

Report of Public Comments

Title:	IDN Variant Issues Project: Draft Integrated Issues Report																				
Publication Date:	2012-02-13																				
Prepared By:	Karen Lentz																				
Comment Period:		Important Information Links																			
Open Date:	2011-12-23	Announcement																			
Close Date:	2012-01-30	Public Comment Box																			
Time (UTC):	23:59	View Comments Submitted																			
Staff Contact:	Kurt Pritz	Email:	kurt.pritz@icann.org																		
Section I: General Overview and Next Steps																					
<p>Gathered feed-back from community on the draft IDN Variant Issues report. Next step is to update the draft report to produce the final IDN Variant Issues Report based on the comments received and produce the project plan for presentation in the next ICANN meeting.</p>																					
Section II: Contributors																					
<p><i>At the time this report was prepared, a total of [number] (n) community submissions had been posted to the Forum. The contributors, both individuals and organizations/groups, are listed below in chronological order by posting date with initials noted. To the extent that quotations are used in the foregoing narrative (Section III), such citations will reference the contributor's initials.</i></p>																					
<u>Organizations and Groups:</u>																					
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 50%;">Name</th> <th style="width: 30%;">Submitted by</th> <th style="width: 20%;">Initials</th> </tr> </thead> <tbody> <tr> <td>Internet Users Contributing Group</td> <td>JFC Morfin</td> <td>IUCG</td> </tr> <tr> <td>At Large Advisory Committee</td> <td>At-Large Staff</td> <td>ALAC</td> </tr> </tbody> </table>				Name	Submitted by	Initials	Internet Users Contributing Group	JFC Morfin	IUCG	At Large Advisory Committee	At-Large Staff	ALAC									
Name	Submitted by	Initials																			
Internet Users Contributing Group	JFC Morfin	IUCG																			
At Large Advisory Committee	At-Large Staff	ALAC																			
<u>Individuals:</u>																					
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 45%;">Name</th> <th style="width: 35%;">Affiliation (if provided)</th> <th style="width: 20%;">Initials</th> </tr> </thead> <tbody> <tr> <td>Chris Dillon</td> <td>University College London</td> <td>CD</td> </tr> <tr> <td>Maneesh Pangasa</td> <td></td> <td>MP</td> </tr> <tr> <td>Timothy Roy</td> <td>Telecommunications Registry Authority</td> <td>TR</td> </tr> <tr> <td>Abdulaziz Al-Zoman</td> <td></td> <td>AZ</td> </tr> <tr> <td>John Klensin</td> <td></td> <td>JK</td> </tr> </tbody> </table>				Name	Affiliation (if provided)	Initials	Chris Dillon	University College London	CD	Maneesh Pangasa		MP	Timothy Roy	Telecommunications Registry Authority	TR	Abdulaziz Al-Zoman		AZ	John Klensin		JK
Name	Affiliation (if provided)	Initials																			
Chris Dillon	University College London	CD																			
Maneesh Pangasa		MP																			
Timothy Roy	Telecommunications Registry Authority	TR																			
Abdulaziz Al-Zoman		AZ																			
John Klensin		JK																			
Section III: Summary of Comments																					
<p><i>General Disclaimer: This section is intended to broadly and comprehensively summarize the comments submitted to this Forum, but not to address every specific position stated by each contributor. Staff recommends that readers interested in specific aspects of any of the summarized comments, or the full</i></p>																					

context of others, refer directly to the specific contributions at the link referenced above (View Comments Submitted).

(IUCG) This comment on behalf of IUCG provided some background on their interest, stated that they were not yet ready to send a full set of comments, and noted the reasons why they were not yet ready.

(CD) This comment highlighted the difficulty with the meaning of “variant” in the report, and suggested that part of the problem may come from the number of scripts covered; in order to address this, perhaps more scripts could be investigated. The comment also included observations about the difficulties with whole-script variants, noting in particular that any dictionary-based system was bound to be extremely difficult and suggesting that, if <color> and <colour> are not whole-string variants, then nothing else should be either (and conversely, if anything else is then <colour> and <color> ought to be).

(MP) This comment concerned topics unrelated to the report.

(TR) This comment discussed how the label generation rules should be established for Arabic characters. It was not clear whether this was meant to be Arabic script or Arabic language, although the comment refers to “Arabic speaking nations/countries.” Since the mechanism for determining such rules is part of follow-on work from this report, this comment is perhaps more relevant to a later stage of work.

(ALAC) This comment contained a number of points:

1. New issues related to variants did not arise in the report, and thus ICANN should move forward quickly on deployment of variants.
2. The scripts studied in this project were few, and thus more study should be undertaken.
3. The team did identify some distinctions, including that different communities seem to have different levels of readiness; therefore, those who are ready with a plan should be given priority.
4. If expert panels are convened to develop zone repertoires and label generation rules, those panels must operate in a transparent manner in order to avoid bias.

(AZ) This comment came from a participant in one of the case studies, and expressed concern about the issue of the Zero Width Non-Joiner. The comment highlighted some discussion from the Arabic case study, and noted that the issue was perhaps not one for the integrated report to outline in such detail, but that the issue should be emphasized to make sure it is not overlooked.

(JK) This was a detailed comment, noting that the review of the report was supported in part by

ICANN. The comment noted that the report is very long and unlikely to be read by many people, and that the Executive Summary is not entirely consistent with the body of the text. The comment raised seven key issues:

1. The question that motivated the report may well have been the wrong one, and a question posed differently could have entailed a different conclusion. One way of evaluating this possibility is to ask whether, for a given variant candidate pair, it is ever reasonable to delegate one of the pair without the other one.
2. The Blocking methodology could be simpler than the report suggests; in addition, we already know how to do this and it is simple. The report does not spend enough time on this option.
3. The term “variant” remains unclear as a result of the report, and the report is not always clear about that. Given that the effort was unable to come up with a single meaning of the term, ICANN should perhaps take that as evidence that the term doesn’t have one meaning, and that therefore things that are “variants” shouldn’t get special treatment.
4. The DNS does not contain words, but mnemonics; the report notes this but does not always attend to the fact.
5. The report, and the entire effort, may be suffering from having started by thinking about the matter from the point of view of the DNS. Perhaps some deeper analysis is needed.
6. The script team approach worked better than many people expected, but it is still problematic: scripts are not always the right category, and sometimes the category is wrong anyway. Moreover, mitigations suggested in the report could be bad for minority languages and developing countries.
7. Some of the mechanisms for establishing label generation rules in the report run a significant risk of creating a *de facto* first come, first served policy in which currently active ICANN participants get an advantage. This might disadvantage those who come later.

The comment drew a few conclusions: First, it seems that variants are simply too complex to be effectively implemented. The comment noted the Katoh report previously urging that nothing like variants be undertaken, and proposed that talk of variants be ended, leaving such determinations up to applicants, while perhaps expanding objection procedures. The comment noted that simplicity is a virtue for the Internet and for ICANN, and that the simplest policy, along the lines suggested, is therefore the best.

Section IV: Analysis of Comments

General Disclaimer: This section is intended to provide an analysis and evaluation of the comments received along with explanations regarding the basis for any recommendations provided within the analysis.

Terminology

A number of comments referenced the difficulties with the term “variant” and its inconsistent usage in the discussions. While it was initially envisaged that a common definition could be established as a result of this project to help inform the discussions on this topic, it became clear throughout the case studies and the development of the integrated report that this was not feasible at the current time. The report takes the approach of acknowledging that the term has acquired many different meanings, and discussing specific cases and the relevant issues identified. The term “variant” is best used with qualifiers that help provide context for the type of case being referred to, and this point is being highlighted in the revision of the report.

Next Steps

Some comments suggested that additional study phases could be helpful given the limited number of scripts studied in the project, and the limited language representation on some of the teams. The script case study teams were very careful to identify areas where they did not have expertise, and this is also noted in the discussion of the scope of the report in section 2.3.

These limitations are understood. While it was not feasible to convene case study teams with expertise in every language worldwide, key users were selected with the goal of illuminating the common and script-specific issues. This was used as a basis for identifying issues and considering solutions across scripts. The limitations of the script teams are also noted in a comment as having possible consequences for the approach taken to label generation rules. Where work proceeds toward setting label generation rules for the root zone, this issue should be taken into account and the widest possible participation and expertise should be sought, to avoid disadvantaging users not represented. Alternatively, the label generation rules adopted should have considerations for incrementally adding languages/language communities as they become interested in developing rules for their use in the root.

A comment suggested that those communities deemed as operationally ready for implementation should proceed toward deployment of IDN variant TLDs. Allowing certain user communities to proceed ahead of others has an impact on the rest of the zone, as the rules adopted in this context would then become the *de facto* rules for the root zone. As noted in a related comment, this essentially creates a “first come first served” process. Definition of rules for the root zone should be an open process. There is no proposal from ICANN to impose any rules deriving from the six case studies or from any of the communities that may have a set of rules.

Whole Script Variants (Whole String Issues)

Comments generally supported the analysis in section 3.5 of the report, while noting an additional risk of opening up new sets of questions outside of the IDN space.

Transparency Requirements

A comment noted that where evaluation panels or committees are established, there is a need for them to operate in a transparent manner in order to avoid bias. Section 4 of the report discusses a range of possible options for label generation rules. It is agreed that transparency principles should be followed with regard to panels that work within the ICANN context. Some of the options discussed in the report involve the processes of third parties which may have different practices in regard to transparency.

Project Scope

A comment notes that the project might have been founded on the wrong question, i.e., working from an assumption that solutions to implement all kinds of IDN variant TLDs would be available. The comment suggests that had the questions been framed differently, the results of the studies may have been different. This point is accepted. However, it should also be noted that the project team did work with the case study teams to challenge their assumptions and explore their thinking, so that it was not assumed that all types of “variant” mechanisms discussed should be accepted and implemented without question.

Blocking

A comment suggests that blocking has been insufficiently considered as a solution to the issues posed. It is acknowledged that a blocking mechanism for the root zone would not be technically complex to implement. This is essentially the approach that has been taken with identified variant TLD labels in the IDN ccTLD Fast Track. As noted in the comment, a major set of issues discussed in the report (section 4) concerns how to establish what are to be considered variant labels: a blocking mechanism could only be implemented after a determination of what were the labels qualified to be blocked. The comment recommends that this can be solved by an objection-based process. Such a process could be considered as a means to implement a blocking mechanism; however, it is likely that at least some minimal set of guidelines would be required for making determinations on these objections, which would necessitate some work in the area of label generation rules, as proposed in section 7 of the report.

Labels as Mnemonics

A comment underscores the point that domain names are mnemonic devices, not necessarily words. This point is accepted, although it should be noted that mnemonics, abbreviations, or acronyms may work in different ways in various scripts. Domain labels are understood as words in some cases, and this affects how issues related to variant labels are understood. The report attempts to make the correct distinctions between labels and words. This is a key distinction when considering the desired outcomes of including multiple scripts and languages in the DNS. The assumption stated in section 1.1, noting that “Recognizing the need for usability of multiple scripts in the DNS, and the desirability of reasonable approximations of natural language usage, it is also assumed that users are not dependent on the ability to use the full natural language without restrictions, and will be able to accommodate certain limitations to the full natural language where necessary” should also be recalled in this context.

Visual Similarity

A comment describes the discussions of visual similarity issues where they occurred in the case study teams, and notes that some scripts must be considered in conjunction with others that are closely related. This is recognized in the report as something to be tackled in additional phases, and this issue is acknowledged as important for future consideration. Visual similarity is also addressed as an area separate from variants in the new gTLD program.

Assignment of Codepoints to Scripts

A comment notes that there is not universal agreement as to which characters are included in particular scripts, and that existing systems may have some elements of arbitrariness. As noted in the comment, this issue has broad consequences and would affect any guidelines or requirements that use such classification systems. This is agreed and is an issue for consideration in the context of label generation rules and processes.

Zero-Width Non-Joiner (ZWNJ)

A few comments raised issues concerning the ZWNJ character. One comment suggested a consortium to arrive at agreement on these issues, while one comment emphasized a set of issues related to the use of the ZWNJ character in TLD labels. It is expected that discussions about the inclusion or non-inclusion of ZWNJ in the root label generation rules (LGR) should happen in a future phase once the LGR process is defined, perhaps as part of the specific script discussions to fill out the LGR.