

ICANN VIP Project: Cyrillic

John C Klensin¹

The team that produced this report is to be commended for keeping focus on the DNS and TLD issues and producing a report that is both useful and succinct. I believe that the fundamental recommendations and projected strategy are sound. Nonetheless, there are several issues with the report that may be worthy of note.

- (1) The most important of those issues is that the report repeatedly calls for further study or further evaluation of various topics. While that is reasonable from some standpoints, in others part of the reason is that neither the Team nor, I assume, any sources they could readily identify. ICANN should elicit specific suggestions from the team about how such studies could be accomplished. See also the comments about “further study” in my “Overview” review (posted to the Latin script team forum).

A key aspect of those “additional study” topics for Cyrillic this report and apparently the expertise that went into it seem to cover the use of Cyrillic by Slavic languages fairly well but to become weaker (less deep or less confident) when other languages are considered. Saying that there could be new issues in languages, possibly identified in the report but not represented on the committee is equivalent to saying that moving forward with Cyrillic policies and registrations would be entering unknown territory. That is, itself, a sort of meta-issue but, unless the Team has a proposal about how to resolve it, the list of issues is inherently incomplete.

Another example of almost the same issue is the disclaimer at the end of Section 1. Apparently, the group believes, or may believe, that there may be a significant number of issues that have not been identified. Presumably some fraction of those hypothetical issues might change the results and conclusions of the report if only they were well-understood. How does the Team recommend getting this situation to the point that the risks associated with unexamined characters are acceptably low? Until that point is reached, does the team believe that its final report should become part of the consolidated report?

¹ This review was prepared at the request of the ICANN Variant Information Project Team and partially supported by ICANN. It reflects the author’s personal views and may not reflect the views of ICANN staff, the members of the VIP teams, or other personnel associated with ICANN. The author had the opportunity to do partial reviews of working drafts of this documents and prepared comments for the team. Some of those comments are reflected in the report as posted and hence do not appear here. Sections of it draw heavily on other work by the author that bears on the issues discussed.

- (2) The “native speakers” count in footnote 1 may be misleading. If one is going to discuss the users of a particular script with a particular language, isn’t some knowledge of literacy rates necessary?
- (3) Cyrillic is one of the scripts that is dependent on Common or Inherited Script characters to form a significant number of strings. See the comments on this subject in my “Overview” review.
- (4) Several comments, e.g., those in Section 3.8, suggest that Unicode may add more precomposed characters where none now exist. Unicode is now discouraging adding new precomposed characters. More important, an important consequence of normalization stability rules is that, should such characters be added later, they will have no consequences for IDNA. Given that situation, it is not clear what this paragraph (and similar ones) mean.
- (5) The first paragraph of Section 6.3 requires that “all user experience implications arising out of variant Cyrillic characters...” as a gating condition for delegation of reserved names. The nature of variants as discussed in that document and of the range of possible user experiences substantially guarantees that meeting this condition requires proving a universal negative. What does the team believe is the actual gating condition? In addition, at least in principle, any plan based on allocation of multiple names in the root to the same party could be broken by a legal attack on ICANN that resulted in separating the names and requiring that they be delegated separately. How do this team, and others, believe that would affect their variant model? Or is the intent simply to try to encourage courts to not do that.
- (6) The report contains a certain amount of hand waving in the form of recommendations that probably cannot be implemented. The discussion immediately above also illustrates an example of this problem. As a better example, Section 7 recommends that “ICANN... only Cyrillic characters that have been vetted by the respective language communities.”. Noting that one of the things that occurred when IDN delegations were first permitted for second-level names in existing TLDs was that registries received requests for strings identified with languages the registries had never heard of and also that the Cyrillic team was unable to recruit experts for non-Slavic languages that use the script, how does the team recommend that ICANN obtain and evaluate an appropriate and adequate vetting process?

(7) A variant strategy depends on blocking of names requires processes to identify the names that are to be blocked, databases to identify those names, and a regime for a registry to reject names requested by a registrar that are blocked rather than already delegated. It raises issues as to whether the blocked variants should be reflected in registry information databases and queries (“whois”). Yet Section 8 says that the team does not envisage any impact on registry/registrar operations. The team should explain that apparent contradiction.