



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

Draft Policy Document For

INTERNATIONALIZED DOMAIN NAMES

Language: PUNJABI



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 PUNJABI	1.5
1.1	22/12/2010	Page No. 5, 6, 7, 20	M	Remove Explicit Halant from the consonant syllable, 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 Declaration of variable.....	4
1.2 ABNF Operators.....	4
1.3 The Vowel Sequence.....	5
1.4 Consonant Sequence.....	5
1.5 ABNF Applied to the IDN.....	7
2. RESTRICTION RULES.....	8
4. LANGUAGE TABLE: PUNJABI.....	12
5. NOMENCLATURAL DESCRIPTION TABLE OF PUNJABI LANGUAGE TABLE	13
6. VARIANT TABLE.....	18
7. EXPERTS/BODIES CONSULTED.....	19
8. PROPOSED ccTLD FOR PUNJABI.....	20

1. AUGMENTED BACKUS-NAUR FORMALISM (ABNF)

1.1 Declaration of variable

Dash	→ Hyphen -
Digit	→ Indo-Arabic digits [0-9]
C	→ Consonant
M	→ Matra
V	→ Vowel
D	→ Bindi / Tippi
X	→ Visarga
Y	→ Avagraha
H	→ Halant
A	→ Addak
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 PUNJABI 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 a Bindi/Tippi (D) or Visarga(X). The number of D or X which can follow a V in PUNJABI may be restricted to one.

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

Examples :

V	ਅ	ਅ
V[D]	ਅੰ	ਅੰ
V[D]	ਅੰ	ਅੰ
V[X]	ਅਃ	ਅਃ

1.4 Consonant Sequence

A consonant sequence admits the following shapes:

1. A single consonant (C)

Example:

C	ਕ	ਕ
---	---	---

2. A consonant optionally followed by dependent vowel sign[M], Bindi/Tippi[D] or Visarga[X]

C[M|D|X]

Example:

C[M]	ਕੀ	ਕਿ
C[D]	ਕੰ	ਕੰ
C[X]	ਕਃ	ਕਃ

3. A CM sequence can be optionally followed by D or X

Example:

CM[D] की की

CM[X] काः काः

4. A sequence of consonants (up to 2) joined by Halanta (CH)C

Example:

CHC ख ख+्+र क+्+र

5. Addak is used for germination. It doubles the following consonant.

Example:

AC¹ क् क् क

6. A sequence of Addak consonant may be followed by Bindi/Tippi[D], Visarga[X] or Matra[M]

Example:

AC[D] क् क् कं

AC[X] क्ः क् कः

AC[M] क्कि क् क्कि

7. A sequence of Addak consonant matra ACM may be followed by Bindi/Tippi[D], Visarga[X]

Example:

ACM[D]² क्कि क् क्किं

ACM[X] क्किः क् क्किः

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

¹ Addak falls on the preceding consonant or a vowel syllable.

² These combinations are extremely rare in punjabi and not encountered very frequently.



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

$(C[N]H) C[N] [D|X|M[D|X]] \mid AC[D|X|M[D|X]]$

Note: D represents either Bindi or Tippi. Combinations of the type CHCMD are possible only in case of Sanskrit words transcribed in PUNJABI such as क्रांती which a client could possibly want to use.

1.5 ABNF Applied to the IDN

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

Vowel-sequence $\rightarrow V [D|X]$

Consonant-sequence $\rightarrow (C[N]H) C[N] [D|X|M[D|X]] \mid AC[D|X|M[D|X]]$

Sequence \rightarrow consonant-sequence \mid vowel-sequence

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 PUNJABI the following rules apply:

1. Tippi will be used in place of bindi if the preceding character is one of the followings:
 1. A consonant
 2. Short e matra ਿ (0A3F)
 3. short u and long u matra ੁ (0A41) ੂ (0A42)
 4. Vowel ਅ (0A05) ਏ (0A07)

2. D, X ,H, A or M cannot occur in the beginning of an IDN domain name

Example:

ਿਕ ਿਕ
ਂਕ ਂਕ
ਕ਼ਕ ਕ਼ਕ
ਗ਼ ਗ਼
ੱਕ ੱਕ

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

3. Halanta

- 3.1. H is not permitted after V, D, X, M, digit and dash

Example:



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

अ्	अ्
कं	कं
कः	कः
क्	क्
1्	1्
-्	-्

3.2 The Halanta shall be permitted only with the following:

Consonant + ्(0A4D) + ञ(0A2F)

Consonant + ्(0A4D) + ञ(0A30)

Consonant + ्(0A4D) + ञ(0A35)

Consonant + ्(0A4D) + ञ(0A39)

4. Addak

4.1 It neither be permitted at the beginning nor at the end of the word.

Example:

ँक

4.2 Addak can not be followed by a halant,visarga and bindi/tippi.

Example:

ँँ

ँःँ

ँँँ

ँँँ

4.3 Addak is used for germination.



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

ਅੱਕਾ ਅਕਕਾ

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

Example:

ਕੰੰ	ਕਂਂ
ਕੀਂਂ	ਕੀਂਂ
ਅੰੰ	ਅਂਂ

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

Example:

ਕੀੀ	ਕੀੀ
-----	-----

7. M is not permitted after V

Example:

ਈਾ	ਈਾ
----	----

8. Nukta can be only allowed after following characters:

ਖ (0A16)

ਗ (0A17)

ਜ (0A1C)

ਫ (0A2B)

ਲ (0A32)

ਮ (0A38)



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

3. EXAMPLES

Following are the some examples to show PUNJABI IDN Label.

Combination	Example	Word With Combination
C	ਕ	ਕਰ
CM	ਰਾ	ਰਾਮ
CD	ਹੰ	ਹੰਸ
CMD	ਹਿੰ , ਨੀਂ	ਹਿੰਦੀ , ਨੀਂਦ
CHC	ਗ੍ਰ	ਗ੍ਰਥ
V	ਅ	ਅਮਨ
VD	ਅੰ	ਅੰਦਰ
AC	ਰੱ	ਰੱਬ
ACM	ੱਕਾ	ਪੱਕਾ



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

4. LANGUAGE TABLE: PUNJABI¹

0A00

Gurmukhi

0A7F

	0A0	0A1	0A2	0A3	0A4	0A5	0A6	0A7
0		ਐ	ਠ	ਰ	ੀ			ੰ
1	ਏ		ਡ		ੳ	ੲ		ੱ
2	ਂ		ਢ	ਲ	ੳ			ੲ
3	ੳ	ਓ	ਣ	ਲ				ੲ
4		ਐ	ਤ					ੳ
5	ਅ	ਕ	ਬ	ਵ				ੳ
6	ਆ	ਖ	ਦ	ਸ਼			ੳ	
7	ਇ	ਗ	ਧ		ੇ		ੳ	
8	ਈ	ਘ	ਨ	ਸ਼	ੈ		ੳ	
9	ਉ	ਙ		ਹ		ਖ਼	ੳ	
A	ਊ	ਚ	ਪ			ਗ਼	ੳ	
B		ਛ	ਫ		ੋ	ਜ਼	ੳ	
C		ਜ	ਬ	੍	ੈ	ੜ	ੳ	
D		ਝ	ਭ		੍		ੳ	
E		ਵ	ਮ	ਾ		ਫ਼	ੳ	
F	ਏ	ਟ	ਯ	ਿ			ੳ	

¹ 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 PUNJABI LANGUAGE TABLE

Unicode Value	Character	Description
BINDI / TIPPI (D)		
0A02	ੰ	GURMUKHI SIGN BINDI
0A70	ੰ	GURMUKHI TIPPI
VISARGA (X)		
0A03	ਃ	GURMUKHI SIGN VISARGA
VOWEL LETTERS (V)		
0A05	ਅ	GURMUKHI LETTER A
0A06	ਆ	GURMUKHI LETTER AA
0A07	ਇ	GURMUKHI LETTER I
0A08	ਈ	GURMUKHI LETTER II
0A09	ਉ	GURMUKHI LETTER U
0A0A	ਊ	GURMUKHI LETTER UU
0A0F	ਏ	GURMUKHI LETTER EE
0A10	ਐ	GURMUKHI LETTER AI
0A13	ਓ	GURMUKHI LETTER OO



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

0A14	ਐ	GURMUKHI LETTER AU
CONSONANTS (C)		
0A15	ਕ	GURMUKHI LETTER KA
0A16	ਖ	GURMUKHI LETTER KHA
0A17	ਗ	GURMUKHI LETTER GA
0A18	ਘ	GURMUKHI LETTER GHA
0A19	ਙ	GURMUKHI LETTER NGA
0A1A	ਚ	GURMUKHI LETTER CA
0A1B	ਛ	GURMUKHI LETTER CHA
0A1C	ਜ	GURMUKHI LETTER JA
0A1D	ਝ	GURMUKHI LETTER JHA
0A1E	ਞ	GURMUKHI LETTER NYA
0A1F	ਟ	GURMUKHI LETTER TTA
0A20	ਠ	GURMUKHI LETTER TTHA
0A21	ਡ	GURMUKHI LETTER DDA
0A22	ਢ	GURMUKHI LETTER DDHA
0A23	ਣ	GURMUKHI LETTER NNA



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

0A24	ਤ	GURMUKHI LETTER TA
0A25	ਥ	GURMUKHI LETTER THA
0A26	ਦ	GURMUKHI LETTER DA
0A27	ਧ	GURMUKHI LETTER DHA
0A28	ਨ	GURMUKHI LETTER NA
0A2A	ਪ	GURMUKHI LETTER PA
0A2B	ਫ	GURMUKHI LETTER PHA
0A2C	ਬ	GURMUKHI LETTER BA
0A2D	ਭ	GURMUKHI LETTER BHA
0A2E	ਮ	GURMUKHI LETTER MA
0A2F	ਯ	GURMUKHI LETTER YA
0A30	ਰ	GURMUKHI LETTER RA
0A32	ਲ	GURMUKHI LETTER LA
0A33	ਲ਼	GURMUKHI LETTER LLA
0A35	ਵ	GURMUKHI LETTER VA
0A36	ਸ਼	GURMUKHI LETTER SHA
0A38	ਸ	GURMUKHI LETTER SA



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

0A39	ਹ	GURMUKHI LETTER HA
0A59	ਖ	GURMUKHI LETTER KHHA
0A5A	ਗ	GURMUKHI LETTER GHHA
0A5B	ਜ	GURMUKHI LETTER ZA
0A5C	ੜ	GURMUKHI LETTER RRA
0A5E	ਫ	GURMUKHI LETTER FA
NUKTA (N)		
0A3C	ੜ	GURMUKHI SIGN NUKTA
VOWEL SIGNS (MATRAS) (M)		
0A3E	ਾ	GURMUKHI VOWEL SIGN AA
0A3F	ਿ	GURMUKHI VOWEL SIGN I
0A40	ੀ	GURMUKHI VOWEL SIGN II
0A41	ੁ	GURMUKHI VOWEL SIGN U
0A42	ੂ	GURMUKHI VOWEL SIGN UU
0A47	ੇ	GURMUKHI VOWEL SIGN EE
0A48	ੈ	GURMUKHI VOWEL SIGN AI
0A4B	ੋ	GURMUKHI VOWEL SIGN OO
0A4C	ੌ	GURMUKHI VOWEL SIGN AU



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

VIRAMA (H)		
0A4D	ॠ	GURMUKHI SIGN VIRAMA
ADDAK (A)		
0A71	ॡ	GURMUKHI ADDAK



6. VARIANT TABLE

The variant table is restricted to characters which are akhand forms and which can be generated out by typing the pertinent consonant followed by the Nukta.

Variants	
ਖ 0A16+0A3C	ਖ 0A59
ਗ 0A17++0A3C	ਗ 0A5A
ਜ 0A1C+0A3C	ਜ 0A5B
ਫ 0A2B+0A3C	ਫ 0A5E
ਲ 0A32+0A3C	ਲ 0A33
ਮ 0A38+0A3C	ਮ 0A36



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

7. EXPERTS/BODIES CONSULTED

The template designed for Gurmukhi was submitted to Dr G.S. Lehal for comments and subsequently refined during the PRSG meeting which Dr Lehal attended as Language Expert.



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

8. PROPOSED ccTLD FOR PUNJABI

India (Bhārat) localized in Punjabi - **ਭਾਰਤ**

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