

SSAC Comment on Examining the User Experience Implications of Active Variant TLDs Report

SAC060

SSAC Comment on Examining the User Experience Implications of Active Variant TLDs Report



A Comment from the ICANN
Security and Stability Advisory Committee (SSAC)
23 July 2013

Preface

This is a Comment to the ICANN Board from the Security and Stability Advisory Committee (SSAC) concerning the report of user experience implications for active variant top-level domains. The SSAC advises the ICANN community and Board on matters relating to the security and integrity of the Internet's naming and address allocation systems. This includes operational matters (e.g., matters pertaining to the correct and reliable operation of the root name system), administrative matters (e.g., matters pertaining to address allocation and Internet number assignment), and registration matters (e.g., matters pertaining to registry and registrar services). SSAC engages in ongoing threat assessment and risk analysis of the Internet naming and address allocation services to assess where the principal threats to stability and security lie, and advises the ICANN community accordingly. The SSAC has no official authority to regulate, enforce, or adjudicate. Those functions belong to others, and the advice offered here should be evaluated on its merits.

A list of the contributors to this Comment, references to SSAC members' biographies and statements of interest, and SSAC members' objections to the findings or recommendations in this Comment are at end of this Comment.

Table of Contents

Executive Summary	4
1. Introduction.....	6
1.1 Conventions and Background	7
1.2 Overview of ICANN’s Root LGR Procedure and User Experience Report.....	7
2. SSAC Recommendations on the Label Generation Rules Procedure	8
2.1 SSAC Comments on the Label Generation Rules Procedure for the Root zone..	8
2.2 SSAC Comment on LGR’s Character Repertoire & Variant Generation Rules..	11
2.3 Comments on LGR’s change process.....	12
3. SSAC Comments on Other User Experience Report Recommendations	14
4. Acknowledgments, Statements of Interests, and Objections and Withdrawals.....	21
4.1 Acknowledgments	21
4.2 Objections and Withdrawals.....	21
4.3 Statements of Interests	22
5. References	22
Appendix A: Summary of Root LGR Procedure.....	24
Appendix B: Summary of Experience Report Recommendations.....	25

Executive Summary

At the request of the Board of Directors of the Internet Corporation for Assigned Names and Numbers (ICANN), the ICANN Security and Stability Advisory Committee (SSAC) provides comments and advice regarding ICANN's Report on Examining the User Experience Implications of Active Variant TLDs.¹

With respect to that Report the SSAC makes the following recommendations:

Recommendation 1: The root zone must use one and only one set of Label Generation Rules (LGR).

Recommendation 2: ICANN must maintain a secure, stable, and objective process to resolve cases in which some members of the community (e.g., an applicant for a TLD) do not agree with the result of the LGR calculations.

Recommendation 3: ICANN should concentrate foremost on the rules for the root zone.

Recommendation 4: ICANN should coordinate and encourage adoption of these rules at the second and higher levels as a starting point by:

- Updating the IDN Implementation Guidelines and recognizing that a modified version of these rules or a review or appeals process must be required to address special cases for the first and second levels;
- Maintaining and publishing a central repository of rules for second-level domain labels (2LDs) for all Top Level Domains (TLDs), encouraging TLD operators to publish their LGRs publicly in the repository maintained by ICANN; and
- Conducting specific training and outreach sessions in cooperation with generic TLD (gTLD) and country code TLD (ccTLD) operators who intend to launch Internationalized Domain Name (IDN) 2LDs or IDN TLDs, with a focus on consistency of user experience. The outreach should include among others registrants, end users, and application developers.

Recommendation 5: Be very conservative with respect to the code points that are permitted in root zone labels.

Recommendation 6: Because the removal of a delegation from the root zone can have significant non-local impact, new rules added to a LGR must, as far as possible, be backward compatible so that new versions of the LGR do not produce results that are incompatible with historical (existent) activations.

¹See Internet Corporation for Assigned Names and Numbers (2013.): Examining the User Experience Implications of Active Variant TLDs at <http://www.icann.org/en/resources/idn/variant-tlds/active-ux-21mar13-en.pdf>.

SSAC Comment on Examining the User Experience Implications of Active Variant TLDs Report

Recommendation 7: Should ICANN decide to implement safeguards, it should distinguish two types of failure modes when a user expects a variant to work, but it is not implemented: denial of service versus misconnection.

Recommendation 8: A process should be developed to activate variants from allocatable variants in LGR.

Recommendation 9: ICANN must ensure that Emergency Back-End Registry Operator (EBERO) providers support variant TLDs, and that parity exists for variant support in all relevant systems and functions associated with new TLD components.

Recommendation 10: The current rights protection regime associated with the Trademark Clearinghouse (TMCH) process is susceptible to homographic attacks. The roles of the involved parties, specifically registrars, registries, and TMCH, related to matching must be made clear.

Recommendation 11: When registries calculate variant sets for use in validation during registration, such calculations must be done against all of the implemented LGRs covering the script in which the label is applied for.

Recommendation 12: The matching algorithm for TMCH must be improved.

Recommendation 13: The TMCH must add support for IDN variant TLDs. Particularly during the TM Claims service, a name registered under a TLD that has allocated variant TLDs should trigger trademark holder notifications for the registration of the name in all of its allocated variant TLDs.

Recommendation 14: ICANN should ensure that the number of strings that are activated is as small as possible.

1. Introduction

Internationalized variant top-level domain (Variant TLD) has been a subject of interest for several years for a number of user communities.² The Internet Corporation for Assigned Names and Numbers (ICANN) Internationalized Domain Name (IDN) Variant TLD Program has been working with subject matter experts in the community to develop solutions to enable the secure and stable delegation of IDN variant TLDs. The Program has recently concluded the work on two key components of the solution: *The Procedure to Develop and Maintain the Label Generation Rules for the Root Zone in Respect of IDNA Labels*³ and the *Report on User Experience Implications of Active Variant TLDs*.⁴

On 11 April 2013 the ICANN Board of Directors passed a resolution directing staff to implement the procedure to develop and maintain the label generation rules for the root zone in respect of IDNA labels, as well as requesting that interested ICANN supporting organizations and advisory committees provide staff with any input and guidance they may have to be factored into implementation of the Recommendations in the user experience report.⁵ This Comment constitutes the response from ICANN's Security and Stability Advisory Committee (SSAC) to the Board's request.

The SSAC has reviewed ICANN's report *Examining the User Experience Implications of Active Variant TLDs* and provides comments on recommendations specifically concerning security and stability issues. Furthermore, the SSAC regards the Procedure to Develop and Maintain the Label Generation Rules for the root zone (hereinafter referred as "LGR procedure"), as it relates to the general discussion of variants and the user experience report, of critical importance to the security and stability of the root zone. Therefore, in addition to commenting on the user experience report, SSAC also reviewed the LGR procedure and provides its comments on that report in this Comment document.

This Comment is organized as follows: In the introduction the SSAC defines the key terminology used in this Comment and provides a summary of the ICANN reports on

²See Frakes, J, et al., "Considerations in the use of the Latin script in variant internationalized top-level domains: Final report of the ICANN VIP Study Group for the Latin script", October 2011; Govind, et al., "Devanāgarī VIP Team Issues Report", October 2011; Hussain, S, et al., "Internationalized Domain Names Variant Issues Project Arabic Case Study Team Issues Report", October 2011; Lee, X., et al., "Report on Chinese Variants in Internationalized Top-Level Domains, October 2011; Segredakis, V., et al., "Study of the Issues Present in the Registration of IDN TLDs in Greek Characters", October 2011; Sozonov, A., et al., "IDN Variant TLDs – Cyrillic Script Issues", October 2011
<http://www.icann.org/en/resources/idn/variant-tlds/reports>.

³See ICANN (2013.): Procedure to Develop and Maintain the Label Generation Rules for the Root Zone in Respect of IDNA Labels (Root LGR Procedure) at: <http://www.icann.org/en/resources/idn/variant-tlds/draft-lgr-procedure-20mar13-en.pdf>

⁴See ICANN (2013.): Examining the User Experience Implications of Active Variant TLDs. Available at: <http://www.icann.org/en/resources/idn/variant-tlds/active-ux-21mar13-en.pdf>.

⁵See ICANN Board of Directors (ICANN Board) (2013): Resolutions on IDN Variant TLD Root LGR procedure and User Experience Study Recommendations at: <http://www.icann.org/en/groups/board/documents/resolutions-11apr13-en.htm#2.a>.

which the SSAC is commenting. Section 2 focuses on the SSAC's comments on the root zone LGR procedure and is further divided as follows: the SSAC's comments on 1) the LGR procedure in general; 2) on the allowable code points and variant generation rules; and 3) on the LGR change process. Finally, Section 3 focuses on the SSAC comments related to the rest of the user experience report recommendations.

1.1 Conventions and Background

This Comment uses the following conventions from the ICANN variant integrated issues report, and the proposed procedure to Develop and Maintain the Label Generation Rules for the Domain Name System (DNS) Root Zone in Respect of IDNA Labels.⁶

IDN Variant -An IDN variant is an alternate code point (or sequence of code points) that could be substituted for a code point (or sequence of code points) in a candidate label to create a variant label that is considered the “same” in some measure by a given community of Internet users.⁷ There is no general agreement of what that sameness requires. Section 3 of the ICANN's Variant Integrated issues report provides a fuller category of variants. It should be noted that ICANN makes a distinction between *code point variants*, where a single character is in some way closely related to – or likely to be confused with – an alternative, and *whole-string variants*, where the token at issue is longer than a character, and may be a morpheme, a full word, or even a phrase, or some meaningful element in a language that uses the script. The current version of the root LGR procedure and user experience documents focus on the code point variants, not the whole-string variants.

When the term “variant” is used in this Comment it includes, if not explicitly written, all kinds of variants, including those that are allocatable, blocked and transitive.

1.2 Overview of ICANN's Root LGR Procedure and User Experience Report

The document *Procedure to Develop and Maintain the Label Generation Rules for the Root Zone in Respect to IDNA Labels* (hereinafter referred as the root zone LGR procedure) describes a mechanism for creating and maintaining the rules with respect to IDN labels for the root zone. This mechanism can be used to determine which Unicode code points are permitted for use in U-labels in the root zone, what variants (if any) are possible to allocate in the root zone, and what variants (if any) are automatically blocked.

The root zone LGR procedure uses two classes of panels to make the determinations: The Generation Panels and the Integration Panel. The Generation Panels are community-based panels that have the task to propose the LGR for the particular script used in each

⁶See ICANN (2012): A Study of Issues Related to the Management of IDN Variant TLDs (Integrated issues report or IIR) at: <http://www.icann.org/en/news/announcements/announcement-20feb12-en.htm>.

⁷See section 1.2 of the Integrated issues report.

SSAC Comment on Examining the User Experience Implications of Active Variant TLDs Report

community. The SSAC specifically thinks it is important that LGR related to a particular script do take into account all uses of that script in all languages where it is in use. While their members require expertise in areas such as linguistics, Internationalizing Domain Names in Applications (IDNA), or DNS, the key feature of the Generation Panels is their make-up of volunteers from the community. The output of the Generation Panels is collected, reviewed, and integrated by the so-called “Integration Panel,” a small panel composed of general experts. The output of the Integration Panel’s work is the label generation rules for the root zone. More information about the procedure can be found in Appendix A.

ICANN’s user experience report 1) summarizes and compares, from a user experience and registry management perspective, variant practices in several ccTLD registries; 2) proposes a set of guiding principles to define an acceptable user experience; 3) identifies how various users communities (e.g., end users, system/network administrators, application developers, registrants, registrars and registries) will be impacted by active variant top-level domains; and 4) based on 1, 2, and 3, proposes 29 recommendations to ensure the security, stability and acceptable user experience for active variant top-level internationalized domains. More information about the recommendations can be found in Appendix B.

2. SSAC Recommendations on the Label Generation Rules Procedure

2.1 SSAC Comments on the Label Generation Rules Procedure for the Root zone.

Recommendation 1: The root zone must use one and only one set of rules for the Root LGR procedure.

Should the root zone have one set of LGR as proposed by the LGR procedure document or should it adopt multiple rules submitted by the applicants as practiced today for the second level?

The SSAC asserts that it should be the former, and would like to reinforce the conclusions reached in the LGR document and user experience report with the following additional rationale:

- The root zone is necessarily shared by everyone on the Internet, and needs a set of LGR that ensures minimal conflict, minimal risk to all users (independent of the language or script they are using and independent of gTLD or ccTLD), and minimal potential for incompatible change over time.
- At the second level, IDN policies can be made based on languages of the community that a registry serves. The root zone, however, does not have any

SSAC Comment on Examining the User Experience Implications of Active Variant TLDs Report

linguistic context. The root zone also lacks other contexts that can be used by a registry to restrict LGR for that particular TLD.⁸

- Today, at the second level, each registry has its own label generation rules for the zones it manages. Sometimes different rules for the same script exist across TLDs. Applying such a model to the root zone would cause stability issues.

Recommendation 2: ICANN must maintain a secure, stable, and objective process to resolve cases where some members of the community (e.g. an applicant for a TLD) do not agree with the result of the LGR calculations.

Although the LGR is defined by experts, it is not fault-proof. It is possible that some portion of the community will disagree with the LGR calculation on certain strings. Such disagreement will likely occur *not* during the formation of the LGR, but in the actual application of the LGR to a given string.

When such a situation arises, it is important to have a secure, stable, and objective process to resolve the issues. Lack of such a process could lead to disagreements between ICANN and a TLD applicant or other community member and could put ICANN under pressure to change certain rules to satisfy the demands of particular parties, and likely lead to ad hoc decisions that would harm the stability of LGR and the root zone.

The spirit of such a process should be to balance (1) the desire/need from certain communities to change certain LGR rules (e.g., additional code points, variant rules, etc.) with (2) the risk that changing the LGR rules without proper review could introduce instability. The SSAC proposes the following straw man process to ICANN for consideration:

- If an applicant or a community disagrees with the LGR calculation, they can appeal to ICANN to reconvene the relevant generation panel to update the LGR. Such an appeal should not focus on the given string, but rather on the LGR as a whole, for a script or a set of scripts.
- Once ICANN receives the request, it should reconvene the relevant label generation panel. The label generation panel should conduct its review by following its *regular* process outlined in the LGR document to consider the specific case (see Section B 2.1.1 of the Root LGR Procedure). Such a process must also undergo integration panel checking and public comment, as per the process already defined.

⁸The context described here in theory could be any kind of policy linked with the TLD string itself.

Recommendation 3: ICANN should concentrate foremost on the rules for the root zone.

The user experience report made the arguments that the LGR used by the root should be used at the second level by default and that deviations from the LGR at the root should be documented and justified. The SSAC agrees in principle with the recommendation but wishes to point out that the root zone is a special case and the approach taken to variant management in the root need not prescribe the approach taken by individual TLD registries. Formulation of TLD registry policy often takes into account specific user context and there may be corresponding reasons for different criteria or rules relating to variant label generation and use, subject to certain minimum requirements necessary for security or stability reasons.

Thus, the SSAC asserts that TLD registry operators should not have an automatic obligation to abide by all the same variant tables and policies used at the root level of the DNS and ICANN should first and foremost concentrate on the rules for the root zone.

Recommendation 4: ICANN should coordinate and encourage adoption of these rules at the second and higher levels for all TLDs as a starting point by:

- Updating the IDN Implementation Guidelines⁹ and recognizing that a modified version of these rules or a review or appeals process must be required to address special cases for the first and second levels;
- Maintaining and publishing a central repository of rules for second level domains (2LDs) for all TLDs, encouraging TLD operators to publish their LGRs publicly in the repository maintained by ICANN; and
- Conducting specific training and outreach sessions in cooperation with gTLD and ccTLD operators who intend to launch IDN 2LDs or IDN TLDs, with a focus on consistency of user experience. The outreach should include among others registrants, end users and application developers.

Although deviations are expected, registry operators should minimize those deviations to ensure consistent user experience across different TLDs. This is because:

- From a security perspective, if two zones have inconsistent allocatable variant code point sets (i.e., two allocatable code points are variants in one zone but are not variants in the other), it could lead to a situation in which a label resolves as a variant in one zone but resolves to a different registrant in the other; such differences in LGRs can be exploited by miscreants or cause other unpredictable experiences.

⁹See ICANN (2011). Guidelines for Implementation of IDNs (version 3.0) at: <http://www.icann.org/en/resources/idn/implementation-guidelines>.

SSAC Comment on Examining the User Experience Implications of Active Variant TLDs Report

- From a usability perspective, if two zones have the same variant code point sets but with different statuses (allocatable versus blocked), it could lead to a situation in which a label resolves as variant for one zone and does not resolve for the other and would make user conclude that the “internet is broken.”

One example of security related confusability that can happen across TLDs and between TLDs and 2LDs is the already existing TLD عمان (xn--mgb9awbf). This is the IDN ccTLD for Oman. In Arabic some characters can be omitted when writing strings, and in this case the character U+064F is not included. The complete string including the character Arabic Dammah should have been عُمان. At the same time if we look at the string for the capital of Jordan, Amman, which is عمان, it can also be written as عمان if the character U+064E (Arabic Fatah) is not explicitly spelled out. This is one example where one could draw the conclusion that all three strings عمان, عُمان and عمان should be included in a variant set. As already explained, we have the string عمان allocated as a TLD for the country عُمان, and further it is important that these three strings are treated as variants in the same set in all TLDs that accept 2LDs in the Arabic script.

It should be mentioned that the variants are not created by only adding or removing a codepoint to/from the original string, as Oman is written as U+0639 **U+064F** U+0645 U+0627 U+0646 while Amman is U+0639 U+0645 **U+064E** U+0627 U+0646. Note that the position of the omitted character (in bold) is different in the two strings.

Finally, both gTLDs and ccTLDs co-exist in the root zone and are relied upon by Internet users around the world. While gTLDs and ccTLDs may involve different operating environments, it is critical that a reliable user experience is produced across the TLD space. Thus, any label generation rules for TLDs will need to be adhered to by both ccTLDs and gTLDs, and to promote consistency, every TLD should maintain and publish to a central ICANN-maintained openly accessible repository rules for second level domains. This includes today’s IDN tables hosted by the Internet Assigned Numbers Authority (IANA). See also Recommendation 10a below regarding TMCH implications with different rules for variant set calculation.

2.2 SSAC Comment on LGR’s Character Repertoire & Variant Generation Rules

Recommendation 5: Be very conservative on code points allowed in the root zone.

The SSAC agrees with the approach in the root LGR procedure document as well as the recommendations in the user experience report. The SSAC would like to reiterate that the inclusion-based approach is preferable over exclusion-based approach for the root zone:

- Inclusion-based approach (preferred): Start with the current restricted Letter, Digit, Hyphen (LDH) American Standard Code for Information Interchange

SSAC Comment on Examining the User Experience Implications of Active Variant TLDs Report

(ASCII) characters (a-z, A-Z, 0-9, -) and then extend it to include relevant, non-problematic¹⁰ "international" characters.

- Exclusion-based approach: Start with the entire Unicode set, and eliminate only characters that can be explicitly demonstrated as being harmful.

The fundamental advantage of the "inclusion-based" model is that it is far easier to restrict something initially and then later relax the restriction than it is to permit something and then later attempt to remove it from use. With respect to the root zone, the inclusion approach should be preferred because it is very difficult to un-delegate a TLD once it is permitted.

In addition, the following guidance given by the user experience report must be provided to both the generation and integration panels:

- The code points allowed for the LGR must include only those minimally needed by a particular script community. For example, the repertoire should not include dead scripts and code points representing archaic characters that are rarely used by a script.
- If the community cannot agree on the need of a code point, the default decision must be to not include it in the repertoire until an agreement is reached.
- Any code point that is optionally written in a script (e.g., some combining marks) must not be included.
- There must be an explicit description and justification for inclusion of each code point that causes a variant, whether directly or in combination with other code point(s), by the community developing the LGR.

Finally, even though the code point variants may be identified at the language level, the root LGR should be the union of all the language level rules. The root LGR should apply at the script level despite the fact that this could generate extra variants in order to promote consistency of use across global end users.

2.3 Comments on LGR's change process

Recommendation 6: Because the implications of removing delegations from the root zone can have significant non-local impact, new rules added to LGR must, as far as possible, be backward compatible so that new versions of the LGR do not produce incompatible results with historical (existent) activations.

It is expected that the LGR for the root zone will be subject to modification from time to

¹⁰Also the term "non-problematic" should be treated in a conservative way. Example of non-problematic characters could be those that do not change with any of the normalization forms (such as NFC, NFD, NFKC and NFKD).

SSAC Comment on Examining the User Experience Implications of Active Variant TLDs Report

time. However, the SSAC recommends that for stability considerations such changes must be based on a defined set of criteria and modifications must be undertaken with the utmost care as it is likely that TLD implementations based on a prior version of the LGR might otherwise become unstable.

At a high level, changes to the LGR can be in the following categories, ordered from least harmful to most harmful to security and stability.

Adding new code points to LGR:

1. Add a code point with [blocked] status which is not considered in an earlier version of LGR and which does not add any variants (e.g. due to its addition in Unicode standard);
2. Add a code point with [allocatable] status which is not considered in an earlier version of LGR and which does not add any variants (e.g. due to its addition in Unicode standard);
3. Add a code point with [blocked] status which is not considered in an earlier version of LGR and which is a variant of an existing code point;
4. Add a code point with [allocatable] status which is not considered in an earlier version of LGR and which is a variant of an existing code point (e.g. due to its addition in Unicode standard).

Changing status of existing code points, causing possible stability issues:

5. Change [blocked] status to [allocatable] status for a code point which is in an earlier version of LGR and which does not add any variants;
6. Change [blocked] status to [allocatable] status for a code point which is in an earlier version of LGR and which is a variant of an existing code point;
7. Move an [allocatable] code point which does not have [allocatable] variant code points to [blocked] status;
8. Move an [allocatable] code point that has [allocatable] variant code points to [blocked] status.

Changing status of existing code points, causing possible security and stability issues:

9. Make two [allocatable] code points variants of each other, where they were not variants earlier and had no variants;
10. Make two [allocatable] code points variants of each other, where they were not variants earlier and had other variants;
11. Change two [allocatable] variant code points to become [allocatable] code points which are not variants anymore and do not have other variants;
12. Change two [allocatable] variant code points to become [allocatable] code points which are not variants anymore and have other variants which will have to be re-grouped with these two (now) non-variant code points.

SSAC Comment on Examining the User Experience Implications of Active Variant TLDs Report

Because the LGR procedure adheres to the principles articulated in the Internet Architecture Board Document RFC6912, “Principles for Unicode Code Point Inclusion in Labels in the DNS”, the chances of cases 7-12 occurring should be limited. Nevertheless ICANN should consider implementing safeguards should these rare cases occur.

When making changes in the LGR, some types of changes (especially 7-12 outlined above) could *adversely impact already allocated and delegated variant labels at TLD and other levels*. Thus, new rules added to LGR should be backward compatible so that new versions of the LGR do not produce incompatible results with historical (existent) activations. The SSAC considers this issue of critical importance.

Recommendation 7: Should ICANN decide to implement safeguards it should seek to distinguish the following two types of failure modes when a user expects a variant to work but it is not implemented:

- Denial of service: the user attempts to visit `http://example.Y`, reading it as being the same Uniform Resource Identifier (URI) as the `http://example.X` that, for example, he or she saw in an advertisement, but the connection does not work (lookup fails) because Y is either blocked, withheld, or X has no variant at all, and `example.Y` is not registered.
- Misconnection: the user attempts to visit `http://example.Y`, reading it as being the same URI as the `http://example.X` that, for example, he or she saw in an advertisement, but arrives at a site controlled by a registrant different to that of `example.X`.

The second case is much more dangerous than the first one. In the first case, the user is frustrated and may conclude that “the Internet does not work,” but no serious harm has arisen. The second case is problematic even if this effect is not the result of malicious work on the part of Y’s operator or `example.Y` registrant. Misconnections to a perfectly legitimate site operating at `example.Y` present issues of possible credential compromise or other accidental disclosure of information in addition to user confusion and frustration.

3. SSAC Comments on Other User Experience Report Recommendations

Recommendation 8: A conservative process needs to be developed to activate variants from allocatable variants in LGR.

Based on the SSAC’s understanding, given the following LGR calculation:

$$\text{LGR}(\text{string}) \rightarrow \text{string1}\{\text{state1}\}, \text{string2}\{\text{state2}\}, \dots, \text{stringN}\{\text{stateN}\}$$

Where `state1`, `state2`, ..., `stateN` is one of the two possible states: allocatable or blocked. A string that is allocatable does not imply automatic activation; rather that it can be allocated. If the string is allocated it is done so "in sync" with the base string that was the

SSAC Comment on Examining the User Experience Implications of Active Variant TLDs Report

input to the LGR. As it is ICANN's role to stipulate this policy, a clear process needs to be developed to avoid ad hoc treatment of new gTLD applications.

The user experience report recommends that ICANN must implement a well-defined and conservative variant TLD allocation process. The SSAC agrees with the recommendations below:

- The approval of a variant TLD must not be automatic, but initiated upon the request of a TLD applicant, explicitly specifying (1) the variant label; (2) the status for which the variant should be evaluated (activated, allocated but not activated, etc.); and (3) the need for the variant (e.g., motivated by linguistic, security, usability and/or other considerations). Unless such an application is initiated, all variants generated against a primary TLD application by the root LGR should remain withheld (and un-allocated).
- TLD variant(s) must be applied for by and allocated¹¹ to the same entity or registry that has applied for the corresponding primary TLD label.
- All requirements for a TLD application approval process also apply to the approval of a variant TLD. These include, for example, requirements for ICANN Governmental Advisory Committee (GAC) and public comments on the label, string similarity evaluation and dispute processes, DNS stability evaluation of the variant TLD label, etc. ICANN must document this process associated with all aspects of variants within the new gTLD ecosystem. The process is needed as the variant in a language may be interpreted as a unique and different label in another language for the same script.
- The registry delegation and re-delegation processes must be extended to include activated variants of a TLD. The registry contract must be updated accordingly.
- The registry fail-over plan should be extended to include activated variants of a TLD. The relevant registry contract must be updated accordingly.

Recommendation 9: ICANN must ensure that EBERO providers support variant TLDs, and that parity exists for variant support in all relevant systems and functions associated with new TLD components.

To maintain the consistency of user experience when a registry is in transition to another operator through the EBERO process, the SSAC recommends that all dispositions /

¹¹As defined in ICANN's integrated variant Issues report. Allocated the term "allocated" refers to a status of some label with respect to a zone, whereby the label is associated administratively to some entity that has requested the label. This term (and its cognates "allocation" and "to allocate") represents the first step on the way to delegation in the DNS. When the registry (zone operator) allocates the label, it is effectively making a label a candidate for activation. Allocation does not, however, affect the DNS at all.

policies of the variant set from the originating registry should be applied to the new registry. The LGRs employed by the source registry must be mirrored as well as registry contents, rendered disposition.

Recommendation 10: In the current design of rights protection related to the TMCH process there is a risk of homographic attacks. The roles of the involved parties, specifically registrars, registries and TMCH, related to matching must be made clear.

From a security and operations perspective, domain names that contain variants of a mark must be protected during the new gTLD sunrise and claims period. Without such protections, miscreants could register variant domain names of a trademark for nefarious purposes.

The ICANN new gTLD program uses a TMCH to be the global repository of existing legal rights for trademarks. The repository is used by registries and registrars for 1) validating eligibility for sunrise processes; 2) notifying applicants for domain names during the trademark claims period of existing legal rights on a particular string; and 3) notifying trademark holders during sunrise and trademark claims period.

To achieve such protections, in principle there are two ways to handle variants and TMCH:

- 1) Variant *calculation* at the registry level, and *checking* TMCH for the existences of marks for variants in the calculated variant set.
- 2) Variant calculation and checking inside the TMCH in addition to the already defined matching algorithm TMCH uses.¹²

The first approach is currently that which is used by the TMCH (see though comment below about the matching algorithm TMCH uses); it relies on the registry to apply LGR for the second level to achieve the desired protections. How this works as currently designed is described below. Supposing that trademark holder submits the mark “example” to the TMCH and that it has a variant:

Sunrise Period¹³

1. A trademark holder submits mark “example” to the TMCH. The trademark information is verified and a signed mark data file is generated indicating that the rights holder is eligible to register the domain name <example.tld> based on those rights.
2. A registry uses an IDN table where “è” is a variant character to “e” and hence

¹²See <http://newgtlds.icann.org/en/about/trademark-clearinghouse/matching-rules-24sep12-en.pdf>.

¹³See <http://newgtlds.icann.org/sites/default/files/main-images/tmch-sunrise-process-1000x898-17jun13-en.png>.

SSAC Comment on Examining the User Experience Implications of Active Variant TLDs Report

“èxample” is a variant label of “example” in the registry. When presented with the signed mark data file (SMD) for registration of <example.tld> the registry has the ability either to: a) block registration of <èxample.tld> or b) allocate <èxample.tld> to the registrant of <example.tld>, according to the registry’s policy.

Trademark Claims Period¹⁴

1. A registrant attempts to register <èxample.tld>.
2. If such registration is not blocked by registry policy, the registry applies its variant rules generating the corresponding variant names, e.g., <example.tld>.
3. For those names that will be ultimately registered, the registry is required to query the list of labels subject to Claims, i.e., it will query both <example> and <èxample> to determine whether there is a match.
4. Since the mark “example” is in the Clearinghouse, a match will be found, and the registry will provide the indicator to the registrar that a Claims notice should be shown to the registrant.
5. A Claims notice is shown to the registrant based on the mark “example” and the registrant can elect whether to proceed.

The benefit of this approach is that it is flexible because:

- The role of the TMCH is to record existing rights, and not make determinations concerning the scope of particular rights and whether certain (variant) strings qualify for the same right.
- There is work in progress in ICANN to establish a set of IDN LGR for the root zone. However, this is designed for a particular purpose (IDNA labels for the root zone) and caution is warranted when applying those rules to other uses (trademark). Whether such rules are appropriate for use by the TMCH should be explored further as the rules could conflict with practices of local law or established registry IDN tables.

The downside of this approach is that:

- Registries could have different IDN tables, even for the same script. Thus it is possible that miscreants can use the different rules to generate and register variants in other TLDs and cause security, stability, or resiliency concerns or result in squatting and other related issues.

¹⁴See <http://newgtlds.icann.org/sites/default/files/main-images/tmch-claims-process-1300x1516-17jun13-en.png>.

SSAC Comment on Examining the User Experience Implications of Active Variant TLDs Report

- In the case that a registry does not have variant handling policy it is very likely that a miscreant would generate and register variants to facilitate inappropriate actions.
- In the case of certain IDN tables, the number of entries in a variant set can be very high and that creates a large number of transactions between the registry and TMCH.

Thus if ICANN continues the current approach for handling variants in the TMCH, the downside mentioned above must be mitigated. Specifically, the SSAC recommends that ICANN should ensure that gTLD operators perform variant generation and check all strings in a variant set, and that this set should be treated as an atomic set. These issues are further explored with Recommendation 11 below.

Another approach for supporting variants is variant calculation and checking inside the TMCH. In this model:

- The TMCH adds support functionalities for IDN variants and define matches not only on what today is defined as “identical match” with a registered string but also if a variant of the registered string matches.
- A name registered that has variants will trigger trademark holder notifications for the registration of the name or its variants.

The benefit of this approach is that it addresses the potential downsides for the previous approach. Centralizing variant generation and checking would bring consistency to the variants generated.

The downside of this approach is that:

- In this case, the role of the TMCH changes from recording existing rights to making determinations concerning the scope of particular rights. Similar to the definition of “identical match” this case already exists today for the TMCH.
- The LGR designed for the particular purpose (IDNA label for the root zone) would be applied for trademarks. This could conflict with practices of local law or established registry IDN tables.
- Generating variants in the TMCH may not be sufficient to address the security concerns related to malicious registrations of names in periods outside of the trademark claims timeframe.

If the ICANN community reaches the conclusion to adopt the second model, the downsides in the approach described above must be addressed.

In both cases, any changes in LGRs for root and second level for all the registries must be timely communicated to ICANN and other relevant parties for variant calculations.

Recommendation 11: When registries calculate variant sets for use in validation during registrations, such calculations must be done against all the implemented LGRs covering that script for which the label is applied.

As practiced today, different registries could use different LGRs for Second Level Domain (SLD) registration for the same script.¹⁵ Thus it is possible that non-overlapping sets of variants are generated (e.g. one registry generates for the second level domain A variants A1, A2, A3 in tldA and another registry generates variants A2, A3, A4 in tldB). In such cases, a miscreant could exploit such a non-overlapping set and cause harm.

The long-term solution is that, if possible, different registries should use consistent LGRs for the SLDs for a given script. In the short run, or given that might not be possible in all cases, when registries calculate variant sets for use in validation during registrations, such calculations must be done against all the implemented LGRs covering that script in which the label is applied for (so both registries tldA and tldB match registrations against the joint set A1, A2, A3 and A4).

Registries have an obligation to perform these comprehensive checks, because they are the last resort in ensuring that these security problems do not happen.

Recommendation 12: The matching algorithm for TMCH should be improved.

Regardless of which one of the strategies is chosen it should be noted that “identical match” as defined by TMCH is not really an identical match as in “bit-by-bit” or “character-by-character comparison” as a transformation stage is included before the actual matching. From a technical standpoint, the transformation stage currently as specified from is unclear and does not take non-ASCII based scripts into account. It should be improved in at least the following areas:

- Include a clear specification that, as specified in IDNA 2008, Unicode Normalization form C (NFC) is to be used on the strings before character-by-character comparison is made.
- Who is responsible for applying the normalization and transformation applied to both the strings registered in TMCH and the strings used for searching.
- Instead of referencing specific non-alphabetic characters that can be removed before matching is made reference the various classes of characters that are defined by the Unicode Consortium.
- Explicitly reference the various derived property values as defined by IDNA2008 (RFC 6452, RFC 5892) and make clear whether the property value

¹⁵For example, there are different tables for ایران (Iran ccTLD), السعودية (Saudi ccTLD) for Arabic script.

SSAC Comment on Examining the User Experience Implications of Active Variant TLDs Report

has impact on if the character is to be transformed or not (for example by being removed before match).

- Include a formal definition of acceptable strings, with the help of for example Augmented BNF for Syntax Specifications (ABNF) as defined by the Internet Engineering Task Force (IETF) in Request for Comment (RFC) 5234.

Recommendation 13: The TMCH must add support for IDN variant TLDs. Particularly during the Trademark Claims service a name registered under a TLD that has variant TLDs should trigger trademark holder notifications for the registration of the name in the TLD and all its allocated variant TLDs.

Currently the TMCH process only considers variants at the SLD level; it is yet to consider variants at the top level (e.g. <example.tld1> and <example.tld2>, where tld1 and tld2 are variants). In this example if an allocated and activated variant TLD's 2LD label results in a hit in the TMCH, then such registration must be reported to the trademark holders for the label. If ICANN approves and delegates variant TLDs, it is important that the TMCH must support such a capability. The SSAC recommends that during the Trademark Claims service a name registered under a TLD that has variant TLDs should trigger trademark holder notifications for the registration of the name in the TLD and all its allocated variant TLDs.

Recommendation 14: ICANN should ensure that the number of strings that are activated is conservative.

Variants introduce a permutation issue both at the top level as well as with combinations of top level and second level:

- At the TLD level, assume a TLD string with four characters, where each character has three variants. Thus the variant set created would be $3^4 = 81$ different strings. The size of the variant sets can grow exponentially.
- At 2LD level, assume a 2LD string with four characters, where each character has three variants, and the same number for top level. Thus the variant set created would be $3^4 \times 3^4 = 72171$.

Such large number of variant strings presents challenges for the management of variant domains at the registry, the registrar and registrant levels. We have seen that some registries have imposed additional rules for variants. One such rule is “no mixing,” e.g., if there are two categories and the characters in those categories are {A,a}, {B,b} and {C,c} respectively, then the variant set we have for the string ABC would be ABC, ABc, AbC, Abc, aBC, aBc and abc, but only ABC and abc are non-mixed versions of the string. Conservatism is also to be used in this case for the root as well.

SSAC Comment on Examining the User Experience Implications of Active Variant TLDs Report

The user experience report recommended that ICANN must implement a conservative variant TLD allocation process. The SSAC agrees with the recommendations below:

- A variant TLD application must be accepted only if the TLD applicant clearly demonstrates the necessity for activating the string. Variants that are not necessary, but are desired, must not be allocated and activated.

4. Acknowledgments, Statements of Interests, and Objections and Withdrawals

In the interest of greater transparency, these sections provide information on three aspects of the SSAC publication process. The Acknowledgments section lists the members who contributed to this particular document. The Statements of Interest section points to the biographies of the Committee members and any conflicts of interest, real, apparent or potential, that may bear on the material in this document. The Objections and Withdrawals section provides a place for individual members to disagree with the content of this document or the process for preparing it.

4.1 Acknowledgments

The committee wishes to thank the following SSAC members and external experts for their time, contributions, and review in producing this Comment.

SSAC members

Bruce Tonkin
Danny McPherson
Lyman Chapin
Patrik Fältström
Ram Mohan
Sarmad Hussain

ICANN staff

Steve Sheng (editor)
Francisco Arias

4.2 Objections and Withdrawals

There were no objections or withdrawals.

4.3 Statements of Interests

SSAC member biographical information is available at:

<http://www.icann.org/en/groups/ssac/biographies-04jun13-en.htm>.

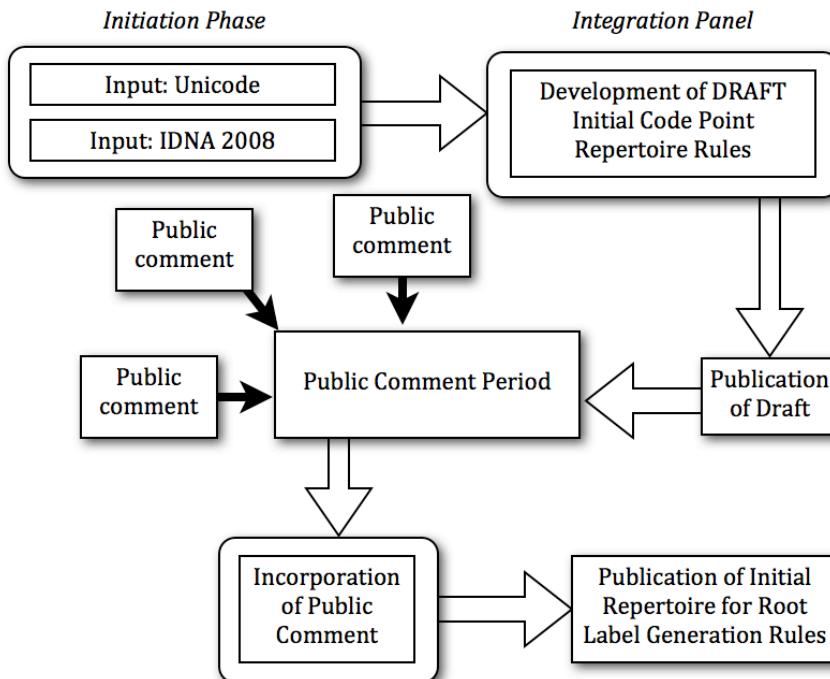
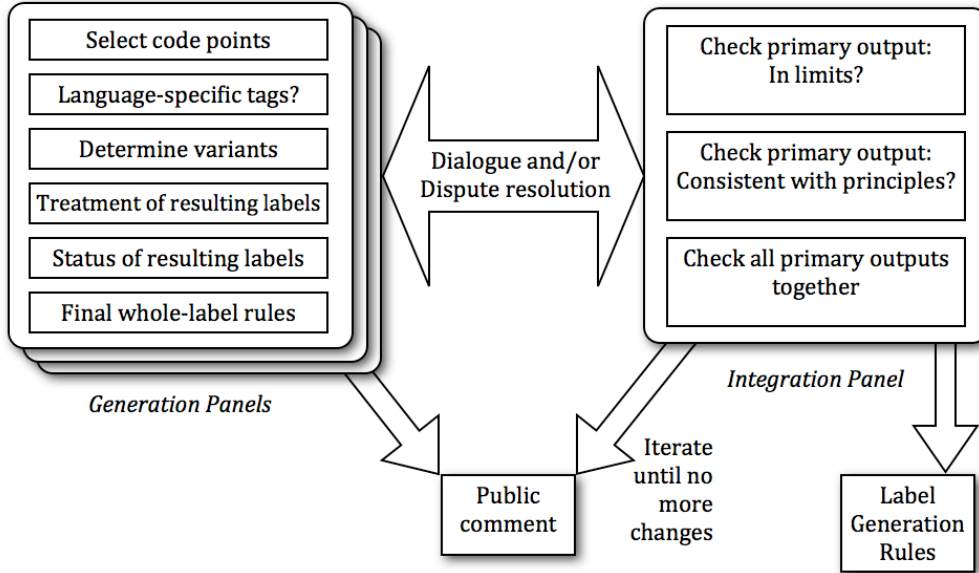
5. References

1. Frakes, J, *et al.*, “Considerations in the use of the Latin script in variant internationalized top-level domains: Final report of the ICANN VIP Study Group for the Latin script,” (Marina del Rey, California: ICANN, October 2011). <http://archive.icann.org/en/topics/new-gtlds/latin-vip-issues-report-07oct11-en.pdf>.
2. Govind, et al., “Devanāgarī VIP Team Issues Report,” (Marina del Rey, California: ICANN, October 2011). <http://archive.icann.org/en/topics/new-gtlds/devanagari-vip-issues-report-03oct11-en.pdf>.
3. Hussain, S, *et al.*, “Internationalized Domain Names Variant Issues Project Arabic Case Study Team Issues Report,” (Marina del Rey, California: ICANN, October 2011). <http://archive.icann.org/en/topics/new-gtlds/arabic-vip-issues-report-07oct11-en.pdf>.
4. Lee, X., *et al.*, “Report on Chinese Variants in Internationalized Top-Level Domains,” (Marina del Rey, California: ICANN, October 2011). <http://archive.icann.org/en/topics/new-gtlds/chinese-vip-issues-report-03oct11-en.pdf>.
5. Segredakis, V., et al., “Study of the Issues Present in the Registration of IDN TLDs in Greek Characters,” (Marina del Rey, California: ICANN, October 2011). <http://archive.icann.org/en/topics/new-gtlds/greek-vip-issues-report-07oct11-en.pdf>.
6. Sozonov, A., et al., “IDN Variant TLDs – Cyrillic Script Issues,” (Marina del Rey, California: ICANN, October 2011). <http://archive.icann.org/en/topics/new-gtlds/cyrillic-vip-issues-report-06oct11-en.pdf>.
7. ICANN Board of Directors (ICANN Board) (2013). Resolutions on IDN Variant TLD Root LGR procedure and User Experience Study Recommendations Available at: <http://www.icann.org/en/groups/board/documents/resolutions-11apr13-en.htm#2.a>.
8. ICANN (2012). A Study of Issues Related to the Management of IDN Variant TLDs (Integrated issues report or IIR). Available at: <http://www.icann.org/en/news/announcements/announcement-20feb12-en.htm>
9. ICANN (2013). Procedure to Develop and Maintain the Label Generation Rules for the Root Zone in Respect of IDNA Labels (Root LGR Procedure). Available at: <http://www.icann.org/en/resources/idn/variant-tlds/draft-lgr-procedure-20mar13-en.pdf>

SSAC Comment on Examining the User Experience Implications of Active Variant
TLDs Report

10. ICANN (2013). Examining the User Experience Implications of Active Variant
TLDs. Available at: [http://www.icann.org/en/resources/idn/variant-tlds/active-ux-
21mar13-en.pdf](http://www.icann.org/en/resources/idn/variant-tlds/active-ux-21mar13-en.pdf)

Appendix A: Summary of Root LGR Procedure



Appendix B: Summary of Experience Report Recommendations

User Experience Report Recommendations		
ICANN	MUST	ICANN must implement a well defined and conservative variant TLD allocation process (Root)
		ICANN must maintain an LGR repository for the root zone and IDN TLDs and make it available to users and programmatically processable
		ICANN must develop, to the extent possible, minimal, simple and consistent LGR for the root zone
		To help ensure that users have a more predictable and consistent experience registering and using primary and variant labels, ICANN must develop, to the extent possible, a minimal, simple and consistent life cycle for the variant TLD sets (across languages and scripts)
		ICANN must define guidelines to evaluate the competence and readiness of the registry to manage variants, to ensure a stable and secure end user experience
		ICANN must require any accredited registrar that supports IDNs with TLD and/or SLD variants to support variants across its registration platform
		ICANN must convene relevant experts involved in domain name disputes to determine any new issues introduced by variants and update existing dispute resolution processes accordingly
		ICANN must define technical requirements and engage with standards organizations, such as the IETF, to determine how the IDN variants should be consistently implemented
	SHOULD	ICANN should require IDN TLD registries with variants to apply the relevant (script) subset of the root zone LGR and state life cycle for variants across second-level domain labels. Deviations should be justified
		ICANN should create educational materials on the use and impact of variants for different user communities
ICANN should develop consistent registration data requirements for variants at root and other levels		
Registry	MUST	Registry that supports variants must make its updated LGR available to ICANN and the community
	SHOULD	Registry should not register any second-level variant labels unless the label registration request has met all approval requirements
		Registry that supports variants should apply the LGR developed for the root across lower-level domains. Deviations from the LGR should be publicly documented and justified
		Registry that supports variants should implement, to the extent possible, state life cycle for the second-level variant recommended by ICANN
		Registry should create educational materials on the use and impacts of variants for different user communities, such as end users, system administrators, etc.
Registry offering variants should require relevant registrars to support IDN variants across their		

SSAC Comment on Examining the User Experience Implications of Active Variant TLDs Report

		registration platforms
Registrar	MUST	Registrar must update its practice to address requirements specific to the registration of IDN variants
		Registrar must support IDN variants across their registration platforms
		Registrar must support registry policies and associated services for collecting and managing registration data of IDN variants
	SHOULD	Registrar should extend linguistic and technical support of IDN variants for registrants
	MAY	Registrar that supports the registration of variants may also update any related services that are impacted by variants
Technical Community	SHOULD	Developers of software tools for the technical community should consider, based on user requirements, to enhance their software to support the administration and management of variants
		Software intended for Internet end users—such as web browsers, email clients, and operating systems—should support variants to the extent necessary to ensure a positive user experience
		To provide end users with a consistent and predictable experience with variants across software applications, developers should, to the extent possible, publicly share best practices and emerging standards in terminology and functionality