



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

Draft Policy Document For

INTERNATIONALIZED DOMAIN NAMES

Language: DOGRI



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	21 January, 2010	Whole Document	M	Language Specific Policy Document for DOGRI	
1.1	22 March, 2011	Whole Document	M	Description of sequence added, Variant removed	
1.2	13 January, 2012	Page5,6,7	M	Inclusion of Vowel Modifier(MODIFIER LETTER APOSTROPHE) [S] after Matra [M]	
1.3	01 January, 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 Consonant Sequence	5
1.5 Sequence	7
1.6 ABNF Applied to the DOGRI IDN	7
2. RESTRICTION RULES	10
3. EXAMPLES:	12
4. LANGUAGE TABLE: DOGRI.....	13
5. NOMENCLATRURAL DESCRIPTION TABLE OF DOGRI LANGUAGE TABLE..	14
6. VARIANT TABLE FOR DOGRI	17
7. EXPERTISE/BODIES CONSULTED	18
8. PROPOSED ccTLD FOR DOGRI	19



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
- Y → Avagraha
- H → Halant
- S → Vowel Modifier
- N → Nukta

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 DOGRI 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) or Vowel Modifier (MODIFIER LETTER APOSTROPHE) (S). The number of D or S which can follow a V in



DOGRI should be restricted to one.

The vowel sequence in DOGRI is therefore V [D|S]

Examples:

Vowel	V	अ
Vowel + Anusvara	V[D]	अं
Vowel + Vowel Modifier	V[S]	अ'

1.4 Consonant Sequence

A consonant sequence admits the following shapes:

1. A single consonant (C)

(The consonant in DOGRI 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] ज्ञ

2. A consonant optionally followed by dependent vowel sign[M], Anusvara[D], Halant [H] or Vowel Modifier [S]

C[M|D|H|S]

Example:

C[M] कि

C[D] कं

C[S] क'

C[H] क् (Pure Consonant)

2.a. A CM sequence can be optionally followed by Anusvara [D] or Vowel Modifier [S].

CM[D|S]

Example:





CM[D] कीं

CM[S] कुंन

3. A sequence of consonants (up to 3) joined by Halant

*2(CH)C

Example:

CHCHC स्र स+्+प+्+र

Subsets

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

Example:

CHC[M] क्की क्की

CHC[D] क्कं क्कं

CHC[S] क्क' क्क'

3.b. *2(CH)CM may be followed by a Anusvara [D], Vowel Modifier(MODIFIER LETTER APOSTROPHE)[S]

Example:

CHCM[D] क्कीं क्कीं

CHCM[S] क्की' क्की'

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

*2(C[N]H)C[N][H|D|S|M[D|S]]



1.5 Sequence

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

1.a A Consonant-sequence can optionally be followed by Avagraha[Y].

1.b A Vowel-sequence can optionally be followed by Avagraha [Y].

1.6 ABNF Applied to the DOGRI IDN

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

Vowel-sequence \rightarrow V [D | S]

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

Sequence \rightarrow consonant-sequence[Y] | vowel-sequence[Y]

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



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|S cannot occur in the beginning of an IDN domain name

Example:

क्
ंक
िक
'क

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.

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

Example:

अ्
कं्
क्ि
1्
-्
'्

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

Example

कंं



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

कीं

अं

क”

की”

अ”

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

Example

कीी

5. 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 DOGRI the following rules apply:

1. For DOGRI nukta shall be allowed after the following characters:

ज (091C)

ड (0921)

ढ (0922)

फ (092B)

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

क्-

क्1

क्ऽ

3. Distribution of S :

The Vowel Modifier 02BC ' shall have the following distribution.

The character forms part of a syllable.

As ABNF shows, it cannot occur at the beginning of a syllable

The Vowel Modifier cannot occur after a Halant i.e. after a pure consonant.

The Vowel Modifier can occur at the end of a syllable.

V [S] इ'यां उ'आं



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

M [S] कु'न

C[S] स'जां-लै न'ग

4. Consecutive hyphens will not be permitted in a domain name.
5. 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).
6. 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 त्त, the number of variants in label such as किताब shall be किताब. किताब generated by adding an extra त् to त्त shall not be permitted. This ensures that over generativity does not take place.
7. 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)



3. EXAMPLES:

Combination	Example	Word With Combination
C	क	कर
CN	ड़	पेड़
CH	ह्	चाह्
CM	ता	ताल
CD	गं	गंगा
CS	स'	स'ज
CMD	हिं	हिंदी
CHC	व्य	काव्य
CHCHC	स्त्र	शास्त्र
V	आ	आन
VD	अं	अंग
VDY	आंऽ	आंऽ
CNMDY	ड़ंऽ	पड़ंऽ



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

4. LANGUAGE TABLE: DOGRI¹

0900

Devanagari

097F

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

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

65

¹ Characters marked in yellow are not applicable to the language.



5. NOMENCLATURAL DESCRIPTION TABLE OF DOGRI LANGUAGE TABLE

Anusvara (D)		
0902	ं	DEVANAGARI SIGN ANUSVARA = bindu
Vowel Modifier 02BC (S)		
02BC	,	Modifier Letter Apostrophe
Independent vowels (V)		
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
0911	ऑ	DEVANAGARI LETTER CANDRA O
0913	ओ	DEVANAGARI LETTER O
0914	औ	DEVANAGARI LETTER AU
Consonants (C)		
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
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
Dependent vowel signs (Matras) (M)		
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
0942	ू	DEVANAGARI VOWEL SIGN UU
0943	ृ	DEVANAGARI VOWEL SIGN VOCALIC R
0947	े	DEVANAGARI VOWEL SIGN E
0948	ै	DEVANAGARI VOWEL SIGN AI
0949	ॉ	DEVANAGARI VOWEL SIGN CANDRA O
094B	ो	DEVANAGARI VOWEL SIGN O
094C	ौ	DEVANAGARI VOWEL SIGN AU
Halant (H)		
094D	्	DEVANAGARI SIGN VIRAMA = halant (the preferred DOGRI name) • suppresses inherent vowel
Nukta (N)		
093C	्	DEVANAGARI SIGN NUKTA
Avagraha (Y)		
093D	ः	DEVANAGARI SIGN AVAGRAHA



6. VARIANT TABLE FOR DOGRI

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	
श्व 0936+094D+0935	श्व 0936+094D+0930+094D+0935	
श्र 0936+094D+0928	श्र 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	
त 0924+094D+0924	त 0924	
द्व 0926+094D+0935	द्व 0926+094D+092C	

Note : श्र and श्ल normally not used in DOGRI 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. Shashi Pathania.



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

8. PROPOSED ccTLD FOR DOGRI

India (Bhārat) localized in Dogri - भारत

Note : You can send your feedbacks to ids-feedback@cdac.in