

Introduction to Tamil Information Technology

Dr. K. Kalyanasundaram
Swiss Federal Institute of Technology,
CH-1015 Lausanne, Switzerland
<chair@infitt.org>



Workshop Outline

- Tamil IT Basics (Dr. K. Kalyanasundaram)
- Tamil IT Projects and vision using Web2.0 and Mobile Platform (Mr. Nagarajan)
- Practical Demo of Tamil in different computers (Mr. TNC Venkatarangan)



Tamil IT Basics (Language Technology)

- Tamil Implementation on computers: fonts and encoding issues
- Tamil Internet: web delivery & features
- Standard and Advanced Applications



What is Tamil Computing?

- Ability to **do everything** that is technologically feasible in computers **natively in Tamil using Tamil scripts**
- கணினியில் தமிழில் தகவல் பரிமாற்றம்
- **Long history:** softwares since mid-eighties (pre-PC era)



Tamil computing for ordinary

(use without knowing the technical details)

- Word-processing and information interchange in Tamil via email, list discussions and blogs
- Accessing websites with Tamil content for current info. (ezines, portal sites, blogs,....)
- Accessing websites for e-resources (Tamil digital library, e-dictionary, online catalogues)
- E-commerce: billing in Tamil, online shops
- E-governance (voter, land registration, marriage certificate,)



Tamil Computing :

Advanced applications

- Data mining of e-resources using search engines
- Auto-recognition and extraction of Tamil text (OCR)
- Voice recognition /text-to-speech synthesis
- Natural Language Processing (Machine-translation of texts, lexical resources, computational linguistics)
- E-commerce: Online stores, billing, inventory
- E-governance
- Online and computer-aided Teaching of Tamil
- Localisation of Free and Open Source Softwares (FOSS)
- Tamil in Mobile and Hand-held/Portable devices



Tamil implementation (Windows, Mac and Unix)



Tamil computing - pre 2000

- **Use existing shrink-wrap softwares made for English and Tamil Fonts**
- **Peggy-riding Tamil on some other language English** (overwrite 7- or 8-bit ASCII space with Tamil characters)
- **Worked** for two decades - even today nearly all computer-set Tamil printing is done this way
- **Not always satisfactory:** problems such as undesirable line-break, collation,...
- **Long term solution:** OS level support (Unicode)



Tamil Computing in 21st C.

- Tamil in multilingual environment
- OS-level support (desirable for advanced applications)

for Unicode (ISO-10646) standard

on all computer platforms



Issues to consider

- Entering the data (keyboard)
- Displaying the data (fonts)
- Storing the Indic Data (encoding)



Information storage & interchange via different protocols

- Text: plain text (.txt), Formatted text (.doc. .rtf, .pdf, .obt)
- Email: SMTP
- Web: HTTP, XML
- Objective: Clean transfer without scrambling of information across platforms (Win, Mac and Unix), applications and protocols
- Protocols that carry encoding/charset information are critical!!



Font encoding

- At the very basic level, information is **stored and exchanged across platforms, across the net as coded characters in binary mode** with reference to some ref. chart
- Till 90s : English was *lingua franca* on computers
- 7-bit scheme ($2^7 = 128$ characters) used known as “PLAIN ASCII” (lower and upper case latin alphabets, punctuation and control characters)
- Extended to 8-bit level (256 character table): extended ASCII / iso-8859-x schemes to cover European languages



Plain ASCII 7-bit chart

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
	00 0000 0000	01 0000 0001	02 0000 0010	03 0000 0011	04 0000 0100	05 0000 0101	06 0000 0110	07 0000 0111	08 0000 1000	09 0000 1001	10 0000 1010	11 0000 1011	12 0000 1100	13 0000 1101	14 0000 1110	15 0000 1111
0	NUL	SOH	STX	ETX	EOT	ENQ	ACK	BEL	BS	HT	LF	VT	FF	CR	SO	SI
	16 0001 0000	17 0001 0001	18 0001 0010	19 0001 0011	20 0001 0100	21 0001 0101	22 0001 0110	23 0001 0111	24 0001 1000	25 0001 1001	26 0001 1010	27 0001 1011	28 0001 1100	29 0001 1101	30 0001 1110	31 0001 1111
1	DLE	DC1	DC2	DC3	DC4	NAK	SYN	ETB	CAN	EM	SUB	ESC	FS	GS	RS	US
	32 0010 0000	33 0010 0001	34 0010 0010	35 0010 0011	36 0010 0100	37 0010 0101	38 0010 0110	39 0010 0111	40 0010 1000	41 0010 1001	42 0010 1010	43 0010 1011	44 0010 1100	45 0010 1101	46 0010 1110	47 0010 1111
2	SP	!	"	#	\$	%	&	'	()	*	+	,	-	.	/
	48 0011 0000	49 0011 0001	50 0011 0010	51 0011 0011	52 0011 0100	53 0011 0101	54 0011 0110	55 0011 0111	56 0011 1000	57 0011 1001	58 0011 1010	59 0011 1011	60 0011 1100	61 0011 1101	62 0011 1110	63 0011 1111
3	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
	64 0100 0000	65 0100 0001	66 0100 0010	67 0100 0011	68 0100 0100	69 0100 0101	70 0100 0110	71 0100 0111	72 0100 1000	73 0100 1001	74 0100 1010	75 0100 1011	76 0100 1100	77 0100 1101	78 0100 1110	79 0100 1111
4	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
	80 0101 0000	81 0101 0001	82 0101 0010	83 0101 0011	84 0101 0100	85 0101 0101	86 0101 0110	87 0101 0111	88 0101 1000	89 0101 1001	90 0101 1010	91 0101 1011	92 0101 1100	93 0101 1101	94 0101 1110	95 0101 1111
5	P	Q	R	S	T	U	V	W	X	Y	Z	[\]	^	_
	96 0110 0000	97 0110 0001	98 0110 0010	99 0110 0011	100 0110 0100	101 0110 0101	102 0110 0110	103 0110 0111	104 0110 1000	105 0110 1001	106 0110 1010	107 0110 1011	108 0110 1100	109 0110 1101	110 0110 1110	111 0110 1111
6	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
	112 0111 0000	113 0111 0001	114 0111 0010	115 0111 0011	116 0111 0100	117 0111 0101	118 0111 0110	119 0111 0111	120 0111 1000	121 0111 1001	122 0111 1010	123 0111 1011	124 0111 1100	125 0111 1101	126 0111 1110	127 0111 1111
7	p	q	r	s	t	u	v	w	x	y	z	{		}	~	DEL



Characters for Tamil Font Encoding

Vowels and vowel diacritics											
அ	ஆ	இ	ஈ	உ	ஊ	எ	ஏ	ஐ	ஒ	ஔ	ஔௌ
a	ā	i	ī	u	ū	e	ē	ai	o	ō	au
Consonants						Grantha letters					
க	k [k, g, x, γ, h]	த	t [t, ḍ, ṭ]	ல	l [l]	ஐ	j [ɟ]				
ங	ṅ [ŋ]	ந	n [n]	வ	v [v]	ஷ	ṣ [ʃ]				
ச	c [tʃ, ʈ, ʃ, s]	ப	p [p, b, β]	ழ	z, ḷ, r [ɻ]	ஸ	s [s]				
ஞ	ñ [ɲ]	ம	m [m]	ள	! [l]	ஹ	h [h]				
ட	ṭ [t, ḍ, r]	ய	y [j]	ற	ṛ, ṛ̣ [r, t, d]	க்ஷ	kṣ [kʃ]				
ண	ṇ [ɳ]	ர	r [r]	ன	ṇ, N [n]						
Numerals											
சு	உ	ந	சு	ரு	கூ	எ	அ	கூ	ய	நா	சூ
Other symbols											
உ	மீ	வா	பு	கூ	ஷ	ரூ	நீ				
day	month	year	debit	credit	as above	rupee	numeral				



Tamil Alphabet Soup (uyirmei letters have distinct glyph forms)

தமிழ் எழுத்துகளின் பட்டியல்																			
எண்	மெய்	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	உயிர்	க்	ங்	ச்	ஞ்	ட்	ண்	த்	ந்	ப்	ம்	ய்	ர்	ல்	வ்	ழ்	ள்	ற்	ன்
1	அ	க	ங	ச	ஞ	ட	ண்	த	ந	ப	ம	ய	ர	ல	வ	ழ	ள	ற	ன்
2	ஆ	கா	ஙா	சா	ஞா	டா	ண்ணா	தா	நா	பா	மா	யா	ரா	லா	வா	ழா	ளா	றா	னா
3	இ	கி	னி	சி	ஞி	டி	ணி	தி	நி	பி	மி	யி	ரி	லி	வி	ழி	ளி	றி	னி
4	ஈ	கீ	நீ	சீ	ஞீ	மீ	ணீ	தீ	நீ	பீ	மீ	யீ	ரீ	லீ	வீ	ழீ	ளீ	றீ	னீ
5	உ	கு	நு	சு	னு	டு	ணு	து	நு	பு	மு	யு	ரு	லு	வு	ழு	ளு	று	னு
6	ஊ	கூ	நூ	சூ	னூ	டூ	ணூ	தூ	நூ	பூ	மூ	யூ	ரூ	லூ	வூ	ழூ	ளூ	றூ	னூ
7	எ	கெ	நெ	செ	னெ	டெ	ணெ	தெ	நெ	பெ	மெ	யெ	ரெ	லெ	வே	ழெ	ளெ	றெ	னெ
8	ஏ	கே	நே	சே	னே	டே	ணே	தே	நே	பே	மே	யே	ரே	லே	வே	ழே	ளே	றே	னே
9	ஐ	கை	நை	சை	னெ	டை	ணை	தை	நை	பை	மை	யை	ரையெ	லையெ	வையெ	ழையெ	ளையெ	றையெ	னையெ
10	ஓ	கொ	நொ	சொ	னொ	டொ	ணொ	தொ	நொ	பொ	மொ	யொ	ரொ	லொ	வொ	ழொ	ளொ	றொ	னொ
11	ஔ	கோ	நோ	சோ	னோ	டோ	ணோ	தோ	நோ	போ	மோ	யோ	ரோ	லோ	வோ	ழோ	ளோ	றோ	னோ
12	ஔ	கௌ	நௌ	சௌ	னௌ	டௌ	ணௌ	தௌ	நௌ	பௌ	மௌ	யௌ	ரௌ	லௌ	வௌ	ழௌ	ளௌ	றௌ	னௌ
ஆயுதம்		ஃ																	
மொத்தம்:		[உயிர்= 12]+[மெய்=18]+[உயிர்மெய்=216]+[ஆயுதம்=1]= 247 எழுத்துகள்																	

Include: numerals (11), granthas (5x13) & symbols (10) => 330



7-bit Tamil Fonts

(get all characters from a minimal set)

- Modelled on Tamil Typewriter layout
- Use vowels, consonants and a set of modifiers
- dead key (f) to get basic “mei” letters from akara-varisai letters
- Except for complex forms (ukara, ukara varisai), use vowel modifiers (kaal, kombu, kokki, kondai,,,) to get uyirmei letters (use font rendering mechanism called “kerning”)

⇒ Minimal set (use <100 glyph forms)



Monolingual 8-bit Tamil fonts for publishing houses

- 7-bit fonts use “kerning” to render ikara, iikara varisai - not suitable to render “outline” forms of characters => Include entire uyirmei letters in 8-bit 256 slot fonts
- Monolingual 8-bit fonts, used in computerized typesetting and in all Tamil printing worldwide.
- 1999 Tamil Internet Conf. : TAM standard



TAM: Monolingual 8-bit encoding

TAM _{xxx} - MONO LINGUAL CODING SCHEME FOR TAMIL														
	32	48	64	80	96	112	128	144	160	176	192	208	224	240
0		0	ௐ	றி	'	ஶீ				ௐ	ௐ	யு	உ	ப
1	!	1	கி	ளி	ஈ	ஹீ			ள்	யு	று	க.	ஊ	ம
2	"	2	ஹி	ளி	தீ	ஶீ	க்			க	ஹு	"	எ	ய
3	#	3	சி	ஹி	தீ	வ.	ங்		ா	ஹு	க.	"	எ	ர
4	\$	4	ஶி	ஹி	யீ	யீ	ச்		ி	டு	யு		ஐ	ல
5	%	5	ஹி	ஹி	யீ	ஶீ	க்	•	•	ஹு	ௐ		ஓ	வ
6	&	6	றி	ஶி	யீ	ஸ்	ட்	ஶீ	•	து	ஶா	ஶா	ஓ	ழ
7	'	7	றி	கீ	ஶீ	ஷ்	ஶ்	ஶீ	•	•	ஶீ	யு	ஶீ	ள
8	(8	யி	ஶீ	ஶீ	ஶீ	த்	ர்	•	ந	ஶா	ஶீ	க	ற
9)	9	யி	ஶீ	ஶீ	ஶீ	த்	ம்	ஶீ	யு	ஶா	ஶீ	ந	ள
10	*	:	யி	ஶீ	ஶீ	ஶீ	ப்	வ்	ஶீ	மு		ஶா	ஶீ	ஸ
11	+	:	யி	[ஶீ	{	ம்	ழ்	ஶீ	யு	ஶா	ஶா	ௐ	ஷ
12	,	<	யி	\	றி		ம்	ள்	ஶீ	ௐ	ௐ	யு	ஶீ	ஶீ
13	-	=	யி]	ஶீ	}			ஶீ	ஶு	ஶு	ஶு	ஶா	ஶு
14	.	>	யி	^	ஶீ	-			ஶீ	யு	ஶு	ஶு	ஶு	ஶு
15	/	?	யி	_	ஶீ	DEL		ந்	ஶீ	ஶு	ஶு	ஶு	ஶு	ஶீ



Bilingual 8-bit fonts for Internet usage

- Use of internet => data transfer involve systems and people not conversant with Tamil
- 1997 Tamil Internet Conf. In Singapore : 8-bit bilingual (English-Tamil) scheme as appropriate for information interchange via Internet (email, www)
- All European Languages use a 8-bit scheme “extended ASCII/Latin-x/ 8859-x” schemes where plain ASCII is in slots 0-127 and select European characters in slots 128-255.



Bilingual 8-bit English-Tamil fonts

- Net-based email, list exchanges in Tamil started in mid-nineties; they use a wide variety of 8-bit schemes, all retaining “plain ASCII” set at 0-127 slots
- Tnet (Naa. Govindasamy), Anjal (Muthu Nedumaran) popular in Malaysia/Singapore
- **TSCII** was developed through Net-based mailing list discussions in 1997/1998 (in 2007 TSCII included in IANA/IETF charset registry)
- Tamilnet 99 Chennai conference introduced **TAB** bilingual 8-bit scheme as a standard.



TAB: Bilingual 8-bit Encoding

TABxxx - BILINGUAL CODING SCHEME FOR TAMIL														
	32	48	64	80	96	112	128	144	160	176	192	208	224	240
0		0	@	P	'	p				கு	ரு		உ	பு
1	!	1	A	Q	a	q		'		நு	று		ஊ	மு
2	"	2	B	R	b	r		'	'	க	ஊ	"	எ	ய
3	#	3	C	S	c	s		"	ஈ	து	உ	"	ஏ	ர
4	\$	4	D	T	d	t		"]	டு	து	'	ஈ	ஸ
5	%	5	E	U	e	u		•	•	ஊ	கு	'	ஒ	வ
6	&	6	F	V	f	v			•	து	து	து	ஒ	ழ
7	'	7	G	W	g	w			•	•	டு	து	ஃ	ள
8	(8	H	X	h	x			•	து	ஊ	டு	க	ற
9)	9	I	Y	i	y			©	டு	து	ஊ	வ	ள
10	*	:	J	Z	j	z			©	டு		து	ஈ	ஸ
11	+	:	K	[k	{			©	டு	து	ஊ	கு	ஈ
12	,	<	L	\	l				©	டு	டு	து	ட	ஊ
13	-	=	M]	m	}				து	டு	து	ஊ	ஊ
14	.	>	N	^	n	~			©	டு	டு	து	த	ஊ
15	/	?	O	_	o	DEL			©	டு	டு	து	த	ஊ



ISCII - Storage code standard for Indic Languages

- ISCII - Indian Scripts Code for Information Interchange
- 8 -bit, upper ASCII (>127) is used for IL
- Umbrella Standard single code for all Indian Languages
- Based on Devanagari
- Co-existence with English
- Cannot have more than one IL at a time
- One collation order for all IL –which is not true in practice



Tamil computing in 21st C.

- **Support** for Unicode /multilingual encoding available at **OS level**
- “**Localisation**” (l12n) (“Internationalization” (i18n for multilingualism))
- “**Localisation**” is not simply translation. Each society has its own cultural traits, often reflected in writing forms (script of the languages)



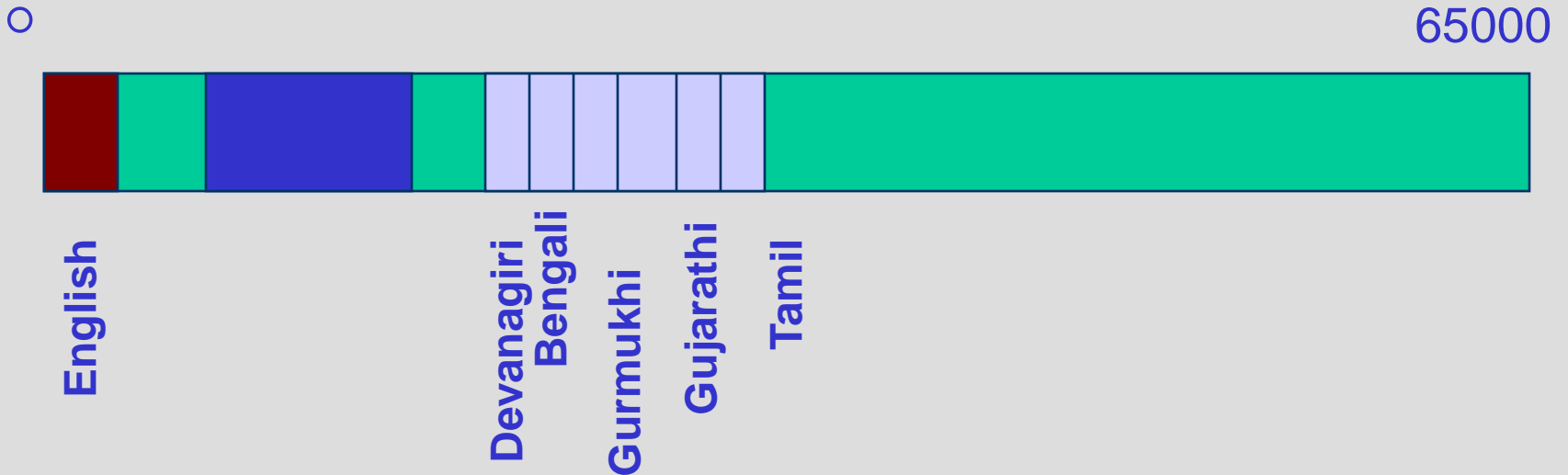
8-bit Indian Script Code ISCII



English

Hindi, Kannada, Tamil, ... any one

Unicode (Multilingual)



Tamil within 16-bit multilingual Unicode

Devanagari : 0900 - 097F

Bengali : 0980 - 09FF

Gurumukhi : 0A00 - 0A7F

Gujarati : 0A80 - 0AFF

Oriya : 0B00 - 0B7F

Tamil : 0B80 - 0BFF

Telugu : 0C00 - 0C7F

Kannada : 0C80 - 0CFF

Malayalam : 0D00 - 0D7F

0B80

Tamil

0BFF

	0B8	0B9	0BA	0BB	0BC	0BD	0BE	0BF
0		ஐ 0B80		ர 0BB0	ீ 0BC0			ய 0BF0
1				ற 0BB1	ி 0BC1			ள 0BF1
2	ஃ 0B82	ஔ 0B8C		ல 0BB2	ூ 0BC2			சு 0BF2
3	ஃ 0B83	ஔ 0B85	ண 0BA3	ள 0BB3				உ 0BF3
4		ஔ 0B84	த 0BA4	ழ 0BB4				ம் 0BF4
5	அ 0B85	க 0B86		வ 0BB5				ஶ 0BF5
6	ஆ 0B86			ஸ 0BB6	ெ 0BC6		ஐ 0BE6	யு 0BF6
7	இ 0B87			ஷ 0BB7	ே 0BC7	ள 0BF7	க 0BE7	ஶ 0BF7
8	ஈ 0B88		ந 0BA8	ஸ 0BB8	ை 0BC8		உ 0BE8	ஷ 0BF8
9	உ 0B89	ங் 0B89	ன 0BA9	ஶ 0BB9			ந் 0BE9	ரீ 0BF9
A	ஊ 0BA8	சு 0BA9	பு 0BAA		ொ 0BCA		சு 0BEA	ரீ 0BFA
B					ோ 0BCB		ரு 0BEB	
C		ஐ 0B8C			ொ 0BCC		சு 0BEC	
D					ஃ 0BCD		எ 0BED	
E	எ 0BE8	ஞ 0BE9	ம 0BAE	ா 0BBE			அ 0BEE	
F	ஏ 0BF8	ட 0BF9	ய 0BAF	ி 0BBF			சு 0BFF	



Tamil Unicode on Windows and Mac OS

- Unicode scheme introduced in 1991 but Tamil block support came only in 2000 (win) & 2005 (Mac OS)
- Microsoft's Windows 2000, Windows XP, Vista
- Apple's Mac OS X 10.4 (aka Tiger)
- Unicode data stored and exchanged in utf-8/utf-16 formats.



Unicode implementation requires “character rendering”

- Unicode encodes the characters of a given script but does not get to “display” (rendering)
- Indic language vowel-consonant combo (abugida /uyirmei) have unique glyph forms
- Truetype fonts simply display stored glyphs
- Open Truetype OTT fonts have font rendering information included
- Windows: UNISCRIBE Macintosh: ATSUI



Tamil Unicode : issues

- Modeled on ISCII (Devanagiri)
- Tamil (all south indian languages) “Dravidian”
language structure different
- Encode “akara varisai uyirmei” and not basic meis: (k = ka + puLLi not as per Tamil grammar)
- Tamil neTungkaNakku order not used (=> extra step needed in collation)
- Problem with definition of a “character”



Alternate 16-bit scheme (TUNE)

- Propose a 512-slot block for Tamil (encode uyir, mei, uyirmei, numerals, Tamil symbols, grantha,...)
- Conceptual differences /definition ?
- Finding a suitable block for Tamil within the BMP



Tamil Text input (Keyboard drivers)

- Keyboard input not linked at all with encoding
- Same keyboard layout can be used with all encodings
- Most keyboard layouts are “intelligent” (interpreted)
- Tamil Typewriter, TamilNet99, Romanized, Mylai,
- Tools available to make your own keyboard (XML based for Mac OS)



Tamil Text Editors

- Many available as freeware (basic features)
- Pro versions with spell-check, grammar check,...
- Many encodings are supported
- Many keyboard layouts are supported
- Text convertors available for flawless conversion of texts between encodings
- Mapping table Unicode \Leftrightarrow TAB, TSCII



Demo

Tamil Text Input on Mac OS X



Tamil Internet



Tamil content on the Net

- Amongst indic languages, after Hindi (400 million speakers), Tamil (70 million speakers) stands on the top with maximum content and data exchange in the Net
- Tamil content is rich and diverse: individual sites, portal sites, ezines, blogs, discussion lists and forums, online education, digital library,



Display of Tamil Webpages



Tamil webpages: delivery modes

- Three common modes of delivery:
 - as « images » (e.g., e-paper of dinakaran)
 - as « text » based on 7/8-bit encodings
 - as « encoded text » for Unicode (utf-8 format)
- webpages can be « static » or « dynamic »

Page content put in permanently (static) or generated upon call from database (dynamic)

- All portals using CMS and blog sites « dynamic »



Display of indic texts

- How does the browser know which font to use to display an indic language text?
- HTML specs require charset defined in the Meta header part of the html file
- HTML specs use « latin-1 » (iso-8859-1 of European languages) as the default encoding
=> in the absence of proper set up in the html file, Tamil page display will be « pure gibberish » (sadly most often this is the case)



Webpage display: font selection by the browser

- HTML specs allowed « font face » tag to enhance typography in web-delivery
- Tamil webpages extensively use this « font face » tag to invoke Tamil fonts
- with introduction of CSS style sheet, HTML 4.0 discourages use of « font face » tags alone
- suggest use of « charset » definition in the Meta header part for non-Latin/Roman language pages
- use of « font face » tag without invoking a charset encoding is incorrect.



Static webpage display: font requirement

- use of propriety font in the « fontface » tag => user has to download and install the font to view the content in Tamil
- till today « font » a major issue for Tamil (indic) language sites, since computer knowledge of most end-users are minimal if any.
- alternative: use « dynamic » fonts (bitstream .pfr or Microsoft .eot)



Dynamic fonts

- dynamic font file (.pfr or .eot) stored at the web-server; upon browser call, content dynamically generated using relevant Tamil font and rendered in Tamil script
- no need for users to have the Tamil font installed locally («elegant solution »)
- BUT.... gross misuse: tendency to use propriety fonts /lack of use of standards



Tamil Font Standards and browser support

- Web browsers depend on the « charset » definitions of meta header to render the pages correctly
- No support for encoding of any 7 or 8-bit indic fonts by any web-browser, web-server of major IT MNCs even after a decade
- Because: none of the indic font encodings are formally registered with IETF/IANA charset registry



Indic language fonts use « hacked » encodings

- ISCII : released only as ISO standard in 1991 => till date no support in any web-browser
- TAB/TAM for Tamil: Govt of Tamilnadu GO in 1999 => till date no support in any web-browser
- Tamil « TSCII » first indic font to be registered with IETF/IANA in 2007 (bitnum code:)



font/charset selection by the browser

- In the absence of officially recognized charsets to use, Tamil webpages invoke « x-user-defined » charset
- Web browser has to be set for this user-defined case (not all browsers give this option)
- softwares invariably invoke « win-1252 » (works ONLY with MS IE browser in Windows). Webmasters do not even check this before uploading the files



Tamil webpages based on Unicode

- Tamil Unicode is enabled at OS level (& many applications) in Win2K/XP/Vista and in Mac OS X (10.4)
- Webpages are encoded and referenced to with “utf-8” format
- Once language settings are set, browser will pick automatically a Tamil unicode font to display Tamil webpages (user “joy”)



Tamil Internet Tour



Email based Mailing List Discussions & Forums

- « soc.culture.tamil » list of Usenet: first online forum for Tamil (early nineties), exchanges in English or romanized Tamil
- « tamil.net » list (mid-nineties): first online forum that allowed email exchanges in Tamil script (8-bit bilingual Inaimathi -> TSCII)
- Many lists at yahoogroups and Googlegroups
- « tamil-ulagam », « agathiyar », « maratthadi », culturally rich



Tamil Internet Tour: Blogs

- Blogs a major part of Internet
- Tamil Blogsites all based on Unicode (use open source packages with searchable databases)
- Over 2000 Tamil bloggers are writing on the Net on a regular basis and the number is increasing constantly (diverse topics, varied level of quality)



தமிழ் வலைப்பதிவுகள் அரங்கம் v 2.0



முகப்பு	இடுகைகள்	பதிவுகள்	மன்றம்	உதவி/தகவல்	"ம" திரட்டி	பூக்கூடை	தமிழ்விழி	கேளிர
---------	----------	----------	--------	------------	-------------	----------	-----------	-------

இல்லம் [முழுப்பட்டியல்](#) [பட்டியலில் சேர்க்க](#) [சேர்க்கை நிலவரம்](#)

தமிழ்மணம் வலைப்பதிவுகளின் புள்ளிவிபரம்

மொத்தப் பதிவுகள்: 1995

கடந்த ஒரு மாதத்தில் எழுதப்பட்ட இடுகைகள்: 4597

கடந்த ஒரு வாரத்தில் எழுதப்பட்ட இடுகைகள்: 1097

கடந்த 24 மணி நேரத்தில் எழுதப்பட்ட இடுகைகள்: 200

ஒரு நாளில் சராசரியாக எழுதப்படும் இடுகைகள்: 153

பின்னூட்ட நிலவரம் காட்டப்படும் பதிவுகள்: 1184

ஒரு நாளில் சராசரியாக இடப்படும் பின்னூட்டங்கள்: 840

- [தமிழில் எழுகலாம் வாருங்கள், வலைபில் பரப்பலாம் பாருங்கள்](#)
- [என் கோடு, உன் கோடு, யுனிகோடு கனி கோடு](#)

தமிழ்மணம் அளிக்கும் தமிழ் அடைப்பலகை(template) மேம்படுத்தும் கருவி (Template Improvement Tool). (மிகு)இச்சேவை பிளாக்கர் அடைப்பலகை(Blogger.com template)

வலைப்பதிவு சேவைகள்

வலைப்பதிவு மென்பொருட்கள்

தமிழ் எழுது கருவிகள்

வலைப்பதிவு திரட்டிகள்

பயனுள்ள வலைத்தளங்கள்

தமிழ் அடைப்பலகைகள்(templates)

படம் வழங்கி சேவைகள்


வலை நூலகங்கள்



Thenkoodu.com - Tamil Blogs Portal

http://www.thenkoodu.com/ RSS Google

தமிழ்மனம் << முகப்பு < தேன்கூடுகாம் திண்ணை மதுரைத்திட்டம் தமிழ் இணையம்...லைக்கழகம் யாழ் இணையம் மரத்தடி.காம் முத்தமிழ் மன்றம் : ஞாலகம் பத்ரிபிள் வலைப்பதிவுகள் புத்தக விமர்சனம் நா. கண்ணன்



தேன்கூடு
TAMIL BLOGS PORTAL

June 29, 2007 | முதல்பக்கம் | என் தேன்கூடு | கட்டளையகம் | வலைப்பதிவினை இணைக்க | உதவி | புதிய கணக்கு

இடுகைகளைப் பதுப்பிக்க

உங்கள் இடுகை உடனடியாக தேன்கூட்டில் தெரிய பிங் செய்யுங்கள்!

[READ/WRITE IN TAMIL](#)

[Templates for Blogger.com*New*](#)

[பெட்டகம்*New*](#)

Username:

Password:

Remember?

[Register](#) [Forgot Pwd?](#)

உங்கள் கவனத்திற்கு...

சிறுதூள் - கோவையில் வெற்றிப் பயணம் - பதிவர் வசந்த

"சற்றுமுன்" பதிவின் கட்டுரைப் போட்டி - பதிவர் சற்றுமுன் குழு

மே 20 - கோவையில் பதிவர் பட்டறை! - பதிவர் யெஸ்.பால பாரதி

பதிவர்களுக்கு அறிவிப்பா? இந்தப் பகுதியில் உங்கள் இடுகை இடம்பெற இங்கு கிளிக்குங்கள்!

பதிவுகளுக்கான உதவிக்குழுவின : தமிழ் வலைப்பதிவர் உதவிப்பக்கம்

தேடல்

தமிழ் அகராதி

Enter the word in தமிழ் / English you want to look up:

இடுகைகள் | திரட்டி | செய்திகள் | மின்னிதழ்கள்

[Tip: பதிவின் பெயருக்கு முன்னர் உள்ள | என்ற எழுத்தை கிளிக் செய்யுங்க, பதிவு விபரங்களைப் பாருங்க! டெக்னோராட்டியில் உள்ள விபரம் கூட தேன்கூட்டில் தெரியும்!! - தேன்கூடு - your top class technology powered portal!]

Jun 29 **கனவுகளைத் தொலைத்தவர்** [பதிவு]

11:46 **Funny quotes Emerging of the new economy(1)** [பதிவு]

INFOSYSism Weekend feast for Guys! 112 Votes 44,524 Views What if Women ruled the world ... 54 Votes 33,159 Views BCCI Hiring Freshers Cricketer... 86 Votes 20,872 Views You have a 1000 poor cows. You put them on a nice campus, & send them one at a time to the US for ...

ஏன் Kingfisher Airlinesல நிறைய பேர் Travel பண்ணாங்க ? ? ? (2) [பதிவு]

Jun 29 **எழில்** [பதிவு]

11:46 **காஷ்மீரில் ராணுவம் சில படங்கள்(0)** [பதிவு]

ஜம்முகாஷ்மீர் விடுதலை முன்னணி போராட்டக்காரர் கோவும் போடுகிறார் விசாரணை செய்யவந்த ராணுவத்தினரை பலாத்காரம் ...

பாண்ட ராம் ஜானகி கோவிலுக்கு தலித் பூசாரி(0) [பதிவு]

Jun 29 **உண்மைத் தமிழன்** [பதிவு]

11:45 **யார் செய்தது குற்றம்..?(1)** [பதிவு]

29-06-2007 என் இனிய வலைத்தமிழ் மக்களே.. என்றும்போல் அன்றும் அலுவலகத்தில் ஆணி பிடுங்கிக் கொண்டிருந்தேன். நண்பர் ...

4 மறுமொழிகள் V.Shabesan [பதிவு]

Jun 29 **i webeelam** [பதிவு]

11:41 **புலிகள் எங்கள் தொப்புள்கொடி உறவு! தமிழ்நூ போராட்டம் வெற்றி பெற வேண்டும்!(0)** [பதிவு]

இந்திய மீளவ சங்கத் தலைவர் திரு ஆண்டினி கோமல் அவர்களை "அவுஸ்திரேலிய ஒலிபரப்பு கூட்டுத்தாபனம்" அண்மையில் ...

Jun 29 **i siru muyarchi** முத்துலெட்சுமி [பதிவு]

11:41 **அன்புடனின் இரண்டாம் பரிசு(0)** [பதிவு]

அன்புடன் குழுவில் எனது காட்சிக்கவிதை இரண்டாவது பரிசுக்குத் தேர்ந்தெடுக்கப்பட்டது. என் கவிதை வரிகள்: ...

Jun 29 **i My thoughts** [பதிவு]

11:40 **தமிழ்த் தாத்தாவும், சுப்பிரமணிய பாரதியும்!(0)** [பதிவு]

தமிழ்த் தாத்தா உ.வே.ஸ்வாமிநாத ஜயரின் காலத்தவர் பாரதியார் என அனைவருக்கும் தெரிந்திருக்கும். சுப்பிரமணிய ...

Jun 29 **i Blogeswari** [பதிவு]

11:40 **SEHAT - Cannes Radio Lion 2007(0)** [பதிவு]

Today, radio is slowly but surely, becoming the marketers' favourite medium. And if the radio spots are

அதிகம் பார்வைபெற்றவை

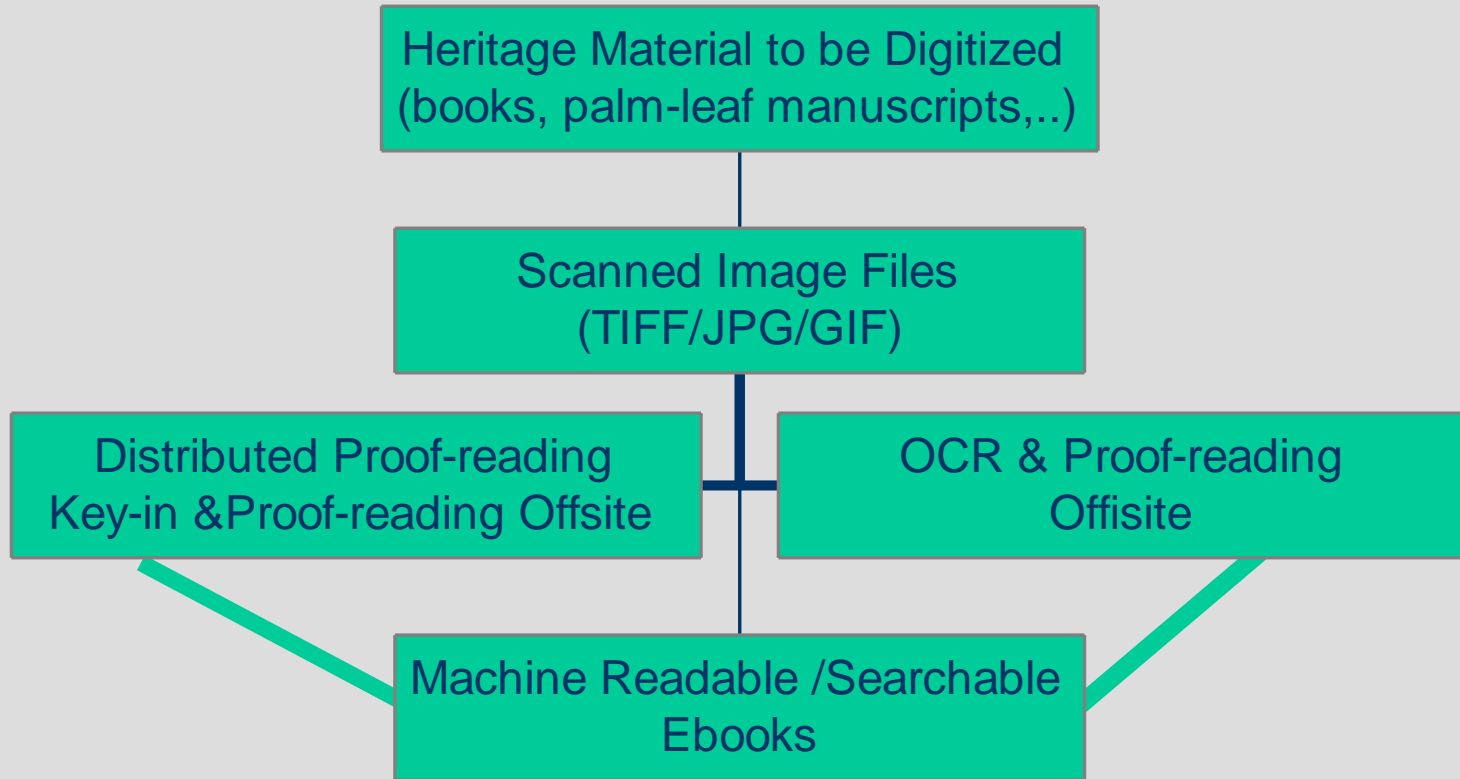
- * பெண்குறியை அறுப்பதை எகிப்து அரசாங்கம் தடைசெய்திருக்கிறது By எழில்
- * பிகரை கரெக்ட் செய்ய முன்று வழிகள் [#20] By சிரிப்பு
- * பார்ப்பதைத் தலைமை - உரையாடல்களுக்கான சில குறிப்புகள் By மிதக்கும் வெளி
- * கவர்ச்சி படங்கள் - Friday Spl By விவசாயி
- * பெண் டி.எஸ்.பி; ஆன் எஸ்.ஐ - புகாருடன் மனைவி By சற்றுமுன்
- * எனக்கு எய்ட்ஸ் இருக்கு By கிமுமத்தூர் எகல்பிரஸ்
- * நிறுத்துங்கள் பிரதீபாவை! By சித்தார்த்தகாரனின் சிந்தனைச்சிந்தல்கள்
- * டி.எம் எல் முதல் கமலஹாசன் வரை...முதல் சந்திப்பு நிலையலைகள்...வானொலி நிகழ்ச்சி By இசைமின் மடியில்
- * அலுவலகத்தில்... By கட்டுமானத்துறை
- * Tamil Magazine Thuglak Cartoon ; ழுன்றாம் அணி !!! By ஊசி aka pin

இன்று மறுமொழியப்பட்டவை (Get your code)

- பெங்களூரு வலைப்பதிவர் சந்திப்பு(6)
- ஏன் இந்த மௌனமோ(புகைப்படம்)(7)
- போலி உண்மைத்தமிழன் யார்?(20)
- யார் செய்தது குற்றம் ?(4)
- மதுரை மே தொகுதி வாக்கு எண்ணிக்கை- Update(7)
- பித்தனின் கிறுக்கல்கள் - 11(3)
- 'கவஜ்' எழுத்துவ எப்படி? - டிப்ஸ்(41)
- என் தனியறை(30)
- கோபித்த காதல்(26)
- சத்தம் போடாதே - இசை வெளிமீட்டு விழா(3)
- Nothing But Wind பல்வாக்கோம்



Digital Archiving of Tamil Heritage Materials



Digital Library of India

- A « mega » effort (part of Million ebooks project of Carnegie-Mellon Univ, PA, USA) to bring online thousands of Indian language books (covering all Indic ones)
- phase I: archive scanned image files; phase II: generate machine readable texts
- searching the catalogue of etexts: now only in romanized format; can be in vernacular/indic languages as well





Digital Library of India

Hosted by: **Indian Institute of Science, Bangalore** in co-operation with
CMU, IIT, NSF, ERNET and **MCIT** for the **Govt. of India** and 21 participating centers.



[Home](#) [Vision](#) [Mission](#) [Goals](#) [Benefits](#) [Content Selection](#) [Current Status](#) [People](#) [Partners](#) [Funding](#) [Copyright Policy](#) [FAQ](#) [RFP](#) [Members](#)

Books Journals Newspapers

Palm-Leaves (Manuscripts)

Title	<input type="text"/>
Author	<input type="text"/>
Year	<input type="text"/> to <input type="text"/>
Subject	Any Subject
Language	Tamil
Scanning Centre	Any Centre
<input type="button" value="Clear"/> <input type="button" value="Search"/>	

[Our Partner Sites for more books Presentations and Report Statistics Report Status Report Feedback](#)

Click [here](#) for total collections **NEW!**

Hits since

0 0 6 6 9 9

[Chemporun](#), 1990020047627. Krishnaveni. 1952. tamil. RELIGION. THEOLOGY. 334 pgs.

[Ramayana Saram](#), 1990020047657. Karpangada Venkatachariar. 1984. tamil. RELIGION. THEOLOGY. 80 pgs.

[Sri Upadesa Sagasthri](#), 1990050006974. Sri Veera Suppaiah Swamigal. 1916. tamil. PRAKARANAM. 168 pgs.

[SriBhagavad Gita vol.2](#), 1990020047785. Narasimhachariar. 1914. tamil. RELIGION. THEOLOGY. 516 pgs.

[Srivachana bhushana Vyakhyanam](#), 1990020047664. No author. 0. tamil. RELIGION. THEOLOGY. 376 pgs.

[Tiruvyamozhi Bhagavad Vishyam vol.2 with various commentaries](#), 1990020047805. Nammalawar. 1913. tamil. RELIGION. THEOLOGY. 558 pgs.

[Tiruvyamozhi Bhagavad Vishyam vol.2 with various commentaries](#), 1990020047806. Nammalawar. 1913. tamil. RELIGION. THEOLOGY. 552 pgs.

[Vaidyam](#), 1990020102896. Pulippani Mahamuni. 1932. tamil. TECHNOLOGY. 98 pgs.

[Yal Nul Thiruvallagkatirettu](#), 2040100049827. Swami Vipulananda. 1947. tamil. . 539 pgs.

[Yatiraja vaibhavam](#), 1990020047668. Vadagunambi. 1976. tamil. GEOGRAPHY. BIOGRAPHY. HISTORY. 46 pgs.

[Yoga Vasistam](#), 1990050006986. S.V. Ganapathy. 1948. tamil. PRAKARANAM. 342 pgs.

[108 srivaishnd-ava tivvateicha vaipavamum raand-a abimaana stalang-kal'um](#), 2990100073320. di en chi viiraraagavan. 1984. tamil. LANGUAGE. LINGUISTICS. LITERATURE. 197 pgs.

[1978 April Sapthagiri tamil](#), 2040100074253. K.Subbarao. 1978. tamil. . 44 pgs.

[1978 August Sapthagiri tamil](#), 2040100074254. K.Subbarao. 1978. tamil. . 44 pgs.

[1978 December Sapthagiri tamil](#), 2040100074255. K.Subbarao. 1978. tamil. . 45 pgs.

[1978 February Sapthagiri tamil](#), 2040100074256. K.Subbarao. 1978. tamil. . 44 pgs.

[1978 January Sapthagiri tamil](#), 2040100074257. K.Subbarao. 1978. tamil. . 45 pgs.

[1978 July Sapthagiri tamil](#), 2040100074258. K.Subbarao. 1978. tamil. . 44 pgs.

[1978 June Sapthagiri tamil](#), 2040100074259. K.Subbarao. 1978. tamil. . 44 pgs.

[1978 March Sapthagiri tamil](#), 2040100074260. K.Subbarao. 1978. tamil. . 44 pgs.

[1978 May Sapthagiri tamil](#), 2040100074261. K.Subbarao. 1978. tamil. . 52 pgs.

[1978 November Sapthagiri tamil](#), 2040100074262. K.Subbarao. 1978. tamil. . 44 pgs.

Search matched: **1617** books with **411723** pages . [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) ... [Next](#)

[0-9](#)

[A-D](#)

[E-H](#)

[I-L](#)

[M-P](#)

[Q-T](#)

[U-Z](#)

Project Madurai



Project Madurai

மதுரை தமிழ் இலக்கிய
மின்மொழித் திட்டம்



- A voluntary Net-based effort devoted to preparation of e-version of Tamil literary works and to distribute them free through the Net
- unique effort amongst Indic languages
- since 1998 PM has produced etexts (machine-readable / searchable) collection (ca. 300) covers all key ancient Tamil works
- All etexts (html & PDF, in TSCII and Unicode) available free on the Net
- Archives mirrored in many global digital library collections





List of electronic texts currently available

[alphabetic listing in TSCII](#)
அகர வரிசைப் பட்டியல் - தகதாம்

[alphabetic listing in unicode](#)
அகர வரிசைப் பட்டியல் - ஒருங்குறி / யூனிகோடு

[chronological release list in TSCII](#)
கால வரிசைப் பட்டியல் - தகதாம்

[Subjectwise listing in TSCII](#)
நூல்வகைப் பட்டியல் - தகதாம்

[Chronological listing in Unicode](#)
காலவரிசைப் பட்டியல் - ஒருங்குறி / யூனிகோடு

• [Etexts in Tamil Unicode format](#)

[கேடுபொறிகள் பக்கம்](#)



Project Madurai

மதுரை தமிழ் இலக்கிய மின்தொகுப்புத் திட்டம்



The content of this Homepage is in Tamil script, as per Unicode, invoking "utf-8" charset in the web-browser. Hence you need to use a Unicode font with Tamil block and select "unicode/utf-8" charset in the browser preferences. Reload the page if necessary to see the tamil text correctly . [Click here for the TSCII version of this Homepage](#)

"மதுரை தமிழ் இலக்கிய மின்தொகுப்புத் திட்டம்"

இது ஒரு உலகலாவிய தமிழர்கள் இணையம்வழி ஒன்றுகூடி தமிழ் இலக்கியங்களின் மின்தொகுப்புக்களை உருவாக்கி அவற்றை இணையம்வழி உலகெங்கிலும் உள்ள தமிழர்களும் தமிழார்வலர்களும் இலவசமாக பெற வசதிசெய்யும் திட்டம்.

எந்த ஒரு சமூகத்திற்கும் இலக்கியங்கள்தான் அக்கலாசாரத்திற்கான ஒரு முக்கிய எடுத்துக்காட்டு. அதை செவ்வனே காத்து உலகலாவிய தமிழர்களுக்கும் ஏனையோருக்கும் பகிர்ந்துகொள்வதும் வரும் சந்ததியினருக்கு கொண்டு செல்வதும் ஒவ்வொரு தமிழரின் கடமை. மதுரைத் திட்டம் இதற்கான ஒரு கூட்டு முயற்சி.

மதுரைத் திட்டம் எந்தவித அரசாங்க (அ) தனியார் நிறுவன உதவியின்றி, எந்தவித வியாபார நோக்கமுமின்றி நடைபெறுகின்ற ஒரு தன்னார்வ (voluntary) முயற்சி. 1998-ம் ஆண்டு தமிழர் திருநாள் (பொங்கல்) அன்று ஆரம்பிக்கப்பட்ட இத்திட்டம் இன்றும் தொடர்ந்து இயன்று வருகின்றது. உலகில் வெவ்வேறு நாடுகளில் வசிக்கும் முன்னூற்றுக்கு மேற்பட்ட தமிழர்களும் தமிழார்வலர்களும் ஒன்றுகூடி இத்திட்டத்தை நடத்தி வருகின்றனர்.

மதுரைத் திட்டம் உலகில் பல நாடுகளில் வசித்து வரும் தமிழர்கள் அவரவர் தங்களது வீடுகளில் தனியார் கணினி கொண்டு கிடைக்கும் நேரங்களில் தமிழ் இலக்கியங்களை கணினியில் உள்ளிட்டு (அ) பிழை திருத்தி மின்தொகுப்புக்கான தயாரிக்கும் ஒரு கூட்டு முயற்சி. தமிழ் இலக்கியங்களை மின்தொகுப்பு பாதுகாத்து மற்றவர்களுடன் இலவசமாக பகிர்ந்து கொள்வதில் விருப்புள்ள அனைவரும் இத்திட்டத்தில் பங்கு பெறலாம்.

மதுரைத் திட்டத்தின் மின்தொகுப்புகள் ஆரம்ப காலத்தில் இணைமதி, மமிலை தமிழ் எழுத்திருக்கிகள் (fonts) கொண்டு தயாரிக்கப்பட்டது. ஆனால் 1999-ம் ஆண்டிலிருந்து இணையம் வழி தமிழ் தகவல் தொழில்நுட்ப பரிமாற்றத்திற்கான என இணையம்வழி நிர்மானிக்கப்பட்ட தமிழ் தகதர (TSCII - Tamil Script Code for Information Interchange) வடிவம் கொண்டு தயாரித்து மின்தொகுப்புகள் வெளியிடப்பட்டு வருகின்றது. மின்தொகுப்புகள் இணையத்தில் இணைய பக்கங்களாகவும் (webpages in html format), PDF வடிவத்திலும் விநியோகிக்கப்பட்டு வருகிறது. 2003-ம் ஆண்டிலிருந்து பல்மொழி ஒருங்குறி குறியீடு (Unicode) முறை தயாரிக்கப்பட்ட மின்தொகுப்புகளையும் வெளியிட்டு வருகிறோம்.

தமிழ் இலக்கியங்களின் சரித்திரம் மிக பழமையானது. முதற் சங்க கால நூல்கள் இரண்டாயிரம் ஆண்டுகளுக்கு முன்பே இயற்றப்பட்டது என்பது வல்லுனர்கள் கருத்து. மதுரைத் திட்டம் காலம், சமயம், தேசப்பிரிவு, இலக்கியப்பிரிவு போன்ற எவ்வித பாகுபாடுமின்றி எல்லாவிதமான தமிழ் நூல்களின் மின்தொகுப்புகளை வெளியிட்டு வருகின்றது. தொன்றுதொட்ட சங்க கால நூல்கள் முதல் தற்கால தமிழ்நூல்கள் வரை அனைத்தும் வெளியிடப்படுகிறது. லோகக் கட்டுப்பாடு பக்கம் வரவில்லை வெளியான நூல்களுக்கான காப்பரிமைகளுக்கான மரியாதை கொடுக்கக் கண்ணியமாக நடப்பது காப்பரிமை இல்லா



நூலகம்

[FONT HELP](#)

நூலகம் திட்டம்

ஈழத்து நூல்களையும் இதழ்களையும் ஆவணப்படுத்திப் பாதுகாத்து அவற்றை அனைவரும் எந்நேரமும் வாசிப்பதற்கும் உசாத்துணைக்கும் பயன்படுத்துவதற்கு இணையத்தில் இலகுவிற கிடைக்கக் கூடியதாக வெளியிடும் ஓர் இலாப நோக்கற்ற தன்னார்வக் கூட்டு முயற்சி.



- நூலகம் திட்டத்தில் ஈழத்து எழுத்தாளர்களின் படைப்புகளே வெளியிடப்படுகின்றன.
- நூலகம் திட்டம் பற்றி அறிந்துகொள்ள [திட்ட அறிமுகப்](#) பக்கத்துக்குச் செல்லுங்கள்.
- நூலகம் திட்டத்துக்குப் பங்களிக்க நூலகம் [மடலாடற் குழுவில்](#) இணையுங்கள்.

- [நூல்கள்](#) (433)
- [இதழ்கள்](#) (290)

- இதுவரை 257 நூல்கள் மின்னூலாக்கப்பட்டுள்ளன.
- 223 நூல்கள் மின்நகல் ஆவணங்களாக்கப்பட்டுள்ளன.
- 290 இதழ்கள் மின்வடிவில் வெளியிடப்பட்டுள்ளன.
- மொத்த மின்னூலாக்கம் $257 + (223-47) + 290 = 723$

நூலகம் திட்டம் தொடர்பான கருத்துக்களை [noolaham'AT'gmail'DOT'com](mailto:noolaham@ATgmail.DOTcom) க்கு எழுதுங்கள்



விக்கிபீடியா
உட்பற்ற அணுகல்கள்

Project Madurai
ஈழநாடு தழிர் இலக்கிய
மின்மொழித் திட்டம்

சென்னை நூலகம்





- « Digital Archive of **global content** for **universal access** »
- Yahoo, H-P, Microsoft, consortium of leading US universities
- “collaborative effort of a group of cultural, technology, non-profit, and non-governmental organizations from around the world to build a permanent archive of multilingual digitized text and multimedia content”.



Online and Computer-aided (multimedia) Teaching of Tamil

- Online teaching of Tamil at various levels now available
- Tamil Virtual Univ. – content still based on 8-bit bilingual TAB (TNGovt websites)
- Tamil Learning Web Portal of UPenn, Philadelphia, USA (based on Unicode)
- Computer-aided teaching of Tamil – widespread in primary & secondary schools of Singapore, Malaysia



Searching Tamil content on the Net : Advantages of Database driven websites



Searching Tamil content in 7-bit, 8-bit font websites

- Several search engines have been proposed
- Auto-recognition of the encoding used?
- Can searches be made in an encoding other than that used by the site?
 - Use APIs (convert query to site encoding, search/retrieve results, convert search results to user-desired encoding)



Searching for data in Tamil Internet

- Univ. of California, Berkeley has archived many many GBs of content of Tamil websites for several years (Prof. Fred Gay)
- Websites use 7-bit Bamini, 8-bit bilingual (TAB, TSCII and many propriety schemes) and Unicode (utf-8)
- No single/universal font encoding standard used => no simple search engine
- Many approaches to search using APIs



Web servers and Database Query

- Linux hosting / Apache : open source GPU
 - Php scripting and MySQL or postgre SQL query
 - Many Free & Open Source Softwares FOSS have built in search modules
 - CMS Drupal, Joomla, MKPortal, ...
 - Blogs Wordpress
 - photogallery: Coppermine
 - Ecommerce: OS commerce
- Windows: Microsoft SQL: ASP, .NET,
- IBM-DB2, SAP DB: (enterprise)
- Oracle: Oracle9i, Oracle 10i (enterprise)



Database based Websites

- Nearly all CMS-based Portals and Blog sites use PHP/MySQL database storage, query (Linux)
- Corporate portals use a variety of enterprise type packages (Microsoft, Oracle, IBM,...)
- Advantage:
 - webpage delivery dynamic (no need for regular revision of webpages)
 - Website content fully searchable !!!



Project Madurai etext archives fully searchable (TSCII & Unicode)

- machine readable version of etexts (not image files) => data searchable
- Search engines (online) available to search for availability of etext of a given work or even occurrence of a given word in one or more works => linguistic (etymology) research now possible



Google search for Tamil content possible in Unicode !!



The image shows a screenshot of a web browser window displaying the Google homepage in Tamil. The browser's address bar shows the URL <http://www.google.com/intl/ta/>. The search bar contains the text "திருக்குறள்" (Tirukkural). Below the search bar, there are several navigation links: "வலை" (Vale), "படங்கள்" (Pattangal), "குழுக்கள்" (Guzhukkal), and "வலைப்பட்டியல்" (Valepattiyal). The main search input field contains the text "திருப்பாவை" (Tiruppavai). To the right of the search bar, there are links for "மேம்பட்ட தேடல் விருப்பங்கள்" (Meempattu Theidal Viruppankal) and "மொழிக் கருவிகள்" (Mozhik Karuvikal). Below the search bar, there are two buttons: "Google தேடல்" (Google Theidal) and "அதிரஷ்டம் என பக்கம்" (Adirashdam Ena Pakkam). At the bottom of the page, there is a link for "Google பற்றிக் தெரிந்துகொள்ளுங்கள் - Google Suisse க்குச் செல்லவும்" (Google Parhik Therintukonlanunangal - Google Suisse kku chchellavum) and a copyright notice "©2007 Google".



Google search in Unicode

திருப்புகழ் - Google Search

http://www.google.com/search?hl=en&client=safari&rls=en&q=திருப்புகழ்&btnG=Search

தமிழ்மணம் < முகப்பு < தேன்கூட்டுதாம் திண்ணை மதுரைத்திட்டம் தமிழ் இணையம்...எலக்கமுகம் யாழ் இணையம் மரத்தடிசாம் முத்தமிழ் மன்றம் :

Web Images Video News Maps Gmail more ▼

Google திருப்புகழ் Search Advanced Search Preferences

Web

History of Tamil Literature - Professor C.R.Krishnamurti
At the beginning of each Thiruppukazh (திருப்புகழ்)song, திருப்புகழ் விரிவுரை. வானதி பதிப்பகம் ...
www.tamilnation.org/literature/krishnamurti/08thathuva.htm - 86k - Cached - Similar pages

திருப்புகழ் " WordPress.com Tag Feed
<http://wordpress.com/tag/திருப்புகழ்/> Feed of posts on WordPress.com tagged "திருப்புகழ்"
Mon, 25 Jun 2007 21:54:14 +0000 ...
wordpress.com/tag/திருப்புகழ்/feed/ - 4k - Supplemental Result - Cached - Similar pages

:Thiruppugazh (Tamil திருப்புகழ் tiruppukal "Glory to ...
Thiruppugazh (Tamil திருப்புகழ் tiruppukal).
www.kerala.com/wiki-Thiruppugazh - 52k - Supplemental Result - Cached - Similar pages

- Sify.com
ஆனால், திருப்புகழ் பாடல்களுக்கு யாராவது நடனமாடிப் பார்த்திருக்கிறீர்களா? ...
tamil.sify.com/pennaenee/feb04/fullstory.php?id=13381409 - 104k - Supplemental Result - Cached - Similar pages

திருப்புகழ் - தமிழ் விக்சனரி ...
திருப்புகழ் என்பது முருகக் கடவுள் மீது அருணகிரிநாதர் இயற்றிய ஒரு பத்தி (பக்தி) நூல். ...
ta.wikipedia.org/wiki/திருப்புகழ் - 18k - Cached - Similar pages

ஆத்திகம்: அருணகிரிநாதர் அருளிய ...
அருணகிரிநாதர் அருளிய திருப்புகழ் -- 21 "பாதி மதி இந்தத் திருவேரகத்துத் திருப்புகழ் ...
aaththigam.blogspot.com/2007/06/21.html - 56k - Cached - Similar pages

ஆத்திகம்: அ.அ.திருப்புகழ் --20 ...
திருப்புகழ் --20 "காமியத் தழுந்தி". அருணகிரிநாதர் அருளிய திருப்புகழ் --20 "காமியத் ...
aaththigam.blogspot.com/2007/05/20.html - 61k - Supplemental Result - Cached - Similar pages
[[More results from aaththigam.blogspot.com](#)]

chennaiLibrary.com - சென்னை நூலகம் - Tamil Literature ...
அங்ஙனம் அருளக் கேட்ட அன்பர் ஆண்டவனை நோக்கிக் கூறியதாகத் திருப்புகழ் படிப்போர் ...
www.chennaiLibrary.com/muruga/thiruppugazh.html - 92k - Cached - Similar pages

Thiruppugazh - Sri AruNagirinAthar's songs of praise for Lord Murugan
திருப்புகழ். Sri AruNagirinAthar's Thiruppugazh ... இருப்பவல் திருப்புகழ் விருப்பொடு படிப்பவர் ...
www.kaumaram.com/thiru/nt0242.html - 19k - Supplemental Result - Cached - Similar pages

Google வலைப்பட்டியல் - World > Tamil > சமூகம் ...
திருமுறைகள், திருப்புகழ், ஆன்மிக குறிப்புகள், ... திருமுறைகள், திருப்புகழ், அபிராமி ...
www.google.com/Top/World/Tamil/சமூகம்/மதம்/வரிந்து/ - 22k - Cached - Similar pages





Digital Bibliography for Tamil Diaspora (DIBTAD) ©K. Kalyanasundaram

Bibliography Lists currently available, in Tamil unicode (utf-8)

Publications in Tamil of early 20th Century

- [Tamil Publications prior to 1900](#)
- [Tamil Publications of 1901 to 1920](#)
- [Tamil Publications of 1921 to 1930](#)
- [Tamil Publications of 1931 to 1940](#)
- [Tamil Publications of 1941 to 1945](#)
- [Tamil Publications of 1946 to 1950](#)
- [Tamil Publications of 1951 to 1955](#)
- [Tamil Publications of 1956 to 1960](#)
- [Tamil Publications of 1961 to 1965](#)

- [Tamil Novels -akara varisai list 1](#)
- [Tamil Novels -akara varisai list 2](#)
- [Tamil Novels -akara varisai list 3](#)
- [Tamil Novels -akara varisai list 4](#)
- [Tamil Novels -akara varisai list 5](#)
- [Tamil Novels -akara varisai list 6](#)
- [Tamil Novels -akara varisai list 7](#)
- [Tamil Novels -akara varisai list 8](#)

Please feel free to leave your comments here

Name:

Email:

Location (city, country):

Url:

Comments:

Sign Guestbook



300 சங்கர்	நீலமணி, கே.பி	1970	136 ப	சென்னை	வானதி பதிப்பகம்
41000 தலைவாங்கிய அராபிய அழகி	இளைய ராஜா	1977	94 ப	மதுரை	கலைப்பொன்னி
512ஆம் நம்பர் பஸ்	புணதாம் பேகர், கோ.ரா	சென்னை	கலைமகள்
2153	தமிழ்வாணன்	சென்னை	மணிமேகலைப் பிரசுரம்
616வது அட்சக்கோடு	அசோகமித்திரன்	1977	268 ப	சென்னை	கலைஞன் பதிப்பகம்
71942	ராஜவேலு, கு	1950	114 ப	சென்னை	இன்ப நிலையம்
821வது அம்சம்	அஸ்வகோஷ்	1979	127 ப	புதுவை	சமூகக்கல்வி மையம்
924 ருபாய் தீவு	சுஜாதா	சென்னை	மோனா
1027 அடி அழகி	புனிதன்	1965	143 ப	சென்னை	வானதி பதிப்பகம்
1127அடி+அழகி	புனிதன்	சென்னை	சென்னை	மணிமேகலைப் பிரசுரம்
1231 இரவுகள்	சண்முகம், ர	1978	340ப	சென்னை	அருணோதயம்
1361 வயதிலே	மெர்னா	1978	87 ப	சென்னை	பூரம்
14A ஃபார் ஆப்பிள் M ஃபார் மர்டர்	ராஜேஷ்குமார்	சென்னை	பாரதி பதிப்பகம்
15A ஃபார் ஆப்பிள் M ஃபார் மர்டர்	ராஜேஷ்குமார்	சென்னை	பாரதி பதிப்பகம்
16PDY 1000	இராமச்சந்திரன், பி.எஸ்	1965	72 ப	சென்னை	எம்.வி. எஸ்
17 ஃபிளைட் 731	தங்கராசன், முல்லை	1972	122 ப	சிவகாசி	முத்து
18 அக்கினி /வங்க / ஆகுள் தாரா சங்கர் பந்த்யோபாத்யாய	குமாரஸ்வாமி, த.நா	1945	268ப	சென்னை	சென்னை
19 அக்கினி சாட்சி	ஆறுமுகம், பூவை எஸ்.	1968	109ப	சென்னை	பத்ம ஜோதி
20 அக்கினி நட்சத்திரங்கள்	இந்துமதி	1981	232ப	சென்னை	திருமகள்
21 அக்கினி வீணை	ஜெகசிற்பியன்	1958	246ப	சென்னை	வானதி பதிப்பகம்
22 அக்கினிக் கோபம்	மணிசேகரன், கோவி	1956	429ப	சென்னை	முத்து நூலகம்
			580,624,802		
23 அக்கினிப் பரீட்சை /ருஷ்ய, அலெக்சேய் தல்ஸ்தோய்	ரகுநாதன்	1976, 1978	ப	மாஸ்கோ	மூ.ப
24 அக்கினிப் பூக்கள்	ராஜேந்திரன், கல்கி	1978	95ப	சென்னை	பிரேம ரத்னம்
25 அகட விகட அதிவினோத கதை	கந்தசாமி பிள்ளை, மதுரை	1907		சென்னை	எம். நரசிம்மலு செட்டி அண்டு கோ
26 அகட விகட ஆனந்தவல்லி	கந்தசாமி பிள்ளை, மதுரை	1907		சென்னை	எம். நரசிம்மலு செட்டி அண்டு கோ
27 அகட விகட கவி அதி அற்புத ஷோக் சுந்தரி சரித்திரம்	கந்தசாமி பிள்ளை	1908	120ப	சென்னை	சி.எஸ். நம்மையா
29 அகதி	மாயாவி	1951	156ப	சென்னை	சாந்தி நிலையம்
28 அகதி	மாயாவி	1955	168ப	தென் ஆற்காடு	
30 அகதி	நீலமணி, கே.பி	1965	62ப	சென்னை	கலைமகள்
31 அகதியின் ஆனந்தம்	ஆச்சார்யா, சி.கே.என்	1966	144ப	சென்னை	ஸ்ரீலட்சுமி விஜயம்
38 அக்னி நதி /உருது, ஆக் கா தார்யா	குர் அதுல் ஜன் ஹைதர், செளரி	1971	594ப	சென்னை	என்.பி.டி / நேஷனல் புக டிரஸ்ட்
32 அக்பர் சால்திரி	தி. ஜானகிராமன்			சென்னை	ஐந்திணைப் பதிப்பகம்
33 அகல் விளக்கு	வரதராசனார், மு	1958	406ப	சென்னை	தாயகம்
34 அகல்யா	சண்முகம், ர	1973	263ப	மதுரை	அறிவு நிலையம் /
35 அகல்யா	சண்முகம், ர	1978		சென்னை	ராணிமுத்து
36 அகல்யா	பாலகுமாரன்	1982	187ப	சென்னை	வானதி பதிப்பகம்
37 அகல்யா பாய்	புஷ்பரதச் செட்டி, ஊ	1885	30ப	சென்னை	எம்.எஸ்.பி.
அகிலன் நாவல்கள் - ஒரு திறனாய்வு (பி.எச்.டி பட்ட					கேரளா பல்கலைக்கொள்ளிந்திய
39 ஆய்வேடு	விஜயலட்சுமி, ஆர்	1977		சைவ சித்தாந்த நூற்பதிப்புக் கழகம்
40 அகிலனின் நாவல்கள் (பி.எச்.டி பட்ட ஆய்வேடு)	குருநாதன், ஆர்	1976	சென்னை	சென்னை பல்கலைக் கழகம்
41 அகிலனின் பொன்மலர்	இரபிசிங், ம.செ.	1976	67 ப	சென்னை	தென்குமரி
42 அகிலா	மணிசேகரன், கோவி	1973, 1977	184ப	சென்னை	வானதி பதிப்பகம்
43 அங்கயற்கண்ணி	மாதவராய முதலியார், எ.ஆர்	1921		சென்னை	டி. ராஜகோபால் முதலியார்
44 அங்கே இங்கே எங்கே	பிரபாகர், பட்டுக்கோட்டை	1981	70ப	சென்னை	மோனா
45 அச்சுத ரங்கம்மா	மணிசேகரன், கோவி	1979	232ப	சென்னை	வ.ப
46 அசட்டுப் பெண்	பாலப்பிரியா	1963	55ப	சென்னை	கலைமகள்
47 அசட்டுப்பிச்சு	ஆர்.வி	சென்னை	கலைமகள்
48 அசடு	காசியபன்	1978	114ப	சென்னை	பாளையங்கோட்டை
49 அசடு வழிந்தது	விக்ர	1956, 1962	42ப	சென்னை	விஜயலக்ஷ்மி
	சித்திலேவை மரைக்காயர், முகமது				மாணவர் இல்லம்



Tamil Computing : Advanced Applications



Tamil Computing :

Advanced users & Developers

- Tamil Databases & Data mining of e-resources of the Net using search engines
- Auto-recognition and extraction of Tamil text (OCR)
- Voice recognition /text-to-speech synthesis
- Natural Language Processing (Machine-translation of texts, lexical resources, computational linguistics)
- E-commerce: Online stores, billing, inventory
- E-governance
- Online and Multimedia-based Teaching of Tamil & Tamil
- Digital Archiving of Heritage Materials (literature books, palm-leaf manuscripts, inscriptions, audio/multimedia storage)
- Localisation of Free and Open Source Softwares (FOSS)
- Tamil in Mobile and Hand-held/Portable devices



Tamil on Mobile platforms

- Tamil SMS first introduced in Jan 2005 by Muthu Nedumaran (“sellinam”)

<http://www.murasu.com/mobile/>

- Singapore Oli Tamil Radio Station accepts Tamil SMS messages <http://www.oli.sg/tamilsms/index.htm>
- Arichchuwadu” Tamil messaging of Microimage, Sri Lanka /AirTel
<http://www.microimage.com/mobile/Tamil/index.asp>
?



Tamil Email/SMS/Chat on Mobile phones



Machine-aided Translation of Tamil texts to other languages

- Very important area of NLP
- MAT requires several modules: e-dictionary, Parts-of-Speech synthesizer
POS, Morphological analyzer
- POS: breaking up the sentence to parts to identify noun, verb, adverb, adjective etc.
- Morph Analyzer: extraction of root verb from inflections (Tamil rich in inflexes)



Tamil Morph Analyser

- Tamil is a morphologically rich language. The root words combine with the morphemes in the form of suffixes.
- The Morphological analyser takes a derived word as input and separates it into root word and associated morphemes.
- It is the basic tool used in spell checker, grammar checker, parser and machine translation systems.
- It is also used for extracting the root words from inflections in Tamil Search Engine and Online Tamil Dictionary.



English Sentence

Language

Eng POS Tagger
morph & chunker
Sentence Parser

Word Sense
Disambiguation
(WSD)

senses

Transfer Grammar
Rule Application

Bilingual

Tense, Aspect
Modality
Lookup

E-I Dictionary
lookup

Indian Language
Generator

IL Sentence

Language

Example of machine translation via morphological/POS analysis

- **mo:han avanutaya kulantayai kanta:n enRu ra:man conna:r.**
- **mohan he(poss) child(acc) see(pst) compl raman say(pst).**
- **(Raman said that mohan saw his child).**

- **si:ta avalai atitta:l enRu kavita conna:r.**
- **sita she(acc) beat(pst) compl kavita say(pst).**
- **(Kavita said that Sita hit her)**



Tamil - Hindi Machine Translation (Anna Univ- KBC Center)

Tamil-Hindi Machine Aided Translation System

http://www.au-kbc.org/research_areas/nlp/demo/mat/

The New York Times... BBC NEWS | News Fro... washingtonpost.com ... Tamil Electronic Libr... 68/ Guardian Unlimited Sign in to Yahoo! The Hindu : Front Pa.

A U - K B C Research Centre

Can't see a text box below?

Tamil-Hindi Machine Aided Translation System

X

Translate

KeyBoard Help

உயிர்	அ	ஆ	இ	ஈ	உ	ஊ	எ	ஏ	ஐ	ஓ	ஔ	ஔ	ஔ	ஔ				
Vowels	a	aa/A	i	ii/I	u	uu/U	e	ee/E	ai	o	oo/O	au	Q					
மெய்	க்	ங்	ச்	ஞ்	ட்	ண்	த்	ந்	ப்	ம்	ய்	ர்	ல்	வ்	ழ்	ள்	ற்	ள்
Consonants	k/g/c	ng	s/ch	nj	T/D	N	t/d	n-	p/b	m	y	r	l	v/w	z	L	R	n

வடமொழி	ஸ்	ஷ்	ஜ்	ஹ்
Grandha Letters	S	sh	j	h

To input Compound characters (உயிர் மெய்), key-in the Pure Consonant (மெய்) followed by the Vowel (உயிர்).
மெய் +உயிர் =உயிர் மெய் .e.g. க் +அ=க.



Tamil – Hindi MAT

- Tamil–Hindi Anusaaraka based MAT
 - Machine–Aided Translation system
 - Lexical level translation
 - In collaboration with IITH & TTU
 - 80–85% coverage
 - User Interfaces: Stand–alone, API, and Web–based on–line
 - Byproducts
 - Tamil morphological analyser
 - Tamil–Hindi bilingual dictionary (~ 36k)



Tamil-Hindi MAT System

AU-KBC Research Centre Natural Language Processing Machine Aided Language Translation : Tamil - Hindi

ரவி மதுரை போயிருக்கிறான் .
அடுத்த மாதம் ரவி ஆயிரம் ரூபாய் சம்பாதித்திருப்பான் .
அவன் எங்களுடைய வீட்டில் தங்கியிருக்கிறான் .
சிறிய வயதில் நான் அடிக்கடி வானொலியில் பாடிக்கொண்டு வந்தேன் .
சாயங்காலம் ரவி கடையில் பேசிக்கொண்டிருப்பான் .
பால்காரன் வந்தாயிற்று .
ரவி வாடகை கொடுத்தாயிற்று .
அவன் இந்த நாவலை படித்தாயிற்று

Translate Text

Reset

Translated Text

रवि मदुरै गया है .
अगला महीना रवि एक हजार रूपया कमाया होगा .
वह हमारे घर में ठहरा है .
छोटी उम्र में मैं बार बार रेडियो में आकर .
शाम को रवि दूकान में बोल रहा होगा .
दूधवाला आ चुका .



English – Tamil MAT

- English – Tamil MAT – A Prototype
 - Includes exhaustive syntactical analysis
 - Limited Vocabulary (100–150)
 - Small set of Transfer rules
- Phase – II
 - Extending the prototype to the full-fledged system
 - Design includes Syntactic and Semantic processing
 - Trilingual system: English ↔ Tamil ↔ Hindi



English-Tamil MAT (Prototype)

A U - K B C Research Centre

[Home](#) [Index](#)

[System Limitations](#)

[Font Download](#)

English - Tamil Machine Aided Translation System (Prototype Version)

she will write a letter

Translate Text

Reset

Given English sentence : she will write a letter.
Translated Tamil sentence : அவள் ஒரு கடிதம் எழுதுவாள் .



English-Tamil MAT System

English - Tamil MAT (IITH-Shakti) System (Prototype Version)

she ate a cake

Translate Text

Reset

TRANSLATED OUTPUT

Roman : avalu oVru kek cAppittAIY

Tamil : அவள் ஒரு கேக் சாப்பிட்டாள்







Tamil computing - why bother at all?



Tamil Computing – Why bother?

- Tamil Diaspora 70 million+, over 10 million live outside India
- Tamil is « **international** »: official language in many countries: India (Tamilnadu), Sri Lanka, Malaysia, Singapore, Canada (?), ...
- >50% Tamil households outside India have PC at home & internet access



Tamil computing - why bother at all?

- Doing things in mother tongue = bringing technology right to your home and putting it to use !!!
- To reach the rural mass who do not speak English (or do not want to use - requirement for e-governance)
- Preparing for younger generation (growing with PC & InfoTech awareness quite high, use of computers integral part of school curriculum (reports, searching net for information,...) in many countries of west and south asia



Tamil Computing – Why bother?

- **User-friendly « native » interface in commerce:**
increasingly important, soon as default (touch-screen, menu-driven; language choice by user, ...)
- banking in real & virtual (e-banking); auto/self ticketing, baggage check-in at airports & train stations; billing, inventory in Tamil; e-commerce (online shops) & e-governance
- Tamil IT profit from abundant IT manpower available at home (TN)



Tamil computing – why bother?

- World is shrinking (India included) !
- Multilingual environment extensive, soon to become default (necessity for India?)
- Globalisation (single universal package) with multilingual support: Microsoft, Yahoo, Google,
- Internationalisation (i18n) and localisation (l12n) very important for MNCs



Language Technology (Tamil) related Research

- Academic Institutions: IIT-Madras, Anna Univ, Anna Univ-KBC Center, IISc, PSG Tech, Bharathidasan Univ.
- Govt. Institutions: CIIL, IIIT, CDAC
- R&D in MNCs (Globalization div): Microsoft, IBM, Oracle, Sun, Yahoo, Google,...
- For diverse Indic languages, R&D effort is largely symbolic => enormous potential to contribute !!!!



Databases based Tamil websites





தமிழ் இணையப் பல்கலைக்கழகம்
TAMIL VIRTUAL UNIVERSITY

புதியன
What's New



Tamil TAB font Installed?
If not **download** now

"தேம" தேமதுரத் தமிழோசை உலகமெலாம்
பரவும் வகை செய்தல் வேண்டும்"



த.இ.ப. ஒரு பார்வை
TVU at a glance
பாடத்திட்டங்கள்
Academic Programmes
நூலகம்
Library
மாதிரித் தளப்பொருள்கள்
Sample Contents
மழலைக்கல்வி
Primer Education
பண்பாட்டுக் காட்சியகம்
Cultural Gallery
பயணியர் தமிழ்
Tourist Tamil
திகழ்கலைக் கல்வி
Performing Arts
பொதுக்கல்வி
General studies
தொடர்பு மையம்
Contact Centre

Feedback

Disclaimer

இலக்கியம்

கலைச்சொற்கள்

கணித்தமிழ்

பிற இணைப்புகள்

Literature

Technical Glossary

Tamil in IT

Other Links



தமிழ் Tamil Language In Context

Home
Overview
Technical Help



- Home
- About Tamil
- Tamil Alphabet
- Tamil Words
- Tamil grammar
- Tamil Culture
- Tamil literature
- Tamilweb Team
- Acknowledgement

Tamil Lesson

Year 1



அறிமுகம், வரவேற்பு, பரிச்சயம், பால்ய நண்பர்
Meeting One Another, invitation, getting to know each other, old friends



விருந்து, விசேஷம், பண்டிகை, சாப்பாடு, ஆட்டோ
Receiving Guests and Hospitality, special occasions, religious ceremonies, hiring an auto



குடும்பம், பேருந்து, வழி, பிரயாணம், சண்டை
Family, Bus, route, travel, fight



போம் பேசு, வெயில் காலம், அஞ்சல், தொலைபேசி உரையாடல்
Bargaining, Summer season, Mailing, Telephone conversations



ஆவல், வழக்குப் பேச்சு, சலவை
Enthusiasm, Idiomatic expressions, Washing etc.



நம்பிக்கை, வழக்கம், பண்பாடு
Beliefs, Habits, Custom etc.

Year 2

2nd year materials are Under Construction



விமர்சனம், வம்பளப்பு, சந்திப்பு
Critiquing, chatting with neighbors, visiting friends, taking a leave of, soliciting & Mother-in-law vs. daughter-in-law



சொல்லிக்கொடுக்கல், சீதோஷணநிலை, மருத்துவரைச் சந்திக்கல்
Teaching a skill, talking about weather, journey & visiting a doctor



என்கிறது, என்று, வந்தோமா, போனோமா!
Use of quotative expressions, idiomatic expressions & Echo expressions



ஒதுவார் & குருக்கள்
Interviews with Oduvar (singer) and kurukkal (Tamil priest)



பரதநாட்டியக்கலைநர், தமிழ் எழுத்தாளர், நாட்டுப்புறக்கலைநர் & குயவர்!
Bharatha Natya artist



லலித் நாடக அகாடமி
Excerpts about Lalit Kala Academy

Tamil Digital Libraries

- Digital Library of Tamil Virtual Univ.
 - has etexts of early sangam period literary works
 - etexts in 8-bit bilingual TAB format
- Microfilm collections of Roja Muttiah Research Library, Chennai (of >30K books)
 - microfilm => digital images => readable text



எண்ணங்கள்

என் எண்ணங்களும், கருத்துகளும். என்னைத் தொடர்பு கொள்ள: [மின்னஞ்சல் முகவரி](mailto:manam.malarattum@gmail.com)

Thursday, June 28, 2007

பண உதவி தேவை: பொறியியல் படிக்கும் மாணவர்களுக்கு

திருப்பத்தூரை அடுத்த மீட்டர் என்ற இடத்தில் சரவணன் என்பவரும் அவரது சில நண்பர்களும் 'மணம் மலரட்டும்' என்ற சமூக சேவை அமைப்பை நடத்திவருகிறார்கள். (இதைப்பற்றி முன்னர் சில பதிவுகளில் குறிப்பிட்டுள்ளேன்.) அந்தப் பகுதிகளில் உள்ள வசதி குறைந்த குடும்பத்தைச் சேர்ந்த மாணவர்களுக்கு பொறியியல்/மருத்துவம் போன்ற படிப்புகளில் சேர பயிற்சி, பண உதவி ஆகியவற்றைச் செய்துவருகிறார்கள்.

மீட்டர், பூங்குளம் பகுதிகளில் இருந்து மூன்று மாணவர்கள் (இரு ஆண், ஒரு பெண்) நல்ல மதிப்பெண்களைப் பெற்றுள்ளனர். அவர்களுக்கு பொறியியல் கல்லூரியில் சேர வாய்ப்புகள் உள்ளன. அவர்களது பெயர்:

1. சந்திரசேகர்
2. நவியரசன்
3. நூர்ஜெஹான்

இவர்கள் பிற்படுத்தப்பட்ட மற்றும் அட்டவணை வகுப்பைச் சேர்ந்த மாணவர்கள்.

ஒருவருக்கு காலேஜ் ஆஃப் எஞ்சினியரிங், கிண்டியிலும் மற்ற இருவருக்கு சுயநிதிப் பொறியியல் கல்லூரிகளில் அரசு கோட்டாவிளும் இடங்கள் கிடைக்கலாம். (முழு விவரம் எஞ்சினியரிங் கவுன்செலிங் முடிந்ததும்தான் தெரியவரும்.)

சுயநிதிப் பொறியியல் கல்லூரியில் சேர்ந்து நான்கு வருடங்கள் படிக்க, கல்விக் கட்டணம், ஹாஸ்டல் செலவு, பிற செலவுகள் என்று வருடம் ரூ. 75,000 ஆகும். மொத்தமாக ரூ. 3 லட்சம்.

கிண்டி கல்லூரியில் படிக்க ஆகும் செலவு சற்றுக் குறையலாம்.

பணம் கொடுத்து உதவ விரும்புவர்கள் சரவணனை அல்லது ஸ்டாலினைத் தொடர்பு கொள்ளுங்கள்.

சரவணன் மின்னஞ்சல்: manam.malarattum@gmail.com



Badri Seshadri
Chennai, Tamil Nadu, IN

[View my complete profile](#)

சில சுட்டிகள்

- என்னைப் பற்றி
- என் பழைய கட்டுரைகள்
- என் பழைய வலையொலிப்படிப்பு ஓடை
- இந்த வலைப்பதிவின் செய்தியோடை - Newsfeed
- My English Blog
- Microcredit in India

துறைகள்

- புத்தகம் (20)
- வர்த்தகம் (16)
- சமூகம் (15)
- கல்வி (14)
- கிரிக்கெட் (12)
- தமிழகம் (10)
- அரசியல் (9)
- நுழைவுத்தேர்வு (9)
- இந்தியா (8)
- உலகக்கோப்பை (7)
- SEZ (4)
- குடியாட்சி (4)
- விவசாயம் (4)
- சினிமா (3)
- தொழில் (3)
- அண்டைநாடு (2)
- தொலைக்காட்சி (2)
- பங்குச்சந்தை (2)
- பட்ஜெட் (2)
- வங்கதேசம் (2)
- இலங்கை (1)
- கணினி (1)
- சீனா (1)
- திரைப்படம் (1)
- தொலைத்தொடர்பு (1)
- நதிகள் (1)
- மதம் (1)
- விடுதலைப் புலிகள் (1)

வலைப்பதிவர்கள்





விக்கிப்பீடியா

கட்டற்ற கலைக்களஞ்சியம்

வழிசெலுத்தல்

- முதற் பக்கம்
- சமுதாய வலைவாசல்
- நடப்பு நிகழ்வுகள்
- அண்மைய மாற்றங்கள்
- குறிப்பில்வழிப் பக்கம்
- உதவி
- நன்கொடைகள்

தேடு

கருவிப் பெட்டி

- இப் பக்கத்தை இணைத்தவை
- தொடர்பான மாற்றங்கள்
- கோப்பைப் பதிவேற்று
- சிறப்புப் பக்கங்கள்
- அச்சுக்குகந்த பதிப்பு
- நிரந்தர இணைப்பு
- இக்கட்டுரையை மேற்கோள் காட்டு

ஏனைய மொழிகள்

- English
- Deutsch
- Français
- Polski
- 日本語
- Nederlands
- Svenska

கட்டுரை உரையாடல் மூலத்தைப் பார் வரலாறு

புதுப் பயனர்களுக்கான உதவி / தமிழ் தட்டச்சு உதவி / Tamil font help

முதற் பக்கம்

கட்டற்ற கலைக்களஞ்சியமான விக்கிப்பீடியாவில் இருந்து.

விக்கிப்பீடியா என்பது எவரும் தொகுக்கக்கூடிய, இலவசமாகப் பயன்படுத்தக்கூடிய ஒரு பன்மொழிக் கலைக் களஞ்சியத் திட்டமாகும். இங்கு நீங்கள்ும் உங்களுக்கும் விருப்பமான தலைப்புகளில் புதிதாக கட்டுரைகள் எழுதலாம்; ஏற்கனவே உள்ள பக்கங்களை திருத்தி எழுதலாம். விவரங்கள் அறிய புதுப் பயனர்களுக்கான விக்கிப்பீடியா அறிமுகப் பக்கத்தை பார்க்கவும்.

கட்டுரைகள் எண்ணிக்கை: **10,917**

கட்டுரைப் பிரிவுகள் : தமிழ் | பண்பாடு | வரலாறு | அறிவியல் | கணிதம் | புவியியல் | சமூகம் | தொழில்நுட்பம் | நபர்கள்

கட்டுரைகள் : அகர வரிசை - துறை வரிசை - புதியன

இன்றைய முதற்பக்கக் கட்டுரைகள்

தேனீக்கள் ஆறுகால்கள் கொண்ட பறக்கும் சிறு பூச்சி இனத்தில் ஒன்றாகும். இவை பூவில் இருந்து பூந்தேனை உறிஞ்சி சேகரித்து தேனடையில் தேனாக சேகரித்து வைக்கின்றன. இவை ஈ பேரினத்தில் ஒரு வகை ஆகும். ஈ பேரினத்தில் இன்று ஏறத்தாழ 20,000 வகைகள் அறியப்பட்டுள்ளன. அவற்றுள் ஏழு இனங்கள்தான் தேனீக்கள் ஆகும். இந்த தேனீக்களில் மொத்தம் 44 உள்ளினங்கள் உள்ளன. அறிவியலில் தேனீக்கள் ஏப்பிடே (Apidae) என்னும் குடும்பத்தில், ஏப்பிஸ் (Apis) என்னும் இனத்தைச் சேர்ந்தவை. தேனீக்கள் பெருங்கூட்டமாக, தேனடை என்னும் பல அறுகோளை அறைகள் கொண்ட, கூடு கட்டி, அதில் தேனை சேகரித்து வாழ்கின்றன. தேனீக்கள் வெளியிடும் மெழுகால் இந்த கூடுகள் அமைக்கப்படுகின்றன.



வெண்பா மரபுச் செய்யுள் வகைகளுள் ஒன்றாகும். தமிழில் மரபுப் பாக்கள், ஒலிப்பியல் அடிப்படையில் அடி, சீர், அசை முதலியவற்றைக் கொண்டு வகை பிரிக்கப்பட்டுள்ளன. அவற்றுள் ஆசிரியப்பா, வெண்பா, கலிப்பா, வஞ்சிப்பா என்பன பரவலாக ஆளப்பட்டுள்ள பழம் பெரும் பாவினங்கள். அவற்றுள் வெண்பா என்னும் வகையில் இரண்டு முதல் பன்னிரண்டு அடிகள் வரைக் கொண்டிருக்கும். வெண்பாக்களுக்கான யாப்பிலக்கணம் ஒரு கட்டுக்கோப்பான இடம் சாரா இலக்கணம் என்று நிறுவப்பட்டுள்ளது.

- மேலும் சில கட்டுரைகள்..

செய்திகளில்

- ஜூன் 27 - டோனி பிளேர் பிரதமர் பதவி துறந்ததைத் தொடர்ந்து கோர்டன் பிரிளான் (படம்) ஐக்கிய இராச்சியத்தின் பிரதமரானார்.
- ஜூன் 22 - அட்லாண்டில் விண்கப்பல் (படம்) பாதுகாப்பாக கலிபோர்னியாவில் உள்ள வான்படையினரின் எட்வேர்ட்ஸ் தளத்தில் பாதுகாப்பாகத் தரையிறங்கியது. முன்னராக புளோரிடாவில் உள்ள கென்டி விண்வெளி மையத்தில் இறங்கவேண்டியிருந்தது, சீரற்ற காலநிலை காரணமாக கடைசி நேரத்தில் ரத்துச் செய்யப்பட்டது.
- ஜூன் 7 - கொழும்பில் விடுதிகளில் தங்கியிருந்த குழந்தைகள், பெண்கள் உட்பட 500-க்கும் மேற்பட்ட தமிழர்கள் இலங்கைக் காவல்துறையினரால் கொழும்பு விடுதிகளில் இருந்து தமிழர் வெளியேற்றப்பட்டனர்.
- மேலும் சில செய்திகள் ...



இன்று..

ஜூன் 29:

- 1864 - கனடாவில் கியூபெக்கில் தொடருந்து விபத்தில் 99 பேர் கொல்லப்பட்டனர்.
- 1995 - அட்லாண்டில் விண்ணோடும் ரஷ்யாவின் மீர் விண்வெளி நிலையக்கூள் மகாக்கூ வாயாக



Tamil Internet Tour: goodies

- Online /on the fly text convertor to convert texts between different Tamil font encodings
 - www.suratha.com/reader.htm
- Online tools to send emails in Tamil



பெங்குதமிழ்

http://www.suratha.com/reader.htm

Google

தமிழ்மணம் < முகப்பு < தேள்கூட்டுகாம் திண்ணை மதுரைத்திட்டம் தமிழ் இணையம்...லைக்கமுகம் யாழ் இணையம் மரத்தடிசகாம் முத்தமிழ் மன்றம் : ஞாலகம் பத்ரியின் வலைப்பதிவுகள் புத்தக விமர்சனம் நா. கண்ணன்

Suratha.Com Tamil News And ++++  

by Suratha Yarlvanan **pls wait loading**.....

பொங்குதமிழ்"Pongku Tamil" ©by Suratha Yazhvanan

யூனிகோட் எழுதியில் புதிதாக பாயினி அமுதம் Tam99 முறைமை சேர்க்கப்பட்டு இலகவான முகவரிக்கு மாற்றம் செய்யப்பட்டுள்ளது.

Best viewed with Internet Explorer 4 and above versions

Indoweb Murasoli Webulagam Thinathanthi Dinamani Thinaboomi Anjal Thatstamil(LIBI)
 Amudham/Dinakaran Mylai Vikatan(old) Tab Tam(kumudam/vikatan) Bamini TSC
 Romanised koeln Anu Graphics (Pallavar) 

எந்த ஒரு தெரிந்த அல்லது தெரியாத எழுத்துருவாகவிறந்தாலும் அதனை மேல் தட்டில் உள்ளிட்டு கீழே உள்ள ஏதாவதொரு துவாரங்களை சுட்டுவதன் மூலம் நீங்கள் உள்ளிட்டதை சீராக தமிழில் வாசிக்கமுடியும்.

யாழ் பல்கலைக்கழக மாணவர்களின் எழுச்சியிகு போராட்டமான பொங்கு தமிழின் நினைவாக இந்த செயலிக்கு பொங்கு தமிழ் எனப் பெயரிடப்படுகிறது.

இந்த தாளிறங்கி எழுத்துருவை பயன்படுத்த அனுமதி தந்த திரு ஆவரங்கால் திரு ஸ்ரீவாஸிற்கு எனது நன்றிகள்

இது புதிது

அனைத்து உங்கள் ஆவணங்களையும் உடனேயே கணினியில்(win95,98மற்றும் ஏனைய இயங்கு தளங்கள்) யூனிகோட் முறையிலோ அல்லது மற்றைய எழுத்துரு வடிவங்களிலோ html வடிவில் சேமிக்க, அச்செடுக்க, திருத்த புதியதோர் செயலி பொங்கு தமிழின் தொடர்ச்சியான [இராவணன்](#) எழுதியாகும்.

உங்கள் கருத்துக்களை suratha@koda.com மெயில்.கொம் என்ற முகவரிக்கு அனுப்புங்கள்



Vidyanidhi: Indian doctoral thesis database

- Vidyanidhi archives indian doctoral theses written in various indic languages
- database is in Unicode
- Theses in Indic languages have two records -one in the Roman script (transliterated) and the other in the vernacular.
- Database web-server
 - MS SQL 2000; A truly multilingual-capable SQL (Achieves satisfactory collation); Front End- ASP; Java script

