



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

# Draft Policy Document For

# INTERNATIONALIZED DOMAIN NAMES

**Language: SINDHI**

**(DEVANAGARI SCRIPT)**



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 February, 2010	Whole Document	M	Language Specific Policy Document for SINDHI (DEVANAGARI)	
1.1	25 April, 2011	Whole Document	M	Formatted to bring in Standard Format.	
1.2	1 March, 2013	All	M,D	Modified character repertoire as per IDNA 2008	1.8



## Table of Contents

1. AUGMENTED BACKUS-NAUR FORMALISM (ABNF) .....	4
1.1 Declaration of variables .....	4
1.2 ABNF Operators .....	4
1.3 The Vowel Sequence.....	4
1.4 The Consonant Sequence .....	5
1.5 Sequence .....	6
1.6 ABNF Applied to the SINDHI (DEVANAGARI SCRIPT) IDN.....	6
2. RESTRICTION RULES .....	7
3. EXAMPLES .....	11
4. LANGUAGE TABLE: SINDHI.....	12
5. NOMENCLATURAL DESCRIPTION TABLE OF SINDHI (DEVANAGARI SCRIPT) LANGUAGE TABLE.....	13
6. VARIANT TABLE FOR SINDHI (DEVANAGARI SCRIPT) .....	17
7. EXPERTISE/BODIES CONSULTED .....	18
8. PROPOSED ccTLD FOR SINDHI (DEVANAGARI SCRIPT).....	19



# 1. 1. AUGMENTED BACKUS-NAUR FORMALISM (ABNF)

## 1.1 Declaration of variables

- Dash → Hyphen -
- Digit → Indo-Arabid digits [0-9]
- C → Consonant
- M → Matra
- V → Vowel
- D → Anusvara
- H → Halant
- N → Nukta

## 1.2 ABNF Operators

Sr. 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 SINDHI (DEVANAGARI SCRIPT) are given.

## 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 Anusvara (D). The number of D which can follow a V in SINDHI (DEVANAGARI SCRIPT) is restricted to one.

The vowel sequence in SINDHI (DEVANAGARI SCRIPT) is therefore V [D]

Examples:



Vowel	V	अ
Vowel + Anusvara	V[D]	अं

## 1.4 The Consonant Sequence

A consonant sequence admits the following shapes:

### 1. A single consonant (C)

(The consonant in SINDHI (DEVANAGARI SCRIPT) shall be treated as co-terminus with the Consonant along with the Nukta sign wherever such a case is pertinent.)

Example:

C क

C[N] क्

### 2. A consonant optionally followed by dependent vowel sign[M] or Anusvara[D] or Halant [H]

C[M|D|H]

Example:

C[M] कि

C[D] कं

C[H] क् (Pure Consonant)

### 2.a. A CM sequence can be optionally followed by [D] (CM)[D]

Example:

CM[D] कीं

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



Example:

CHCHC      न्त्र न्त्र न्त्र न्त्र

3.a. The combination may be followed by M, D or H

Example:

CHC[M]      क्की    क्की

CHC[D]      क्कं    क्कं

CHC[H]      क्क्    क्क्

3.b. \*2(CH)CM may be followed by a D

Example:

CHCM[D]      क्कीं    क्कीं

The final canonical structure of the consonant sequence in IDN can be defined in ABNF as: \*2(C[N]H)C[N][H|D|M[D]]

## 1.5 Sequence

1. A sequence can be made up by Consonant-sequence or Vowel-sequence.

## 1.6 ABNF Applied to the SINDHI (DEVANAGARI SCRIPT) IDN

The formalism can be applied to create/validate IDN labels. So a valid IDN label can be defined as follows.

Vowel-sequence → V [D]

Consonant-sequence → \*2(C[N]H)C[N][H|D|M[D]]

Sequence → consonant-sequence | vowel-sequence

IDN-label → ( sequence | digit ) \* ([dash] (sequence |digit))



## 2. Additional Examples putting more light on ABNF:

Below are some of the examples which will help a casual reader understand some of the rules ABNF puts in place. These are just given for reference purposes and are not meant to be comprehensive.

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

Example:

क्

ंक

िक

2. 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.

3. H is not permitted after V, D, M, digit and dash

Example:

अ्

कंक्

क्ि

1क्

-क्

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

Example

कंं

कींं



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

अं

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

कीी

6. M is not permitted after V  
Example

ई





## 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 SINDHI (DEVANAGARI SCRIPT) the following rules apply:

1. Nukta shall be allowed after the following characters:

क (0915)

ख (0916)

ग (0917)

ज (091C)

ड (0921)

ढ (0922)

फ (092B)

Nukta shall not be allowed for any other consonant.



2. A consonant sequence that is intended to end with Halant [H] can only be followed by Hyphen or digit. Thus following combinations are permissible.

क्-

क्1

3. Consecutive hyphens will not be permitted in a domain name.
4. The number of identical consonants joined by a Halant within a label shall not exceed two. Thus त (ta+halant+ta) is permitted but not त्त (ta+halant+ta+halant+ta).
5. Wherever a variant is present in a given label, the variants shall be in a relationship of transitivity but the generation of the variant table shall be limited only to the relationship existing between the two variants. Thus given a variant त and त्, the number of variants in label such as किताब shall be किताब. कित्ताब generated by adding an extra त् to त shall not be permitted. This ensures that over generativity does not take place.
6. A label containing not more than three "akshara", which have got variants shall be permitted. As an example let us consider a, b, c and d as four aksharas in a given label having a', b', c' and d' as variants in which case such a label will be disallowed. (E.g. of disallowed label - abcd, acdb, cdaba and so on)



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

### 3. EXAMPLES

Combination	Example	Word With Combination
C	भ	भर
CN	ज्ञ	साज्ञ
CH	म्	भारतम्
CM	ता	तारो
CD	गं	गंगा
CMD	सिं	सिंधी
CHC	ष्क	पुष्करु
CHCHC	न्त्र	यन्त्रकार
V	ई	ई
VD	अं	अंग



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

### 3. LANGUAGE TABLE: SINDHI<sup>1</sup>

0900

Devanagari

097F

	090	091	092	093	094	095	096	097
0		ऐ 0910	ठ 0920	र 0930	ी 0940	ॐ 0950	ऋ 0960	० 0970
1	ॐ 0901	ऑ 0911	ड 0921	र 0931	ॐ 0941	ॐ 0951	ॐ 0961	ॐ 0971
2	ं 0902	ओ 0912	ढ 0922	ल 0932	ॐ 0942	ॐ 0952	ॐ 0962	ँ 0972
3	ः 0903	ओ 0913	ण 0923	ळ 0933	ॐ 0943	ॐ 0953	ॐ 0963	
4	ॐ 0904	औ 0914	त 0924	ळ 0934	ॐ 0944	ॐ 0954	। 0964	
5	अ 0905	क 0915	थ 0925	व 0935	ॐ 0945		॥ 0965	
6	आ 0906	ख 0916	द 0926	श 0936	ॐ 0946		० 0966	
7	इ 0907	ग 0917	ध 0927	ष 0937	ॐ 0947		१ 0967	
8	ई 0908	घ 0918	न 0928	स 0938	ॐ 0948	क 0958	र 0968	
9	उ 0909	ङ 0919	त 0929	ह 0939	ॐ 0949	ख 0959	३ 0969	
A	ऊ 090A	च 091A	प 092A		ॐ 094A	ग 095A	४ 096A	
B	ऋ 090B	ॠ 091B	फ 092B		ॐ 094B	ज 095B	५ 096B	ग 097B
C	ॠ 090C	ज 091C	ब 092C	ॠ 093C	ॐ 094C	ड 095C	६ 096C	ज 097C
D	ँ 090D	झ 091D	भ 092D	ॠ 093D	ॐ 094D	ढ 095D	७ 096D	१ 097D
E	ॐ 090E	ज 091E	म 092E	ा 093E		फ 095E	८ 096E	ड 097E
F	ए 090F	ट 091F	य 092F	ि 093F		य 095F	९ 096F	ब 097F

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



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

## 7. 5. NOMENCLATURAL DESCRIPTION TABLE OF SINDHI (DEVANAGARI SCRIPT) LANGUAGE TABLE

<b>Anusvara (D)</b>		
0902	ं	DEVANAGARI SIGN ANUSVARA = bindu
<b>Independent vowels (V)</b>		
0905	अ	DEVANAGARI LETTER A
0906	आ	DEVANAGARI LETTER AA
0907	इ	DEVANAGARI LETTER I
0908	ई	DEVANAGARI LETTER II
0909	उ	DEVANAGARI LETTER U
090A	ऊ	DEVANAGARI LETTER UU
090B	ऋ	DEVANAGARI LETTER VOCALIC R
090F	ए	DEVANAGARI LETTER E
0910	ऐ	DEVANAGARI LETTER AI
0913	ओ	DEVANAGARI LETTER O
0914	औ	DEVANAGARI LETTER AU
<b>Consonants (C)</b>		
0915	क	DEVANAGARI LETTER KA
0916	ख	DEVANAGARI LETTER KHA
0917	ग	DEVANAGARI LETTER GA



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

0918	घ	DEVANAGARI LETTER GHA
0919	ङ	DEVANAGARI LETTER NGA
091A	च	DEVANAGARI LETTER CA
091B	छ	DEVANAGARI LETTER CHA
091C	ज	DEVANAGARI LETTER JA
091D	झ	DEVANAGARI LETTER JHA
091E	ञ	DEVANAGARI LETTER NYA
091F	ट	DEVANAGARI LETTER TTA
0920	ठ	DEVANAGARI LETTER TTHA
0921	ड	DEVANAGARI LETTER DDA
0922	ढ	DEVANAGARI LETTER DDHA
0923	ण	DEVANAGARI LETTER NNA
0924	त	DEVANAGARI LETTER TA
0925	थ	DEVANAGARI LETTER THA
0926	द	DEVANAGARI LETTER DA
0927	ध	DEVANAGARI LETTER DHA
0928	न	DEVANAGARI LETTER NA
092A	प	DEVANAGARI LETTER PA
092B	फ	DEVANAGARI LETTER PHA



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

092C	ब	DEVANAGARI LETTER BA
092D	भ	DEVANAGARI LETTER BHA
092E	म	DEVANAGARI LETTER MA
092F	य	DEVANAGARI LETTER YA
0930	र	DEVANAGARI LETTER RA
0932	ल	DEVANAGARI LETTER LA
0935	व	DEVANAGARI LETTER VA
0936	श	DEVANAGARI LETTER SHA
0937	ष	DEVANAGARI LETTER SSA
0938	स	DEVANAGARI LETTER SA
0939	ह	DEVANAGARI LETTER HA
097B	ग	DEVANAGARI LETTER GGA
097C	ज	DEVANAGARI LETTER JJA
097E	ड	DEVANAGARI LETTER DDDA
097F	ब	DEVANAGARI LETTER BBA
<b>Dependent vowel signs (Matras) (M)</b>		
093E	ा	DEVANAGARI VOWEL SIGN AA
093F	ि	DEVANAGARI VOWEL SIGN I • stands to the left of the consonant
0940	ी	DEVANAGARI VOWEL SIGN II
0941	ु	DEVANAGARI VOWEL SIGN U



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

0942	ॄ	DEVANAGARI VOWEL SIGN UU
0943	ॆ	DEVANAGARI VOWEL SIGN VOCALIC R
0947	ॐ	DEVANAGARI VOWEL SIGN E
0948	ॐ	DEVANAGARI VOWEL SIGN AI
094B	ॐ	DEVANAGARI VOWEL SIGN O
094C	ॐ	DEVANAGARI VOWEL SIGN AU
<b>Halant (H)</b>		
094D	्	DEVANAGARI SIGN VIRAMA = halant (the preferred SINDHI (DEVANAGARI SCRIPT) name) • suppresses inherent vowel
<b>Nukta (N)</b>		
093C	ृ	DEVANAGARI SIGN NUKTA





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

## 6. VARIANT TABLE FOR SINDHI (DEVANAGARI SCRIPT)

VARIANTS		
द्र 0926+094D+0917	द्र 0926+094D+0930	द्र 0926+094D+0928
द्व 0926+094D+0927	द्व 0926+094D+0918	
ष्ट 0937+094D+091F	ष्ठ 0937+094D+0920	
श्च 0936+094D+0935	श्च 0936+094D+0930+094D+0935	
श्न 0936+094D+0928	श्न 0936+094D+0930+094D+0928	
श्च 0936+094D+091A	श्च 0936+094D+0930+094D+091A	
श्ल 0936+094D+0932	श्ल 0936+094D+0930+094D+0932	
त्त 0924+094D+0924	त्त 0924	
द्व 0926+094D+0935	द्व 0926+094D+092C	

Note : श्न and श्ल normally not used in SINDHI (DEVANAGARI SCRIPT)  
but are introduced since some browsers may display these shapes.



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

## 7. EXPERTISE/BODIES CONSULTED

Expertise provided by Dr. K.L. Lekhwani CIIL.



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

## 8. 8. PROPOSED ccTLD FOR SINDHI (DEVANAGARI SCRIPT)

India ( Bhārat) localized in Sindhi - भारत

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