

ॐ असतो मा सद्गमय । तमसो मा ज्योतिर्गमय ॥

मृत्योर्मांमृतं गमय । ॐ शान्ति शान्ति शान्ति ॥

ॐ असतो मा सद्गमय । तमसो मा ज्योतिर्गमय ॥

मृत्योर्मांमृतं गमय । ॐ शान्ति शान्ति शान्ति ॥

ॐ असतो मा सद्गमय । तमसो मा ज्योतिर्गमय ॥

**DEVANĀGARĪ VIP TEAM ISSUES  
REPORT**

**DEVANĀGARĪ VIP GROUP**

मृत्योर्मांमृतं गमय । ॐ शान्ति शान्ति शान्ति ॥

ॐ असतो मा सद्गमय । तमसो मा ज्योतिर्गमय ॥

मृत्योर्मांमृतं गमय । ॐ शान्ति शान्ति शान्ति ॥

ॐ असतो मा सद्गमय । तमसो मा ज्योतिर्गमय ॥

मृत्योर्मांमृतं गमय । ॐ शान्ति शान्ति शान्ति ॥

ॐ असतो मा सद्गमय । तमसो मा ज्योतिर्गमय ॥

मृत्योर्मांमृतं गमय । ॐ शान्ति शान्ति शान्ति ॥

## विषय-सूची

0. प्रारंभिक तैयारियाँ.....	4
1. पृष्ठभूमि और सिंहावलोकन.....	4
2. संरचना.....	5
1. आधार तत्त्व.....	6
2. देवनागरी: एक सिंहावलोकन.....	7
2.1. देवनागरी: एक ऐतिहासिक संदर्भ.....	7
2.2. लिखित देवनागरी की संरचना.....	7
2.3. मूलभूत इकाई: अक्षर.....	9
3. विषय.....	12
3.1. भाषा बनाम लिपि संबंधी मुद्दे.....	12
3.2. देवनागरी लिपि में भिन्न रूप.....	12
3.2.1. भ्रामक रूप से एक समान एकल वर्ण.....	13
3.2.2. भ्रामक रूप से एक समान संयुक्त वर्ण.....	13
3.2.3. वर्णों को संयुक्त करने की वजह से उत्पन्न भिन्न रूप.....	14
3.3. डोमेनों के प्रदर्शन के संबंध में सॉफ्टवेयर व्यवहार से संबंधित मुद्दे:.....	14
3.3.1. ब्राउजर संबंधी मुद्दे.....	14
3.3.2. ईमेल पतों का रिजोल्यूशन.....	16
3.4. देवनागरी लिपि में 02BC का मामला:.....	16
3.5. संपूर्ण लिपि भ्रांतियोंसे युक्त.....	17
4. विषयेतर विचार.....	18
4.1. संकर लिपि वर्ण मिश्रण.....	18
4.2. वर्तनी के माध्यम से उत्पन्न समस्वन.....	22
4.3. शून्य चौड़ाई वाले योजक यानी जीरो विड्थ ज्वाइनर ZWJ और जीरो विड्थ नॉन-ज्वाइनर ZWNJ.....	22
4.4. प्रशासनिक मुद्दे.....	23
4.5. बहुभाषी gTLD (जीटीएलडी) का प्रबंधन.....	25
5. रजिस्ट्रार और रजिस्ट्री संदर्भ.....	26
5.1. DNS : (डीएनएस) प्रौद्योगिकी और ऑपरेशंस परिप्रेक्ष्य.....	27
5.2. सुरक्षा और स्थिरता.....	28
5.3. उपयोगकर्ता परिप्रेक्ष्य.....	28
5.4. सिस्टम एडमिनिस्ट्रेटर संबंधी परिप्रेक्ष्य.....	29

5.5.	अंतिम उपयोगकर्ता संबंधी परिप्रेक्ष्य.....	29
5.6.	WHOIS (हूईज़) संबंधी मुद्दे.....	29
5.7.	पंजीकरण प्रक्रिया संबंधी मुद्दे.....	30
5.8.	DNSSEC (डीएनएसएसईसी) संबंधी मुद्दे .....	31
6.	ग्रन्थ सूची का चयन .....	32
	परिशिष्ट सूची.....	35

## 0. प्रारंभिक तैयारियाँ

### 1. पृष्ठभूमि और सिंहावलोकन

इस रिपोर्ट का लक्ष्य देवनागरी लिपि के लिए विशेष संदर्भ के साथ टीएलडी से संबंधित मुद्दे हैं। हालाँकि एक व्यापक संदर्भ या दृष्टिकोण के भीतर इन मुद्दों की स्थिति का पता लगाने के लिए इस सामान्य पृष्ठभूमि और सिंहावलोकन की व्यवस्था की गयी है।

ICANN (आईकैन) द्वारा लैटिन के अलावा अन्य लिपियों को खोलने की नीति के आशीर्वाद से नयी भाषाओं और लिपियों का एक बाढ़-गेट खुल गया है और डोमेन नाम स्वाभाविक रूप से वास्तव में बहुभाषी हो जाएगा। इस नई नीति का लाभ उठाते हुए भारत ने भारत की 22 आधिकारिक भाषाओं के लिए भारतीय लिपियों और भाषाओं में आईडीएन प्रदान करने की चुनौती ग्रहण की है (इन भाषाओं की एक सूची परिशिष्ट II में सामान्य जानकारी के लिए उपलब्ध कराई गई है। देवनागरी का इस्तेमाल करते हुए आधिकारिक भाषाओं को प्रकाशित किया गया है)।

22 आधिकारिक भाषाओं में अंतर्राष्ट्रीयकृत डोमेन नामों को उपलब्ध कराने के लिए भारत के लिए एक नीति दस्तावेज को तैयार करने में लगभग 5 साल लगे हैं। यह सुनिश्चित करने के लिए सालों में 2005 में शुरू होने वाली इस नीति को विस्तृत किया गया है कि अंतिम उपयोगकर्ताओं के पास यथासंभव सुरक्षित वातावरण प्राप्त होगा जब वे अपनी मूल लिपि का इस्तेमाल करते हुए एक भारतीय भाषा में अपने नामों को रजिस्टर कराते हैं।

7 भारतीय भाषाओं (हिंदी, तमिल, तेलुगु, गुजराती, बंगला, उर्दू और पंजाबी) को पहले ही ICANN (आईकैन) और आईएनए के लिए प्रस्तावित किया गया है और इन भाषाओं में देश नाम "भारत" के लिए सीसीटीएलडी को पहले ही डीएनएस मूल क्षेत्र में अनुमोदित और प्रत्यायोजित कर दिया गया है।

चूंकि लिपियाँ एक समान संरचना नियमों का हिस्सा नहीं हैं और उनका अपना "संरचना व्याकरण" होता है; इसलिए यह चीजों के लिए अनुकूल था कि ICANN (आईकैन) ने महसूस किया कि छः लिपियों में "टेस्ट केसों" के सृजन से समस्याओं के साथ-साथ संबंधित मुद्दों की एक बेहतर धारणा प्राप्त होगी। अध्ययन के लिए चुनी गई लिपियाँ (लैटिन के अलावा): यूनानी, सिरिलिक, अरबी, देवनागरी, चीनी वास्तव में दुनिया की चार प्रमुख लेखन प्रणालियों को प्रतिबिंबित करती हैं - आबुजिदा (यूनानी और सिरिलिक), अब्जद् (अरबी), अक्षर या अल्फासिलेबरी (देवनागरी) और ध्वन्यात्मक-सिमेटिक (चीनी)।

इस संदर्भ के भीतर ईमेल के माध्यम से लगातार कई चर्चाओं की शुरुआत की गई। देवनागरी (परिशिष्ट I देखें) के लिए एक टीम का गठन किया गया जिसने देवनागरी लिपि (परिशिष्ट II देखें) का इस्तेमाल करते हुए न केवल हिंदी को बल्कि अन्य प्रमुख भाषाओं को भी गले लगाया। इन चर्चाओं का अंत जून में सिंगापुर में सभी समूहों की एक बैठक में और जुलाई में पुणे में देवनागरी समूह की एक अन्य बैठक में हुआ। इसके बाद सितम्बर में पुणे में टीम की एक एकदिवसीय बैठक हुई जिसमें ड्राफ्ट रिपोर्ट पर चर्चा की गई और उसे परिष्कृत किया गया।

लगातार कई चर्चाओं में विषय का अध्ययन करने वाली टीम के निर्माण से पहले और उसके बाद, धीमी गति से आम सहमति बनाने की प्रक्रिया विकसित हुई है और इस प्रक्रिया की दिशा में एक प्रमुख कदम एक प्रारंभिक मसौदा है जिसमें प्रत्येक लिपि से संबंधित समस्याओं खास तौर पर इसकी लेखन संरचना के संदर्भ और उससे उत्पन्न होने वाले भिन्न रूपों की धारणा के साथ मुद्दों को निरूपित किया गया है।

इन्हीं चिंताओं और मुद्दों को यह रिपोर्ट संबोधित करती है। यह रिपोर्ट देवनागरी का इस्तेमाल करते हुए अंतर्राष्ट्रीय भाषागत डोमेन नामों के सृजन के लिए विभिन्न मुद्दों के साथ-साथ लेखन प्रणाली की पृष्ठभूमि को भी स्थापित करने का प्रयास करती है। यह पालन किए जाने वाले प्रस्ताविक कार्य मसौदे पर पहुँचने के लिए पिछले कुछ महीनों में हुई चर्चाओं, दूरसंवादों, ईमेल आदान-प्रदान के साथ-साथ दस्तावेज निर्माण का परिणाम है।

## 2. संरचना

इस रिपोर्ट के निम्नलिखित अनुभाग हैं जिसके मूल लेआउट को पुणे में आयोजित मामले का अध्ययन करने वाली टीम की एक बैठक में अंतिम रूप दिया गया था:

भाग 1 बुनियादी तत्वों की स्थापना करता है जो हमारी राय में देवनागरी के लिए मुद्दा संबंधी रिपोर्ट की नींव की ईंट है।

भाग 2 देवनागरी, देवनागरी का इस्तेमाल करने वाली भाषाओं और भाषा की लेखन प्रणाली के एक संक्षिप्त स्केच के विकास का एक सिंहावलोकन प्रदान करके संदर्भ में चीजों को सेट करने का प्रयास करता है।

चूंकि इस दस्तावेज का लक्ष्य आईडीएन भिन्न रूपों के सभी पहलुओं के लिए प्रासंगिक मुद्दों पर प्रकाश डालना है: भाषाई, तकनीकी, सामाजिक, वित्तीय, और प्रशासनिक, इसलिए इन मुद्दों पर एक क्रमबद्धरीतिसे में प्रकाश डाला गया है<sup>1</sup>। भाग 3 विवादस्पद विषय के लिए प्रासंगिक प्रमुख मुद्दों की एक सूची है और यह सभी कोणों से समस्याओं की जाँच करता है।

कुछ निश्चित संख्या में कुछ ऐसे क्षेत्र हैं जिनका भिन्न रूप संबंधी मुद्दों पर कोई सीधा असर नहीं है लेकिन उनकी अंतर्भूत स्वभाव की वजह से ये परोक्ष रूप से सर्वेक्षण के तहत संदिग्धता से जुड़े हुए हैं। ये भाग 4 में सूचीबद्ध हैं।

चूंकि रजिस्टर आईडीएन में एक महत्वपूर्ण भूमिका निभाता है इसलिए इस क्षेत्र को एक विशेष अनुभाग, भाग 5 समर्पित किया गया है।

सहायक जानकारी प्रदान करने वाले परिशिष्टों की एक निश्चित संख्या रिपोर्ट को पूरा करती है।

---

<sup>1</sup> चूंकि इनमें से कुछ रोचक हैं लेकिन भिन्न रूपों के मुद्दे पर इनकी प्रत्यक्ष प्रासंगिकता नहीं है, इसलिए उन्हें परिशिष्ट V में सूचीबद्ध किया गया है

## 1. आधार तत्त्व

आईडीएन प्रयोजनों के लिए, ICANN (आईकैन) द्वारा प्रासंगिक प्रौद्योगिकियों की स्वीकृति के बारे में कुछ अनुमान करना है। इन अनुमानों का गठन बुनियादी निर्विवाद तत्वों से होता है जो देवनागरी भिन्न रूप मुद्दों पर इस रिपोर्ट का आधार है। ये तत्व निम्नानुसार हैं:

1.1. कोई अन्य प्रासंगिक वैश्विक कोडित वर्ण सेट उपलब्ध न होने की वजह से ही यूनिकोड स्वीकार्य है। यूनिकोड के स्वीकरण में इसके सामान्यीकरण मॉडल और सामान्यीकरण के संदर्भ के साथ यूनिकोड की स्थिरता नीति का स्वीकरण शामिल है।

1.2. यूनिकोड गुणों से संबंधित IDNA2008 (आईडीएनए2008) की व्याख्या और संस्करण विकास मॉडल सहित IDNA2008 (आईडीएनए2008) स्वीकार्य है।

1.3. सटीक लुकअप (ज्ञात आइटम या मद खोज) पर DNS (डीएनएस), भाषा विशिष्ट जानकारी और भाषा विशिष्ट या लिपि विशिष्ट लुकअप या मिलान क्रियाविधियों की अनुपस्थिति और उपनाम जिसमें संदर्भ नहीं होता है या जो DNS (डीएनएस) ट्री में कहीं से भी संकेत कर सकते हैं, सहित DNS (डीएनएस) स्वीकार्य है।

1.4. TLD (टीएलडी) नाम केवल "अक्षरों" तक ही सीमित हैं। एक TLD (टीएलडी) लेबल के भीतर अंक और हाइफन के साथ-साथ ZWJ U+200C, ZWNJ U+200D की अनुमति नहीं होगी।

1.5. DNS (डीएनएस) के विषय सामग्रियाँ स्मरक के बारे में हैं न कि "शब्दों" या विशेष भाषाओं में लंबे वाक्यों के बारे में। सच्चाई यह है कि एक विशेष भाषा में कुछ लिखा जा सकता है या यहाँ तक कि उसे उसके शब्दकोष में देखा जा सकता है, जिसका मतलब यह नहीं है कि DNS (डीएनएस) में दिखाई देने वाला स्ट्रिंग को प्राप्त करने के लिए एंटाइटेल्मेंट हो।

1.6. किसी भी डोमेन नाम ट्री में अलग, प्रशासकीय ढंग से भिन्न, पंजीकरण और रखरखाव और प्रशासनिक व्यवस्था के साथ अधीनस्थ क्षेत्र हो सकते हैं।

1.7. यह मुद्दा रिपोर्ट केवल आईडीएन भिन्न रूप टीएलडी (देवनागरी के लिए विशेष संदर्भ के साथ) तक सीमित है और अधीनस्थ क्षेत्रों के तहत पंजीकरण पर लागू नहीं हो सकता है हालाँकि रिपोर्ट में चर्चित मुद्दे उन अधीनस्थ क्षेत्रों के कामकाज में लाभकारी अंतर्दृष्टि प्रदान कर सकते हैं।

## 2. देवनागरी: एक सिंहावलोकन

देवनागरी का यह सिंहावलोकन देवनागरी का एक भाषाई परिचय है। इसका आरंभ देवनागरी के ऐतिहासिक विकास के साथ होता है और अनुभाग 2.1. में देवनागरी की संरचना का वर्णन है। अनुभाग 2. 2. अन्तर्निहित केंद्रक की धारणा को विकसित करता है: अक्षर और आगे चलकर भिन्न रूपों के लिए प्रासंगिक कुछ अक्षर संरचनाओं की ओर ध्यान आकर्षित करता है। सरल लिप्यंतरण का इस्तेमाल उदाहरणों के उच्चारण के लिए एक मार्गदर्शिका के रूप में किया गया है।

### 2.1. देवनागरी: एक ऐतिहासिक संदर्भ

देवनागरी भारत गणराज्य की आधिकारिक भाषाओं के रूप में मान्यता प्राप्त इंडो-आर्यन भाषाओं हिंदी, मराठी, मैथिलि, डोगरी, बोरो, संथाली, संस्कृत और नेपाली की मुख्य लिपि है। इस लिपि को अन्य देशों जैसे फिजी (हिंदी) और नेपाल (नेपाली) के साथ भी शेर किया गया है। यह संबंधित इंडो-आर्यन भाषाओं बघेली, भीली, भोजपुरी, हिमाचली बोलियों, मगही, नेवारी और राजस्थानी भाषा की भी एकमात्र लिपि है। यह प्राचीन भाषाओं संस्कृत और प्राकृत से काफी करीब से जुड़ी हुई है। यह कश्मीरी (हिंदू वक्ताओं द्वारा), सिंधी और संथाली के लिए एक वैकल्पिक लिपि है। अरुणाचल प्रदेश, बिहार और अंडमान एवं निकोबार द्वीपसमूह की आदिवासी भाषाओं को बोलने वालों द्वारा इसके इस्तेमाल में वृद्धि हो रही है।

यह सुविख्यात है कि देवनागरी जनक लिपि ब्राह्मी से विकसित हुई है जिसके आरंभिक ऐतिहासिक रूप में अशोकन ब्राह्मी के नाम से जाना जाता है जो लगभग चौथी सदी ई.पू. की बात है। ब्राह्मी का पता सर जेम्स प्रिंसेप ने 1837 में लगाया था। ब्राह्मी और उसके विकास के अध्ययन से पता चला है कि इसने भारत में कई लिपियों को जन्म दिया है जिसका उल्लेख ऊपर किया जा चुका है और कुछ भारत के बाहर जैसे श्रीलंका, म्यांमार, कम्पूचिया, थाईलैंड, लाओस, और तिब्बत देशों में इससे कई लिपियाँ उत्पन्न हुई हैं।

वर्तमानकालीन देवनागरी में ब्राह्मी के विकास में मध्यवर्ती रूप शामिल थे जो अन्य लिपियों के लिए आम थे जैसे उत्तर में गुप्ता और शारदा और दक्षिण में ग्रन्थ और कदंबा। कहा जा सकता है कि देवनागरी का विकास कुटिल लिपि से हुआ है जो गुप्ता लिपि का वंशज है जिसके परिणामस्वरूप यह ब्राह्मी का एक वंशज है। *कुटिल* शब्द का अर्थ "वक्र" है जिसका इस्तेमाल ब्राह्मी की सीधी रेखाओं की तुलना में लिपि की वक्र आकृतियों को अभिलक्षित करने के लिए एक वर्णनात्मक शब्द के रूप में किया जाता था। ब्राह्मी से देवनागरी के विकास पर नजर डालने से पता चलता है कि किस तरह इंडिक लिपियों में विविधता आई है: नक्षाकारों और लेखकों के हस्त कार्य जो विभिन्न प्रकार के स्ट्रोक या कारीगरी का इस्तेमाल करते थे जिसके फलस्वरूप विभिन्न क्षेत्रीय शैलियों का निर्माण हुआ (सिंह2006 का संदर्भ लें)।

ब्राह्मी का उल्लेख करने वाली सारी चीजें ऐतिहासिक कारणों के लिए हैं और किसी भी तरह से इस रिपोर्ट को ब्राह्मी से संबंधित प्रस्तुति नहीं मानना चाहिए क्योंकि इसका मुख्य केंद्र बिंदु केवल देवनागरी रही है।

### 2.2. लिखित देवनागरी की संरचना

देवनागरी एक अल्फासिलेबरी है और लेखन प्रणाली का केंद्र शब्दांश या अक्षर है। यही वह इकाई है जिसे लिपि के उपयोगकर्ताओं द्वारा सहज रूप से मान्यता प्राप्त है। अक्षर की धारणा को समझने के लिए अनुभाग 2.2. में लेखन प्रणाली का एक संक्षिप्त सिंहावलोकन दिया गया है और अनुभाग 2.3. में खुद अक्षर का गहराई से वर्णन किया जाएगा।

2.2 देवनागरी की लेखन प्रणाली को निम्नलिखित से निर्मित प्रणाली के रूप में अभिव्यक्त किया जा सकता है:

### 2.2.1. व्यंजन वर्ण

देवनागरी व्यंजन वर्णों में एक अन्तर्निहित तटस्थ मध्य स्वर वर्ण /ə/ होता है। पारंपरिक वर्गीकरण के अनुसार उन्हें उनके ध्वन्यात्मक गुणों के अनुसार वर्गीकृत किया गया है। 5 (वर्ग) समूह और एक गैर वर्ग समूह है। प्रत्येक वर्ग में पाँच व्यंजन वर्ण होते हैं जिन्हें उनके गुणों के अनुसार वर्गीकृत किया गया है। प्रथम चार व्यंजन वर्णों को आवाज और महाप्राण ध्वनि उच्चारण के आधार पर वर्गीकृत किया गया है और अंतिम व्यंजन वर्ण समरूपी अनुनासिक वर्ण है।

वर्ग	अनुच्चरित		उच्चरित		अनुनासिक
	-Asp	+Asp	-Asp	+Asp	
1 कंठ्य	क	ख	ग	घ	ङ
2 तालु	च	छ	ज	झ	ञ
3 मूर्धन्य	ट	ठ	ड	ढ	ण
4 दंत्य	त	थ	द	ध	न
5 द्वि-ओष्ठ	प	फ	ब	भ	म

### गैर - वर्ग

य	र	ल	ळ	व	श	ष	स	ह
---	---	---	---	---	---	---	---	---

### 2.2.2. अंतर्निहित स्वर हंता: हलंत<sup>2</sup>

सभी व्यंजन वर्णों में एक अन्तर्निहित स्वर चिह्न (तटस्थ मध्य स्वर) होता है। यह दर्शाने के लिए एक विशेष चिह्न की जरूरत पड़ती है कि यह अन्तर्निहित स्वर खंडित है। इसे हलंत (◌) के नाम से जाना जाता है। इस प्रकार हलंत दो व्यंजन वर्णों को जोड़ता है और संयुक्ताक्षरों का निर्माण करता है जो 2 से 3 व्यंजन वर्णों का संयोजन हो सकता है (उपरोक्त 1.2. का संदर्भ लें)।

### 2.2.3. स्वर वर्ण

सभी स्वर वर्णों के अलग-अलग प्रतीक होते हैं जिन्हें एक स्वर ध्वनि के शुरू में या उसके बाद स्वतंत्र रूप से उच्चारित किया जाता है। अन्तर्निहित ध्वनि के अलावा अन्य स्वर ध्वनि की ओर संकेत करने के लिए व्यंजन वर्ण के साथ एक स्वर आपरिवर्तक (मात्रा) जुड़ा रहता है। चूंकि व्यंजन वर्ण में पूर्व निर्मित स्वर चिह्न होता है इसलिए अ को छोड़कर सभी स्वर वर्णों की मात्रा समान होती है।

सहसंबंध को निम्नानुसार दर्शाया गया है:

अ	आ	इ	ई	उ	ऊ	ऋ	ए	ऐ	ओ	औ
	ा	ि	ी	ु	ू	ृ	े	ै	ो	ौ

अंग्रेजी से ली गई ध्वनियों को दिखाने के अलावा, देवनागरी का इस्तेमाल करने वाली कुछ भाषाएँ जैसे हिंदी, मराठी और कोंकणी भी 2 स्वर वर्णों और उनके समरूपी मात्राओं को निम्नानुसार स्वीकार करते हैं:

<sup>2</sup> यूनिकोड यूनिकोड 3.0 और उससे ऊपर का संदर्भ लें शब्द (विराम को वरीयता देता है। इस रिपोर्ट में दोनों शब्दों का इस्तेमाल उस वर्ण को दर्शाने के लिए किया गया है जो अन्तर्निहित स्वर को दबा देता है।



एँ ओँ

एँण्ड /and/ और /or/

मराठी में एँ की जगह अँ का इस्तेमाल किया जाता है

2.2.4. अनुस्वार /ँ/

अनुस्वार एक सम-जैविक अनुनासिक अक्षर का प्रतिनिधित्व करता है। यह विशेष वर्ग से संबंधित एक अनुनासिक व्यंजन+हलंत+व्यंजन के एक संयुक्त समूह को प्रतिस्थापित करता है। गैर-वर्ग व्यंजन से पहले अनुस्वार एक अनुनासिक ध्वनि का प्रतिनिधित्व करता है। आधुनिक हिंदी, मराठी और कोंकणी में समरूपी अर्द्ध-अनुनासिक के लिए अनुस्वार को वरीयता देते हैं:

सन्त बनाम संत /sənt/, चम्पा बनाम चंपा /tʃəmpa/

2.2.5. अनुनासिकता: चंद्रबिंदु ँँ

चंद्रबिंदु/अनुनासिक आँख /ākʰ/ की तरह पूर्ववर्ती स्वर की अनुनासिकता को दर्शाता है। वर्तमानकालीन हिंदी उपयोगकर्ता चंद्रबिंदु की जगह अनुस्वार का इस्तेमाल करना चाहते हैं।

2.2.6. नुक्ता ी

मुख्य रूप से हिंदी में इस्तेमाल किए जाने वाले नुक्ता चिह्न को फ़ारसी-अरबी से लिए गए शब्दों को प्रदर्शित करने के लिए व्यंजन वर्णों के नीचे रखा जाता है। यह दिखाने के लिए इसे क ख ग ज फ के साथ जोड़ा जा सकता है कि एक नुक्ता वाले इन व्यंजन वर्णों से बने शब्दों का उच्चारण फ़ारसी-अरबी शैली में किया जाना है।

जैसे फ़िरोज़ /firoz/

फड़फड़ाती ध्वनियों का संकेत देने के लिए हिंदी में ड ढ के नीचे भी इसका इस्तेमाल किया जाता है।

फड़फड़ाहट के अपवाद के साथ, आधुनिककालीन हिंदी उपयोगकर्ता शायद ही आजकल नुक्ता वर्णों का इस्तेमाल करते हैं।

2.2.7. विसर्ग ः और अवग्रह ऽ

विसर्ग का इस्तेमाल अक्सर संस्कृत में किया जाता है और यह काफी हद तक /h/ की ध्वनि का प्रतिनिधित्व करता है, जैसे दुःख /duːkh/

अवग्रह ऽ पूर्ववर्ती स्वर पर अतिरिक्त जोर डालता है और इसका इस्तेमाल संस्कृत ग्रंथों में किया जाता है। इसका इस्तेमाल शायद ही देवनागरी का इस्तेमाल करने वाली अन्य भाषाओं में किया जाता है।

### 2.3. मूलभूत इकाई: अक्षर

देवनागरी वर्णों के इस वर्गीकरण को बेकस-नौर रीतिवाद ISCA '91(आईएससीआईआई '91) पर आधारित "संयोजनात्मक व्याकरण" में कम किया जा सकता है जो अक्षर के सुनिर्माण को सुनिश्चित करता है। रीतिवाद लिपि की नोडल इकाइयों का वर्णन करता है: व्यंजन और स्वर और निर्धारित करता है कि इनमें से प्रत्येक नोडल इकाई के साथ किन तत्त्वों को संयुक्त किया जा सकता है। अक्षर के मौलिक गुणों की परिभाषा निम्नलिखित है:

अक्षर देवनागरी की ग्राफीमिक इकाई है। शब्दांश और अक्षर के बीच बस इतना अंतर है कि जहाँ एक ओर शब्दांश में एक या एक से अधिक पश्च स्वर व्यंजन होते हैं वहीं दूसरी ओर अक्षर में ऐसा नहीं होता है जिसे हम निम्नलिखित उदाहरणों से बखूबी समझ सकते हैं:

स्वनिम रूप	शब्दांश इकाइयाँ	अक्षर इकाइयाँ
cha:ruḷəta:	CV.CV.CV.CV	CV. CV. CV. CV
e:k	VC.	V. C
upka:r	VC. CVC	V. C. CV. C
indira	VC. CV. CV	VC. CV. CV
əst	VCC	V. CC
əkʃər	VC. CVC	V. CCV. C

सारणी 1: बोले जाने वाले रूपों के शब्दांश और अक्षर विभाजन

जैसा कि सारणी 1 में हम देख सकते हैं कि लिखित और बोले जाने वाले शब्दांश के बीच एक चिह्नित अंतर है खास तौर पर जहाँ तक शब्दांश सीमा जैसे /upka:r/ में व्यंजन समूहों के विभाजन का सवाल है।

पश्च स्वरित व्यंजन बनाम अक्षर के बारे में सामान्यीकरण का एकमात्र अपवाद अनुस्वार है जो अगले विराम के साथ सह-जैविक के रूप में सतही अन्तर्निहित अनुनासिक व्यंजन है। अनुस्वार को ग्रफीम का एक हिस्सा माना जाता है। अनुस्वार के वर्तनी विषयक और ध्वन्यात्मक रूप निम्नलिखित हैं:

बिंदी	[bindi:]	'point'
कंबल	[kəmbəl]	'blanket'
डंडा	[dəŋdɑ:]	'stick'
खंजर	[kʰəŋjər]	'knife'
कंघी	[kəŋgʰi:]	'comb'

सारणी 2: देवनागरी में अनुस्वार का प्रदर्शन

स्वर वर्ण शब्द के शुरु में और पश्च स्वरित रूप से अक्षर की एक स्वतंत्र इकाई है।

अ	आ	इ	ई	उ	ऊ	ए	ऐ	ओ	औ
ə	a	ɪ	i:	ʊ	u:	e:	æ:	o:	əu

सारणी 3: स्वतंत्र स्वर अक्षर

स्वर और व्यंजन वर्णों को अलग-अलग प्रकार की इकाई माना जाता है और इसलिए इन्हें ग्रफीम में प्रदर्शित किया जाता है जब व्यंजन के बाद स्वर का इस्तेमाल होता है। निम्नलिखित अक्षरों में केवल एक व्यंजन वर्ण और उसके बाद एक स्वर्ण वर्ण है:

क	का	कि	की	कु	कू	के	कै	को	कौ
ə	ka	kɪ	ki:	kU	ku:	ke:	kæ:	ko:	kəu

सारणी 4: देवनागरी CV अक्षर

जैसा कि सारणी 3 में प्रथम ग्राफीम में देखा जा सकता है, तटस्थ स्वर /ə/ को एक व्यंजन वर्ण में निहित माना जाता है। स्वर वर्ण का उच्चारण कुछ संदर्भों में, उदाहरण के लिए, पल/pal/ में प्रथम ग्राफीम में शुरु में और बीच में शब्द की तरह किया जाता है। निहित तटस्थ स्वर वर्ण का उच्चारण अंत में या कुछ संदर्भों में मध्य में शब्द की तरह नहीं किया जाता है।

दो व्यंजन वर्णों का समूह

निम्नलिखित विशेषताओं को दो व्यंजन समूहों से संबंधित माना जा सकता है:

1. आम तौर पर, प्रथम व्यंजन के अक्षर का आधा हिस्सा दूसरे व्यंजन के पूर्ण अक्षर के पहले आता है: जैसे, स्क<sk>, प्त<pt>, क्ल<kl> इत्यादि। वैकल्पिक रूप से, हलंत के नाम से जाने जाने वाले अप्रकाशित व्यंजनों के लिए विशेषक को निर्दिष्ट करने की प्रथा का इस्तेमाल प्रथम व्यंजन जैसे द्भ<dbh> उद्भव /udbhəv/ के लिए किया जाता है।
2. जैसा कि ऊपर उल्लिखित है, C+r के लिए, /r/ को एक सबस्क्रिप्ट या उप लिपि के रूप में निर्दिष्ट किया जाता है जो देखने में एक इंसक्रिप्ट की तरह लगता है: क्र<kr>, ख<khr>, फ्र<phr>
3. r+C समूहों के लिए, /r/ को ग्राफीम के ऊपर एक सुपरस्क्रिप्ट के रूप में निर्दिष्ट किया जाता है जैसे, र्म<rm>, र्ज<rdʒ>
4. निम्नलिखित दो व्यंजन समूहों के मामले में, एक नए युक्ताक्षर समूह का निर्माण किया जाता है। ये हैं: त्र<tr>, क्ष<kʃ>, ज्ञ<jn>, श्र<ʃr>, क्त<kt>

तीन व्यंजन वाले समूह:

इनकी विशेषताएँ निम्नलिखित हैं:

1. आम तौर पर, प्रथम दो व्यंजनों को उनके अक्षरों के आधे हिस्से के लिए निर्दिष्ट किया जाता है और तीसरे को पूरी तरह से निर्दिष्ट किया जाता है जैसे स्प्ल<spl>। इस सम्मलेन का इस्तेमाल आम तौर पर उधार स्वरूप लिए गए शब्दों के लिए किया जाता है।
2. C+C+r समूहों के लिए, और r+C+C समूहों के लिए, जो अत्यधिक प्रतिबंधित या सीमित हैं, दो व्यंजन समूहों के लिए समागम लागू होता है, जैसे स्त्र<str>

### 3. विषय

मुद्दों पर चर्चा और समाधान की पहचान करने के लिए लिपिगत दृष्टिकोण के आधार पर निम्नलिखित व्यावहारिक विचारों पर ध्यान रखा जाना आवश्यक है :

आधारभूत प्रतिनिधिमंडल: आधारभूत क्षेत्र की दो व्यापक श्रेणियां हैं : ccTLDs (सीसीटीएलडीज़) तथा अन्य कभी-कभी gTLDs (जीटीएलडीज़) कहलाता है। इन विभिन्न प्रकार के क्षेत्रों के विभिन्न नीति शासी प्रतिनिधिमंडल हो सकते हैं; लेकिन आधारभूत क्षेत्र से किसी भी TLD (टीएलडी) का नीति शासी प्रतिनिधिमंडल हमेशा एक ही होनी चाहिए। आधारभूत क्षेत्र में लेबल अक्षरों तक ही सीमित हो जाता है : हाइफ़न, और अन्य गैर अक्षर ZWJ (ज़ीरो विड्थ ज्वाइनर) तथा ZWNJ (ज़ीरो विड्थ नॉन-ज्वाइनर ) ग्राह्य नहीं हैं।

स्पूफ़िंग, फार्मिंग और फ़िशिंग में कमी लाने के लिए भाषा टेबिल की धारणाओं का परिचय, उत्तम गठन पर आधारित प्रतिबंधित नियम तथा संस्करण जरूरी है। इस प्रकार देवनागरी पर आधारित भाषाएं, जो अक्षरों से संचालित हैं, के उत्तम आकार को संभालने के लिए एक रूपवाद विकसित करने की आवश्यकता है।

संभावित क्षेत्र जहां इस तरह के कारक लागू होते हैं, वे हैं:

3.1 भाषा बनाम लिपि

3.2 संस्करण

3.3 डोमेन के प्रदर्शन के संबंध में सॉफ्टवेयर व्यवहार से संबंधित मुद्दे:

3.4 सम्पूर्ण लिपि भ्रामक

3.5 देवनागरी में 02BC के लिए केस

इस प्रकार इन्हें विकसित किया जाएगा।

#### 3.1. भाषा बनाम लिपि संबंधी मुद्दे

ccTLD (सीसीटीएलडी) के भीतर .भारत के लिए भाषा बनाम लिपि के मुद्दों के विरोधाभास (कुछ मुद्दों को रजिस्ट्री स्तर पर काम में लाने के लिए) संभाले जा सकते हैं क्योंकि इंटरनेट पर प्रत्येक व्यक्ति द्वारा जरूरी तौर पर आधारभूत क्षेत्र को देखा जाता है, आधारभूत क्षेत्र में उपयोगकर्ता के कुछ टूटने के आधार पर एक उपयोगकर्ता की भाषा का उचित अनुमान लगाना असंभव है इसलिए, लेबल में अक्षरों की इस लिपि से उसके संस्करणों पर निर्णय लेने का प्रयोजन प्रासंगिक हो जाएगा।

यह आशा है कि लिपियों का परिचय जैसे कि देवनागरी का, विशेष रूप से भाषाओं की एक बड़ी संख्या का प्रतिनिधित्व करने के लिए इस्तेमाल किया जाता है, जिससे भाषाओं से निपटने के लिए एक उपयुक्त तंत्र का विकास हो जाएगा।

#### 3.2. देवनागरी लिपि में भिन्न रूप

केवल दो या दो से अधिक अक्षर या वर्ण संयोजन को संस्करण रूप में समझा जाएगा अगर :

a. वे प्रोटोकॉल मान्य हैं और

b. यूनिकोड सामान्य नियमों IDNA 2008 (आईडीएनए 2008) में सूचीबद्ध नहीं हैं, जिसका मतलब है कि वे अक्षर या वर्ण संयोजन जो यूनिकोड सामान्य रूप NFC (एनएफसी) और/ या NFKC (एनएफकेसी) के तहत स्थिर नहीं हैं।

देवनागरी में संस्करण दो प्रकार के होते हैं : एकल वर्ण जो लेबल के भीतर एक जैसे दिखते हैं तथा लिगचुरल आकृतियाँ जो दृष्टिगत रूप में भ्रामक हैं और एक दूसरे के परिपेक्ष्य में गलत हो सकते हैं।

### 3.2.1. भ्रामक रूप से एक समान एकल वर्ण

ये एकल अक्षर हैं जो भ्रामक रूप से समान आकार के हैं और एक दूसरे के परिपेक्ष्य में भ्रामक हो जाते हैं। संस्करणों की इन श्रेणियों को .भारत ccTLD (सीसीडीएलडी) नीति के अंतर्गत नहीं माना जाता है चूंकि वहाँ एक संभावना यह हो जाती है कि यह दृष्टिकोण बहुत ज्यादा प्रतिबंधात्मक साबित हो जायेगा।

उदाहरणार्थ

घ U+0918	ध U+0927
भ U+092D	म U+092E

सारणी 5

इस तालिका में केवल एक नमूना सूची शामिल है। अधिक विस्तृत सूची परिशिष्ट IV में उपलब्ध है।

### 3.2.2. भ्रामक रूप से एक समान संयुक्त वर्ण

चूंकि देवनागरी ही संयुक्तता को सहयोग देती है इसलिए संयुक्तता जिसके एक समान दिखने तथा ब्राउज़र के छोटे URL (यूआरएल) बार में आसानी से भ्रमित हो जाने की संभावना प्रबल हो जाती है।

उदाहरणार्थ

द्र U+0926 U+094D U+0917	द्र U+0926 U+094D U+0930	द्न U+0926 U+094D U+0928
द्ध U+0926 U+094D U+0927	द्ध U+0926 U+094D U+0918	
ष्ट U+0937 U+094D U+091F	ष्ट U+0937 U+094D U+0920	
द्ध U+0926 U+094D U+ 0935	द्ध U+0926 U+094D U+092C	

सारणी 6

इस तालिका में केवल एक नमूना सूची शामिल है। अधिक विस्तृत सूची परिशिष्ट IV में उपलब्ध है।

### 3.2.3. वर्णों को संयुक्त करने की वजह से उत्पन्न भिन्न रूप

देवनागरी लिपि के लिए यह उप-प्रकार प्रासंगिक नहीं है चूंकि चीजों की वर्तमान स्थिति में ऐसा कोई मामला प्रकट नहीं होता। हालांकि, अगर कोड ब्लॉक U+0900 का भविष्य का संस्करण होगा, और अगर उन संस्करणों को देवनागरी के लिए सामान्य नियमों द्वारा संभाला नहीं गया तो, इन्हें प्रस्तुत करने की आवश्यकता होगी।

### 3.3. डोमेनों के प्रदर्शन के संबंध में सॉफ्टवेयर व्यवहार से संबंधित मुद्दे:

DNS (डीएनएस) विशेष रूप से न केवल वेब के बारे में है बल्कि अन्य क्षेत्रों जैसे कि ईमेल उपयोगकर्ता एजेंट, कैलेंडरिंग प्रोग्राम आदि को भी प्रभावित करता है। हालांकि अध्ययन के अनुसार, ब्राउज़रों तथा ईमेल ग्राहकों (देवनागरी लिपि से संबंधित विशेष संदर्भ के साथ) से संबंधित मुद्दों को लिया जाएगा। यहां जिन मुद्दों पर प्रकाश डाला जा रहा है शायद उनका डोमेन के प्रदर्शन के संबंध में अन्य सॉफ्टवेयर व्यवहार के लिए लागू किया जा सकता है। इस बात पर ध्यानाकर्षण करना आवश्यक है कि ब्राउज़रों के मामले में, यह पुनर्निर्देशन प्रोटोकॉल की एक विशेषता हो सकती है जिसका इस्तेमाल किया जा रहा है (यानी http अनुप्रेषित)। अन्य सॉफ्टवेयर (जैसे ईमेल ग्राहकों, कैलेंडर, और कुछ मोबाइल अनुप्रयोगों के रूप में) के साथ ऐसा नहीं है, वहां जहां कोई अंतर्निहित प्रोटोकॉल अनुप्रेषित तंत्र नहीं है।

#### 3.3.1. ब्राउज़र संबंधी मुद्दे

ब्राउज़र के URL (यूआरएल) बार में डोमेन नाम का प्रतिनिधित्व करने के लिए ब्राउज़र अंतर्निहित OS (ओएस) प्रतिपादन इंजन पर भरोसा करते हैं। इस प्रकार OS (ओएस) के प्रतिपादन इंजन से जुड़े मुद्दें ब्राउज़र में निहित होते हैं। जो फ्रॉन्ट्स यूआरएल बार में लागू होते हैं वो ब्राउज़र्स द्वारा चुना जाता है जो अंतर्निहित OS (ओएस) द्वारा डोमेन नाम की लिपि के लिए डिफ़ॉल्ट फ्रॉन्ट के रूप में उपलब्ध कराया जाता है।

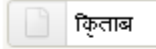
ब्राउज़रों की इन विशेषताओं से संबंधित मुद्दों की दो व्यापक श्रेणियां हैं :

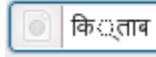
3.3.1.1. प्रतिपादन इंजन से संबंधित मुद्दे

3.3.1.2. फ्रॉन्ट संबंधित मुद्दे

3.3.1.1. प्रतिपादन इंजन से संबंधित मुद्दे

जब भी यूनिकोड सक्षम अनुप्रयोग में देवनागरी लिपि के कुछ पाठ प्रस्तुत किए जाते हैं, प्रतिपादन इंजन उस पाठ को शब्दांशों के रूप में विभक्त करता है। इन अक्षर गठन नियमों को मानकीकृत नहीं किया गया है, और न ही यूनिकोड ने इससे संबंधित किसी भी विशिष्ट नियमों को प्रस्तुत किया है। इस प्रकार विभिन्न प्रतिपादन इंजन के व्यवहार अलग हैं तथा आरोपित पाठ की भाषा/ लिपि की समझ पर निर्भर करता है जो शायद ही कभी पूर्ण रूप से सही होता है। इसका उदाहरण सैद्धांतिक रूप में नीचे दिया जाता है जो बताता है कि कैसे अलग वातावरण के तहत एक ही ब्राउज़र एक या अधिक अशुद्धपूर्ण गठित अक्षर प्रदर्शित नहीं करता है :

 → Windows + Firefox - Malformed word rendered as well-formed

 → Fedora + Firefox - Malformed word rendered as mal-formed

उपरोक्त सैद्धांतिक उदाहरण एक वैध लेबल के साथ दिया गया है : किताब (पुस्तक) से पता चलता है कि ऑपरेटिंग सिस्टम का प्रतिपादन इंजन किस प्रकार अशुद्धपूर्ण गठित अक्षर के होने के बावजूद उसे शुद्धपूर्ण गठन रूप में प्रस्तुत होने की अनुमति देता है। परीक्षण डोमेन में प्रथम शब्दांश : कि के बाद हलंत/ विराम के साथ किताब/ kitab/ लिखा गया। विंडोज के अंतर्गत फायरफोक्स से पता चलता है कि इस अशुद्धपूर्ण गठित अक्षर अशुद्धपूर्ण गठन के रूप में प्रस्तुत नहीं हुआ है। फेडोरा के अंतर्गत उसी समान स्थिति में शब्दांश के अशुद्धपूर्ण गठन को उजागर करता है। यह स्पष्ट करना जरूरी है कि यहाँ उठाये गये मुद्दे में संयुक्त आकृतियों के प्रस्तुतिकरण की समस्या को संबोधित किया गया है, जो या तो इस तरह (ओएस के देसी फ्रॉन्ट के कारण) से प्रस्तुत किया गया है कि संस्करण में दृष्टिगत रूप से प्रासंगिक न हो या फिर साधारणतया अशुद्धपूर्ण रूप में ही प्रस्तुत किया गया हो। यह आशा की जाती है कि सॉफ्टवेयर डेवलपर्स इस तरह के मुद्दों के बारे में अधिक जागरूक होंगे, इससे ये सही हो जायेगा। हालांकि जो मुद्दा उठाया गया है उस पर विचार किया जाना चाहिए। लेकिन इस आश्वासन के बारे में एक प्रकार से कहा जा सकता है कि सॉफ्टवेयर सही शोधन लाते हैं ।

### 3.3.1.2. फ्रॉन्ट संबंधित मुद्दे

ब्राउज़र्स में डोमेन नाम के प्रतिपादन के मामले में, ब्राउज़र के एड्रेस बार के अंतर्गत डोमेन नाम में जो फ्रॉन्ट लागू किया जाता है वह प्रमुख भूमिका निभाता है। प्रत्येक ऑपरेटिंग सिस्टम का एक विशिष्ट फ्रॉन्ट है जो ओएस समर्थित हर लिपि / भाषा के लिए डिफ़ॉल्ट फ्रॉन्ट के रूप में कार्य करता है। ब्राउज़र एड्रेस बार में डोमेन नाम प्रदर्शित करने के लिए ओएस द्वारा उपलब्ध करायी गयी डिफ़ॉल्ट फ्रॉन्ट का प्रयोग करता है।

इसी प्रकार प्रतिपादन इंजन के लिए, फ्रॉन्ट कार्यान्वयन भी विक्रेता-विक्रेता में भिन्न होता है। और इसीलिए एक ही तरह के डोमेन नाम फ्रॉन्ट गुण, फ्रॉन्ट द्वारा अपनाये गये ओर्थोग्राफी, इशारा, वजन, कर्निंग आदि पर निर्भर करता है, अलग रूपों में भी देखा जा सकता है। जैसा कि नीचे दिए गए उदाहरण के रूप में देखा जा सकता है जहां हिंदी और बंगाली के फ्रॉन्ट समान होने के बावजूद अलग दिख रहे हैं : हिंदी बंगाली की तुलना में अधिक पठनीय है।



Word "kitab" written in Devanagari script rendered legibly



Word "kitab" written in Bengali script rendered illegibly

देवनागरी लिपि में लिखित शब्द किताब स्पष्ट द्रष्टव्य है

बंगाली लिपि में लिखित शब्द किताब स्पष्ट द्रष्टव्य नहीं है

चूंकि वहाँ कोई केंद्रीय सत्ता नहीं है जो आम सहमति से कार्यान्वयन को सुनिश्चित कर सके। यह आशा है कि उपयोगकर्ता सहयोगी सॉफ्टवेयर अनुप्रयोगों जो दावा करते हैं कि वे देवनागरी समर्थित क्षमता सम्पन्न हैं, के पास एक सूचीबद्ध सेट होना चाहिए, जो उपयोगकर्ता के अनुभव में सुधार लाने और उसे युक्तिसंगत बनाने की ओर एक लंबा रास्ता तय कर सके।

### 3.3.2. ईमेल पत्तों का रिज़ोल्यूशन

यहाँ उठाए गए प्रश्न .भारत के लिए प्रासंगिक हैं लेकिन यह अन्य रजिस्ट्री के कुछ मामलों में भी लागू हो सकता है।

संस्करण (उपरोक्त 3.2 का संदर्भ लें) के लिए उठायी गयी उचित समस्याओं के समाधान के लिए एक संकल्प के रूप में चिह्नित किया है। उदाहरण .भारत के लिए लिया जाएगा लेकिन अन्य किसी भी TLD (टीएलडी) लेबल के लिए समान रूप से लागू हो सकते हैं।

इस तरह के ईमेल जैसे कि

वित्त-मंत्रालय.भारत : वित्त मंत्रालय

एड्रेस के मालिक भी क्या संस्करण के वारिस होंगे: वित्त-मंत्रालय.भारत

यह देखा जाता है कि त U+0924 U+094D U+0924 संस्करण के रूप में त U+0924 उत्पन्न होता है

वित्त U+0935 U+093F U+0924 U+094D U+0924
वित्त U+0935 U+093F U+0924

कुछ सवाल उठते हैं वे हैं:

- जब दोनों ईमेल मान्य हैं, क्या वहाँ एक अलियासिंग तंत्र होगा ?

यह मुद्दा भी ईमेल को रिज़ोल्व करते मेल सर्वर के साथ बारीकी से बंधा हुआ है।

### 3.4. देवनागरी लिपि में 02BC का मामला:

U+02BC Modifier Letter Apostrophe (संशोधक अक्षर एपोस्ट्रोफी) " ' ", जो एक स्वर चिह्न या लंबे चिह्न के रूप में कार्य करते हैं, इनका प्रयोग अक्सर बोरो, डोगरी, मैथिली आदि भाषाओं में, जो देवनागरी लिपि आधारित हैं तथा बंगला जोकि बंगाली लिपि आधारित है, किया जाता है। डोगरी से एक उदाहरण के तौर पर, जहां 02BC का प्रयोग संकोचन मार्कर के रूप में किया जाता है, इस मुद्दे को स्पष्ट करेंगे:

करा'रदा । (अर्थ : सम्पन्न किया)

U+02BC संशोधक अक्षर " ' " ब्लॉक U+02B0-U+02FF से आता है, जबकि सभी अक्षर जो देवनागरी लिपि से संबंधित हैं ब्लॉक U+0900-U+097F से आये हैं। यदि नीतिगत निर्णय के रूप में gTLDs (जीटीएलडीज़) में लिपि मिश्रण की अनुमति नहीं है, इन अक्षरों को तब भी एक अपवाद के रूप में अनुमति दी जानी चाहिए क्योंकि इस अक्षर के बिना भाषा का प्रतिनिधित्व सम्पूर्ण नहीं किया जा सकेगा<sup>3</sup>। यह नोट किया जाए कि भाषाओं (बोरो, डोगरी, मैथिली) के लिए उपलब्ध डिवाइस जो इस अक्षरों का उपयोग कर रहे हैं, इन अक्षरों को प्रवेश करने का माध्यम प्रदान करते हैं जो कि इन भाषाओं में अपेक्षाकृत उच्च आवृत्ति में इस्तेमाल किया जाता है। आशा है कि 02BC एक स्वीकार्य देवनागरी कोड सेट बनाएगा, छठी परिशिष्ट देवनागरी कोड प्वाइंट के लिए सिफारिश की एक सूची प्रदान करता है।

<sup>3</sup> इस मामले पर वाद-विवाद होना आवश्यक है चूंकि विशेषज्ञों में मतभेद हैं।



### 3.5. संपूर्ण लिपि भ्रांतियोंसे युक्त

यह एक परिदृश्य में संदर्भित है, जहाँ दो स्ट्रिंग्स, S1 और S2 मौजूद होते हैं। S1 में हर अक्षर की लिपिगत संपत्ति P1 है। S2 में हर अक्षर की लिपिगत संपत्ति P2 है। इसलिए प्रत्येक स्ट्रिंग एक एकल लिपि में उपस्थित है। लेकिन दो स्ट्रिंग्स कुछ भाषाओं के किसी भी सक्षम वक्ता के लिए भ्रामक है<sup>4</sup>। यूनिकोड का उदाहरण यह है कि उस यूनिकोड सेट से लैटिन अक्षरों वाले स्ट्रिंग तथा जिसमें पूरी लिपि सिरिलिक में भ्रामक (लोअरकेस केवल) हो सकता है।

देवनागरी के मामले में, एक मामला देखा जा सकता है जहाँ देवनागरी में एक पूरी स्ट्रिंग मौजूद है और किसी अन्य लिपि में एक पूरी स्ट्रिंग के अनुरूप हैं, जैसे कि यह स्पूफिंग को प्रेरित कर सकता है। इस प्रकार एक ब्राउज़र के URL के भीतर निम्न दो उदाहरण एक जैसे देखे जा सकते हैं :

मोरी (देवनागरी लिपि) भोरी (गुजराती लिपि) [morī]

दो लिपियों में बड़ी संख्या में एक समान अक्षर मौजूद हैं, केवल एक अंतर "शीरोरेखा" का अभाव है जो अक्षरों के ऊपर खींची जाती है जो देवनागरी में होती है लेकिन जिसका गुजराती में अभाव है। हालांकि ब्राउज़र के छोटे प्वाइंट साइज में, दोनों समान दिखते हैं, जिससे स्पूफिंग को बढ़ावा मिलता है। हालांकि इन लेबलों को यूनिकोड भ्रामकीय परीक्षण (देखें <http://www.unicode.org/reports/tr39/>, 4.2 अनुभाग) को प्रस्तुत नहीं किया गया है, प्रत्येक मामले में 4 कोड प्वाइंट के मैनुअल जांच से पता चलता है कि यह यूनिकोड भ्रामक सूची : <http://unicode.org/Public/security/revision-02/confusables.txt> में सूचीबद्ध नहीं हैं।

हालांकि यूनिकोड की भ्रामक सूची अत्यंत व्यापक है, जिसमें "गुजराती अक्षर" के लिए केवल 43 भ्रामक शब्दांश मौजूद हैं, जिनमें से कोई भी देवनागरी सेट से मैच नहीं करता है। सभी भ्रामक शब्दांश सूचीबद्ध होने की क्रिया को मान्य करने के लिए सूची की पूरी तरह से जाँच की आवश्यकता है।

इंडिक लिपियों में अक्षरों की एकरूपता को देखते हुए और साथ ही यह देखते हुए कि पूरी लिपि में ऐसे सभी अक्षरों की पहचान करना बहुत मुश्किल है, इस अभ्यास को अगर अनुमति मिली तो स्पूफिंग को बढ़ावा मिलेगा और इसकी अनुमति नहीं दी जानी चाहिए।

इसके अलावा वास्तविक मामलों में जहां दो लेबल दो लिपियों में अलग हैं, यह रिजोल्यूशन के मुद्दों को विवादित बना सकता है। इस प्रकार यदि मोरी दोनों ही लिपियों में एक वैध लेबल है और सम्पूर्ण लिपि भ्रामक होने के कारण इसमें से एक (जैसे की गुजराती) को अनुमति प्राप्त नहीं है, इसके कारण रिजोल्यूशन के समाधान के मामले में विवाद उत्पन्न हो सकता है।

<sup>4</sup> देखिये यूनिकोड की रिपोर्ट , <http://www.unicode.org/reports/tr39/> , विशेष रूप से 4.1 धारा।

## 4. विषयेतर विचार

ये जैसे विचार है, जिनका भिन्न मुद्दों पर कोई सीधा असर नहीं है, लेकिन उनके आंतरिक स्वभाव की वजह से, ये सर्वेक्षण के तहत परोक्ष रूप से समस्याग्रस्त होते हैं और इसलिए उन्हें इस रूप में नाम दिया जाता है।

### 4.1. संकर लिपि वर्ण मिश्रण

टीएलडी में, एक वर्गीकरण के भीतर लिपियों के मिश्रण की अनुमति की कोई संभावना नहीं होती।

यदि लिपि मिश्रण की अनुमति दी जाती है, तो हमारी राय है कि इसके परिणामस्वरूप बड़े पैमाने पर चकमा देने (स्पूफिंग), फ़िशिंग (किसी बड़े संस्थान के किसी व्यक्ति से व्यक्तिगत जानकारी हासिल करने की नीयत से ईमेल भेजने की प्रक्रिया) और धोखाधड़ी हो सकती है, क्योंकि लिपियों के भीतर कई विभक्तियां एक दूसरे का भ्रम पैदा करती हैं। यूनिकोड स्तर पर भी इसे माना गया है<sup>5</sup>। भारत के लिए नीति ccTLD (सीसीटीएलडी) कोड ब्लॉक मिश्रण की अनुमति नहीं देती।

हालांकि प्रति-भाषी दृश्य समानताओं की एक सूची के नीचे प्रदान की गई है। इस पर ध्यान दिया जाना चाहिए कि ऐसी समानताएं एकल अक्षर तक सीमित हैं और संयुक्ताक्षरों के लिए नहीं। इन विभिन्न कोड ब्लॉकों से वर्णों के मिश्रण से स्पूफिंग संभव हो सकती है। सूची में किसी भी तरह से व्यापक नहीं है, लेकिन निहित खतरे का संकेत करने के लिए पर्याप्त हैं।

देवनागरी लिपि	सजात लिपि	सजात लिपि में कूट-संकेत बिंदु
<b>स्वर</b>		
उ U+0909	बांग्ला	ঔ U+0993
उ U+0909	गुरुमुखी	ੜ U+0A24
ऋ U+090B	गुजराती	ઋ U+0AE0
<b>व्यंजन</b>		
क U+0915	बांग्ला	ক U+0995
ग U+0917	गुजराती	ગ U+0A97

<sup>5</sup> <http://www.unicode.org/reports/tr39/>, विशेष रूप से अनुभाग 5.

ग U+0917	गुरुमुखी	ग U+0A17
घ U+0918	गुरुमुखी	घ U+0A2C
घ U+0918	गुजराती	ध U+0A98
ङ U+0919	गुजराती	ड U+0A99
छ U+091B	गुजराती	छ U+0A9B
ज U+091E	गुजराती	झ U+0A9E
ट U+091F	गुरुमुखी	ट U+0A1F
ठ U+0920	गुजराती	ढ U+0AA0
ठ U+0920	गुरुमुखी	ठ U+0A20
ड U+0921	गुजराती	ड U+0AA1
ढ U+0922	गुरुमुखी	ढ U+0A2B
त U+0924	गुजराती	त U+0AA4
ध U+0927	गुजराती	ध U+0AA7

न U+0928	गुजराती	न U+0AA8
न U+0928	बांग्ला	न U+09A8
न U+0928	बांग्ला	न U+09A3
प U+092A	गुजराती	प U+0AAA
प U+092A	गुरुमुखी	प U+0A17
प U+092A	गुरुमुखी	प U+0A2A
प U+092A	गुरुमुखी	प U+0A6B
म U+092E	गुरुमुखी	म U+0A38
म U+092E	गुजराती	म U+0AAE
य U+092F	गुजराती	य U+0A9A
र U+0930	गुजराती	र U+0AAE
र U+0930	गुरुमुखी	व U+0A15
ल U+0932	बांग्ला	न U+09B2
व	गुजराती	व

U+0935		U+0AB5
श U+0936	गुजराती	शल U+0AB6
ष U+0937	गुजराती	षल U+0AB7
स U+0938	गुजराती	सल U+0AB8
ह U+0939	गुजराती	हल U+0AB9
<b>नुक्ता अक्षर</b>		
ग U+095A or U+0917 U+094D	गुरुमुखी	गल U+0A5A
ढ U+095D Or U+ 0922 U+094D	गुरुमुखी	ढल U+0A5E

सारिणी 7

## 4.2. वर्तनी के माध्यम से उत्पन्न समस्वन

देवनागरी आधारित भाषाओं में, समस्वानिक पाठान्तर, जो दो समस्वन (अंग्रेजी के *color-colour* जैसा वर्तनी पाठांतर) जैसे हिंदी और हिंदी/ Hindi हो सकते हैं<sup>6</sup>, लेकिन ऐसे पाठांतरों के नियम गलत रूप से परिभाषित हैं और इससे विकृति के अवसरों को बढ़ावा मिल सकता है। भारत के लिए ccTLD (सीसीटीएलडी) नीति के दायरे में ऐसे पाठांतरों पर विचार नहीं किया जाता। भारत जैसे वेरिफ़ाई पर विचार नहीं किया गया है।

## 4.3. शून्य चौड़ाई वाले योजक यानी ज़ीरो विड्थ ज्वाइनर ZWJ और ज़ीरो विड्थ नॉन-ज्वाइनर ZWNJ

ZWJ (ज़ीरो विड्थ ज्वाइनर) (U+0200D) और ज़ीरो विड्थ नॉन-ज्वाइनर कोड बिंदु है, जो यूनिकोड मानक द्वारा उस कड़ी को जोड़ने के लिए प्रदान किये गये हैं, जहां लिपि को संयुक्त और असंयुक्त वर्णों के बीच का विकल्प होता है। इन नियंत्रण कोड के उपयोग के बिना, स्ट्रिंग या कड़ी को एक इच्छित वैकल्पिक रूप में बदला जा सकता है। तकनीकी तौर पर टीएलडी इनमें से किसी को अनुमति नहीं देता और यह एक तरह का गैर मुद्दा लग सकता है। हालांकि दो मामलों पर विचार की आवश्यकता है।

### 4.3.1. Eye-lash ra का मामला<sup>7</sup>

=	=
U+0930 U+094D U+200D	U+0931 U+094D

यूनिकोड 21 0 Eye-lash ra के प्रतिनिधित्व के लिए RA+VIRAMA+ ZWJ (ज़ीरो विड्थ ज्वाइनर) का उपयोग निर्धारित करता है। यह सेक्शन 9 की तत्कालीन धारा R5 (जो अब नियम R5a है) के तहत प्रगृहीत है। यूनिकोड 31 0/4. 0 ISCI की पसंद प्रतिबिंबित करता है, जो कि अब नियम R5 है : “*In conformance with the ISCI standard, the half-consonant form rrah is represented as eyelash-ra. This form of ra is commonly used in writing Marathi...*” (“ISCI मानक के अनुरूप आधे व्यंजन रूप rrah का प्रतिनिधित्व Eye-lash ra के रूप में होता है। ra का यह फार्म आमतौर पर मराठी लेखन में प्रयोग किया जाता है (यूनिकोड 3.0)”)।

तो /दर्या/ "घाटियों शब्द को यूनिकोड सिद्धांतों में U+0926 U+0930 U+094D U+200D U+092F U+093E (दर्या) साथ ही साथ U+0926 U+0931 U+094D U+092F U+093E (दर्या) के रूप में लिखा जा सकता है।

नेपाल में नेपाली उपयोगकर्ता RRA U+0931 लिपिचिह्न को मान्यता नहीं देते, क्योंकि र, उनकी फ्रॉन्ट सूची में नहीं दिखती। Eye-lash ra पैदा करने के लिए संयोजन के उपयोग का अभ्यास दिखता है : Ra (U+0930)

<sup>6</sup> उपरोक्त भाग 2 का संदर्भ लें

<sup>7</sup> आईलैश ra का इस्तेमाल कोंकणी, नेपाली और मराठी में होता है। कुछ भाषाविदों द्वारा र (रेफ) से अलग माना जाता है। हालांकि पिछले वाले को एक फड़फड़ाहट के रूप में व्यवहार किया जाता है लेकिन बाद वाला एक प्रतल ध्वनि है (उपरोक्त कल्याण काले और अंजलि सोमन 1986 का संदर्भ लें)। मराठी में नाममात्र के जोड़ों में इस तरह मामले नजर आते हैं जैसे आचार्यास “शिक्षक को” बनाम आचार्यास “बावर्ची को” या दर्या /darya/ “सागर” बनाम दर्या /darya/ “घाटी”। इसी तरह के मामले कोंकणी और नेपाली में भी नजर आ सकते हैं। चूंकि इसे एक परमाण्विक वर्ण के रूप में नहीं देखा जा सकता इसलिए एक चित्र का इस्तेमाल किया गया है।

+ VIRAMA (U+ 094D + ZWJ (ज़ीरो विड्थ ज्वाइनर)(U+200D)। लेकिन जब से ZWJ (ज़ीरो विड्थ ज्वाइनर)को TDL (टीएलडी) लेबल में अनुमति नहीं है, नेपाली उपयोगकर्ताओं के लिए टीएलडी लेबल के भीतर eye-lash ra को दिखाने के लिए वैकल्पिक तरीकों को खोजने की आवश्यकता होगी।

#### 4.3.2. संज्ञा मानदंड का केस:

नेपाली जैसी भाषाओं में ZWNJ (ज़ीरो विड्थ नॉन-ज्वाइनर ) का उपयोग के कुछ संज्ञा मानदंडों की सही पीढ़ी की अनुमति देता है।

श्रीमान् + को = श्रीमान्को

श्रीमान् + ZWNJ (ज़ीरो विड्थ नॉन-ज्वाइनर ) + को = श्रीमान्को

श्रीमान् शब्द [shrīmān] विराम पर समाप्त होता है। इससे प्रत्यय को [ko] जोड़ने से एक गलत रूप पैदा होता है, जहां प्रत्यय और मूल रूप संयुक्त श्रीमान्को उत्पन्न करता है। यह उपयोगकर्ता समुदाय के लिए अस्वीकार्य होगा, जो दोनों का श्रीमान्को के रूप में सीमांकन चाहता है। मूल रूप और प्रत्यय के स्पष्ट रूप से संकेत को सुनिश्चित करने के लिए ZWNJ (ज़ीरो विड्थ नॉन-ज्वाइनर ) को डाला जाता है, जैसा कि ऊपर के उदाहरण में दिखाया गया है।

इस मामले में निरोधक नियमों को लागू नहीं किया जा सकता, क्योंकि ऐसे मानदंडों की संख्या बहुत बड़ी है और इन्हें एक शब्दकोश में खोजना संभव नहीं होगा।

#### 4.4. प्रशासनिक मुद्दे<sup>8</sup>

ये मुद्दे डोमेन (अनुक्षेत्र) नाम, आरक्षित नाम खोलने, परस्पर विरोधी प्रस्तावों और साथ-साथ शुल्क संरचना के अनुक्षेत्र में भारत सरकार द्वारा अपनायी जाने वाली नीति के लिए प्रासंगिक हैं।

यहां कुछ मुद्दे उठते हैं, जिनमें से काफी कुछ कानूनी और आर्थिक नीतियों की प्रकृति वाले हैं।

##### 4.4.1. आरक्षित नाम सूची

आरक्षित नाम सूचियां संवेदनशील नामों के लिए बनाई जाती हैं, जिनकी किसी भी देश को रक्षा करनी होती है। निम्नलिखित मुद्दे, विशेष रूप से gTLD (जीटीएलडी) के संबंध में पैदा हो सकते हैं।

1. क्या gTLD (जीटीएलडी) को आरक्षित सूची की ज़रूरत है? क्या सरकार राजनीतिक संवेदनशीलता वाले आरक्षित नामों की एक सूची भेजेगी? यदि ऐसा है तो क्या इसमें भुगतान के मुद्दे शामिल हैं? (जिस मामले में पाठांतर के लिए विशिष्ट प्रक्रियाओं की आवश्यकता हो सकती है)।

2. एक आरक्षित gTLD (जीटीएलडी) के सभी पाठांतरों को रोकने के लिए अनुरोध करना चाहिए ?

##### 4.4.2. विवाद समाधान

---

<sup>8</sup> वास्तव में भिन्न रूप मुद्दा परियोजन दायरे में न होने के बावजूद यहाँ प्रस्तुत मुद्दे बहस को विस्तारित कर सकते हैं और इसलिए उन्हें कायम रखा जाता है।

यह कानूनी नीतियों और तंत्रों का एक क्षेत्र है, जिसे उसी से निपटने के लिए विकसित किया गया है, विशेष रूप से जब बहुभाषी लेबल की शुरुआत की जरूरत हो। जबकि "धोखाधड़ी की नीयत" और "साइबर स्क्वेटिंग (बाद में बेचने के लिए नामी कंपनियों के नाम रजिस्टर कराना)" जैसे क्षेत्रों में पहले से ही कानूनी निवारण तंत्र है, बहु-भाषावाद के समक्ष उसके स्वयं के मुद्दे होते हैं: बहुभाषी अलग अलग भाषाओं में लेबल युक्त बंडल हैं। यहां निम्नलिखित प्रमुख मुद्दों की पहचान की जा सकती है:

1. शिकायतकर्ता किस तरह एक पूरे लेबल के अधिकारों का दावा करता है ?
2. अगर एक शिकायतकर्ता को पता है कि एक पार्टी ने एक डोमेन नाम दाखिल किया है, जिसमें शिकायतकर्ता के दावे वैध हैं तो क्या एक शिकायत दर्ज की जा सकती है ?
3. निर्णय लेने की क्रियाविधि  
क्या उदाहरणों की अनुमति दी गई ? और अगर ऐसा है तो उदाहरणों के लिए कौन सा तंत्र विकसित किया जाएगा ?  
बहुभाषी स्वामित्व में क्या तंत्र के एक अलग सेट को शामिल करने की आवश्यकता होगी ?  
किसी विवाद को हल करने की विशेषज्ञता एक महत्वपूर्ण मुद्दा है। सीधे शब्दों में कहें, एक बहुभाषी विवाद के क्षेत्र में कौन कहेगा कि एक शिकायत वैध है। क्या मामला राज्य सरकार या एक सक्षम भाषा प्राधिकार के पास भेजा जायेगा ?
4. अंतरराष्ट्रीय ट्रेडमार्क संकल्प:  
जब एक ट्रेडमार्क या डोमेन नाम पर दो देशों ने दावा किया है तो किस प्रक्रिया का पालन होगा ?  
जैसे तमिल को श्रीलंका और भारत दोनों एक आधिकारिक भाषा मानते हैं। अगर एक देश में एक कंपनी के लिए तमिल में एक ट्रेडमार्क अन्य देश में एक कंपनी के ट्रेडमार्क जैसा दिखता है, तो क्या होगा ? क्या कानूनी प्रक्रिया के दौरान लेबल को अचल कर दिया जायेगा और उसे न्यायाधीन माना जायेगा ?
5. सरकार बनाम एक व्यक्ति या एक कॉर्पोरेट निकाय:  
एसे मुकदमेबाजी के मामले में क्या व्यक्तिगत दावे पर सरकार को प्राथमिकता दी जाएगी ?
- 4.4.3. भुगतान संबंधी मुद्दे  
बहुभाषी लेबलों के निर्माण और प्रत्येक से उत्पन्न पाठांतरों के संबंध में भुगतान के कुछ मुद्दे उभर सकते हैं:
  1. क्या पाठांतर प्रदान करने और पंजीकरण के लिए एक शुल्क लगेगा ?
  2. क्या पंजीकरण के इच्छुक से उसे दिये गये पाठांतर को हटाने के लिए शुल्क लिया जायेगा (साइबर-स्क्वेटिंग के मुद्दे)
  3. क्या एकाधिक भाषाओं में एक लेबल पंजीकृत कराने वाले को रियायत दी जा सकती है ?



#### 4.5. बहुभाषी gTLD (जीटीएलडी) का प्रबंधन

यहां उठाए गए मुद्दे gTLD (जीटीएलडी) से संबद्ध हैं, जहां TLD (टीएलडी) का प्रबंधन एक देश के कानून के बाहर प्रबंधित होता है।

इस क्षेत्र में कुछ मुद्दों की चर्चा की जानी चाहिए:

1. इनका आवंटन कैसे किया जायेगा, खासकर जब एक से अधिक देश एक ही भाषा का प्रयोग करते हैं ?
2. क्या किसी देश के सामाजिक और राजनीतिक रूप से संवेदनशील नामों को पंजीकृत करने के लिए विशिष्ट आरक्षण दिया जायेगा, जैसे राजनीतिक दलों, भाषा आदि के नाम के रूप में ?
3. और उपरोक्त के परिणामस्वरूप पाठांतरों बनाने के लिए किस नीति को लागू किया जायेगा? क्या पंजीकरण कराने वाले को , जो संभव हों, उन पाठांतरों को बाहर रखने की अनुमति दी जाएगी? इसके वित्तीय निहितार्थ क्या होंगे ?
4. यदि एक निगमित निकाय विभिन्न लिपियों में एक gTLD (जीटीएलडी) दर्ज कराने का इच्छुक है, तो कौन सी नीति लागू होगी? यह सुझाव दिया गया है कि मुद्दे के हल के लिए प्रत्येक लिपि/ भाषा के लिए निर्धारित नीति को लागू किया जायेगा।

## 5. रजिस्ट्रार और रजिस्ट्री संदर्भ

एक रजिस्ट्री के भीतर, तब एक महत्वपूर्ण तकनीकी विचार उभरता है, जब अंतर्राष्ट्रीय डोमेन नाम दर्ज किये जाते हैं। डोमेन नाम दोनों स्क्रिप्ट संकेत और एक भाषा संकेत के साथ टैग किया हुआ होना चाहिए। इस लक्ष्य को हासिल करने के लिए, एक रजिस्ट्री के लिए कुछ नीतियों की स्थापना करनी होगी, जिन्हें एक डोमेन नाम रजिस्टर करने के लिए मिले अनुरोध के बाद अवश्य लागू किया जायेगा। इन नीतियों के विकास में जिन तकनीकी मुद्दों पर विचार किया जाना है, वे निम्नानुसार हैं।

कुछ मामलों में, एक डोमेन नाम को या तो उसकी लिपि या उसकी भाषा के साथ टैग करना पर्याप्त हो सकता है। उदाहरण के लिए, गुरुमुखी लिपि केवल गुरुमुखी भाषा के लिए प्रयोग की जाती है। इस मामले में रजिस्ट्री भाषा का अनुमान कर सकती है, जब वह गुरुमुखी लिपि टैग के साथ एक डोमेन नाम प्राप्त करती है।

इसी तरह, केवल तमिल लिपि ही तमिल भाषा का समर्थन करती है। इस प्रकार जब एक डोमेन नाम तमिल भाषा के साथ टैग किया जाता है, रजिस्ट्री तमिल लिपि टैग का अनुमान कर सकती है।

हालांकि, या तो देवनागरी या फारसी -अरबी लिपि सिन्धी भाषा का समर्थन कर सकती है। इस मामले में जब रजिस्ट्री पंजीकृत किये जाने वाले एक डोमेन नाम को प्राप्त करती है, तो उसे दोनों भाषाओं और अपनी लिपियों के साथ टैग होना चाहिए।

इसके अलावा, देवनागरी लिपि का उपयोग कई भाषाओं, जैसे, हिंदी या नेपाली का समर्थन करने के लिए किया जा सकता है। इस मामले में जब रजिस्ट्री पंजीकृत किये जाने वाले एक डोमेन नाम प्राप्त करती है, तो उसे अपनी दोनों भाषाओं और अपनी लिपि के साथ टैग होना चाहिए।

तकनीकी मुद्दा यह है कि gTLD (जीटीएलडी) रजिस्ट्रियों और ccTLD (सीसीटीएलडी) रजिस्ट्रियों द्वारा इस्तेमाल किये जाने वाले मानक EPP (ईपीपी) प्रोटोकॉल में ऐसा करने का कोई मानक तरीका नहीं है, जो ICANN (आईसीएनएन) सिफारिशों का पालन करना पसंद करते हैं। इन मूल्यों में प्रत्येक को शामिल करने के लिए एक परिभाषित विस्तार है, मगर एक साथ दोनों के लिए नहीं। यह मामला वर्तमान में IETF (आईईटीएफ) के साथ अपनाया जा रहा है।

यह मुद्दा दो मायनों में पंजीयकों को भी प्रभावित करता है। उस हद तक, जहां कोई मानक नहीं है, एक रजिस्ट्रार को सभी EPP (ईपीपी) विस्तारों को लागू करना होगा, जो विभिन्न रजिस्ट्रियां इस समस्या के हल के लिए निर्दिष्ट कर चुन सकती हैं। उनके लिए, जो EPP (ईपीपी) पंजीयकों का उपयोग नहीं करते, उन्हें ccTLD (सीसीटीएलडी) के समर्थन के लिए जिसकी भी जरूरत हो उसे लागू करना होगा।

इसके अलावा, जब रजिस्ट्रार मौजूद हैं तो वे पंजीयन चाहने वाले के लिए अंतरफलक हैं। कई लिपियों और भाषाओं के समर्थन का चयन करने वाले पंजीयकों को यूजर इंटरफेस (उपयोगकर्ता अंतरफलक विकसित करने की जरूरत होगी, जिससे पंजीयन चाहने वाले द्वारा प्रयुक्त लिपि और भाषा को सुविधाजनक व सरल बनाया जा सकेगा।

अंत में, एक इच्छित पाठांतर के मुद्दे के संबंध हमारी चर्चा में इसका उल्लेख हुआ कि सामान्यतः किसी पाठांतर को किसी दूसरे पाठांतर की तुलना में वरीयता नहीं दी जाती है। हालांकि, RFC 3743 (आरएफसी 3743) की जरूरत होती है कि भाषा के स्तंभ में पसंदीदा पाठांतर तालिका कम से कम एक कोड बिंदु को निर्दिष्ट किया जाये। देवनागरी लिपि के संदर्भ में इस बात को प्राथमिकता दी जायेगी कि अगर पाठांतर तालिका को तब तक खाली छोड़ा जा सकता है, जब तक एक पंजीयन कराने वाला वांछित कोड बिंदु का चयन नहीं कर लेता। उस समय, क्रियाशील रूप से एक रजिस्ट्री को पंजीकरण प्रक्रिया<sup>9</sup> के बाकी अंश को पूरा करने से पहले पसंदीदा पाठांतर स्तंभ में चुने हुए कोड बिंदु को डाला जा सकता है।

### 5.1. DNS : (डीएनएस) प्रौद्योगिकी और ऑपरेशंस परिप्रेक्ष्य

यह ध्यान में रखना महत्वपूर्ण है कि DNS (डीएनएस) तकनीकी रूप से एक शुद्ध खोजी प्रोटोकॉल है: एक डोमेन नाम के साथ अनुक्रमित एक विशिष्ट जानकारी (DNS (डीएनएस) रिकॉर्ड प्रकार का) के लिए किया गया आग्रह प्रतिक्रिया में लौट आता है। अंतर्राष्ट्रीय डोमेन नामों के मामले में, डोमेन नाम के अनुरोध को A-LABEL (ए-लेबल) का होना आवश्यक है। शायद अधिक महत्वपूर्ण बात यह है कि DNS (डीएनएस) भाषा और लिपि के मामले में अज्ञेयवादी है, क्योंकि यह सूचना न तो DNS (डीएनएस) में संग्रहीत किया जाता है और न ही सीधे वैश्विक DNS (डीएनएस) बुनियादी संसाधनों के किसी हिस्से में उपलब्ध है। उस संदर्भ में, एक विशुद्ध रूप से तकनीकी नजरिये से देखा जाये तो अंतर्राष्ट्रीय डोमेन नाम DNS (डीएनएस) के संचालन में किसी तरह की विलक्षण चुनौती पेश नहीं करते।

हालांकि, अंतर्राष्ट्रीय डोमेन नाम TLDs (टीएलडी) के संदर्भ में चर्चा की एक आम बात है एक TLD (टीएलडी) के साथ दूसरे के "उपनाम" की इच्छा। वांछित "उपनाम" व्यवहार की बारीकियां विभिन्न तरह की हो सकती हैं, पर धारणात्मक रूप से इरादा यही है कि एक TLD (टीएलडी) में एक डोमेन नाम की खोज वैसा ही परिणाम दिखाये, जैसा "उपनाम" वाले TLD (टीएलडी) में बाद वाली खोज में दिखता है। समानता वाले दो डोमेन नाम का आम तौर इरादा यह होता है कि वे एक दूसरे का पाठांतर हों और यही विवाद का प्रमुख बिंदु है। "पाठांतर" की परिभाषा में कोई आम सहमति नहीं है।

संभावित परिभाषाओं के संबंध में एक पूर्ण चर्चा इस टिप्पणी के दायरे से परे है। हालांकि, यह ध्यान देना महत्वपूर्ण है कि आज DNS (डीएनएस) प्रौद्योगिकी के साथ सभी तरह की परिभाषाएं पूरी तरह से कार्यान्वित और लागू नहीं की जा सकतीं। "उपनामीकरण" के संबंध में इसका रजिस्ट्री नीतियों पर असर पड़ेगा। महत्वपूर्ण अंतर है कि DNS (डीएनएस) व्यवहार के बारे में नीतियां DNS (डीएनएस) पदानुक्रम में एक स्तर से परे लागू नहीं की जा सकतीं, जिसके आधार पर नीति परिभाषित है। विशेष रूप से, एक रजिस्ट्री नीति की स्थापना का चयन कर सकती है, जहां सभी संभावित पाठांतर DNS (डीएनएस) पदानुक्रम के TLD (टीएलडी) स्तर पर समान (DNS (डीएनएस) में वैसी ही प्रतिक्रिया दिखाये) व्यवहार करे। हालांकि यह TLD (टीएलडी) स्तर पर कई मामलों में काम कर सकता है, लेकिन DNS (डीएनएस) TLD (टीएलडी) में दूसरे

---

<sup>9</sup> RFC (आरएफसी) 3743 इस बिंदु पर असंगत है। अनुभाग 3.2.1 इस रूप में निश्चित नहीं है कि एक अशक्त प्रविष्टि की अनुमति है या नहीं हालांकि इससे संभवतः यह संकेत मिलता है कि एक प्रविष्टि की आवश्यकता है। अनुभाग 5.1 ABNF (एबीएनएफ) एक अशक्त प्रविष्टि की अनुमति नहीं देता है। हालांकि, अनुभाग 5.2 एक अशक्त प्रविष्टि की उपस्थिति के बारे में बोलता है। यह दस्तावेज मानती है कि ABNF (एबीएनएफ) निश्चित विनिर्देश है।

स्तर के डोमेन प्रत्याजित नामों पर इस नीति को लागू नहीं कर सकते हैं। यह उपयोगकर्ता के अनुभव पर एक नाटकीय प्रभाव छोड़ता है।

### 5.2. सुरक्षा और स्थिरता

पाठांतर नीतियों और उनके कार्यान्वयन के मूल्यांकन के लिए लॉग करने, समीक्षा और DNS (डीएनएस) क्वेरी यातायात के विश्लेषण एक सुझाव दिया जाता है। विशेष रूप से, अनुप्रयोगों और सेवाओं के व्यवहार और कभी-कभी उनका उपयोग करने वाले उपयोगकर्ता DNS (डीएनएस) प्रश्नों और प्रतिक्रियाओं के क्रमों में मिले यातायात पैटर्न से अनुमान लगा सकते हैं। उदाहरण के लिए, रजिस्ट्रियां अस्तित्वविहीन डोमेन (यानी, NXDOMAIN (एनएक्सडोमेन) प्रतिक्रियाओं की समीक्षा वाले DNS (डीएनएस) के संदर्भ में) के प्रश्नों के लिए TLD (टीएलडी) के DNS (डीएनएस) यातायात के समीक्षा कर सकेंगी। इन कार्य संपादनों के विश्लेषण से संकेत मिल सकता है कि भाषा तालिकाएं अधूरी हैं या पाठांतर का प्रयोग उतना प्रत्याशित नहीं है। उपयोगकर्ता के लिए एक सुसंगत, समान और अचरज नहीं पैदा करने वाला (यानी, उपयोगकर्ता की उम्मीद के मुताबिक) अनुभव स्थिरता का एक अनिवार्य घटक है। DNS (डीएनएस) हस्तांतरण लॉग उपयोगकर्ता की उम्मीदों के संबंध में कुछ अंतर्दृष्टि प्रदान करते हैं। और इस प्रकार इस बात की पुष्टि करने की क्वेषमता दिखती है कि उपयोगकर्ता समुदाय की जरूरतों को पूरा किया जा रहा है। कुछ TLDs (टीएलडी) DNS (डीएनएस) पदानुक्रम में निचले स्तर पर विश्लेषण जारी रखने के लिए दूसरे स्तर डोमेन धारकों के साथ साझेदारी पर विचार की इच्छा जता सकते हैं।

### 5.3. उपयोगकर्ता परिप्रेक्ष्य

अंतर्राष्ट्रीय डोमेन नाम शुरू करते हुए उपयोगकर्ता के नजरिए से दो मुद्दों पर विचार हो सकता है: अंतर्राष्ट्रीय डोमेन नामों का प्रस्ताव और उनकी प्रस्तुति। दो अंतर्निहित तकनीकी मुद्दे हैं। सबसे पहले, क्या एक उपयोगकर्ता सिस्टम में वांछित यूनिकोड कोड बिंदु दर्ज कर सकता है? इसका जवाब कुछ हद तक हार्डवेयर (प्रयोग में लाया जा रहा कुंजीपटल कोड बिंदु उपलब्ध करा सकता है) और सॉफ्टवेयर (क्या सॉफ्टवेयर एक मान्य प्रविष्टि के रूप में कोड बिंदु मान को स्वीकार करेगा) पर निर्भर करता है। दूसरा, उपयोग वाला सिस्टम इस तरीके से यूनिकोड कोड बिंदु प्रदर्शित करेगा, जो उपयोगकर्ता पहचान सके? इसका जवाब कुछ हद तक एक उपयुक्त फ्रॉन्ट तालिका की उपलब्धता पर निर्भर करता है, जो एक नक्काशी के प्रतिनिधित्व वाले मूल्य के साथ एक कोड बिंदु द्वारा अनुक्रमित हो और जो दिखाये जाने पर उपयोगकर्ता पहचान सके। ये मुद्दे एक स्थानीय संदर्भ में हल करने के लिए मुख्यतः सरल तरीका अपनाते हैं, लेकिन जब एक वैश्विक संदर्भ में विचार हो रहा हो, तो वे तब चुनौतीपूर्ण हो जाते हैं, जब आप समझते हैं कि कैसे एक उपयोगकर्ता उनकी संरचनाओं को बनाए रखने की उम्मीद करता है, जैसे यह सभी मामलों में "काम" करता है। इस संदर्भ में "काम करने" का अर्थ है कि उपयोगकर्ता का अनुभव को समान और सतत रहता है, अर्थात् उपयोगकर्ता किसी भी प्रविष्टि या प्रस्तुति घटना से हैरान नहीं होता। विशेष रूप से, एक वेब ब्राउज़र के मामले पर विचार करें।

वेब ब्राउज़र आजकल आमतौर पर क्षेत्रीय रूप से पैक किये हुए हैं, जिसका अर्थ है कि उनके लिए एक ब्राउज़र प्राप्त करना संभव है, जिसका डिफ़ॉल्ट व्यवहार क्षेत्रीय लिपियों या प्रयोग हो रही भाषाओं के लिए अनुकूलित है। हालांकि, इसके लिए यह जरूरी है कि ब्राउज़र (और उपयोगकर्ता के लिए) के समर्थन के लिए उचित

हार्डवेयर और सॉफ्टवेयर उपलब्ध है। इसके अलावा, एक उपयोगकर्ता द्वारा ब्राउजर का उपयोग अक्सर क्षेत्रीय दायरे से बाहर चला जाता है, जिसका अर्थ है कि एक उपयोगकर्ता को उन वेब साइटों या वेब साइटों पर सूचना (दस्तावेज़) का सामना करना पड़ सकता है, जो अतिरिक्त विन्यास (हार्डवेयर या सॉफ्टवेयर या दोनों में परिवर्तन) के बिना स्थानीय वातावरण में प्रदर्शित या इस्तेमाल नहीं किये जा सकते। महत्वपूर्ण सवाल यह है कि कैसे स्थानीय ढांचा (हार्डवेयर और सॉफ्टवेयर) बदलते प्रवेश और प्रस्तुति की जरूरतों या आवश्यकताओं की उपस्थिति में बनाए रखा जा सकता है ?

#### 5.4. सिस्टम एडमिनिस्ट्रेटर संबंधी परिप्रेक्ष्य

भूमिका के रूप में सिस्टम व्यवस्थापक एक स्थानीय संरचना बनाए रखने के लिए जिम्मेदार है। एक उद्यम की स्थिति में अधिक कौशल की उपस्थिति की उच्च संभावना होती है और इस प्रकार स्थानीय संरचना को बहाल रखने रास्ते में संसाधनों (जैसे, कर्मचारी या पैसे) द्वारा विवश होने की हालत आ सकती है। हालांकि, कई उपयोगकर्ताओं के पास मोबाइल उपकरण या अन्य निजी संसाधन होते हैं, जिसके लिए वे सिस्टम व्यवस्थापक और उत्पाद उपयोगकर्ता की दोहरी भूमिका निभाते हैं। इन उपयोगकर्ताओं के पास संभावित सर्वोत्तम उपयोगकर्ता अनुभव प्राप्त करने के लिए आवश्यक स्थानीय संरचना को ठीक तरह से बनाए रखने के कौशल की कमी की संभावना होगी।

#### 5.5. अंतिम उपयोगकर्ता संबंधी परिप्रेक्ष्य

पंजीकरण: यह ध्यान में रखना महत्वपूर्ण है कि प्रयोक्ताओं की एक विशाल संख्या एकलभाषी है और कई मामलों में भाषा और लिपि लैटिन आधारित नहीं हैं। DNS (डीएनएस) की आवश्यकता यह है कि अंतर्राष्ट्रीय डोमेन नामों के प्रश्न नाम के A-LABEL (ए-लेबल) के साथ निष्पादित करना होता है, जो उपयोगकर्ताओं के लिए एक बोझ बन जाते हैं। नाम का A-LABEL (ए-लेबल) रूप एक संकेतक है, जो नाम को इस तरह बदल (एक पलट देने वाले मानचित्रण का उपयोग कर) देता है, जिसमें केवल US-ASCII (यूएस-एस्की) वर्ण ही शामिल होते हैं। यह बदलाव यह सुनिश्चित करता है कि अंतर्राष्ट्रीय डोमेन नाम का उपयोग मौजूदा DNS (डीएनएस) बुनियादी सुविधाओं के साथ पीछे ले जाने के लिए संगत है। A-LABEL (ए-लेबल) फार्म के साथ कार्य करना कई उपयोगकर्ताओं के लिए एक हद तक बोझ है, क्योंकि संकेतक स्वयं को US-ASCII (यूएस-एस्की) वर्णों का एक यादृच्छिक अनुक्रम प्रस्तुत करता है, क्योंकि प्राथमिक तौर पर इसके साथ काम करना अप्राकृतिक है, यहां तक उनके लिए जो US-ASCII (यूएस-एस्की) से परिचित हैं। उपयुक्त सॉफ्टवेयर का प्रयोग इस बोझ को कम कर सकता है, जिसका परिणाम यह होता है कि उपयोगकर्ता को अपने स्थानीय हार्डवेयर और सॉफ्टवेयर तक ही सीमित रहते हैं।

#### 5.6. WHOIS (हूर्इज़) संबंधी मुद्दे

IDNs (आईडीएन) को लागू करने के रास्ते में आ रहे महत्वपूर्ण WHOIS (हूर्इज़) मुद्दे का सच यह है कि मानक WHOIS (हूर्इज़) प्रोटोकॉल (RFC (आरएफसी) 3912 के रूप में परिभाषित) का अंतरराष्ट्रीयकरण नहीं किया गया है, जिसका मतलब है कि यह इंगित करने के लिए कोई मानक तरीका नहीं है कि उपयोग में वांछित भाषा या लिपि है या वास्तविक भाषा या लिपि। WHOIS (हूर्इज़) प्रोटोकॉल एक साधारण अनुरोध और प्रतिक्रिया में आया कार्य संपादन है: एक डोमेन नाम एक सर्वर में डाला जाता है और नतीजा वापस दिखता है। आजकल इंटरनेट पर प्रमुख संकेतक US-ASCII (यूएस-एस्की) है, लेकिन अंतर्राष्ट्रीयकरण की कमी का एक परिणाम

यह है कि स्थानीय, क्षेत्रीय और स्वामित्व समाधानों की संख्या बढ़ती जा रही है, जिससे अंतर्राष्ट्रीयकरण की कमी की ओर संकेत करने का प्रयास माना जा सकता है। इस परिवर्तनशीलता का उपयोगकर्ता के अनुभव पर एक नाटकीय रूप से प्रतिकूल प्रभाव पड़ता है।

उदाहरण के लिए, WHOIS (हूईज़) प्रतिक्रिया में सूचना टैग करने के लिए इस्तेमाल किये गये लेबल मुख्य रूप से US-ASCII (यूएस-एस्की) में संकेत किये गये हैं। यह सीधे-सीधे विश्वास किया जाता है कि लेबलों को उसी भाषा और लिपि में प्रदर्शित किया जाना चाहिए, जिसमें डेटा है, लेकिन यह मानक WHOIS (हूईज़) प्रोटोकॉल के साथ संभव नहीं है।

इस मुद्दे से जुड़ा दूसरा मामला यह है कि क्या प्रदर्शित करना है, यह नीतिगत मामला है, जो कि कुछ अंश तक पाठांतर पंजीकरण नीति द्वारा निर्देशित है। निम्नलिखित प्रश्नों पर विचार करें।

1. यदि एक पाठांतर डोमेन नाम रजिस्ट्री डेटाबेस में मौजूद है, लेकिन DNS (डीएनएस) (यानी, डोमेन नाम आरक्षित है) में मौजूद नहीं है, तो क्या एक डोमेन नाम के लिए WHOIS (हूईज़) अनुरोध रेफरल के साथ वापस आयेगा, जिसमें संकेत होगा कि नाम एक उच्चकोटि के नाम का पाठांतर है या वह उच्चकोटि के नाम के लिए प्रतिक्रिया वापस करेगा? प्रतिक्रिया से क्या संकेत मिलना चाहिए कि नाम मौजूद नहीं है ?
2. क्या भिन्न डोमेन नामों को उनसे जुड़ी भिन्न-भिन्न WHOIS (हूईज़) जानकारीयों की अनुमति दी जानी चाहिए? इस सवाल का जवाब एक अंश तक इस बात पर निर्भर है कि क्या अलग-अलग मालिकों को भिन्न-भिन्न डोमेन नाम रजिस्टर करने की अनुमति दी जानी चाहिए।
3. यदि एक भिन्न डोमेन नाम अपने उच्चकोटि के समकक्ष डोमेन नाम से भाषा या लिपि के मामले में भिन्न है, तो यह कैसे उपयोगकर्ता के लिए प्रस्तुत किया जायेगा, यदि उपयोगकर्ता (या शायद नहीं प्रदर्शित कर सकता) उच्चकोटि के डोमेन नाम की भाषा या लिपि नहीं समझता ?
4. यदि एक WHOIS (हूईज़) अनुरोध भिन्नता के साथ एक डोमेन नाम के लिए है, तो क्या प्रतिक्रिया में इन भिन्नताओं को शामिल किया जाना चाहिए? क्या होगा यदि अनुरोध के बाद पाठांतर की भाषा या लिपि उपयोगकर्ता द्वारा समझ में न आ सके या प्रदर्शित हो सके।

### 5.7. पंजीकरण प्रक्रिया संबंधी मुद्दे

IDNs (आईडीएन) और पाठांतरों के पंजीकरण के रास्ते में जिस महत्वपूर्ण तकनीकी मुद्दे का सामना करना सचाई है, वह है कि पंजीकृत किये जाने वाले डोमेन नाम के उपयोग में भाषा, लिपि या दोनों का संकेत करने वाले EPP (ईपीपी) प्रोटोकॉल में कोई मानक तरीका नहीं है। जैसा कि रजिस्ट्री और रजिस्ट्रार के संदर्भ में वर्णित है, यह जो पंजीयन चाहने वाले उपयोगकर्ता अंतराफलक पर प्रभाव छोड़ता है साथ ही यह जानने की क्षमता पर भी असर होता है कि बहुत सारे पाठांतरों के सेट में से किस डोमेन नाम का पंजीकरण करना है। इसके बाद का मुद्दा है कि एक रजिस्ट्री के पास यह बताने की नीति की जरूरत होगी कि वह कैसे भावी डोमेन नाम पंजीकरणों के पाठांतरों से निपटेगी। निम्नलिखित प्रश्नों पर विचार करें।

1. क्या तुल्यता की एक उचित परिभाषा के लिए डोमेन नाम पाठांतरों के लिए बराबर माना जायेगा ?
2. यदि पाठांतर बराबर हैं, तो तब क्या सभी का पंजीयन (DNS (डीएनएस) प्रतिनिधिमंडल सहित) किया जायेगा, जब एक पहले वाले को प्रस्तुत किया गया हो? क्या पाठांतर आरक्षित (DNS (डीएनएस) प्रतिनिधिमंडल शामिल नहीं) रहेगा और और केवल अनुरोध पर प्रत्यायोजित किया जायेगा ?
3. यदि पाठांतर अनुरोध पर पंजीकरण के लिए आरक्षित रखे जाते हैं तो किसको पंजीकरण के अनुरोध की अनुमति होगी? पहले पंजीकृत किये गये पाठांतर के मालिक को या या किसी को भी, जो अनुरोध करे ?

तुल्यता के प्रश्न को लेकर एक महत्वपूर्ण तकनीकी मुद्दा है DNS (डीएनएस) पर प्रभाव जैसा कि DNS (डीएनएस) तकनीक और संचालन परिप्रेक्ष्य में वर्णित है। DNS (डीएनएस) पदानुक्रम में उस सीमा के बाहर DNS (डीएनएस) व्यवहार को लागू नहीं किया जा सकता है, जिस पर नीति को परिभाषित किया गया है। यह उपयोगकर्ता के अनुभव पर एक नाटकीय प्रभाव डाल सकता है। अंत में, व्यापारिक के नजरिए से एक रजिस्ट्री के लिए एक नीति निर्दिष्ट करने की जरूरत है, जिसमें यह स्पष्ट हो कि यह कैसे यह पंजीकृत डोमेन नामों के पाठांतरों के लिए शुल्क (या शुल्क नहीं लगेगा) लगेगा।

#### 5.8. DNSSEC (डीएनएसएसईसी) संबंधी मुद्दे

यहाँ IDN (आईडीएन) या पाठांतर से जुड़े विशिष्ट मुद्दे नहीं हैं जो DNSSEC (डीएनएसएसईसी) की तैनाती को प्रभावित करते हैं।

DNSSEC (डीएनएसएसईसी) के दृष्टिकोण से, एक IDN (आईडीएन) या पाठांतर TLD (टीएलडी) साधारणतः एक अलग क्षेत्र है। DNS (डीएनएस) प्रौद्योगिकी और संचालन परिप्रेक्ष्य वाली चर्चा को याद करते हुए कहा जा सकता है कि एक क्षेत्र में लेबलों के उद्देश्य या मूल्य के मामले में DNS (डीएनएस) का कोई संदर्भ नहीं होता। DNS (डीएनएस) तकनीकी रूप से एक विशुद्ध खोजी प्रोटोकॉल है। चर्चा का एक आम बिंदु है TLD (टीएलडी) "उपनामीकरण" के मुद्दे का कुंजी प्रबंधन के मुद्दों से सहसंबंध, जिसे साधारणतः तब हल किया जाना चाहिए, जब DNSSEC (डीएनएसएसईसी) को तैनात किया जाये। यह युग्मन एक अनावश्यक जटिलता पैदा करता है, क्योंकि कुंजी प्रबंधन को लागू करने से संबंधित प्रश्नों को DNS (डीएनएस) और DNSSEC (डीएनएसएसईसी) के संदर्भ में ही उत्तर दिया जाना चाहिए जैसे एक IDN (आईडीएन) या पाठांतर को DNS (डीएनएस) और DNSSEC (डीएनएसएसईसी) का मात्र एक "लेबल" होना चाहिए।

## 6. ग्रन्थ सूची का चयन

इस रिपोर्ट का प्रारूपण में नीचे दी गई और विषयगत रूप से सजायी गई ग्रंथ सूची, दस्तावेज, किताबें, लेख और वेबोग्राफीज को देखा गया है।

### लेखन प्रणालियां

Dillinger. D., The Alphabet. A Key to the History of Mankind. 3rd Edition in 2 Volumes. Hutchison. London. 1968.

### देवनागरी

Agrawala, V. S. (1966). The Devanāgarī script. In: Indian Systems of Writing. (Pp. 12-16) Delhi: Publications Division.

Agyeya, Sacchindanand Hiranand Vatsyayan. 1972. Bhavanti. Delhi: Rajpal and Sons.

Beames, John. 1872-79. A Comparative Grammar of the Modern Aryan Languages of India. 3 vols. London, Trubner and Co. [Reprinted by Munshiram Manoharlal, New Delhi, 1966.]

Bhatia, Tej K. 1987. A History of the Hindi Grammatical Tradition: Hindi-Hindustani Grammar, Grammarians, History and Problems. Leiden/New York: E. J. Brill.

Bright, W. (1996). The Devanāgarī script. In P. Daniels and W. Bright (eds), The World's Writing Systems. (Pp. 384-390). New York: Oxford University Press.

Cardona, George. 1987. Sanskrit. In The World's Major Languages. Bernard Comrie (ed.). London: Croom Helm. 448-469.

Dwivedi, Ram Awadh. 1966. A Critical Survey of Hindi Literature. Delhi: Motilal Banarsidass.

Faruqi, Shamsur Rahman. 2001. Early Urdu Literary Culture and History. Delhi: Oxford University Press.

Guru, Kamta Prasad. 1919. Hindi Vyakaran. Varanasi: Nagari Pracharini Sabha. (1962 edition).

Kachru, Yamuna. 1965. A Transformational Treatment of Hindi Verbal Syntax. London: University of London Ph.D. dissertation (Mimeographed).

Kachru, Yamuna. 1966. An Introduction to Hindi Syntax. Urbana: University of Illinois, Department of Linguistics.

Kalyan Kale and Anjali Soman, 1986. Learning Marathi. Shri Vishakha Prakashan, Pune :

McGregor, R. S. (1977). Outline of Hindi Grammar. 2nd ed. Delhi: Oxford University Press.

McGregor, R. S. 1972. Outline of Hindi Grammar with Exercises. Delhi: Oxford University Press.

McGregor, R. S. 1974. Hindi Literature of the Nineteenth and Early Twentieth Centuries. Wiesbaden: Harrassowitz.

McGregor, R. S. 1984. Hindi Literature from Its Beginnings to the Nineteenth Century. Wiesbaden: Harrassowitz.

Pandey, P. K. (2007). Phonology-orthography interface in Devanāgarī for Hindi. Written Language and Literacy, 10 (2): 139-156. 2007.

Rai, Amrit. 1984. A House Divided. The Origin and Development of Hindi/Hindavi. Delhi: Oxford University Press.

Sharad, Onkar. 1969. Lohiya ke Vicar. Allahabad: Lokbharati Prakashan.

Singh, A. K. (2007). Progress of modification of Brāhmī alphabet as revealed by the inscriptions of sixth-eighth centuries. In P.G. Patel, P. Pandey and D. Rajgor (eds), The Indic Scripts: Paleographic and Linguistic Perspectives. (Pp. 85-107). New Delhi: DK Printworld.

Sproat, R. (2000). A Computational Theory of Writing Systems. Cambridge University Press.



Tiwari, Pandit Udaynarayan. 1961. Hindi Bhasha ka Udgam aur Vikas [The Origin and Development of the Hindi Language]. Prayag: Leader Press.  
Verma, M. K. 1971. The Structure of the Noun Phrase in English and Hindi. Delhi: Motilal Banarsidass.

## बहुभाषावाद

### सामान्य

Multilingual Internet Names Consortium. MINC.  
Dam, Mohan, Karp, Kane & Hotta, IDN Guidelines 1.0, ICANN, June 2003  
Martin J. (December 20, 1996). "URLs and internationalization". World Wide Web Consortium. IDN TABLES: <http://www.iana.org/domains/idn-tables/>

### भाषा विशिष्ट

#### 1. भारतीय लिपियाँ और भाषाएँ

IS 10401: 8-bit code for information interchange. 1982  
IS 10315: 7-bit coded character set for information interchange. 1985  
IS 12326: 7-bit and 8-bit coded character sets-Code extension techniques. 1987  
ISO 15919, Information and documentation - Transliteration of Devanāgarī and related Indic scripts into Latin characters. 2001  
ISO 2375: Procedure for registration of escape sequences. 2003  
ISO 8859: 8-bit single-byte coded graphic character sets - Parts 1-13. 1998-2001  
IDN POLICY [http://mit.gov.in/sites/upload\\_files/dit/files/India-IDN-Policy.pdf](http://mit.gov.in/sites/upload_files/dit/files/India-IDN-Policy.pdf)

#### भारतीय लिपियों का रोमनीकरण

Library of Congress. Romanization Standards.. USA. 2002  
Stone. Anthony., <http://homepage.ntlworld.com/stone-catend/trind.htm>

#### 2. चीनी

CHINESE: Chinese Domain Name Consortium". CDNC. 2000-05-19

#### 3. उर्दू

URDU: <http://urduworkshop.sdnpk.org>

## आरएफसी

RFC 2181, Clarifications to the DNS Specification: section 11 explicitly allows any binary string  
RFC 2870 Root Name Server Operational Requirements June 2000  
RFC 3490 Internationalizing Domain Names in Applications (IDNA) March 2003  
RFC 3492, Punycode: A Bootstring encoding of Unicode for Internationalized Domain Names in Applications (IDNA), A. Costello, The Internet Society (March 2003)  
RFC 5890 "Internationalized Domain Names for Applications (IDNA): Definitions and Document Framework"  
RFC 5891 "Internationalized Domain Names in Applications (IDNA): Protocol"

RFC 5892 "The Unicode Code Points and Internationalized Domain Names for Applications (IDNA)" August 2010  
RFC 5893 "Right-to-Left Scripts for Internationalized Domain Names for Applications (IDNA)"

## यूनिकोड

Unicode Consortium. Unicode ver.3.0.

---. Unicode ver.3.2.

---. Online version of Unicode ver.4.1 . (archived).

----. Online version of Unicode ver.5.0 & 5.1. ([www.unicode.org](http://www.unicode.org))

----. Online version of Unicode ver.6.0 ([www.unicode.org](http://www.unicode.org))

## परिशिष्ट सूची

परिशिष्ट I: देवनागरी टीम के सदस्य.

परिशिष्ट II. भारत की राजभाषाओं की सूची:

परिशिष्ट III: देवनागरी में भ्रम पैदा करने वाले वर्णों की सूची:

परिशिष्ट IV: देवनागरी में भ्रम में डालने वाले साहित्य की सूची:

परिशिष्ट V: भिन्न विषयों वाले प्रोजेक्ट के असंबद्ध विषय, लेकिन रोचक समझे जाने वाले.

परिशिष्ट VI: देवनागरी कोड बिंदुओं अनुशंसित सूची

## परिशिष्ट ।

### देवनागरी टीम के सदस्य

सदस्य	भूमिका
डॉ. गोविंद	केस अध्ययन समन्वयक
तूलिका पांडेय	टीम के सदस्य
महेश कुलकर्णी	टीम के सदस्य
के बी नारायणन	टीम के सदस्य
डॉ. जेम्स गाल्विन	टीम के सदस्य
अमरदीप सिंह चावला	टीम के सदस्य
जितेंदर कुमार	टीम के सदस्य
राजीव कुमार	टीम के सदस्य
भाविन तुरखिया	टीम के सदस्य
शशि भारद्वाज	टीम के सदस्य
प्रो. प्रमोद पांडेय	टीम के सदस्य
रायोमोंड डॉक्टर	टीम के सदस्य
डॉ. कल्याण काले	टीम के सदस्य
प्रभाकर श्रोत्रिय	टीम के सदस्य

मनीष दलाल	टीम के सदस्य
बसंता श्रेष्ठ	टीम के सदस्य
बाल कृष्ण बाल	टीम के सदस्य
सत्येन्द्र कुमार पांडेय	टीम के सदस्य
नेहा गुप्ता	टीम के सदस्य
अक्षत जोशी	टीम के सदस्य
<b>स्टाफ सदस्य</b>	
सदस्य (स्टाफ)	भूमिका
फ्रांसिस्को आरियस	विषय विशेषज्ञ (रजिस्ट्री ऑफ्स)
नाएला सारास	केस स्टडी सम्पर्क
निकोलस ओस्टलर	विषय विशेषज्ञ (भाषाविज्ञान)
स्टीव शेंग	विषय विशेषज्ञ (नीति)
एंड्रयु सलिवन	विषय विशेषज्ञ (प्रोटोकॉल)
किम डेविस	विषय विशेषज्ञ (सुरक्षा)

## परिशिष्ट II:

### भारत की राजभाषाओं की सूची

भारत एक भाषाविद् के लिए बहुत उर्वर भूमि है, क्योंकि यहां 4 प्रमुख भाषा परिवारों के साथ 6616 भाषाओं (2001 की भारत का जनगणना ) और 20,000 + बोलियों की पहचान की गई है<sup>10</sup> SIL (एसआईएल) रिपोर्ट। इस विशाल विविधता का सामना करने के लिए भारत के संविधान ने सबको सहेजने के लिए उल्लेखनीय प्रावधान किये है, जिसने राष्ट्रीय सरकार के लिए हिंदी और अंग्रेजी के रूप में दो सरकारी संचार भाषाएं निर्धारित की हैं। इसके अलावा 22 अनुसूचित भाषाओं के एक सेट की पहचान की गई है, जो ऐसी भाषाएं हो सकती हैं।

- प्रशासनिक प्रयोजनों के लिए विभिन्न राज्यों द्वारा आधिकारिक तौर पर अपनायी गई,
- राष्ट्रीय और राज्य सरकारों के बीच संचार के एक माध्यम के रूप में,
- विश्वविद्यालय में परीक्षाओं और साथ-साथ सरकारी स्तर के लिए ।
- मतदाता सूचियों, अद्वितीय पहचान संख्या कार्यक्रम UIDAI (यूआईडीएआई) आदि जैसे राष्ट्रीय डेटाबेस के लिए।

22 अनुसूचित भाषाएं प्रतिनिधित्व तालिका के साथ निम्नांकित हैं। देवनागरी लिपि का उपयोग करने वाली भाषाओं लाल रंग में लिखा गया है।

भाषा	आईएसओ	आधिकारिक भाषा	परिवार	लिपि
असमी	asm	असम	इंडो - आर्यन	असमी
बंगला	ben	त्रिपुरा और पश्चिम बंगाल	इंडो - आर्यन	बांग्ला
मणिपुरी	mni	मेइती	तिब्बती - बर्मन	बांग्ला मइतेई - मेयक
बोरो	brx	असम	तिब्बती - बर्मन	देवनागरी संशोधित:
डोगरी	dgo	जम्मू और कश्मीर	इंडो - आर्यन	देवनागरी संशोधित:
हिंदी	hin	अंडमान और निकोबार द्वीप समूह, बिहार, चंडीगढ़, छत्तीसगढ़, दिल्ली, हरियाणा, हिमाचल प्रदेश, झारखंड, मध्य प्रदेश, राजस्थान, उत्तर प्रदेश और उत्तरांचल	इंडो - आर्यन	देवनागरी
कोंकणी	kok	गोवा	इंडो - आर्यन	देवनागरी रोमन (लैटिन)

<sup>10</sup> [http://www.ethnologue.com/show\\_country.asp?name=in](http://www.ethnologue.com/show_country.asp?name=in)

मैथिली	mai	बिहार	इंडो - आर्यन	देवनागरी
मराठी	mar	महाराष्ट्र	इंडो - आर्यन	देवनागरी
नेपाली	nep	सिक्किम	इंडो - आर्यन	देवनागरी
संस्कृत	San	अखिल भारतीय	इंडो - आर्यन	देवनागरी
गुजराती	guj	दादरा और नगर हवेली, दमन और दीव, और गुजरात	इंडो - आर्यन	गुजराती
पंजाबी	pan	पंजाब	इंडो - आर्यन	गुरुमुखी
कन्नड़	kan	कर्नाटक	द्रविड़ियन	कन्नड़
मलयालम	mal	केरल और लक्षद्वीप	द्रविड़ियन	मलयालम
संथाली	sat	झारखंड	मुंडा	ओल चिकी
उड़िया	ori	उड़ीसा	इंडो - आर्यन	उड़िया
कश्मीरी	kas		इंडो - आर्यन	फारसी - अरबी देवनागरी
सिंधी	snd	अखिल भारतीय	इंडो - आर्यन	फारसी - अरबी देवनागरी गुजराती रोमन (लैटिन)
उर्दू	urd	जम्मू और कश्मीर	इंडो - आर्यन	फारसी - अरबी
तमिल	tam	तमिलनाडु और पांडिचेरी	द्रविड़ियन	तमिल
तेलुगू	tel	आंध्र प्रदेश	द्रविड़ियन	तेलुगू

हालांकि ये 22 भाषाएं 4 विशिष्ट भाषा परिवारों से संबंध रखती हैं: इंडो - आर्यन, द्रविड़ियन, मुंडा और तिब्बती - बर्मन , जहां तक लेखन प्रणाली का संबंध है, दो प्रमुख परिवारों पहचाना जा सकता है: वे भाषाएं जिनकी लेखन प्रणाली ब्राह्मी जैसे हिंदी , बांग्ला, पंजाबी और सभी द्रविड़ भाषाओं से निकली है।

- भाषाएं जिनकी लेखन प्रणाली प्रकृति में फारसी - अरबी है। ये केवल संख्या में तीन हैं: कश्मीरी, सिंधी और उर्दू। इनमें से सिंधी और कश्मीरी एक ब्राह्मी आधारित लेखन प्रणाली अर्थात् देवनागरी का उपयोग कर भी लिखी जा सकती है।  
लेखन प्रणालियों के छोटे उप सेट भाषाओं के खांचे में देखे जा सकते हैं, जैसे मइतेई और OI Ciki जिसके पास स्वदेशी लिपि प्रणाली है।

परिशिष्ट III

देवनागरी में भ्रमित करने वाले वर्णों की सूची

वर्ण 1	वर्ण 2
उ U+0909	ऊ U+090A
ड U+0919	ढ U+0921
ज U+091C	झ U+091E
ब U+092C	व U+0935
ऋ U+090B	ॠ U+0960
थ U+0925	य U+092F
प U+092A	ष U+0937
भ U+092D	म U+092E
इ U+0907	ई U+0908
ए U+090F	ऐ U+0910
ओ U+0913	औ U+0914



क U+0915	फ U+092B
ट U+091F	ठ U+0920
त U+0924	ल U+0932
र U+0930	र U+0931
ल U+0932	ळ U+0933

परिशिष्ट IV :

देवनागरी में भ्रम में डालने वाले संयुक्ताक्षरों की सूची

1. एक जैसे दिखने वाले जोड़े

यूनिकोड	H1	यूनिकोड	H2
U+0915 U+094D U+0915	क्क	U+0932 U+094D U+0932	ल्ल
U+0915 U+094D U+092F	क्य	U+092B U+094D U+092F	फ्य
U+0915 U+094D U+0932	क्ल	U+092B U+094D U+0932	फल
U+0924 U+094D U+0924	त्त	U+0932 U+094D U+0932	ल्ल
U+0924 U+094D U+092A	त्प	U+0932 U+094D U+092A	ल्प
U+0924 U+094D U+0935	त्व	U+0932 U+094D U+092A	ल्व
U+0924 U+094D U+092F	थ्य	U+092F U+094D U+092F	र्य
U+092A U+094D U+092A	प्प	U+0937 U+094D U+092A	षष
U+092A U+094D U+092E	प्य	U+0937 U+094D U+092E	ष्य
U+092C U+094D U+092C	ब्ब	U+0935 U+094D U+0935	व्व
U+092C U+094D U+092F	ब्य	U+0935 U+094D U+092F	व्य
U+092D U+094D U+092F	भ्य	U+092E U+094D U+092F	म्य
U+0937 U+094D U+091F	ष्ट	U+0937 U+094D U+0920	ष्ठ
U+0936 U+094D U+0935	श्ध	U+0936 U+094D U+0930 U+094D U+0935	श्च
U+0936 U+094D U+0928	श्न	U+0936 U+094D U+0930 U+094D U+0928	श्ल
U+0936 U+094D U+0932	क्ष	U+0936 U+094D U+0930 U+094D U+0928	क्षल

2. एक जैसे दिखने वाले त्रिक

यूनिकोड	H1	यूनिकोड	H2	यूनिकोड	H3
U+0918	घ	U+0927	ध	U+0926 U+094D U+092f	घ
U+0918 U+094D U+092F	घ्य	U+0927 U+094D U+092F	ध्य	U+0926 U+094D U+092f	घ
U+091F U+094D U+091F	ट्ट	U+0920 U+094D U+0920	ठ्ठ	U+0922 U+094D U+0922	ढ्ढ
U+0924 U+094D U+092F	त्य	U+0932 U+094D U+092F	ल्य	U+0924 U+094D U+0924 U+094D U+092F	ल्य

## परिशिष्ट V

भिन्न मुद्दों वाली परियोजना के अप्रासंगिक विषय, लेकिन रुचिकर प्रतीत होते हैं।

भिन्न मुद्दों वाली रिपोर्ट के लिए अप्रासंगिक मुद्दे, लेकिन जिनमें पाठांतर शामिल हैं, यहां प्रस्तुत किये जा रहे हैं।

### 1. रजिस्ट्री प्रबंधन

एबीएनएफ का रजिस्ट्री प्रबंधन<sup>11</sup>, पाबंदी वाले नियम, भाषा तालिकाएं और पाठांतर तालिकाएं। ऊपर देवनागरी लेबलों के प्रत्यायोजन से उत्पन्न होने वाले मुद्दों पर विचार - विमर्श किया गया। ये एक तरीके से पैदा हुए मुद्दों से संबद्ध हैं, जिसमें भाषा और पाठांतर तालिकाओं को रजिस्ट्री द्वारा प्रबंधित किया जाएगा। यह चर्चा भारत, के लिए नीति तक सीमित है। हालांकि ये मुद्दे उठाये गये हैं, क्योंकि ये एक खास वर्ग के हैं और उनका बड़ा असर पड़ सकता है।

इस संबंध में उठने वाले कुछ बड़े मुद्दे निम्नानुसार हैं:

1. देवनागरी के मामले में, एक बड़ी संख्या में भाषाएं कोड ब्लॉक U+900 का उपयोग करती हैं। भारत की रजिस्ट्री के लिए भाषावार समाधान प्रदान करना होगा कि कैसे रजिस्ट्री भाषा तालिका को बरकरार रखेगी ?
2. उपरोक्त के परिणामस्वरूप क्या रजिस्ट्री प्रत्येक भाषा के लिए भिन्न तालिका का समर्थन करेगी? हिंदी पाठांतर तालिका में केवल दो प्रकार के पाठांतर हैं, जबकि मराठी, कोंकणी और नेपाली भी पाठांतर तालिका के तीसरे प्रकार को (उपरोक्त अनुभाग 2.2 का संदर्भ लें) स्वीकार करती हैं
3. टीएलडी के मामले में भारत, के अलावा कौन से नियम लागू होंगे? यह सुझाव दिया जाता है कि इस मामले में ICANN (आईकैन) को प्रत्येक लिपि / भाषा के लिए परिभाषित पाठांतर नियमों और तालिकाओं को लागू करना चाहिए।

### 2. WHOIS (हुईज) का "स्थानीयकरण"

"स्थानीयकरण" शब्द WHOIS (हुईज) लिए इस्तेमाल किया जाता रहा है, लेकिन मुद्दे इससे काफी आगे तक असर डालते हैं। दो मामलों की पहचान की जा सकती है:

1. लेबल का कोई पाठांतर नहीं है। उस मामले में प्रमुख मुद्दा यह है कि सूचना प्रदर्शित कैसे होगी। क्या जानकारी भाषा / लिपि में प्रदर्शित होनी चाहिए। यहां भाषा की प्राथमिकता होती है। एक कोंकणी वक्ता हिंदी और उसके उलट प्रदर्शित जानकारी पसंद नहीं करना चाहता। WHOIS (हुईज) से संबंधित स्थानीयकरण और भाषा - वार सूचना एक प्रमुख मुद्दा बन जाता है
2. यह मान लें कि पंजीयन चाहने वाले को पाठांतर आबंटित है (शुल्क के भुगतान के साथ/उसके बिना), तो इस आबंटन से निम्नलिखित मुद्दे उभरते हैं:
  1. एक परिदृश्य में, जहां एक उपयोगकर्ता एक पाठांतर की जांच करता है तो क्या उस पाठांतर से संबद्ध सभी पाठांतरों को प्रदर्शित किया जाना चाहिए। यह स्वीकार किया जाता है कि जेडडब्ल्यूजे /ज़ीरो विड्थ नॉन-ज्वाइनर के मामले में यह खास तौर पर महत्वपूर्ण हो जाता है, क्योंकि स्क्रीन पर दोनों एक जैसे दिखते हैं।

<sup>11</sup> उपरोक्त फुटनोट का संदर्भ लें। एबीएनएफ इंडिक अक्षर को हैंडल करने के लिए विकसित संवर्धित बेकस-नौर रीतिवाद का एक संक्षिप्त रूप है। नियम शासी अक्षरों (L) के अलावा यह हाइफन (H) और डिजिट या अंक (D) को भी हैंडल करता है।

- जैसे गड्डा जैसे एक लेबल के मामले में  
गड्डा (ज़ीरो विड्थ नॉन-ज्वाइनर के बिना) गड्डा (ज़ीरो विड्थ नॉन-ज्वाइनर के साथ)  
एक ही तरह के दृश्य परिणाम देते हैं।
2. उपरोक्त के परिणामस्वरूप क्या WHOIS (हुईज) जानकारी किसी दिए गए लेबल और उसके पाठांतर के समान होना चाहिए या इसे अलग होना चाहिए? जो भी पसंद होगी, उसका रजिस्ट्री के कामकाज पर असर होगा।
3. ऐसी स्थिति में, जहां एक पाठांतर बाद के चरण में अवमूल्यित या जोड़ा जाता है, तो रजिस्ट्री कैसे ऐसी जानकारी को प्रदर्शित करेगा। क्या एक रजिस्ट्री के पास एक व्यवस्थित "पुनः अनुक्रमण" प्रणाली है और यदि हां, तो अर्थशास्त्र और सुप्रचालन विज्ञान के मामले में इससे होने वाले खर्च का क्या होगा ?
4. उपरोक्त मामले के परिदृश्य (1-3) उन पाठांतरों के लिए हैं, जो स्वीकार कर लिये गये हैं। टाइप 2 पाठांतरों, जहां पाठांतर स्वचालित रूप से अवरुद्ध है, तो क्या रजिस्ट्री को ऐसे पाठांतरों का प्रदर्शन करना चाहिए ?

परिशिष्ट VI:

देवनागरी कोड अंक की अनुशंसित सूची:

Continuous Range / Character	Validity Status <sup>12</sup>	Unicode Name
0901..0903	PVALID	# DEVANAGARI SIGN CANDRABINDU..DEVANAGARI SIGN VISARGA
0905..0928	PVALID	DEVANAGARI LETTER A..DEVANAGARI LETTER NA
092A..0933	PVALID	DEVANAGARI LETTER PA..DEVANAGARI LETTER LLA
0935..0939	PVALID	DEVANAGARI LETTER VA..DEVANAGARI LETTER HA
093A..093B	PVALID	DEVANAGARI VOWEL SIGN OE..DEVANAGARI VOWEL SIGN OOE
093C..094D	PVALID	DEVANAGARI SIGN NUKTA.. DEVANAGARI SIGN VIRAMA
094F	PVALID	DEVANAGARI VOWEL SIGN AW
0950..0952	PVALID	DEVANAGARI OM..DEVANAGARI STRESS SIGN ANUDATTA
0956..0957	PVALID	DEVANAGARI VOWEL SIGN UE..DEVANAGARI VOWEL SIGN UUE
0960..0963	PVALID	DEVANAGARI LETTER VOCALIC RR..DEVANAGARI VOWEL SIGN VOCALIC LL
0966..096F	PVALID	DEVANAGARI DIGIT ZERO..DEVANAGARI DIGIT NINE
0971..0972	PVALID	DEVANAGARI SIGN HIGH SPACING DOT..DEVANAGARILETTER CANDRA A
0973..0977	PVALID	DEVANAGARI LETTER OE..DEVANAGARI LETTER UUE
097B..097C	PVALID	DEVANAGARI LETTER GGA..DEVANAGARI LETTER JJA
097E..097F	PVALID	DEVANAGARI LETTER DDDA..DEVANAGARI LETTER BBA
02BC	PVALID	MODIFIER LETTER APOSTROPHE

<sup>12</sup> <http://www.unicode.org/Public/idna/6.0.0/IdnaMappingTable.txt>