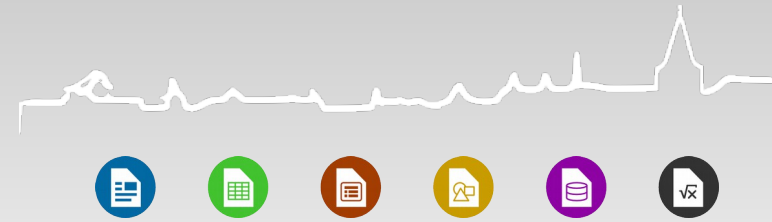


Localising LO to Tamil & Its Deployment in Malaysian Tamil-Medium Public Schools

- ▼ Elanjelian Venugopal
- ▼ tamiliam@gmail.com

- ▼ Arun Kumar
- ▼ thangam.arunx@gmail.com

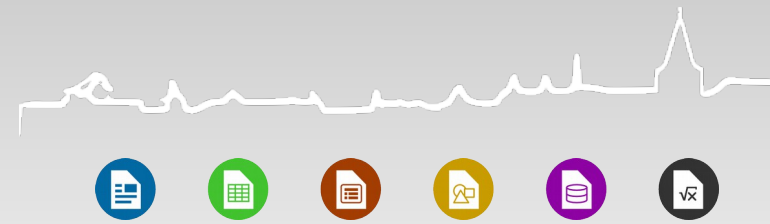




Project's guiding principles:

- ▼ To make LO accessible to Tamil speakers
- ▼ High quality and Standard Translation
- ▼ Wide applicability:
 - ▼ Government agencies
 - ▼ Schools (teachers & students)
 - ▼ General public

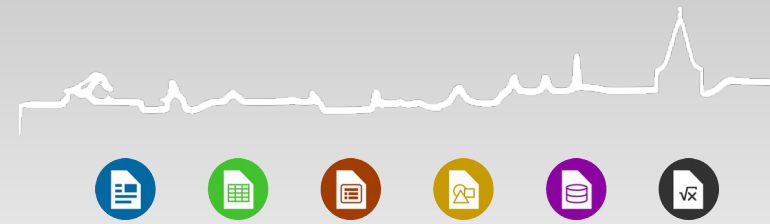




- ▼ **Two key challenges:** 1) Terminology & 2) Context Sensitivity
 - ▼ Terminology: The standardisation of technical glossary in Tamil is still a work in progress, & unstable. Currently:
 - ▼ Tamil Virtual Academy technical glossaries
 - ▼ FUEL initiative
 - ▼ Private efforts by Free Tamil Computing volunteers
 - ▼ The spoken and written Tamil is often vastly different, thus making the translated strings artificial:
 - ▼ Eg. *slide* vs *villai*; *label* vs *vilakkaccheettu*
 - ▼ Challenge is to find a balance that is acceptable



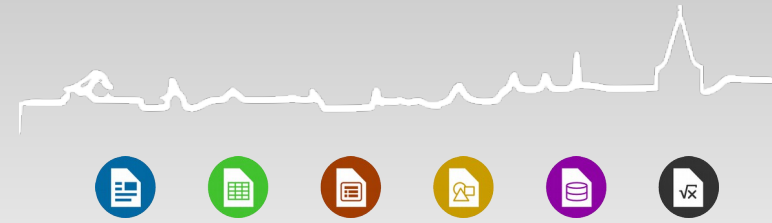
L10N Challenges in Tamil



- ▼ Context Sensitivity: In Tamil, it is particularly difficult to translate merely with glossary.
 - ▼ Noun-Verb confusion:
 - ▼ **Design** (n) = '*vadivamaippu*'; Design (v) = '*vadivamai*'
 - ▼ **View** (n) = '*paarvai*'; View (v) = '*paar*'
 - ▼ Ambiguity in English words:
 - ▼ Yes = '*aam*' vs '*vEndum*'
 - ▼ No = '*illai*' vs '*vEndam*'
 - ▼ Practical issues:
 - ▼ While translations could be made context-sensitive by checking the terminology or sentence (or sentence fragment) usage – perhaps by using KeyID features – it is difficult to teach volunteers
 - ▼ Problem repeats when new strings are introduced



L10N Challenges in Tamil



▼ An Example:

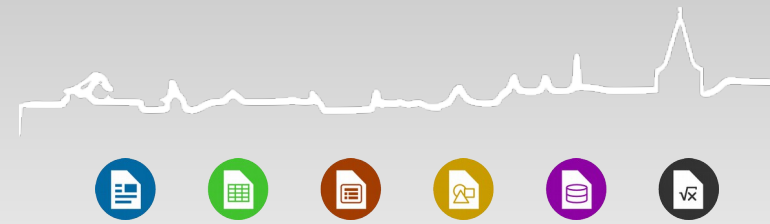
▼ Tamil - Subject + Object + Verb(SOV)

▼ *\$overall\$* இல் *\$current\$* ஆவணம்

▼ English - Subject + Verb + Object(SVO)

▼ *document \$current\$ of \$overall\$*

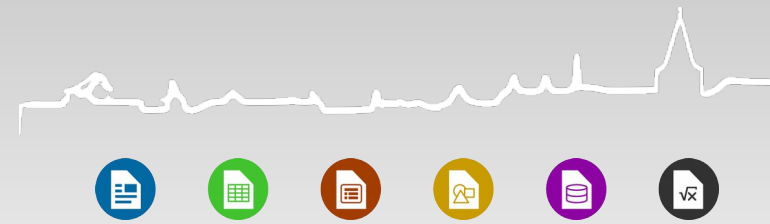




▼ Dominance of English among Educated Tamils

- ▼ Many Tamils consider themselves sufficiently fluent in English and do not show much interest to use software (let alone LO) in Tamil
- ▼ Since 1990, many attempts have been made to teach Microsoft Office in English to Tamil School students. Results are mixed. Only very basic mastery was achieved.

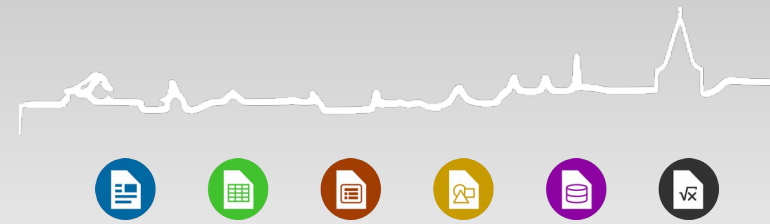




▼ Titian Digital (Digital Bridge) Experiment

- ▼ In 2009, the Tamil L10N team in Malaysia initiated a Foundation-funded project to:
 - ▼ Assess the effectiveness of teaching ICT using native language;
 - ▼ Create a market for Tamil UI; and
 - ▼ Customise applications to process/handle Tamil.
- ▼ The project used Edubuntu linux distribution that include LO as default office application. Also includes:
 - ▼ Specific curriculum with supporting workbooks
 - ▼ Regular teacher training and coaching



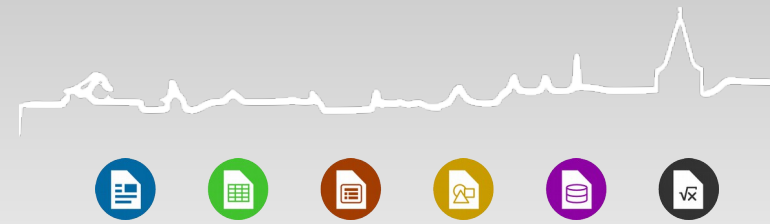


▼ Titian Digital today

- ▼ Since 2009, the Project has grown--
 - ▼ From 2 schools to 72 schools, and has secured funds for 20 more;
 - ▼ From serving 430 students/year to around 26,000 students/year.
- ▼ Other developments:
 - ▼ The Malaysian government is major contributor since 2012
 - ▼ Many free software enthusiasts have started to emerge
 - ▼ Students are now training to participate in software competitions
- ▼ Next Audacious Goal: to convince Tamil Nadu Government on the merits of teaching IT through Tamil



To Break Challenges



▼ **And, finally**

Hello from Malaysia!

▼ `/home/tulashi/Downloads/Libre Office.mp4`



Aarhus 2015 CONFERENCE



Thank You!

- About us: <http://thamiZha.org>
- LibreOffice: <http://ta.libreoffice.org>



All text and image content in this document is licensed under the Creative Commons Attribution-Share Alike 3.0 License (unless otherwise specified). "LibreOffice" and "The Document Foundation" are registered trademarks. Their respective logos and icons are subject to international copyright laws. The use of these therefore is subject to the trademark policy.

