

Date: 1st April 2011

COMMENTS

The comments are divided into the following parts:

GENERIC COMMENTS ON THE DOCUMENTS:

1. The Hindi Document has issues of transparency insofar as the translation is concerned. An overt example is the use of चर for variant. The term implies a variable in mathematics or in programming and is not necessarily a variant. The use of such terms leads to lack of transparency and ease of access for the common man.

SUGGESTION: The document be suitably modified for ease of reading and accessibility.

2. The French Document has the following error in the PDF:

Error! Bookmark not defined.

SUGGESTION: The document be recompiled and the contents be rebuilt for an accurate PDF.

SPECIFIC COMMENTS ON THE POLICY

1. Definition of a Variant

PROBLEM STATEMENT

A variant is defined as :

variant characters occur where a single conceptual character can be identified with two or more different Unicode Code Points with graphic representations that may or may not be visually similar.

While this definition does hold true partly, it may create confusion in the mind of linguists to whom the task of determine a variant table is given. The latter part of the definition is particularly problematic.

graphic representations that may or may not be visually similar.

While this may hold true for German where the umlaut characters (a, e, o) and the β can have dual representations;

it is not necessarily pertinent to Arabic or Indic or Spanish where such cases do not arise and such definition would lead the linguist to identify two characters with variant spellings, thereby moving from graphic representation to spelling norms.

SUGGESTION

At some point the gray area between “graphical identity” and “orthography” needs to be clearly demarcated.

It would be advisable to amplify the definition of a variant and provide examples

2. WORK PLAN

Insofar as the work-plan is concerned, items 3-8 will be the focus of a follow-on by ICANN policy development and implementation guidelines.

However Item 4 closely meshes with Items 1 and 2 and for all these (1-2-4), more parameters need to be defined in addition to the ones already proposed.

ITEM 1

Create a commonly understood glossary of terms and ensure that such terms are accurate and vetted with appropriate technical and linguistic communities

ITEM 2

Identify the set of challenges of working with IDN variant TLDs that are based on (a) linguistic accuracy, (b) technical feasibility and accuracy, (c) usability, (d) accessibility, and (e) security and stability

ITEM 4

Analyze and arrive at rules where possible, or guidelines where rules are not possible, that address the challenges of working with IDN variant TLDs outlined in task 2

SUGGESTIONS

In the case of Item 1 a clear-cut set of terms be identified and determined. While this may seem easy, in certain communities, the use of certain TLD's may be unacceptable. The use of .xxx may not be acceptable to a large number of Asian communities

Vetting for culture and acceptability within an ethno-linguistic group should be advocated as a caveat.

Insofar as Item 2 and 4 are concerned, more parameters be added to the guide-lines. This especially important in countries such as India or South Africa where a large number of official languages using different scripts are constitutionally recognised and the use of a term in one language may lead to an entirely different meaning in another

As an example (for transparency German and English are used: German *mist* and English *mist* have entirely different meanings)

The following are proposed:

1. *The term chosen must as far as possible be pan-country : accessible and common to as many languages as possible. Using a cognate which is spread across all cultures of the country would be a good idea.*
2. *The term must be evocative and easy to remember. Using a difficult word would not be a good candidate, since the term has to be transparent to the common man.*
3. *The name chosen should be short or permitting abbreviation. Acronyms or shortened forms would be possible where typing the first three letters automatically generates out the full TLD..*
4. *The term chosen should be as far as possible not clashing with a proper name. e.g. vaidya in Hindi or Marathi for medicine would be a bad choice, since it clashes with a common surname: Vaidya*
5. *When all fails use Transliteration but once more ensure that the transliteration is "Ethniified" i.e.*

- rendered as would be pronounced in the language of that culture.*
6. *The term should have as little variants as possible, which would avoid spoofing. In complex scripts as few ligatures as possible.*
 7. *Finally the term should pass the test of community approval.*
 8. *In case of Follow-on task 4, the term “technical implementation area” should be clearly defined so as to include the “user applications” like email clients, browsers in it. Although this has been mentioned in the subsequent paragraphs the clear connection does not get established. This would call for explicit opinions from the case study teams about how they want the applications to function.*

Finally while the notion of a problem statement is absolutely pertinent, no clear guidelines have been provided as to how such a statement may be created.

Some statements of problems such as “homographic identity”, “pertinence” are generic to all cultures.

Others such as ligatural identity and ligatural variants depending on the font used within the browser are more pertinent to complex scripts. The same would possibly apply to CJK scripts. A committee be appointed to cull out such criteria, identify the same and which would function as guidelines for all communities as a whole. This would also help in Phase 4 of “**Synthesis of Issues**”