

Open Source Localization Toolkit

A project of **Open Forum of Cambodia**

being carried out under a

grant from **PAN ASIA ICT R&D Grants Programme** with reference 0401A4_L17

provided by: APDIP, ISOC, IDRC, AMIC and APNIC.

First Interim Report to the PAN ASIA ICT R&D Grants Programme

(Six month after start of project)

20 January 2005

INDEX

1	Synthesis.....	1
2	The problem being addressed.....	2
3	Preliminary findings	2
4	Change of Strategy	3
5	Project Output and Dissemination.....	3
6	Continuation of the project.....	5
7	Overall preliminary assessment.....	5
8	Financial Report	6
8.1	Forecast of expenditure for the next period.....	6
9	Signatures	7

1 Synthesis

In July 2004 Open Forum was awarded a grant by the **PAN ASIA ICT R&D Grants Programme** (donors: APDIP, ISOC, IDRC, AMIC and APNIC) to work on the development of an **Open Source Localization Toolkit (OSLT)**.

The **main objective** of the project is to reduce as much as possible the necessary research, work and expertise that a country or linguistic minority needs to undertake the project of creating a complete set of free computer programs that work entirely in the local language. At a more technical level, the project involves the localization (software preparation and translation) of Open Source.

More specifically, the project involves the development of an **Open Source Localization Toolkit**, formed by a **Manual** and a **CD**. The **Manual** will consist of a series of “How to” sub-manuals that include everything that a prospective localizer needs to learn: from management issues such as “How to decide what you need and how to write-up a Localization project” to the most technical parts, including installation of translation tools, finding and processing the files that need to be translated and – finally - building (physically creating) applications in local language. The toolkit will be based on the real work and experience of a localization team.

The Toolkit project is specially addressed to small countries that do not have the necessary expertise and knowledge to start a localization project by themselves, as well to national minorities that use their own script or language and wish to use for its people to enter the digital age.

Nevertheless, we believe that the information contained in the Toolkit will be of use to all Asia/Pacific localization projects.

The **Toolkit** project has very strong synergies with the **KhmerOS** project (Khmer software Initiative), as it gather experience from KhmerOS, and produces information that can be use also in future stages of KhmerOS.

2 The problem being addressed

Looking at the Open Source Software that is available round the world, it is clear that – only with Open Source - it is possible to gather the necessary set of programs to fulfill the needs of normal computer users without using any proprietary software (programs such as word processor, spreadsheet, e-mail, etc.).

All this Open Source software can be translated and localized (adjusted for local standards and tastes), meaning that it is possible to create a complete set of free programs in the local language that covers the needs of most users.

For those countries or regions that before did not have any programs in their own language, experience shows that this translation produces an important push in the use of computers, as it strongly reduces the training needs (training can be done in the local language), allows teaching from an earlier age and permits computerization of the Government in its local language.

The problem is that the Open Source community is sometimes highly technical and endogamic, so it takes a lot of time to gather the necessary information about the steps to be followed to localize text, and then, in some cases the procedures to be followed are not well documented. Gathering the necessary information to start a complete localization project, understanding the procedures, evaluating the necessary resources and preparing a plan can sometimes take up to six months of a trained computer scientist.

The **Open Source Localization Toolkit** attempts to provide the necessary information and resources to strongly reduce the time and technological knowledge necessary to prepare and carry on such a plan.

3 Preliminary findings

The project was first conceived as an umbrella project, gathering information about different OpenSource projects, their localization methods and their integration. Once the project was started, some findings made necessary a change in methodology.

- No localization information was available for the versions of the programs that will be available when this project comes out (such as OpenOffice 2.0, which will probably see the light in April 2005).
- For other projects, the information and methodology used are extremely technical, involving the use of development tools and high security methodology, which does not fit well with the spirit of the toolkit.
- Some projects with very low documentation level were key to coordinated localization projects (projects that localize several different programs).
- There was nothing written of how to develop a localization program.
- There was nothing written about strategies to follow in a localization process.

4 Change of Strategy

Without changing the goals of the project, a change of strategy was done, which involved direct action in the world of OpenSource, instead of placing an umbrella on top of it without affecting anything.

The reason for this change was that it is becoming clear that coordinated localization projects (projects that are social-goals oriented) can be highly improved if localization of all projects can be done using the same programs to localize all projects. In order to use the same program, the files that need to be localized need to be in the same format... and this is what the South Africa based **Translate Toolkit** project does: converting the files that need to be translated (for some of the key OpenSource projects) into a unique standard format (PO format, already used by some of the projects), and convert them back to the application's format once the translation is done.

Translation is therefore done with specialized editors (PO editors) that simplify translation and facilitate the work of translation teams, as they do not need to learn how to use different tools for translating different programs

A very important issue is that PO editors work well with Indic and left-to-right languages, making it especially interesting for Asia. .

5 Project Output and Dissemination

The main problem with the Translate Toolkit was that the available documentation was very deficient. AS good reference documentation of this software was very important for the **Open Source Localization Toolkit** project, developing such documentation became the first technical task of our project. Of course it was not so easy, as formal documentation revealed many issues (cases in which the programs did not work accordingly to the new documentation) and many fixes and interactions were necessary. The documentation developed by **Open Forum** covers at present the state of the **Translate Toolkit** project.

As a follow up to this, it was now necessary to link (through documentation) major OpenSource projects to the Translate Toolkit. This was easy to do with **OpenOffice**, as no specific way of translating it existed, and the OpenOffice project decided to integrate the toolkit as part of the build process, so that OpenOffice would directly produce PO files that could be translated with a PO editor. The translation back to OpenOffice format was done using the Translate Toolkit.

As OpenOffice 2.0 is coming out soon, no localization documentation was available. Open Forum started working on this documentation, linking it with the Translate Toolkit, and making reference to it. The documentation covered technical aspects of localization (modification of files, adding local information) as well as translation. This documentation is now almost complete, waiting for information that needs to be gathered about new parts of OpenOffice that might require localization (such as the installer). OpenOffice has requested that this documentation be checked at the OpenOffice website as the Localization information for OpenOffice 2.0. This fulfills our goals of assuring that a simple and well documented procedure exists for the localization of OpenOffice 2.0, and it is integrated with the rest of an

overall localization project. In short, Localization documentation developed by the **Open Source Localization Toolkit** will become the official localization documentation for **OpenOffice 2.0**, simplifying the localization procedure for this product and linking it to a global localization procedure.

While the **Open Source Localization Toolkit** is not yet a single unit, dissemination of the different parts that have already been developed is assured by the fact that they have been integrated into the applications and tools that they apply to, and can therefore be used not only as part of the **Open Source Localization Toolkit**, but also as independent pieces of information. These documentation are already being used by localization projects in Asia (Nepal, Lao)

On a less technical side, two more efforts are being carried at the heart of this project:

- A detailed manual on how to plan and write a localization project for a given country and language. It covers all aspects of the project, from defining its social goals to understanding the work-load and preparing a plan and a budget.
- A study on the role of localization in the acceptance of OpenSource software by final users. The goal of this study is to determine the best possible strategies for the planning and distribution of localized software, and integrating the in the above document.

These two projects are much less technical and require careful study and planning. They are both developed at a 50%, but they will not yet be available for a few months.

Even if these documents are not finished, they have already been used by an intern from East Timor (staying in Cambodia for three months) to develop a strategic localization plan for Tetum language in East Timor.

In order to keep close contact with important projects and acquire information, Javier Solá participated in the 2004 conference of OpenOffice in Berlin, participated in the review process of a similar project in Bombay being carried by CDAC and contacted some distributions to understand details of the packaging process, meeting one of them in Spain.

As part of the dissemination activities, Javier Solá will also participate in the AsiaSource event in Bangalore, India, at the end of January 2005, where he will act as a facilitator on localization, trying to facilitate the localization process for several different countries in the region.

6 Continuation of the project

The next six months will be dedicated to:

- Finish the above mentioned documents, providing a manual that will allow small economies to prepare their own localization plans and strategies.
- Finish the OpenOffice 2.0 documentation, including all aspects that will appear in the next two months, before the launch of OpenOffice 2.0.
- Work on a simple localization process for Mozilla products. The Mozilla Foundation has changed its philosophy. In the Mozilla product, Internationalization was considered a key issue. Mozilla was probably the most advanced internationalized product in the world of OpenSource. Unfortunately, the Foundation has now completely changed its priorities, giving priorities to trade-mark issues and its contracts with third parties (Google), forcing all their localizations to go through their CVS system, and catering only to western countries, without any concerns for other types of languages or for computers that cannot reach a 1024x768 resolution (most of the computers in developing countries). As the old (simple) mechanisms for localization are still there, we are planning to develop documentation that will explain how these products can be localized without having to forcefully go through the Foundation's new approach to OpenSource, using simple tools and making use of the excellent localization possibilities of the programs.
- Advance on the list of programs that are used by users, trying to figure out equilibrium between the amount of programs to localize and the percentage of users that will find in this list 100% of the programs that they use.

7 Overall preliminary assessment

While its goals stay the same, the **Open Source Localization Toolkit** project has become a real player in the world of OpenSource, helping produce necessary documentation that has value not only inside this project, but also outside of it, as part of specific localization projects.

A unified localization methodology was not part of what we planned at the beginning, as at the time the world of FOSS localization was much more scattered. Our work on it has not only simplified the documentation that we have to do, it has also simplified the world of localization itself. Through it we are not only reaching better the goals of our project, but also providing added value to existing localization projects.

Our belief is that we are right on track, serving the international community in a way that we could never have expected when we started this project.

8 Financial Report

The first installment of the grant, converted to US\$, the currency normally used in Cambodia, came out to be \$14,724.00, and was received in August 2004.

During the first six months the expenses have included 100% of the Capital equipment required by the project (\$1,232), 25% of the salaries budgeted (\$6,600), \$1,573 on traveling expenses and \$7 charged by the bank on reception of the grant transfer.

The expenses match budgeted amounts for this period. The only disparity is that the amount spent on traveling to Europe was higher than expected; it is – nevertheless - compensated by the fact that other traveling expenses that added value to this project (such as trips to India and Poland) were funded by other projects.

As \$14,724 were received, no interests were generated, and \$9,112 were spent in the first six month period, Open Forum still has \$5,612 belonging to the project that are not yet spent.

No.	Interim Financial Report 1			
1	CAPITAL EQUIPMENT			
	Breakdown	Budgeted	Semester 1	Available
	1.1 Development level computer	\$1,232	\$1,232	\$0
2	SALARIES			
	Breakdown	Budgeted	Spent	Available
	2.1 Project Manager (25% dedication)	\$12,000	\$3,000	\$9,000
	2.2 Computer Scientist (100% dedic.)	\$13,200	\$3,300	\$9,900
3	INTERNATIONAL TRAVEL			
	Breakdown	Budgeted	Spent	Available
	5.1 Travel within Asia	\$2,100	\$0	\$2,100
	5.2 Travel to Europe or the US	\$1,200	\$1,573	-\$373
4	BANK EXPENSES			
	Breakdown	Budgeted	Spent	Available
	4.1 Expenses related to Grant transfer	\$0	\$7	-\$7
	TOTALS (in US\$)	\$29,732	\$9,112	\$20,620

8.1 Forecast of expenditure for the next period.

During the next six month, the project will require an estimated amount of **\$7,300**, which cover \$6,300 in salaries (as budgeted) and an estimated US\$1,000 on traveling.

9 Signatures

As required, this report is hereby signed by the project manager of the OpenSource Localization Toolkit and the Financial Officer of the Open Forum of Cambodia.

A handwritten signature in black ink, appearing to be 'F. Solá', written over a long horizontal line.

Javier Solá
Project Manager

A handwritten signature in blue ink, appearing to be 'Nin Sokmalen', written over a long horizontal line.

Nin Sokmalen
Financial Officer



Certification

20 January 2005

The Open Forum of Cambodia NGO hereby certifies that:

- It has received in its FOREIGN TRADE BANK OF CAMBODIA bank account - from the **PAN ASIA ICT R&D Grants Programme** - the amount of **US\$14,724.00**.
- No interests have been credited by the FOREIGN TRADE BANK OF CAMBODIA in this period to the account in which the funds where stored
- The amount of **US\$9,112** has been spent and accrued to the Open Source Localization Toolkit Project up to date.

A handwritten signature in blue ink, appearing to read 'Nin SokMalen', with a long horizontal line extending to the right.

Nin SokMalen
Financial Officer
Open Forum of Cambodia