Pay it Forward

The Tamil Digital Heritage Project

Chitra Sankaran charts the history of a community-driven project that has seen the digitisation of 350 homegrown Tamil literary works.

Associate Professor Chitra Sankaran teaches in the Department of English Language and Literature at the National University of Singapore. She is the author of several academic books as well as a murder mystery novel set in Singapore and India. She is a key member of the Tamil Digital Heritage Project team.

beavers are a suborder of rodents found mainly in North America. They are natural conservationists and famously hardworking, using whatever available tools at hand to improve their living environment and make a better future for themselves. People will be surprised to hear that of late, another suborder of beavers has been identified in Singapore. Only, these beavers – the members of the Tamil Digital Heritage Project (TDHP) team – have been working hard to digitise important literary works in Tamil.

How it All Started

In 2013, a small group of concerned Singaporeans came together to think of a way to commemorate the nation's 50th birthday and do something that would not only benefit the community, but would last well beyond the SG50 celebratory period. Thus was born the idea of the Tamil Digital Heritage Collection – a project to digitise Singapore-related Tamil literary works published over the last 50 years.

Just as beavers identify a problem and then work studiously to set it right, this group of far-sighted people realised that many valuable Tamil publications had disappeared from the public domain due to various reasons, including the death of an author, copies not being deposited at the National Library, and so on.

The TDHP team realised that something has to be done to address this problem. The market for Tamil books in Singapore is not robust, and in fact, most of these books are not even available in local bookshops. While such books could be useful literary resources for schools, they are infrequently used due to their unavailability. Given these circumstances, the TDHP team realised that the best way forward was to digitise the books so that they would be made accessible not only to Singaporeans but also to interested readers from all over the world.

As a first step, the team identified the following seven criteria to select books for the digitisation project:

1. Written in Tamil;
2. Literary creations such as poetry anthologies, novels, short stories, literary essays, etc;
3. Authored by either Singaporeans or Singapore permanent residents;
4. Feature Singapore or relating to Singapore for books written by foreigners;
5. Published between 1965 – the year Singapore became an independent republic – and 2015, when it celebrates 50 years of independence;
6. Part of the National Library Board's collection; and
7. Where written permission had been granted by authors to digitise the book.

Based on these criteria, approximately 350 books were selected, ranging from novels to short stories and poems, and featuring 80 prominent Tamil writers and poets, including K.T.M. Iqbal, P. Krishnan, M. Ilankannan and Jayanthi Shankar.

The real work began at the end of 2013. The first step was to obtain support from the various authors. Most came forward willingly, realising the immense benefits the project would bring to their works and also towards preserving the literary heritage of Singapore-Tamil books. Most importantly, the books would have a permanent home that would provide easy access to anyone interested in Tamil literary works. A secondary motivation was that the project would be pivotal in connecting the Singapore-Tamil literary community to the larger Tamil diasporic world. Fortunately, only a handful of authors declined to join the project.

Community Support and Funding

The TDHP team had the support of four public organisations from the very start – the National Library Board, National Heritage Board, National Arts Council and the National Book Development Council. With the help of these organisations, the team was able to proceed with the project. They were supported, over time, by a pool of nearly 250 volunteers, mostly from the Tamil teaching community, who undertook various tasks – much like beavers dedicated to their specific tasks.

The first task was to obtain official consent from authors whose books had been selected for digitisation. Although these books were physically available in the National Library Board's (NLB) collection, copyright belongs to the authors and related parties, and special permission was required to digitise these books and make them accessible in an online repository.

The next task was to create annotations for the texts. In order for digitised books to be searchable online, they need to have clear descriptions with appropriate keywords. A group of nearly 50 volunteers, including teachers, authors and other professionals, volunteered their time and services for this important task. They attended workshops conducted by team members to learn the rudiments of annotating texts. Based on the knowledge acquired at these workshops, the volunteers beavered away and managed to complete annotating all the books within the short span of three months.

The NLB offered to take on the expense of digitisation. Various vendors were invited to submit quotations and proposals, which were studied and evaluated by NLB, with the TDHP team providing technical advice on the Tamil language. Eventually, two vendors with the best technological know-how and expertise were selected. Finding the right service provider was not easy as the project needed to tap on the expertise of a company that not only satisfied the technical requirements but was also well versed in the Tamil language. There were few vendors who demonstrated these abilities.

The TDHP team, along with another large group of willing volunteers, then took on the painstaking task of proofreading and matching digitised copies to the originals. Computers are not foolproof; the electronic conversion of text and images into digital copy through Optical Character Recognition (OCR) software does inadvertently introduce errors in the scanned copy that can only be picked out by human effort.

Proofreaders involved in the Tamil Digital Heritage Project team hard at work.

Concurrently, experts began looking into ways in which the digitised books could be made searchable using appropriate and effective search engines that not only facilitate the search of a specific text but also articles and other materials related to the text. But that was not the end of it. The project team also wanted users to be able to copy, paste and collate information using the digitised collection.

It was patently clear to all involved that the scale of the project was unprecedented – this was the first time that such a massive and concerted effort at digitising an entire national collection of Tamil works was taking place. Another notable aspect of this project is the community element: it is a people-driven project, initiated by and for the benefit of the community. The push for the TDHP did not come from top leaders or community organisations, but from well-meaning citizens who initiated the idea, sought community support and obtained the patronage of S. Iswaran, Minister in the Prime Minister's Office and Second Minister for Home Affairs and Trade and Industry. The TDHP clearly debunks the idea that Singaporeans are apathetic and lack initiative, and that all large projects are top-down efforts that emanate from the government.

In order to underscore this "common-man" approach, the TDHP team also made a special effort to solicit funding of small amounts from a large number of people rather than large amounts from a small group of wealthy benefactors. The team launched a campaign to collect just $50 per adult and $2 per student. The ongoing target is to secure donations from at least 1,000 members of the public. The names of donors will be featured on the TDHP website.

Launch and Future Plans

The project will benefit Tamil-language speakers from all walks of life, including students, teachers, authors, researchers and academics as well as laypeople, who can use the collection as a source for research, as social historical documents, as inspirations for their own creative works and in a multitude of other ways. Indeed, not only Singaporeans but people living in any corner of the world can now access valuable Tamil publications with a mere click of the mouse.

More than a year and a half after its inception and many hours of painstaking work, the Tamil Digital Heritage Collection was launched on 22 August at the National Library Building. The occasion was graced by Deputy Prime Minister and Minister for Finance Tharman Shanmugaratnam, and by Minister S. Iswaran, the project's patron. The evening's entertainment was based entirely on content taken from the digitised collection and featured a unique showcase of music, dance, recitals and songs featuring talented artists from the Singaporean Tamil community.

In his speech, Minister Tharman hoped that some of the digitised books could "be translated into our other languages so that our whole nation can benefit from the memories of a small community". He added that Singapore's policy of multiculturalism, which gives equal opportunities to all races, has meant that even minorities have felt their culture can be nurtured and developed.

Minister Tharman likened the work of the TDHP team to the labour of a great Tamil scholar more than a century ago, Swaminatha Iyer, popularly known as "Thamizh Thatha" or "the Grandfather of Tamil". He collected all the ancient Tamil classical texts which were in the form of decaying palm leaves scattered around the state of Tamil Nadu and brought them to print. This was his lifelong work over five decades.

The TDHP has come to fruition through the handiwork of ordinary Singaporeans who have dedicated themselves to the task of preserving their linguistic heritage. They are truly the beavers of Singapore. An interesting fact about beavers is that they grow in size throughout their entire lives. The Singapore "beavers" behind the TDHP see their efforts as only the starting point of a project that will flourish through the generations and gradually include all genres of Tamil books, perhaps even books published before 1965.

It is also envisaged that the project will serve to inspire communities to initiate similar projects in their own countries. It is not inconceivable that one day in the distant future a child will go online and research a school project on "Singapore in the 21st Century Through the Eyes of Singaporean Tamil Writers". We hope that when that day comes, the body of Tamil literature deposited in the Tamil Digital Heritage Collection would have grown exponentially. ◆

To access the Tamil Digital Heritage Collection, visit NLB's BookSG website at http://eresources.nlb.gov.sg/printheritage/browse/Tamil_Digital_Heritage_Collection.aspx

Deputy Prime Minister Tharman Shanmugaratnam launching the Tamil Digital Heritage Collection. Looking on are S. Iswaran, Second Minister for Home Affairs and Trade and Industry, and Elaine Ng, Chief Executive Officer, National Library Board.