



Department of Information Technology,  
Ministry of Communications and Information Technology,  
Government of India, New Delhi

# **Draft Policy Document For**

# **INTERNATIONALIZED DOMAIN NAMES**

**Language: TAMIL**



Department of Information Technology,  
Ministry of Communications and Information Technology,  
Government of India, New Delhi

## RECORD OF CHANGES

**\*A - ADDED M - MODIFIED D - DELETED**

VERSION NUMBER	DATE	PAGES AFFECTED	A* M D	TITLE OR BRIEF DESCRIPTION	COMPLIANCE VERSION OF MAIN POLICY DOCUMENT
1.0	19/11/09	Whole Document	M	Language Specific Policy Document for TAMIL	1.5
1.1	22/11/20 10	Page No 8, 17	A	Restriction rule added, ccTLD added	1.6



Department of Information Technology,  
Ministry of Communications and Information Technology,  
Government of India, New Delhi

1. AUGMENTED BACKUS-NAUR FORMALISM (ABNF).....	4
1.1 Naming of Variables:.....	4
1.2 ABNF Operators:.....	4
1.3 The Vowel Sequence.....	4
1.4 Consonant Sequence.....	5
1.5 ABNF Applied to TAMIL IDN.....	6
2. RESTRICTION RULES.....	7
3. EXAMPLES.....	9
4. LANGUAGE TABLE: TAMIL .....	10
5. NOMENCLATURAL DESCRIPTION TABLE OF TAMIL LANGUAGE TABLE.....	11
6. VARIANT TABLE.....	15
7. EXPERTS/BODIES CONSULTED.....	16
8. PROPOSED ccTLD FOR TAMIL.....	17

# 1. AUGMENTED BACKUS-NAUR FORMALISM (ABNF)

## 1.1 Naming of Variables:

Dash → Hyphen -  
Digit → Indo-Arabic digits [0-9]  
C → Consonant  
V → Vowel  
M → Matra  
X → Visarga/Aytham  
H → Halant/Virama

## 1.2 ABNF Operators:

S. No.	Symbols	Functions
1	“/”	Alternative
2	“[ ]”	Optional
3	“*”	Variable Repetition
4	“()”	Sequence Group

In what follows the Vowel Sequence and the Consonant Sequence pertinent to Tamil are given. To facilitate use , equivalents in Devanagari are provided.

### 1.3 The Vowel Sequence

A vowel sequence is made up of a single vowel. It may be followed but not necessarily (optionally ) by an Visarga (X) The number of X which can follow a V in Tamil may be restricted to one.

The vowel sequence in Tamil is therefore V [X ]

Examples :

V	அ	அ
V[X]	அஃ	அஃ

### 1.4 Consonant Sequence

A consonant sequence admits the following shapes:

1. A single consonant (C)

Example:

C	க	க
---	---	---

2. A consonant optionally followed by Dependent Vowel sign[M] or Visarga[X] or Halant/Virama [H]

C[M|X|H]

Example:

C[M]	கி	கி
C[X]	கஃ	கஃ
C[H]	க்	க் (Pure Consonant)



Department of Information Technology,  
Ministry of Communications and Information Technology,  
Government of India, New Delhi

3. A sequence of consonants (up to 3) joined by Halant/Virama \*2(CH)C

Example:

CHCHC      க்ஷய → க ஂ ஷ ஂ ய      क्षय क् ष य

## 1.5 ABNF Applied to TAMIL IDN

Consonant Sequence → \*2(CH) C [H|X|M]

Vowel Sequence → V[X]

Sequence → Consonant Sequence | Vowel Sequence

IDN-Label → (Sequence | digit)\*([dash] (Sequence | digit))

## 2. RESTRICTION RULES

The ABNF is generic in nature and when applied to a specific language/script certain restriction rules apply. In other words, in a given language some of the Formalism structures do not necessarily apply. To take care of such cases restriction rules are set in place. These restrictions will help to fine-tune the ABNF. In the case of Tamil the following rules apply:

1. H | M | X cannot occur in the beginning of an IDN domain name

Example:

ஃக	ஃக்
ிக	ிக்
஁க	஁க்

As can be seen they will result automatically in a “golu” marking an invalid character. This is an intrinsic property of the Indic syllable and is quasi automatically applied wherever supported by the OS.

2. H is not permitted after V, X, M, digit and dash

Example:

அஃ	அஃ
க஁ஃ	க஁ஃ
கிஃ	கிஃ
1ஃ	1ஃ
-ஃ	-ஃ

3. Visarga/Aytham[X] is permitted after consonant-sequence or vowel-sequence is restricted to one



Department of Information Technology,  
Ministry of Communications and Information Technology,  
Government of India, New Delhi

Example:

क०००० कः

अ०००० अः

4. Visarga/Aytham[X] is not permitted after Matra.

की० किः

5. Number of M permitted after consonant-sequence is restricted to one

Example:

कीी कीी

6. M is not permitted after V

Example:

न०ा ई०ा

7. A consonant syllable that is intended to end with Halant [H] can only be followed by Hyphen or digit.





Department of Information Technology,  
Ministry of Communications and Information Technology,  
Government of India, New Delhi

### 3. EXAMPLES

Following are the some examples to show TAMIL IDN Label.

Combination	Example	Word With Combination
C	ப	படி
CH	க்	காக்கா
CM	கி	கிடங்கு
CX	எஃ	எஃகு
CHC	ம்ம	அம்மா
CHCHC	ர்க்	சேர்க்கை
V	இ	இப்ப
VX	அஃ	அஃறிணை



Department of Information Technology,  
Ministry of Communications and Information Technology,  
Government of India, New Delhi

## 4. LANGUAGE TABLE: TAMIL <sup>1</sup>

	0B80	Tamil						0BFF
	0B8	0B9	0BA	0BB	0BC	0BD	0BE	0BF
0		ஐ		ர	ீ	ஒ		ய
1				ற	ு			ா
2	ஃ	ஓ		ல	஁			சு
3	ஃ	ஓ	ண	ள				உ
4		ஔ	த	ழ				மீ
5	அ	க		வ				ஶ
6	ஆ			ஸ	ெ		ஃ	யு
7	இ			ஷ	ே	ௌ	க	ங்
8	ஈ		ந	ஸ	ை		உ	ஷை
9	உ	ங	ன	ஹ			ந	று
A	ஊ	ச	ப		ொ		ச	நீ
B					ோ		ரு	
C		ஐ			ெள		சு	
D					஁		எ	
E	எ	ஞ	ம	ா			அ	
F	ஏ	ட	ய	ரி			கூ	

<sup>1</sup> Characters marked in yellow are not applicable to the language.



Department of Information Technology,  
Ministry of Communications and Information Technology,  
Government of India, New Delhi

## 5. NOMENCLATURAL DESCRIPTION TABLE OF TAMIL LANGUAGE TABLE

Hex Entity	Character	Character Name
<b>VISARGA/AYTHAM (X)</b>		
0B83	ஃ	TAMIL SIGN VISARGA
<b>TAMIL VOWEL LETTERS (V)</b>		
0B85	அ	TAMIL LETTER A
0B86	ஆ	TAMIL LETTER AA
0B87	இ	TAMIL LETTER I
0B88	ஈ	TAMIL LETTER II
0B89	உ	TAMIL LETTER U
0B8A	ஊ	TAMIL LETTER UU
0B8E	எ	TAMIL LETTER E
0B8F	ஏ	TAMIL LETTER EE
0B90	ஐ	TAMIL LETTER AI
0B92	ஒ	TAMIL LETTER O
0B93	ஔ	TAMIL LETTER OO



Department of Information Technology,  
Ministry of Communications and Information Technology,  
Government of India, New Delhi

0B94	ஓள	TAMIL LETTER AU
<b>TAMIL CONSONANTS (C)</b>		
0B95	க	TAMIL LETTER KA
0B99	ங	TAMIL LETTER NGA
0B9A	ச	TAMIL LETTER CA
0B9C	ஜ	TAMIL LETTER JA
0B9E	ஞ	TAMIL LETTER NYA
0B9F	ட	TAMIL LETTER TTA
0BA3	ண	TAMIL LETTER NNA
0BA4	த	TAMIL LETTER TA
0BA8	ந	TAMIL LETTER NA
0BA9	ன	TAMIL LETTER NNNA
0BAA	ப	TAMIL LETTER PA
0BAE	ம	TAMIL LETTER MA
0BAF	ய	TAMIL LETTER YA
0BB0	ர	TAMIL LETTER RA



Department of Information Technology,  
Ministry of Communications and Information Technology,  
Government of India, New Delhi

0BB1	ற	TAMIL LETTER RRA
0BB2	ல	TAMIL LETTER LA
0BB3	ள	TAMIL LETTER LLA
0BB4	ழ	TAMIL LETTER LLLA
0BB5	வ	TAMIL LETTER VA
0BB6	ப	TAMIL LETTER SHA
0BB7	ஷ	TAMIL LETTER SSA
0BB8	ஸ	TAMIL LETTER SA
0BB9	ஹ	TAMIL LETTER HA
<b>TAMIL VOWEL SIGNS (MATRAS) (M)</b>		
0BBE	ா	TAMIL VOWEL SIGN AA
0BBF	ி	TAMIL VOWEL SIGN I
0BC0	ீ	TAMIL VOWEL SIGN II
0BC1	ு	TAMIL VOWEL SIGN U
0BC2	ூ	TAMIL VOWEL SIGN UU
0BC6	ெ	TAMIL VOWEL SIGN E



Department of Information Technology,  
Ministry of Communications and Information Technology,  
Government of India, New Delhi

0BC7	ே	TAMIL VOWEL SIGN EE
0BC8	ை	TAMIL VOWEL SIGN AI
0BCA	ொ	TAMIL VOWEL SIGN O
0BCB	ோ	TAMIL VOWEL SIGN OO
0BCC	ௌ	TAMIL VOWEL SIGN AU
<b>TAMIL VIRAMA (H)</b>		
0BCD	ஃ	TAMIL SIGN VIRAMA



Department of Information Technology,  
Ministry of Communications and Information Technology,  
Government of India, New Delhi

## 6. VARIANT TABLE

VARIANT	
<p>ଓୱା</p> <p>0B92+0BB3</p>	<p>ଓୱା</p> <p>0B94</p>



Department of Information Technology,  
Ministry of Communications and Information Technology,  
Government of India, New Delhi

## 7. EXPERTS/BODIES CONSULTED

Expertise provided by C-DAC Thiruvananthapuram.





Department of Information Technology,  
Ministry of Communications and Information Technology,  
Government of India, New Delhi

## 8. PROPOSED ccTLD FOR TAMIL

India (Bhārat) localized in Tamil - இந்தியா

**Note:** You can send your feedbacks to [ids-feedback@cdac.in](mailto:ids-feedback@cdac.in)