



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

# **Draft Policy Document For INTERNATIONALIZED DOMAIN NAMES**

**Language: GUJARATI**



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 GUJARATI	1.5
1.1	22/11/2010	Page No 9, 16, 18	A, D	Restriction rule added, Variant deleted, 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).....	5
1.1 Declaration of variable.....	5
1.2 ABNF Operators.....	5
1.3 The Vowel Sequence.....	6
1.4 Consonant Sequence.....	6
1.5 ABNF Applied to the GUJARATI IDN.....	8
2. RESTRICTION RULES.....	9
3. EXAMPLES.....	11
4. LANGUAGE TABLE: GUJARATI.....	12
5. NOMENCLATURAL DESCRIPTION TABLE OF GUJARATI LANGUAGE TABLE.....	13
6. VARIANT TABLE .....	17
7. EXPERTS/BODIES CONSULTED.....	18
8. PROPOSED ccTLD FOR GUJARATI .....	19

# 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 → Anusvara  
B → Chandrabindu (Used very rarely in Gujarati)  
X → Visarga  
Y → Avagraha  
H → Halant

## 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 Gujarati 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 Anusvara (D), Chandrabindu (B) or a Visarga (X)  
The number of D,B or X which can follow a V in Gujarati may be restricted to one.

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

Examples :

V	અ	अ
V[D]	અં	अं
V[B]	अँ	अँ
V[X]	अः	अः

Standard Gujarati does not use Chandrabindu, although the same is used for Sanskrit words.

### 1.4 Consonant Sequence

A consonant sequence admits the following shapes:

1. A single consonant (C)

Example:

C	ક	क
---	---	---

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

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

Example:

C[M]	કી	की
------	----	----



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

C[D]	डं	कं
C[B]	डँ	कँ
C[X]	डः	कः
C[H]	ड्ह	क् (Pure Consonant)

2.a. A CM sequence can be optionally followed by \*1[D|B|X]  
(CM) \*1[D|B|X]

Example:

CM[D]	डीं	कीं
CM[B]	डँ	काँ
CM[X]	वीः	वीः

3. A sequence of consonants (up to 4) joined by Halanta \*3(CH)C

Example:

CHCHCHC क्च्य → क् + च् + र् + य् न+क्+र+य

Subsets

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

Example:

CHC[M]	क्की	क्की	क्की
CHC[D]	क्कं	क्कं	क्कं
CHC[B]	क्कँ	क्कँ	क्कँ
CHC[X]	क्कः	क्कः	क्कः
CHC[H]	क्क्	क्क्	क्क्



3.b. \*3(CH)CM may be followed by a \*1[D|B|X]

Example:

CHCM[D] ક્કી ક્કીં વ્કી વ્કીં

CHCM[B] ક્કૃ ક્કૃં વ્કૃ વ્કૃં

CHCM[X] ક્કીઃ ક્કીંઃ વ્કીઃ વ્કીંઃ

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

**\*3(CH)C [H|D|B|X |M|D|B|X]]**

It should be noted that in the case of Sanskrit the possible sequence could be:

**\*4(CH)C [H|D|B|X |M|D|B|X]]**

as exemplified in the word:

કાત્સ્ન્ય કાત્સ્ન્ય

Although this is possible, it is not used in Gujarati unless a potential client chooses it as a possible IDN.

## 1.5 ABNF Applied to the GUJARATI 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 | B |X]

Consonant-sequence → \*3(CH)C [H|D|B|X |M|D|B|X]]

Sequence → consonant-sequence [Y] | vowel-sequence [Y]

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 Gujarati the following rules apply:

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

Example

્ક	્ક
િક	િક
ંક	ંક
ૅક	ૅક
ઃક	ઃક

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, D,B, X, M, digit and dash

Example

અ્	અ્
કં્	કં્
કૅ્	કૅ્
કઃ્	કઃ્
ફી	ફિ
1્	1્





-ॢ

-ॣ

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

Example

कं

कं

कँ

कँ

कः

कः

कीं

कीं

कीः

कीः

अं

अं

अँ

अँ

अः

अः

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

Example

कीी

कीी

5. M is not permitted after V

Example

ईी

ईी

6. The combination of Anusvara + Visarga[DX], Chandrabindu + Anusvara[BD], chandrabindu + Visarga[BX] and vice-versa is not permissible

Example

कंः कः

कँ कं

कः कँ

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



### 3. EXAMPLES

Following are the some examples of GUJARATI IDN.

Combination	Example	Word containing example
C	ર	રમ
CH	ફ	વાફ (Rarely used)
CM	તુ	તુ
CD	પં	પંપ
CX	શઃ	અક્ષરશઃ Rarely used
CMD	માં	માં
CMB	કાં	કાંચ
CMX	દુઃ	દુઃખ
CHC	ક	કમ
CHCHC	સ્પ્ર	સ્પ્રે
CHCHCHC	ક્ષ્ત્ર	એક્ષ્ત્રા
V	આ	આજ
VD	અં	અંક
VB	આં	આંખ
VX	આઃ	આઃ



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

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

### GUJARATI PERMISSIBLE CHARACTER SET

	0A8	0A9	0AA	0AB	0AC	0AD	0AE	0AF
0		ૐ	ઠ	ર	ી	ૐ	જ	
1	ૐ	ઑ	ડ		ુ		ૐ	ૐ
2	ં		ઢ	લ	ૂ		ૐ	
3	ઃ	ઑ	ણ	ળ	ૃ		ૐ	
4		ઑ	ત		ૃ			
5	અ	ક	થ	વ	ે			
6	આ	ખ	દ	શ			ૐ	
7	ઈ	ગ	ધ	પ	ે		ૐ	
8	ઈ	ઘ	ન	સ	ૈ		ૐ	
9	ઉ	ડ		હ	ૌ		ૐ	
A	ઊ	ચ	પ				ૐ	
B	જ	છ	ફ		ૌ		ૐ	
C	ૐ	જ	બ	ૐ	ૌ		ૐ	
D	ઐ	ઝ	ભ	ડ	ૃ		ૐ	
E		ઝ	મ	ા			ૐ	
F	ઐ	ટ	ય	િ			ૐ	

<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 GUJARATI LANGUAGE TABLE

<b>Candrabindu (B)</b>		
0A81	ँ	GUJARATI SIGN CANDRABINDU
<b>Anusvara (D)</b>		
0A82	ं	GUJARATI SIGN ANUSVARA
<b>Visarga (X)</b>		
0A83	ઃ	GUJARATI SIGN VISARGA
<b>Vowels (V)</b>		
0A85	અ	GUJARATI LETTER A
0A86	આ	GUJARATI LETTER AA
0A87	ઇ	GUJARATI LETTER I
0A88	ઈ	GUJARATI LETTER II
0A89	ઉ	GUJARATI LETTER U
0A8A	ઊ	GUJARATI LETTER UU
0A8B	ૠ	GUJARATI LETTER VOCALIC R
0A8D	એ	GUJARATI VOWEL CANDRA E
0A8F	૨	GUJARATI LETTER E
0A90	૩	GUJARATI LETTER AI
0A91	૪	GUJARATI LETTER CANDRA O
0A93	૫	GUJARATI LETTER O
0A94	૬	GUJARATI LETTER AU
<b>Consonants (C)</b>		



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

0A95	ક	GUJARATI LETTER KA
0A96	ખ	GUJARATI LETTER KHA
0A97	ગ	GUJARATI LETTER GA
0A98	ઘ	GUJARATI LETTER GHA
0A99	ઙ	GUJARATI LETTER NGA
0A9A	ચ	GUJARATI LETTER CA
0A9B	છ	GUJARATI LETTER CHA
0A9C	જ	GUJARATI LETTER JA
0A9D	ઝ	GUJARATI LETTER JHA
0A9E	ઞ	GUJARATI LETTER NYA
0A9F	ટ	GUJARATI LETTER TTA
0AA0	ઠ	GUJARATI LETTER TTHA
0AA1	ડ	GUJARATI LETTER DDA
0AA2	ઢ	GUJARATI LETTER DDHA
0AA3	ણ	GUJARATI LETTER NNA
0AA4	ત	GUJARATI LETTER TA
0AA5	થ	GUJARATI LETTER THA
0AA6	દ	GUJARATI LETTER DA
0AA7	ધ	GUJARATI LETTER DHA
0AA8	ન	GUJARATI LETTER NA
0AAA	પ	GUJARATI LETTER PA
0AAB	ફ	GUJARATI LETTER PHA



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

0AAC	બ	GUJARATI LETTER BA
0AAD	ભ	GUJARATI LETTER BHA
0AAE	મ	GUJARATI LETTER MA
0AAF	ય	GUJARATI LETTER YA
0AB0	ર	GUJARATI LETTER RA
0AB2	લ	GUJARATI LETTER LA
0AB3	ળ	GUJARATI LETTER LLA
0AB5	વ	GUJARATI LETTER VA
0AB6	શ	GUJARATI LETTER SHA
0AB7	ષ	GUJARATI LETTER SSA
0AB8	સ	GUJARATI LETTER SA
0AB9	હ	GUJARATI LETTER HA
<b>Dependent Vowel Signs (MATRAS) (M)</b>		
0ABE	ા	GUJARATI VOWEL SIGN AA
0ABF	િ	GUJARATI VOWEL SIGN I
0AC0	ી	GUJARATI VOWEL SIGN II
0AC1	ુ	GUJARATI VOWEL SIGN U
0AC2	ૂ	GUJARATI VOWEL SIGN UU
0AC3	ૃ	GUJARATI VOWEL SIGN VOCALIC R
0AC5	ૈ	GUJARATI VOWEL SIGN CANDRA E
0AC7	ે	GUJARATI VOWEL SIGN E
0AC8	ૐ	GUJARATI VOWEL SIGN AI



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

0AC9	ૌ	GUJARATI VOWEL SIGN CANDRA O
0ACB	ો	GUJARATI VOWEL SIGN O
0ACC	ૌ	GUJARATI VOWEL SIGN AU
<b>Avagraha (Y)</b>		
0ABD	ઽ	GUJARATI SIGN AVAGRAHA
<b>Halant (H)</b>		
0ACD	્	GUJARATI SIGN VIRAMA



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

## 6. VARIANT TABLE

VARIANTS		
झ		झ
0AAB+ 0AAF		0AAB+ 0ACD+ 0AAF
झ		झ
0AA6+ 0ACD+0AA7		0AA6+ 0ACD+0A98
झ		झ
0AA6+ 0ACD+0AAC		0AA6+ 0ACD+0AB5
झ	झ	झ
0AA6+ 0ACD+0AB0	0AA6+ 0ACD+0AA8	0AA6+ 0ACD+0A97





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

## 7. EXPERTS/BODIES CONSULTED

Mr. Ashok Karania C.E.O. Magnet Technologies in consultation with Gujarati Sahitya Parishad .



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

## 8. PROPOSED ccTLD FOR GUJARATI

India (Bhārat) localized in Gujarati - ભારત

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