

Santhosh Thottingal

Language engineer & typeface designer

santhosh.thottingal@gmail.com

thottingal.in

github.com/santhoshtr

Profile overview

I am a software engineer with 12+ years experience with special interest and expertise in language computing. I am also a digital typeface designer and designed two well known typefaces for Malayalam script.

- Experience in building language technology infrastructure for low resource, complex script languages using open source technology by building developer community around it.
- Built core language tools for Malayalam during the early days of digital exposure and they are used by hundreds of thousands of users daily.
- Strong knowledge in opentype engineering with complex scripts, maintains a dozen of widely used fonts for Malayalam, Tamil scripts.
- Digital typeface designer with unique aesthetic sense on the complex scripts. Fonts designed are widely used by thousands of users for various purposes.
- Experience in technology leadership to architect, build and deploy massively multi lingual language tools to millions of users.

Education

Bachelor of Technology Degree in Computer science and Engineering, University of Calicut - 2005

Professional Experience

Senior Software Engineer, Wikimedia Foundation Since 2011

- Responsible for Mediawiki internationalization and localization support and Wikimedia interface localization tools and workflows
- Designed and developed an input tool system for 150+ languages with the help of community and available in all wikipedias.
- Lead engineer for language technology tools and systems in Language team of Wikimedia Foundation

- Architected the Wikipedia article translation system which so far created more than 200,000 new wikipedia articles in more than 280 languages. Uses machine translation and multiple translation automation tools.
- Designed and developed Universal Language Selector project of WMF, a language selection tool and utility library

Software Engineer, Infosys Limited. 2005 - 2011

- Worked as J2EE specialist for various clients in Aerospace and Automotive industry domains.

Designer Portfolio

- Designer of well known Malayalam fonts Chilanka and Manjari.
- Maintainer of 12+ fonts for Malayalam and Tamil. I have deep understanding and experience in opentype engineering for Indian languages.

Language technology projects and Research

- Malayalam morphology analyser using Finite State Transducer technique. In progress. Also defining and implementing syllable models for phonetic analysis of text related to this.
- Developer and Maintainer of Malayalam Spelling Checker dictionary. Part of many GNU/Linux Distributions.
- Defined Hyphenation Patterns of 11 Indian Languages. Used in various typesetting applications such as LaTeX. Also in Android and browsers
- Designed and developed LibIndic Indic language computing framework and library. This is collection of language processing libraries and utilities for Indian languages.
- Developer and Maintainer of Swanalekha transliteration based phonetic input method for Malayalam, Hindi, Kannada, Bengali, Oriya, Gujarati, Punjabi, and Telugu. Available in GNU/Linux distros. Used by thousands of people.
- Designed and developed general purpose input method library jquery.ime used in Wikipedia, which later ported to Android and Firefox OS
- Developed various small opensource utilities around Indic language processing.

Open Source contributions and community experience

- Developed Dhvani Indian Language Text to Speech System - A TTS that can speak Hindi, Punjabi, Gujarati, Bengali, Oriya, Telugu, Kannada, Tamil, Malayalam, Marathi and Pashto. Won FOSS India Award 2008.
- Defined and implemented Collation rules for Malayalam, Tamil, Sourashtra languages in for GNU C Library.
- Participated in localization of Debian, GNOME, KDE, Firefox projects.

Mentoring and leadership

- Project Administrator, Swathanthra Malayalam Computing(SMC)
- Mentor, Google Summer of Code 2007, 2013, 2014 and 2016.
- Language Committee member of Wikimedia foundation.
- Former coordinator for Openoffice native language consortium, India region
- Represents Wikimedia foundation in Unicode consortium.
- Panel Member, Neo Brahmi script generation panel, ICANN

June 10, 2017.