Report of Public Comments

Title:	IDN TLDs - LGR Procedure Implementation - Guidelines for Designing Script-Specific Label Generation Rules (LGR)		
Publication Date:	20 March 2015		
Prepared By:	IDN Program Team		

Comment Period:		Important Information Links	
Comment Open Date:	15 January 2015	Announcement	
Comment Close Date:	27 February 2015	Public Comment Box	
		View Comments Submitted	

Staff Contact:	Sarmad Hussain	**Email:**	sarmad.hussain@icann.org

Section I: General Overview and Next Steps

To support IDN variants in the root zone, the ICANN community, at the direction of the Board, undertook several projects to study and make recommendations on their viability, sustainability and delegation. One of these projects is the implementation of the Procedure [PDF, 772 KB] allowing for the development of Label Generation Rules (LGR) for the Root Zone. The LGR for the Root Zone is a mechanism for creating and maintaining rules with respect to IDN labels for the root. This mechanism will be used to determine which Unicode code points are permitted for use in U-Labels in the root zone, what variants (if any) are allocatable and what variants (if any) are automatically blocked.

The work defined by the Procedure will be carried out by community based Generation Panels that produce LGR proposals for their scripts, to be reviewed and integrated by an Integration Panel. To facilitate the work of the Generation Panels ICANN has produced a set of guidelines and introductory documents.

These documents consist of Guidelines for Developing Script-Specific Label Generation Rules (LGR) for Integration into the Root Zone Label Generation Rules [PDF, 946 KB] (Guidelines), as well as a set of Considerations for Designing a Label Generation Ruleset for the Root Zone [PDF, 193 KB] (Considerations) that Generation Panels might apply in making and documenting their decisions. Finally, the documents contain a summary of the Requirements for LGR Proposals [PDF, 198 KB] (Requirements) on format and contents for submitting an LGR proposal.

As result of the public comments the guidelines and associated documents will be finalized and made available to Generation Panels to support their work.

Section II: Contributors

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.

Organizations and Groups:

Name	Submitted by	Initials
SaudiNIC	Raed I Al Fayez	

Individuals:

Name	Affiliation (if provided)	Initials
Marc Blanchet		MB

Section III: Summary of Comments

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

Comments made by Raed I. Al Fayez (SaudiNIC)

Item I — Master key algorithm
Describes the use of a Master Key algorithm in managing variant sets.

Item II — Positional Variants
This comment refers to the fact that Arabic code points that may be identical if they occur in the middle of a label, for example, may be distinct if they occur at the end. Keeping track of this is recommended as a way to reduce the number of variants.

Item III — Input sets for defining variants
Recommends dividing variants into subsets corresponding to keyboard layouts, such that any variant label should only contain variants from a single layout. This is recommended as a means to cut down on allocatable variants.

Item IV — Step-by-step adoption
Recommends phasing the development of the Arabic LGR to include better known and establish languages first.

Item V — Study variants across the whole Arabic script
Recommends to base the Arabic LGR on a complete study of the variant issues for all languages, so as to not create the risk of incompatible changes as future languages are added.

Item VI — Variants types
Discusses various types of variants, including "interchangeable variants" which term refers to variants that are free spelling variants.

Item VIII — International Reachability
Discusses the issue of accessing labels from localized environments that may not support the same code points on their keyboards.

Comments made by Marc Blanchet
Discusses the coordination process required among Generation Panels for related scripts or overlapping repertoires and suggests ways to streamline that process.

Section IV: Analysis of Comments

General

One of the set of comments submitted does not address the guidelines so much as to give input for the Arabic Generation Panel in their deliberations. This comment will be passed on for consideration. The second set of comments addresses the process, and the guidelines will be updated to satisfy it.

In response to comments made by Raed I. Al Fayez (SaudiNIC)

Item I — Master key algorithm

This is an implementation / tooling issue that has more to do with operating the registration database and providing collision detection for labels with variants, which is out of scope for the current project.

For the purposes of collision detection, any set of permuted variant labels can be represented by a canonical element – as long as the set is guaranteed to be symmetric and transitive (that is, no matter which label in the set you use to start the permutation you generate all of them).

The information to build the variant label set from any of its member labels is already present in the LGR so the technique does not require explicit support during LGR development.

One thing that this algorithm does not do is to reduce the number of allocatable variants.

Item II — Positional Variants

Positional variants are challenging and potentially complex. The opposite approach, to make them all variants independent of position is the conservative answer — at least in the case of blocked variants.

In case of allocatable variants, giving only those positional variants that have matching forms the allocatable type might result in some reduction of the numbers for allocatable labels. While this might be a good strategy for minimizing allocatable variant labels, and should be considered by the Arabic Generation Panel, it is not general enough to warrant including in the overall guidelines.

Item III — Input sets for defining variants

This essentially boils down to classing the variants by keyboard layout. Any two variant code points that do not coexist on the same (standard) keyboard layout would not be allowed to co-occur in the same label. This is a potentially useful strategy for reducing allocatable variant labels, and should be considered by the Arabic Generation Panel, it is not general enough to warrant including in the overall guidelines.

Item IV — Step-by-step adoption

Following the procedure, the root is limited to languages that are in widespread common use; which is perhaps not quite the same as the "mature" languages alluded to in the comment. The TLD IDN LGR process does not support languages for which no stable, standardized orthographies exist (or which are effectively not (yet) used for electronic communications by their communities). The procedure requires the panels to consider the risk from any phased approach to LGR developments, lest there are incompatibilities from later additions. There is

also a stated process goal of being free of bias.

The comment does not indicate how the proposal avoids these risks or would not conflict with the stated process goal.

Item V — Study variants across the whole Arabic script

This corresponds with the mandate from the Procedure to reduce risk and increase stability by evaluating the whole problem as far as practical.

Item VI — Variants types

As they are, in essence, free spelling variations, the case of interchangeable variants are an issue that needs discussion in the context of other spelling variations in other scripts. It will fall to the Arabic script Generation Panel to address this issue and to decide whether this is worth addressing by the variant mechanism; it is not general enough to warrant including language about it in the overall guidelines.

Item VIII — International Reachability

The international reachability is an issue that is shared with other scripts, for example, for European not all Latin gTLDs or ccTLDs would normally be directly accessible from all keyboards in the region – users would be expected to be able to type the correct string using alternate means. Based on the analysis of the variants issue study group for Latin, no use of variants was contemplated for this case.

It will fall to the Arabic Generation Panel to recommend the best treatment of this for the Arabic script and to explain their recommendation in the rationale for their LGR proposal.

In response to comments made by Marc Blanchet
Whenever scripts are related and especially if repertoires overlap between Generation Panels, interdependencies among the LGRs will exist. To resolve them requires coordination between the panels. The recommendation is for ICANN to take a more active role in facilitating that coordination. ICANN will review this issue and update the guidelines to reflect any mechanisms added as part of such facilitation.