

Noolaham Foundation

Project Report



Project Title	Digitization and Dissemination of Contemporary Tamil Newspapers Pilot Project
Project Number	NF/PG/2010/0002
Project Location	Colombo, Jaffna and Eastern Sri Lanka
Sector	Digital Library, Collection Development and Print Publications Archive
Implementing agency and contribution	Noolaham Foundation
Grant Agency and Contribution	Noolaham Foundation
Total Budget and Expenditure	Not Recorded
Project Period	2010
Responsible Stakeholders	Mr. Shaseevan Ganeshanathan, Mr. Santhirasegaramoorthy Ramanan

Summary

The **Digitization and Dissemination of Contemporary Tamil Newspapers Pilot Project** was carried out by Noolaham Foundation in 2010 to digitize the current issues of contemporary newspapers. It was a collaborative project undertaken by Noolaham Foundation in coordination with various publishers. The project aimed at collecting, preserving, and disseminating digital copies of printed newspapers. The main objectives this project were to provide open access of contemporary newspapers, preserve and create wide global visibility of the valuable information in these papers. The project was implemented as a pilot project to evaluate feasibility, time, cost, adverse events, and effect size (statistical variability) prior to a full-scale project.

Totally 246 valuable contemporary newspapers under four titles were digitized through this project and made available online through Noolaham Foundation's Digital Library at www.noolaham.org. These collections included 165 issues of *Thinakkathir*, 42 issues of *Thisai*, 10 issues of *Nihari* and 29 issues of *Aathavan* newspaper; the last two were published by Ravaya publishers, Colombo. These publications analyze social, political, cultural and judicial issues related to Sri Lanka. A web portal was also created so that all contemporary magazines can be accessed easily.

Introduction and Background

Noolaham Foundation is a non-profitable, collective and voluntary endeavor aimed at preserving books, magazines and other publications related to the Tamil speaking communities in Sri Lanka in e-format in order to make them available to and for the benefit of all those who are able to access the internet. Therefore Noolaham Foundation has a digital library which provides free public access of e-documents online. Until 2009 Noolaham Foundation did not have bigger scanners to digitize newspapers. In 2010, A3

sized scanner was obtained so that bigger publications such as newspapers could be scanned.

Newspapers are typically published daily or weekly. A newspaper is a periodical publication containing typically published news articles and feature articles on national and international news. The news includes political events and personalities, business and finance, crime, weather, and natural disasters, health and medicine, science and technology, sports and entertainment, society, food and cooking, clothing and home fashion, and arts. Printed newspapers have been, for centuries, the most popular source of information for the public and a feature of democracy. They are connected with the freedom of expression of the people, political parties, social and financial groups, scientists and artists. Even after the introduction of broadcasting technology, newspapers remained a very important primary source of news. Newspapers still greatly influence the formulation or the expression of the public opinion by providing the necessary podium for journalists to express their personal views and to analyze political, economic, business or cultural events.

As Mahatma Gandhi said, 'One of the objects of a News newspaper is to understand the popular feelings and give expression to them; another is to arouse among the people certain desirable sentiments and the third is to fearlessly expose popular defects.' The term "newspaper" covers surprisingly wide range of publication of comprehensive and objective information on all aspects of the country's social economic, political and cultural life. Newspapers put everyone in touch with the active world. The newspaper, in the present era, is the most powerful media in educating, informing as well as entertaining people. Published with a huge volume and wide variety of information, it makes itself distinct from other media. It is used regularly by people from each and every strata of the society. Besides, academicians and professionals from diverse fields considerably use it retrospectively for their research and/ or application activities. However, it has been observed that due to the absence of an efficient and effective newspaper information retrieval system in Sri Lanka, the retrieval of such information is being hampered, and the multidimensional development activities have been severely affected by the same. The libraries in Sri Lanka are not easy to give a precise definition of reference work.

Today in the digital world, digitization of library information resources has advanced rapidly. Digitization solves problems faced by traditional libraries like conservation, preservation, storage and space, and provides effective services to cater to the user's requirements. That being said, due to the sharp influence of technology, which has reshaped the lives of thousands of Tamil people, the ethics of the information have changed and many prefer to read displayed digitized papers to perusing an old-fashion printed newspaper. The computer technology avalanche has even changed the news reporting ethics. Digital newspapers are capable of constantly updating their contents and report news almost instantly as the right of making available to the public. Among the Sri Lankan Tamil speaking community there is no systematic funding for the digitization of historic and contemporary newspapers. These were considered when the project was initiated by Noolaham foundation. The rationale for the entire project is to make information and knowledge available to future generations that will base internet as one of the primary sources for intellectual purposes.

Nihari and *Aathavan* were published by *Ravaya* publishers, Colombo. Both *Aathavan* and *Nihari* were published on a weekly basis. *Thinakkathir* is a daily newspaper based in Eastern Sri Lanka and published by the World Voice publishers at Eastern graphics Printing Press. *Thisai* was a weekly newspaper based in Jaffna, Sri Lanka and published by the New Era publishers. These four titles were chosen by NF for this project.

Objective/s and Achievements or Results

The main purposes for the 'Digitization and Dissemination of Contemporary Tamil Newspapers Pilot Project' were to promote wider access to information among a large number of users, to preserve and conserve the national, regional and cultural heritage through digitization of back volume of newspapers, which constitute the raw materials of history, to provide user-friendly interface to access the news content and enhancing indexing and search ability of the newspapers, to search across the different newspaper titles and to draw together materials relating to a wide range of research and learning topics and to access to all or part of the content of different newspapers on the World Wide Web. Especially this project focused on Noolaham's objective of Collection Development.

This project was the first large scale newspaper digitization project in Sri Lankan Tamil. During the project, the original paper copies were used as a source for digital images during the scanning process. Scanning was conducted in NF using A3 size flatbed scanners. Each scanned newspaper page was treated as a separate image and saved as an archival master at 300 dpi. During the project NF followed guidelines based on digital library standards and best practices.

Totally 246 valuable contemporary newspapers under the four titles were digitized through this project and made available online through Noolaham Foundation's Digital Library at www.noolaham.org. These collections included 165 numbers of 'Thinakkathir', 42 numbers of 'Thisai', 10 numbers of 'Nihari' and 29 number of 'Aathavan' newspapers. Web portals were also created for each paper for easy access and could be downloaded in the user friendly interface in pdf format. During the project period huge numbers of newspapers were identified and collected by NF. Altogether 1070 newspapers were collected under seven titles. Those are 496 of Thinamurasu, 214 of Sarinihar, 114 of Puthiya Poomi, 29 of Aathavan, 10 of Nihari, 165 of Thinakkathir, and 42 of Thisai. Among them Aathavan, Nihari, Thinakkathir, and Thisai titles were chosen for this project because this project was approached as a pilot project. Because of the magnitude of the newspapers collections, Thinamurasu, Sarinihar and Puthiya Poomi newspapers were digitized under separate projects.

NF strongly believes **this** project will build good knowledge among the Sri Lankan Tamil speaking community and improve its aspects on social policies. At its best, a digitized newspaper makes its information searchable and accessible to anyone with internet access. Further some other secondary functions will be enhanced through this project like campaign for desirable civic products and to help eliminate undesirable conditions, to give the readers a portion of entertainment through such devices a comic strips, and cartoons,

and to publish informative and analytical features often written by special columnists on issues of topical interest. In addition digitalizing the collections will enhance access and improve preservation. It offers opportunities to preservers and provides access to this content. Digitalization initiatives will allow users to search collections rapidly and comprehensively from anywhere at any time.

Constraints / Challenges and lesson learned

During this digitization project NF has been confronted by a number of challenges

- Offering scanned images without any fully searchable text would only partially fulfill users' needs, as many (if not all) desire to search all of the texts available to them. It is necessary to consider providing navigation within the digital data through the use of indexing software, even though NF provided metadata and content as well.
- During the project period huge numbers of newspapers were identified and collected but all of them would not be digitized through this project because of insufficient funds and lack of human resources.

Suggestions and Recommendations

- Once the digital image is created it needs to be made searchable and accessible. The images are processed with optical character recognition software to make the books text searchable. But there is no such software available in the market which satisfactorily provided the required results in Tamil language. Descriptive metadata and content must be created for each title, issue, image, and possibly article to help the users find what they are looking for.
- NF should digitize the precious collections when funds and human resources are available.
- Noolaham Foundation should collaborate more and more with other publishers and groups in creating digital content to share information and resources.

Acknowledgements

Noolaham Foundation would like to thank everyone for their support to implement this project successfully. These contributors participated in various activities such as making contact with publishers and repositories, provide their newspaper collections to NF for digitization, finding any missing newspapers, providing A3 scanners to scan the newspapers, participate in scanning, providing technical support and launching the digital copies on the digital library.

