6 ur-IN
  - Thunderbird 1.0.7 ur-IN
  - OpenOffice.org 2.0.3 ur-IN
- Online terminology translations

Terminology glossary compiled for Urdu localization
http://www.crulp.org/software/localization/OSS/glossary/searchGlossary.asp
Translation Process

- Translation resources
- Translate
- Look up new terminology
- Update core glossary
- Complete translation
- Core glossary
Translation Exercise

- Failed to remove this account.
- Filters associated with this folder will be updated.
- Horizontal scrolling.
- New languages can be configured using the Languages Panel.
Translation & Review

- Translation by linguists
- Review and updates by developers
  - Typos
  - Incorrect translations
  - Insertion of placeholders & tags
    - Are you sure you want to unsubscribe from %S?

کیا آپ واقعی قوت دہ سے رکنیت ختم کرنا چاہتے ہیں؟

کیا آپ واقعی %S سے رکنیت ختم کرنا چاہتے ہیں؟
Translation Issues

- Translation selection
- Translation consistency
- Distinct translation mapping
  - Search/Find
  - Display/Show
- No capitalization in Urdu
Testing & Quality Assurance

- User testing
- Software specific testing processes
- Error categorization by severity level
  - High priority: software feature not usable
  - Low priority: software feature not user friendly
Build & Release

- Full installer release
- Language pack release

Urdu language pack available at http://psi-im.org/download/
Thank You