



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

# **Draft Policy Document For**

# **INTERNATIONALIZED DOMAIN NAMES**

**Language: TELUGU**



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	20/11/09	Whole Document	M	Language Specific Policy Document for TELUGU	1.5
1.1	22/11/201 0	Page No 10, 23	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.....	5
1.4 Consonant Sequence.....	5
1.5 ABNF Applied to TELUGU IDN.....	7
2. RESTRICTION RULES.....	8
3. EXAMPLES.....	11
4. LANGUAGE TABLE: TELUGU.....	12
5. NOMENCLATURAL DESCRIPTION TABLE OF TELUGU LANGUAGE TABLE .....	13
6. VARIANT TABLE.....	20
7. EXPERTS/BODIES CONSULTED.....	22
8. PROPOSED ccTLD FOR TELUGU.....	23

# 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
D	→ Anusvara/Sunna
B	→ Chandrabindu/Arasunna
X	→ Visarga
H	→ Halant/Virama
Y	→ Avagraha

## 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 Telugu 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 Anuswara (D), Candrabindu(B) or a Visarga (X)  
The number of D,B or X which can follow a V in Telugu may be restricted to one.  
The possibility of a Visarga following a Chandrabindu or Anuswara is ruled out.

The vowel sequence in Telugu is therefore V [D | B | X ]

Examples:

V	అ	అ
V[D]	అం	అం
V[B]	అఁ	అఁ
V[X]	అః	అః

## 1.4 Consonant Sequence

A consonant sequence admits the following shapes:

1. A single consonant (C)

Example:

C	క	క
---	---	---

2. A consonant optionally followed by dependent vowel sign[M] or anuswara[D] or Candrabindu[B] or visarga[X] or Halanta [H]

C [M|D|B|X|H]

Example:

C[M]	కి	కి
C[D]	కం	కం

C[B]	कँ	कँ
C[X]	कः	कः
C[H]	क	क् (Pure Consonant)

3. A sequence of consonants (up to 3) joined by Halanta \*2(CH)C  
 Example:

CHCHC      कृक → कृ + क + य      क्र

Subsets

3.a. The combination may be followed by H,M , D,B or X  
 \*2(CH)C[H|M|D|B|X]

Example:

CHC[M]	कृम	क्की   क् क ी
CHC[D]	कृड	क्कं   क् क ं
CHC[B]	कृब	क्कँ   क् क ौ
CHC[X]	कृः	क्कः   क् क ः
CHC[H]	कृ	क्क्   क् क क्

3.b. \*2(CH)CM may be followed by a D,B or X  
 \*2(CH)CM[D|B|X]

Example:

CHCM[D]	कृड	क्कीं   क् क ी ं
CHCM[B]	कृब	क्कीँ   क् क ी ौ
CHCM[X]	कृः	क्कीः   क् क ी ः



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

The final canonical structure of the consonant sequence in IDN can be defined in ABNF as:

$*2(CH) C [H|D|B|X|M[D|B|X]]$

## 1.5 ABNF Applied to TELUGU IDN

Consonant Sequence  $\rightarrow *2(CH) C [H|D|B|X|M[D|B|X]]$

Vowel Sequence  $\rightarrow V[D|B|X]$

Sequence  $\rightarrow \text{Consonant Sequence}[Y] \mid \text{Vowel Sequence}[Y]$

IDN-Label  $\rightarrow (\text{Sequence} \mid \text{digit}) * ([\text{dash}] (\text{Sequence} \mid \text{digit}))$

## 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 Telugu the following rules apply:

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

Example:

ఁక	క్
ిక	క్
ంక	క్
క	క్
ఃక	క్

As can be seen they will result automatically in a “golū” 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, D, B, X, M, digit and dash

Example

అ	అ
కం	కం
క	క





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

कः	कः
क	कि
क	क
क	क

3. Number of D,B or X permitted after consonant-sequence or vowel-sequence or M is restricted to one

Example:

कः	कः
कः	कः
कः	कः
कः	कः
कः	कः
कः	कः
कः	कः
कः	कः

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

Example:

कः	कः
----	----



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

5. M is not permitted after V

Example:

ॐ ईा

6. The combination of Anuswara+Visarga, Visarga+Anuswara, chandrabindu+anusvara and vice-versa is not permissible

Example:

ॐः कंः

ॐँ कँ

ॐः कः

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



### 3. EXAMPLES

Following are the some examples of TELUGU IDN Label.

Combination	Example	Word containing example
C	క	కల
CH	త్	భారత్
CM	కా	కాకి
CD	కం	కందకం
CX	నః	పునః దర్శనం
CMD	కాం	కాంకర్యం
CMX	దుః	దుఃఖ
CHC	ర్త్	బెర్త్
CHCHC	స్త్	అస్త్
V	అ	అల
VB	ఉః	ఉః
VD	అం	అంకె



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

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

0C00

Telugu

0C7F

	0C00	0C01	0C02	0C03	0C04	0C05	0C06	0C07
0	అ	ఐ	ఈ	ఊ	ఋ		ౠ	
1	ఌ		ఋ	ౠ	ౡ		ౢ	
2	ౣ	౤	౥	౦	౧		ౡ	
3	ౢ	ౣ	౤	౥	౦		౧	
4		ౡ	ౢ	ౣ	౤			
5	౥	౦	౧	ౡ		ౢ		
6	౦	౧	ౡ	ౢ	ౣ	౤	౥	
7	౥	౦	౧	ౡ	ౢ	ౣ	౤	౥
8	౥	౦	౧	ౡ	ౢ	ౣ	౤	౥
9	౥	౦	౧	ౡ	ౢ	ౣ	౤	౥
A	౥	౦	౧	ౡ	ౢ	ౣ	౤	౥
B	౥	౦	౧	ౡ	ౢ	ౣ	౤	౥
C	౥	౦	౧	ౡ	ౢ	ౣ	౤	౥
D	౥	౦	౧	ౡ	ౢ	ౣ	౤	౥
E	౥	౦	౧	ౡ	ౢ	ౣ	౤	౥
F	౥	౦	౧	ౡ	ౢ	ౣ	౤	౥

82

The Unicode Standard 5.1, Copyright © 1991-2008 Unicode, Inc. All rights reserved.

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

## 5. NOMENCLATURAL DESCRIPTION TABLE OF TELUGU LANGUAGE TABLE

Hex Entity	Character	Character Name
<b>CANDRABINDU (C)</b>		
0C01	◌̣	TELUGU SIGN CANDRABINDU
<b>ANUSVARA (D)</b>		
0C02	◌̣̣	TELUGU SIGN ANUSVARA
<b>VISARGA (X)</b>		
0C03	◌̣̣̣	TELUGU SIGN VISARGA
<b>VOWEL LETTERS (V)</b>		
0C05	అ	TELUGU LETTER A
0C06	ఆ	TELUGU LETTER AA
0C07	ఇ	TELUGU LETTER I
0C08	ఈ	TELUGU LETTER II
0C09	ఉ	TELUGU LETTER U



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

0C0A	ఉ	TELUGU LETTER UU
0C0B	ఋ	TELUGU LETTER VOCALIC R
0C0E	ఎ	TELUGU LETTER E
0C0F	ఏ	TELUGU LETTER EE
0C10	ఐ	TELUGU LETTER AI
0C12	ఒ	TELUGU LETTER O
0C13	ఓ	TELUGU LETTER OO
0C14	ఔ	TELUGU LETTER AU
0C60	ఋ	TELUGU LETTER VOCALIC RR
<b>CONSONANTS (C)</b>		
0C15	క	TELUGU LETTER KA
0C16	ఖ	TELUGU LETTER KHA
0C17	గ	TELUGU LETTER GA



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

0C18	ఘ	TELUGU LETTER GH A
0C19	ఙ	TELUGU LETTER NG A
0C1A	చ	TELUGU LETTER CA
0C1B	ఛ	TELUGU LETTER CH A
0C1C	జ	TELUGU LETTER JA
0C1D	ఝ	TELUGU LETTER JH A
0C1E	ఞ	TELUGU LETTER NY A
0C1F	ట	TELUGU LETTER TT A
0C20	ఠ	TELUGU LETTER TT HA
0C21	డ	TELUGU LETTER DD A
0C22	ఢ	TELUGU LETTER DD HA
0C23	ణ	TELUGU LETTER NN A



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

0C24	త	TELUGU LETTER TA
0C25	థ	TELUGU LETTER THA
0C26	ద	TELUGU LETTER DA
0C27	ధ	TELUGU LETTER DHA
0C28	న	TELUGU LETTER NA
0C2A	ప	TELUGU LETTER PA
0C2B	ఫ	TELUGU LETTER PHA
0C2C	బ	TELUGU LETTER BA
0C2D	భ	TELUGU LETTER BHA
0C2E	మ	TELUGU LETTER MA
0C2F	య	TELUGU LETTER YA
0C30	ర	TELUGU LETTER RA





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

0C31	ఱ	TELUGU LETTER RRA
0C32	ల	TELUGU LETTER LA
0C33	ల్ల	TELUGU LETTER LLA
0C35	వ	TELUGU LETTER VA
0C36	ష	TELUGU LETTER SHA
0C37	ష్శ	TELUGU LETTER SSA
0C38	స	TELUGU LETTER SA
0C39	హ	TELUGU LETTER HA
0C58	తా	TELUGU LETTER TSA
0C59	డా	TELUGU LETTER DZA
<b>VOWEL SIGNS (MATRAS) (M)</b>		
0C3E	ా	TELUGU VOWEL SIGN AA
0C3F	ి	TELUGU VOWEL SIGN I



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

0C40	ం	TELUGU VOWEL SIGN II
0C41	ు	TELUGU VOWEL SIGN U
0C42	ూ	TELUGU VOWEL SIGN UU
0C43	ృ	TELUGU VOWEL SIGN VOCALIC R
0C44	ౄ	TELUGU VOWEL SIGN VOCALIC RR
0C46	ె	TELUGU VOWEL SIGN E
0C47	ే	TELUGU VOWEL SIGN EE
0C48	ై	TELUGU VOWEL SIGN AI
0C4A	ొ	TELUGU VOWEL SIGN O
0C4B	ో	TELUGU VOWEL SIGN OO
0C4C	ౌ	TELUGU VOWEL SIGN AU
<b>HALANT (H)</b>		
0C4D	్	TELUGU SIGN VIRAMA



















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

AVAGRAHA (Y)		
0C3D	2	TELUGU SIGN AVAGRAHA



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

## 6. VARIANT TABLE

VARIANTS			
 0C17+0C4D+0C26	 0C17+0C4D+0C27		
 0C17+0C4D+0C17	 0C17+0C4D+0C32		
 0C26+0C4D+0C26	 0C26+0C4D+0C27		
 0C30+0C4D+0C26	 0C30+0C4D+0C2 7	 0C30+0C4D+0C25	 0C30+0C4D+0C23
 0C38+0C4D+0C26	 0C38+0C4D+0C27		
 0C30+0C4D+0C2C	 0C30+0C4D+0C2D		
			



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

0C30+0C4D+0C1 A	0C30+0C4D+0C1B	
ర 0C30+0C4D+0C2 A	ర 0C30+0C4D+0C2B	ర 0C30+0C4D+0C38
చ 0C1A+0C4D+0C1 A	చ 0C1A+0C4D+0C1B	
ప 0C2A+0C4D+0C2 A	ప 0C2A+0C4D+0C2B	ప 0C2A+0C4D+0C38
బ 0C2C+0C4D+0C26	బ 0C2C+0C4D+0C27	బ 0C2C+0C4D+0C32
స 0C38+0C4D+0C2 A	స 0C38+0C4D+0C2B	స 0C38+0C4D+0C38



सत्यमेव जयते  
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 TELUGU

India (Bhārat) localized in Telugu - భారత్

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