

Noolaham Foundation
Project Report

Project title	Text Digitization Project in Colombo 2007
Project number	NF/PG/2007/0002
Project location	Colombo
Sector	Print Publications Archive
Implementing agency and contribution	Noolaham Foundation
Grant Agency and Contribution	Contributors : Mr. T. Kopinath, Mr. V. Jasikaran, Mr. R. Pathmanaba Iyer, Mr. T. Thevathas
Total Budget and Expenditure	Expenditure: LKR 107,220.00
Start and End date	2005 - 2007
Responsible Stakeholder/s	Mr. Kopinath Thillainathan

Summary:

The Text Digitization Project in Colombo 2007 was a digitization project which was initiated to convert physical books into machine readable formats. It was one of the first funded projects carried out by Noolaham Foundation. Prior such funded projects, most digitization tasks were done by volunteers and they did not progress quickly. Some volunteers who were participating in the activities of Noolaham Foundation proposed that they would like to make monetary contributions and a project should be coordinated to digitize books, as they would not find time, expertise and resources to type the books manually. Thus the funded projects were started.

A total of 44 books were digitized through this project and made available online through Noolaham Foundation's Digital Library www.noolaham.org. This project was a successful initiative and received special appreciation from Sri Lankan Tamil writers and publishers.

Introduction and Background:

Project Noolaham was started in 2005 adopting the Project Gutenberg approach of involving volunteers contribute to the creation of an online ebook library. But as there was no practical OCR available for Tamil language, we had to type the selected books from cover to cover. The process of adding new e-books was quite slow because only a few individuals were able to type the books fast.

Moreover, only smaller books were added to the Noolaham digital library. Some volunteers who were participating in Noolaham Foundation's activities proposed that as they would not find time, expertise and resources to type the books manually, they would like to make

monetary contributions and a project should be coordinated to digitize books. Thus this funded project was started.

In the project, books were typed by professional typists, so more money was required as we expected. Thus Mr. S. Sritharasingh (year 2005) and Mr. P. Eelanathan (year 2006) made the first monetary contribution towards this project. In 2007, Noolaham appealed to volunteers to support such projects and more donations were received from individuals including Mr. R. Pathmanabhalyer, Mr. Kanaga Sritharan and Mr. Param Piratheepan. Even though this project has been documented under the year 2007, it should be noted that some components of the project were done in 2005 and 2006.

As Unicode typing was new to professional typists, books were typed in bamini font and then converted to Unicode. After proof reading, they were converted to the html format and uploaded at Noolaham's Digital Library.

This project was implemented in Colombo. T. Kopinath coordinated this project. He contributed towards quality control, proof reading, and conversion to Unicode and htm formats. Suratha's Pongu Thamizh convertor was used to convert the bamini text into Unicode.

Objective/s and Achievements or Results:

The objective of the Text Digitization Project in Colombo 2007 was to convert physical books into machine readable formats and to make them available online through Noolaham Digital Library (www.noolaham.net).

Forty four books of various kinds (About 3,600 A4 pages in Bamini font size 12) were digitized through this project. 44 books (49, 51, 88, 226, 229, 321, 348-380 44 books) were added to Noolaham Foundation's Digital Library (www.noolaham.org).

Constrains / Challenges and lesson learned:

- When Bamini font was converted to Unicode, the convertor automatically converted English texts into Tamil letters. Thus English words had to be checked and retyped again. This required us to proof read the books once more, and resulted in delays.. This is the second project in which Noolaham Foundation faced this issue and we were unable to get any volunteers to proofread. Thus Noolaham discontinued projects that involved typing.
- Up to this project, Noolaham functioned as a fully virtual endeavor. But having a dedicated office space made this project one that could be coordinated easily. We realized the usefulness of having an office space.

Suggestion and Recommendation:

- Adequate budget and human resources should be made available to enable quality outcomes and to help finish the project within the proposed period., It will help us in our future activities.
- The need for an office space was understood in this project.