

# Open-Tamil - திறமூல தமிழ் நிரல் தொகுப்பு

அருளாளன், சையது அப்தாகிர், பரதன் தியாகலிங்கம், சூனிவாசன், சத்தியா மகாதேவன்,  
அருண்ராம், மற்றும் முத்து அண்ணாமலை.  
அனுகூலம் மின்னஞ்சல்: [ezhillang@gmail.com](mailto:ezhillang@gmail.com), நாள்: ஜூலை 1, 2020.

## 1. அறிமுகம்

ஒப்பன் தமிழ் என்பது ஒரு திறமூல நிரல் தொகுப்பு திட்டம். இது எழில் கணினி மொழியில் ஆக்கத்தை  
தொடர்ந்து தமிழில் பலரும் எளிதாக கணினி செயலிகளை பைத்தான் மொழியில் உருவாகவேண்டும் என்ற  
நோக்கில் எழிலின் ஒரு கூற்றாகப் பிறப்பெடுத்தது. இந்த நிரல் திட்டம் முதலில் பைத்தான் மொழியில்  
வெளிவந்தது - பின்னர் சில சேவைகள் மட்டும் ஜாவா, ரூபி மொழிகளில் வழங்கப்பட்டன - எனினும்  
பெரும்பாலான வசதிகள் பைத்தான் மொழியின் வாயிலாகவே பெறமுடியும்.



படம். 1: தமிழ் பேசு திட்டத்தின் சின்னம்.

## 2. கட்டமைப்புகள்

இந்த நிரல்தொகுப்பிலுள்ள மொட்யூல்களாவன கீழே. இவற்றின் முழு விவரங்களையும் காண  
[http://tamilpesu.us/static/sphinx\\_doc/\\_build/html/sphinx\\_doc/](http://tamilpesu.us/static/sphinx_doc/_build/html/sphinx_doc/) இங்கு செல்லலாம்.

	Module	பயன்பாடுகள்/சார்புகள்
1	tamil	Tamil tokenization, word ordering, encoding converters, numerals, text summarizer.
2	ngram	corpus modeling classes

3	solthiruthi	Tamil spelling checker algorithms
4	spell	Tamil spelling checker application
5	tamilmorse	Morse code generation, decoding for Tamil
6	tamilsandhi	Tamil sandhi-checker - packaged with Open-Tamil but developed independently by Nithya and Shrinivasan.
7	transliterate	Tamil transliteration tools
8	tamilstemmer	This module is new in version 0.96 and provides access to simple stemmer functions originally created by Damodharan Rajalingam
9	tabraille	Tamil Braille generation following Barathia Braille standard
10	kural	Thirukkural source text and English translation

Open-Tamil source code examples like numeral to audio generation, ngram generation, corpus analysis etc. see link [here](#).

### 3. வெளியீடு, உரிமம், நிறுவுதல்

2015-இல் முதல் வெளியீடு (வரிசை எண் 0.4) கண்டு பின்னர் இந்த ஆண்டு ஜூன் 12-இல் சமீபத்திய (ஒன்பதாம்) வெளியீடு (வரிசை எண் 0.97) கண்டது. இந்த நிரல் தொகுப்பு MIT உரிமம் வழியாக நீட்சி செய்தும், பகிர்ந்து மறுசெயல்பாட்டிலும் உபயோகிக்கலாம்.

சமீபத்திய வரிசை எண் 0.97-இல் வெளிவந்த புதிய அம்சங்களானவையாவன:

1. மாத்திரை கணித்தல் – தமிழ் உரையில் உள்ள சொற்களின் மாத்திரை அளவை கணிக்க புதியசார்பு 'tamil.utf8.total\_maaththirai()' என்று திரு. பரதன் தியாகலிங்கம் அவரால் பங்களிக்கப்பட்டது.
2. வடமொழி சொல்பட்டியல் மோனியர்-வில்லியம்ஸ் அவரது அகராதியில் இருந்து திரிக்கப்பட்டு இங்கு சேர்க்கப்பட்டது
3. 'tabraille' என்ற module-இல் கண்பார்வை குறை உள்ளவர்களினால் தமிழ் பாரத பிரெயில் என்ற தரத்தை கையாளும் வகை சில உத்திகள் உள்ளன.
4. 'kural' என்ற module-இல் திருக்குறளை நேரடியாக கையாள சில உத்திகள் உள்ளன. இது 2013-இல் வெளிவந்த 'libkural' என்பதன் மூலப்பதிவாகும்.

இதனை நிறுவ இப்படி கட்டளை கொடுக்கலாம்,

```
$ pip install open-tamil
```

ஏற்கனவே நிறுவியிருப்பின் புதிய அத்தியாயத்தில் நிறுவ, என்றும் கொடுக்கலாம்.

```
$ pip install --upgrade open-tamil
```


## 4. வளர்ச்சி

ஒப்பன்-தமிழ் திட்டம் இதனைக்கொண்டு பல மென்பொருட்கள் இன்று இயங்கிவருகின்றன - இவற்றில் முக்கியமானவை <http://tamilpesu.us> என்ற வலைத்தளம். இந்த நிரல்தொகுப்பில் இருந்து செயல்பாடுகளை மொத்தமாக வலைவழியாக தமிழ் ஆர்வலர்கள் கணிமை செய்யாமல் பயன்படுத்த இது உதவும்.

Open Tamil Web

Eng Documentation

ஆவனமாக்கம்



பயன்பாடுகள்

உரை ஒலி மாற்றி

சொல் தேடல் - குறுக்கெழுத்து

தமிழ் எண்கள்

Tamil Stemmer

தமிழ் சந்திப்பிழை திருத்தி

பெருக்கல் அட்டவணை

சொல் எண்ணி

ஒலிபெயர்ப்பி

தமிழ் எழுத்துப் பிழைதிருத்தி

ஒலிவழி (புனிகோட்) மாற்றி

### வாய்பாடு

இந்த பயன்பாடு உங்களுக்கு பெருக்கல் அட்டவணை தமிழ் மொழியில் உருவாக்க உதவி செய்யும்

எண்

பெருக்கல் அட்டவணை

7

10

calc

ஏழு X ஒன்று = ஏழு [ 7 x 1 = 7 ]  
ஏழு X இரண்டு = பதினான்கு [ 7 x 2 = 14 ]  
ஏழு X மூன்று = இருபத்தொன்று [ 7 x 3 = 21 ]  
ஏழு X நான்கு = இருபத்தெட்டு [ 7 x 4 = 28 ]  
ஏழு X ஐந்து = முப்பத்தைந்து [ 7 x 5 = 35 ]  
ஏழு X ஆறு = நாற்பத்திரண்டு [ 7 x 6 = 42 ]  
ஏழு X ஏழு = நாற்பத்தொன்பது [ 7 x 7 = 49 ]  
ஏழு X எட்டு = ஐம்பத்தாறு [ 7 x 8 = 56 ]  
ஏழு X ஒன்பது = அறுபத்து மூன்று [ 7 x 9 = 63 ]  
ஏழு X பத்து = எழுபது [ 7 x 10 = 70 ]

படம் 2: ஒப்பன்-தமிழ் வழி உருவாக்கப்பட்ட தமிழ்பேசு வலைதளத்தில் உள்ள பெருக்கல் அட்டவணை செயலி.

ஒப்பன் தமிழ் கொண்டு பல தமிழ்இயல்மொழி ஆய்வுகள் (உதாரணமாக Tamil NLP, PyTamil) என்ற திட்டங்களும் செயல்படுகின்றன. இது எங்களுக்கு தெரித்தவை மட்டுமே!

## 5. பங்களிப்பாளர்கள்

மற்ற திற மூல மென்பொருட்களைப்போலவே ஒப்பன்-தமிழ் இதன் உருவாக்கம், மற்றும் வளர்ச்சி கிட வலைத்தளத்தில் வழியாக நிர்வாகிக்கப்படுகிறது. இதன் சுட்டி - <https://github.com/Ezhil-Language-Foundation/open-tamil> வழில் மொழி அறக்கட்டளையின் பார்வையில் இது மேம்படுத்தப்பட்டாலும், இதன்வழியாக பத்துக்கும் மேற்பட்ட பங்களிப்பாளர்கள் உள்ளனர்.இந்த திட்டம் ஏரக்குறைய 800 பங்களிப்புகளை பெற்றும், 114 வழு/திறணாம்சங்களையும் முடிவுபடுத்தியும், மேலும் 82 திறணாம்சங்களை ஒழுங்கு செய்தும் வடிவமைப்புக்காக குறிக்கப்பட்டுள்ளன.

இந்த திட்டத்தை அனைவரும் தொடர்ந்து பயன்படுத்தியும், ஆதரிக்குமாறும் கேட்டுக்கொள்கிறோம்.