The IDN Variant Issues Project:

A Study of Issues Related to the Delegation of IDN Variant TLDs

20 April 2011

Project Proposal updated based on comments received during the Public Comment period held from 21 February 2011 to 6 April 2011

The IDN Variant Issues Project:

A Study of Issues Related to the Delegation of IDN Variant TLDs

Table of Contents

Executive Summary

The IDN Variant Issues Project will undertake work - through case studies - to identify issues associated with the beneficial and safe delegation of IDN variant TLDs.

Specifically ICANN proposes to conduct as many as six case studies in the following scripts (Arabic, Chinese (Traditional and Simplified), Cyrillic, Devanagari, Greek, and Latin) to investigate the set of issues that need to be resolved to facilitate a good user experience for IDN variant TLDs. From these case studies, an Issues Report will be created. Implementation of each case study is subject to having the required level of support from the relevant community and the necessary expertise.

Managing IDN variants is a complex and important subject, and the success of the project is dependent on significant community expertise input and cooperation in doing the work.

1. Introduction

Language communities that use variant characters are affected by decisions about how variants are managed and implemented in new TLDs. This is of concern in both IDN gTLD and IDN ccTLD implementations.

In 2009, an independent implementation working team was formed after discussions during the ICANN meetings in Mexico City and Sydney to study these issues. The team included linguistic and technical experts from various language communities, and was co-chaired by two ICANN Board Directors who are well-versed in the fields of IDN and DNS. The team recommended that variants not be delegated as TLDs at that time, and that if desired variants are to be delegated, certain conditions must be fulfilled[1].

To develop potential solutions for the delegation of IDN variant TLDs, the ICANN Board in its 2010 meeting[2] in Norway directed the CEO to:

> … develop (in consultation with the ICANN Board ES-WG) an issues report identifying what needs to be done with the evaluation, possible delegation, allocation and operation of IDN gTLDs containing variant characters, as part of the new gTLD process in order to facilitate the development of workable approaches to the deployment of gTLDs containing variant characters IDNs. The analysis of needed work should identify the appropriate venues (e.g., ICANN, IETF, language community, etc.) for pursuing the necessary work. The report should be published for public review.

This document describes ICANN's proposed plan to develop an initial Issues Report.

As many as six case studies in the following scripts (Arabic, Chinese (Traditional and Simplified), Cyrillic, Devanagari, Greek, and Latin) are planned to identify the set of issues that, if resolved, may enable the delegation of IDN variant TLDs for the benefit of the respective user communities. It is expected that the results of the case studies will play a crucial role in the identification of issues, roadblocks and potential solutions towards the handling of IDN variants TLDs. It is worth noting that the implementation of each case study is subject to having the required level of support from the relevant community and the necessary expertise.

---

[1] Definitions accessible at http://www.icann.org/en/topics/new-gtlds/idn-implementation-working-team-report-final-03dec09-en.pdf
[2] See ICANN Board of Directors. (2010) Adopted Board Resolutions. Trondheim, Norway. Retrieved November 30, 2010, from http://www.icann.org/en/minutes/resolutions-25sep10-en.htm#2.5

Section two of the document provides some background on IDN variants; section three defines the project scope; section 4 details the overall work plan, work team composition, sharing of information, preliminary timeline, budget philosophy, etc.
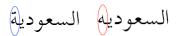
An earlier version of this proposal was posted for public comment from 21 February 2011 to 6 April 2011. A total of ten community submissions were posted to the forum (http://www.icann.org/en/public-comment/#idn-variant-tld). The IDN variant team analyzed these comments and provides responses in section 5.

2. Background

There is no universally agreed definition of a "variant". One possible definition[3] is that 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. IDN variant TLDs contain one or more characters that have such variants.

Possible examples of IDN variant TLDs[4]:

   a.   Arabic Example:

السعوديه السعودية

   b.   German Example:

      "koeln" and "köln"

   c.   Chinese Example:

中国 vs. 中國

3. Project Scope

There are several identified issues with the delegation of IDN variant TLDs. Excluding a variant of the TLD may disenfranchise communities that use the characters in the excluded TLD strings, while allowing the delegation of variants without carefully

---

[3] See RFC 3743, Joint Engineering Team (JET) Guidelines for Internationalised Domain Names (IDN) Registration and Administration for Chinese, Japanese and Korean. Available at: http://www.ietf.org/rfc/rfc3743.txt. The IDN Implementation Working Team Final Report also uses this definition, see < http://www.icann.org/en/topics/new-gtlds/idn-implementation-working-team-report-final-03dec09-en.pdf>

[4] These are for illustration purposes only, it does not imply that these are requirements for variants.

considering its impact could lead to inconsistent user experience as well as security and stability issues.

The expected user experience if using delegated IDN variant TLDs as compared to using the related base label TLD may vary from case to case. In general, to ensure the success delegation of IDN variant TLDs, the following tasks needs to be completed:

Project Tasks:
1. Create a commonly understood glossary of terms and ensure that such terms are accurate and vetted with appropriate technical and linguistic communities and are used consistently throughout the project to improve the dialogue among participants
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

The IDN Variant Issues Project focuses on question (1) and (2).

Follow-on Tasks:
3. Determine the circumstances (where they exist) where certain types of IDN variant TLDs might be eligible for delegation
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
5. Arrive at rules and guidelines, both in the registry operational requirement area and the technical implementation area
6. Determine the responsibilities of TLD operators who would be responsible for managing such delegated IDN variant TLDs
7. Determine what kind of compliance programs may be necessary to ensure that IDN variant TLDs operate according to the arrived at rules and guidelines
8. Identify viable and sustainable outreach mechanisms to communicate and interact with the community on the issues report

Tasks (3) through (8) will be the focus of follow-on projects by ICANN policy development, implementation guidelines produced by ICANN staff in consultation with the community, and relevant technical work by IETF, and other interested organizations.

The IDN Variant Issues Project goal is to identify the "problem statement", by having diverse sets of teams identify use cases and describe the foreseeable problems for their linguistic communities that IDN variant TLD policy should seek to address. Groups should explain how they expect IDN variant TLDs to work from the perspective of how users use and interact with domains. For example, how domains should function when entered into web browsers, how domains should be entered in email addresses, how

domains should function during the domain registration process, how domains should be entered into server configurations, and so on. It should consider what would be undesirable outcomes in these different areas as well, with a particular emphasis on security and stability issues.

Case study teams should avoid describing solutions, but rather focus on the problem statement. A clear vision on what properly functioning IDN variant TLDs would look like to domain users can then be turned into a unified list of issues to address and goals. This process should produce a clear description of scope and use cases, which should make it easier to assess whether potential solutions will provide satisfactory results.

Once this agreed set of issues is harmonized across the different case study teams, along with settling on common terminology, work on analyzing potential solutions can proceed as a second phase. The work on identifying potential solutions will consider linguistic accuracy, technical feasibility, usability, accessibility, and security and stability. It is anticipated the solution space will be a mixture of rules for registries and operational requirements that a TLD operator can achieve through operational practices where possible, or through protocol mechanisms.

4. The Project

Individual case studies will be undertaken to focus on the known situations where variants are in common use and may be part of applications requesting the delegation of IDN variant TLDs. Each study, carried out by a volunteer community study team with help from ICANN staff, will concentrate on specific issues that are relevant to their communities. This section describes the work plan, the composition of case study teams, how information is to be shared among the case study teams and with external entities, the project management and coordination, the timeline, and the project budget philosophy.

In order to inform the discussion inside the case study teams, a survey of current registry policy around IDN variants at the second level will be done in parallel with the work of the case studies. The survey will help understand how variants are managed in registries below the root to avoid reinventing solutions. It will contain questions related to, for example, the type of variants that are managed by the registry, the policy and technical requirements for the registrants of such variants, etc. Team members will assemble a questionnaire to be sent to TLD (gTLD and ccTLD) operators that currently offer IDN registrations. Responses will be summarized and synthesized in a report that will be publicly available.

4.1   Work Plan

There are four steps in the plan: define objectives, form study teams, undertake case studies, synthesize issues, and create issues report.

1. **Define objectives:** The first step in the work plan is to determine the objectives of the study teams. Proposed objectives are outlined above. The ICANN project team will propose a set of objectives, to be reviewed by the new Board IDN Variant Working Group (BV-WG), and published for community input. (NOTE, this task is already completed with the publication of the draft proposal)

2. **Formation of case study teams:** The second step is the establishment of the case study teams to undertake the case studies. The proposed composition of the study teams is described below. Each team will be led by a community case study coordinator, identified by ICANN in consultation with the relevant community.

3. **Case studies:** Each case study team will develop a detailed study plan to achieve the set objectives. The output of each Case Study will include:

   - An issues report for that case. It is expected that some issues will be germane to all cases while some will be case specific.
   - A set of issues necessary for the goals (1) and (2) listed in Section 3.

4. **Synthesis of issues:** A coordination team comprised of representatives from the case study teams and ICANN staff will develop a single issues report. The report will be divided into two sections. One section will be comprised of common issues germane across the cases studied. The second section will be comprised of issues particular to specific cases.

5. **Issues report:** The issues report will describe each of the general and case-specific issues to be resolved for the cases studied. It will also provide a detailed roadmap that can be used for studying additional cases. The experience gained from these initial case studies should enable the efficient development of similar issues reports for additional cases.

4.2   Composition of the Case Study Teams

Each team will be comprised of the following: team coordinator, community representatives, linguistic experts, DNS and IDNA experts, security & scalability experts, policy experts and registry/registrar operations experts – see Table 1 below. Successful completion of the work depends, above all else, upon finding community resources to perform the substantive work.

Table 1: Composition of each Case Study team

Area of Expertise	Description
Case Study Coordinator	Team coordinator
Community Representation	Experts with understanding of local culture, customs, and practices.
Linguistics	Experts in the specific languages/script, ideally with some knowledge in Unicode .
DNS & IDNA	Experts knowledgeable in protocol-level technical details of DNS and IDNA.
Security and Stability	Experts knowledgeable in security and stability implications of variants on Internet protocol usage.
Policy	Experts knowledgeable of ICANN's bottom-up process.
Registry/Registrar Operations	Experts familiar with local registry operations, standards and local registration policies.

The ICANN project staff responsible for recruiting and supporting the teams includes:

Project leader: Dennis Jennings
Management oversight: Kurt Pritz
DNS & IDNA, and security & stability: Kim Davies
Policy coordination: Steve Sheng
Registry operations: Francisco Arias
Linguistic outreach: Baher Esmat
Community liaison: Naela Sarras
Project management: Anand Mishra

4.3   Project Management and Coordination

To help implement IDN Variant TLDs project ICANN will provide Project Management services.

These services will consist of:
- Planning - project planning, work to develop a plan with the case study teams
- Project schedule – develop and maintain a working integrated schedule with agreed upon milestones
- Communications – internal and external communication, status reports, progress updates, coordination of activities

- Execution – working with the teams to facilitate timely deliverables
- Coordination – collecting and synthesizing information, providing meeting minutes and other support as needed

4.4   Sharing of Information

The subject of variants in domain names is being increasingly broadly discussed in the community among groups such as JIG[5], language-specific initiatives[6], IRD[7], and IETF. To ensure the successful completion of this project, all those that are working on variants or related subject matters should stay closely coordinated.

There are two types of information sharing needed: sharing within and across the case study teams, and sharing of the case study team's work with the external entities mentioned above.

To facilitate information sharing within and across the case studies, ICANN will ask that:

1. The leaders of the case study teams hold regular calls (e.g. monthly) to update each other of their group's progress and also to discuss if any general principles, or synergies, can be applied.
2. An internal wiki is created for the project and one for each case study team that can be viewed by members of the other study teams.
3. The ICANN Project Leader, or a member of the ICANN Project Team, attends all the calls for each study group.

To facilitate information sharing with external entities, ICANN will ask:

1. That there is an overlap of membership of case study team members with external entities working on these areas. For example a technical expert working on name aliases in IETF should be on at least one of the case study teams.
2. The case study teams as a whole provide quarterly updates to other external groups, and the other external working groups will be asked to also keep the case study teams informed of their progress.
3. Staff to create an informational wiki (accessible to the public) that documents all of the different groups' work on domain name variants to date.
4. The output of this work be socialized by publishing for public comment and considered as input to IETF's relevant working groups.

---

[5] Refers to the Joint ccNSO-GNSO IDN working group.
[6] See Arabic Script IDN Working Group <http://asiwg.org/wiki/>, Chinese Domain Name Consortium < http://www.cdnc.org/>
[7] Refers to the Joint GNSO-SSAC Working Group on Internationalised Registration Data.

4.5   Timeline

This is not a complex project in terms of interdependencies or task loading. However, recruiting for this type of effort has always been problematic as budgeted or planned for positions remain unfilled and there are a limited number of technical experts available to participate. Therefore, slightly more time is provided than might be anticipated to recruit skill sets and perform work (because skill sets might have to be shared among the case studies). There are limited numbers of IDNA and DNS experts and the timeline described here is contingent on securing requisite skills for each case study.

Regular progress reports will be made to the new Board IDN Variant Working Group (BV-WG). Table 2 shows the proposed timeline for the project.

Table 2: Preliminary Timeline for the Project

Task	Estimated Completion Date
Objective setting	31 Mar 2011
Recruiting Case Study teams	6 June 2011
Complete Case Studies	30 Sept 2011
Synthesize issues across studies	30 Nov 2011
Issues report publication	15 Dec 2011

4.6   Budget Philosophy

To the extent possible and in order to minimize incremental expense, the case study effort will be performed by ICANN volunteers and supported by ICANN staff members that are cross-utilized in various cases. However, in order to complete the work in a timely manner, expenditures will be made. Resources on the critical path to success should be retained (DNS and linguistic expertise) in some fashion in order to ensure their availability and the completion of their work. The Finance Committee will be consulted to determine sources of funds. The Board will be consulted if reprioritization of budgeted tasks is required.

5. Community Consultation

An earlier version of this proposal was posted for public comment from 21 February 2011 to 6 April 2011. A total of ten community submissions were posted to the forum (http://www.icann.org/en/public-comment/#idn-variant-tld). The ICANN IDN Variant Issues Project Team (hereinafter "The Project Team") identified six key points, and provides responses to each below.

1. Proposal to add a Greek Case Study:

One commenter recommended that ICANN add a sixth case study for Greek based on their experience with variants.

Response: ICANN's rational for selecting the case studies are: 1) language/scripts that are used by a large number of people in the world, and/or 2) IDNs are already deployed at second level and where there are known needs or issues with IDN variants. In light of the comments from the Greek registry and its operational experience with IDN variants, and upon reviewing our selection criteria, the IDN variant team recommend that the ICANN Board Variant Working Group to consider including a Greek case study, subject to sufficient support and participation from the Greek community as determined by the Project Team.

The Project Team would like to emphasize that the goal of the project is not to solve variants in all languages/scripts, but merely to identify a set of issues in relation to the delegation of IDN variant TLDs, using a small number of representative case studies. We intend that the approach and results will be useful to other parties that are considering studying variants in their languages/scripts.

2. Question of whether the project is about Language-level or Script-level variants:

One commenter raised the question whether the project is about language-level or Script-level variants.

Response: The Project Team agrees that this is an important question that needs to be answered. However, we consider this is a question for the case studies to answer and recommend this as one of the questions that the case studies address.

3. Recommendation to better define the scope of the Indic case:

Several commenters raised the issue that the Indic case needs to be more clearly defined.

Response: The Project Team agrees, and based on the response, we recommend the Indic case be refined to be about the Devanagari script.

4. Visual similarity between scripts:

Several commenters raised the issue of string similarity between related scripts such as Greek, Cyrillic and Latin.

Response: The Project Team agrees the case study teams will be expected to include

visual similarity issues related to variants. For issues of pure visual similarity between proposed TLDs, ICANN will rely, for now, on mechanisms already in place to handle string similarity in the new gTLD program and the IDN ccTLD Fast Track Process. For example, in the new gTLD program, the string similarity review involves a preliminary comparison of each applied for string against existing TLDs, reserved names and other applied for strings. A string similarity panel conducts this review.

5. Proposed timeline for the Case Study teams:

Several commenters called for allowing some cases to proceed ahead of others if they are ready, others called for not delegating any variant TLDs until all the technical and policy related issues of delegating variant IDN TLDs are identified and vetted through the ICANN stakeholder model.

Response: The Project Team acknowledges and understands both views. As such, we intend to keep the original approach identified in the project proposal to work towards creating a synthesized issues report based on the individual reports from each case study before making a decision on any delegation of IDN variant TLDs.

6. Size of teams vs. community involvement and visibility of the work

Several respondents expressed views on the size of the case study teams, suggesting that the teams be kept at a manageable size while at the same time making sure that the work of the teams is visible to the community and allows experts in the community to provide feedback.

Response: The Project Team agrees with the community and will work on a mechanism to facilitate participation. ICANN is considering approaches used by other organizations, such as the IETF, for facilitating participation. ICANN staff will refine the approach and provide the details to the case study teams and community members.