

JIG Initial Report on IDN TLD Variant Policies for the Root Zone

Date: March 8, 2011

Version: DRAFT 0.4

This report is intended to be a document to solicit input from the community. The document is a preliminary stocktaking of policy considerations as well as possible approaches to address such considerations for the implementation of application, allocation and delegation policies for IDN Variants in the root zone for IDN gTLDs and IDN ccTLDs.

The JIG (Joint ccNSO-GNSO IDN Working Group) was created to discuss issues of common interest between the ccNSO and the GNSO on IDNs (Internationalized Domain Names), especially IDN TLDs. The JIG has identified 3 issues of common interest to date:

1. Single Character IDN TLDs
2. IDN TLD Variants
3. Universal Acceptance of IDN TLDs

This report is specific to issue 2. IDN TLD Variants.

I. Background & Related Works

The issue of IDN Variants and its nature as a policy matter has been recognized since the first publishing of the IDNA standards. In conjunction with the first set of RFCs (RFC3490, 3491 & 3492), an IESG Statement on IDN (<http://www.ietf.org/iesg/statement/idn.html>) was issued to explain that *“For some human languages, there are strings of characters that have equivalent or near-equivalent meanings. When someone registers a name containing such a string, the registry might want to automatically generate a list of semantically or visually equivalent strings and suggest that they also be registered.”*

During the IDN discussion the issue of variants was already identified. Subsequent to the publishing of the first set of standards track RFCs, the Chinese, Japanese and Korean language communities jointly developed and published the JET IDN Guidelines (<http://www.ietf.org/rfc/rfc3743.txt>), describing a general framework for IDN Variants. The Chinese language community further specified an IDN Variant framework suitable for Chinese domain names in the CDNC Guidelines (<http://www.ietf.org/rfc/rfc4713.txt>).

The Arabic script community also studied the issue of IDN Variants, and the Arab Working Group on Arabic Domain Names (AWG-ADN), established by the League of Arab States (LAS) has issued an informational RFC on Linguistic Guidelines for the Use of the Arabic Language in Internet Domains (<http://www.ietf.org/rfc/rfc5564.txt>).

The ICANN community has also recognized the issue of IDN Variants, and has published and continued to revise the ICANN IDN Guidelines (<http://www.icann.org/en/topics/idn/implementation-guidelines.htm>). The issue of IDN Variants has been described in all versions of the guidelines. The latest version of the ICANN IDN Guidelines (<http://www.icann.org/en/topics/idn/idn-guidelines-26apr07.pdf>) explains and specifies that, *“A domain registry will publish the aggregate set of code points*

that it makes available in clearly identified IDN-specific character tables, and will define equivalent character variants if registration policies are established on their basis.”

During the deliberations of the New gTLD PDP, a GNSO IDN WG was formed in November 2006 (http://gns0.icann.org/issues/idn-tlds/idn_working_group-18nov06.htm) to address policy issues that may arise from the introduction of Internationalized Domain Names at the top level (IDN TLDs). The IDN WG produced a final Outcomes Report (<http://gns0.icann.org/drafts/idn-wg-fr-22mar07.htm>) in March 2007. Recommendations from the Outcomes Report were eventually incorporated into the GNSO Final Report on the Introduction of New gTLDs (<http://gns0.icann.org/issues/new-gtlds/pdp-dec05-fr-part-08aug07.htm>).

The Outcomes Report specified 2 areas of broad agreement that is directly relevant to the issue of IDN Variants:

4.1.4. One String per new IDN gTLD:

Agreement that the approach of the New gTLD PDP with one string for each new IDN gTLD application is relevant, except in the rare cases when there is a need to cover script-specific character variants of an IDN gTLD string.

4.1.8. Suggested Approach towards Aliasing:

Agreement to address aliasing as a policy issue, rather than in terms of any specific technical mode for implementation of such a feature.

In considering an IDN ccTLD Fast Track, the ccNSO council put forward a charter (<http://ccnso.icann.org/workinggroups/idnc-charter.htm>), which was approved by the ICANN board at the Los Angeles meeting in October 2007, for the establishment of the IDNC Working Group, comprised of members of the GNSO, ccNSO, GAC, ALAC, SSAC, representative of the technical community, and ICANN staff. The IDNC produced a Final Report (and Board Proposal) on the Fast Track Process for IDN ccTLDs in June 2008 (<http://ccnso.icann.org/workinggroups/idnc-wg-board-proposal-25jun08.pdf>).

One of the overarching requirements of the IDNC work included to “Comply with the IDNA protocols and IDN guidelines”, which describes IDN Variants (as mentioned above). The IDNC Final Report also explained that prospective Fast Track IDN ccTLDs should have an established IDN Language Policy in accordance with the ICANN IDN Guidelines:

4. Prepare language table

...The language/script table to be used by the IDN ccTLD may already exist i.e. has been prepared by another Territory using the same language/script and was already submitted. In this case the selected delegate should indicate its intention to use that language/script table...Territories using the same script are encouraged to cooperate in developing a language/script table, in accordance with IDN guidelines.

In the ongoing implementation of the IDN ccTLD Fast Track, the issue of IDN Variants at the top-level was further examined. The IDN Implementation Working Team Final Report (<http://www.icann.org/en/topics/new-gtlds/idn-implementation-working-team-report-final-03dec09-en.pdf>) makes the recommendation that:

Recommendation 1

1.1 Desired Variant Labels must be indicated by the applicant;

- 1.2 *Desired Variant Labels will be allocated, and may also be delegated, to the applicant;*
- 1.3 *The delegation of Desired Variant Labels, where otherwise appropriate, is contingent upon the applicant agreeing to conform to the Rules outlined in Section 3.3¹;*
- 1.4 *Undesired Variant Labels will be neither allocated nor delegated, and will be blocked.*

Subsequently, a proposal for Synchronized IDN ccTLDs was developed (<http://www.icann.org/en/topics/idn/fast-track/proposed-plan-synchronized-idn-cctlds-22mar10-en.pdf>) and acknowledged by the ICANN Board at its March 21010 meeting in Nairobi meeting (<http://www.icann.org/en/minutes/resolutions-12mar10-en.htm#10>).

At its meeting on June 25, 2010 in Brussels, the ICANN Board passed 2 resolutions (<http://www.icann.org/en/minutes/resolutions-25jun10-en.htm#2> and <http://www.icann.org/en/minutes/resolutions-25jun10-en.htm#4> respectively) to approve the delegations of 2 IDN ccTLDs (“.中国” and “.台灣” respectively) under the IDN ccTLD Fast Track process along with their IDN Variants (“.中國” and “.台灣” respectively). In each of the resolutions, a specific clause described the considerations on the issue of IDN Variants:

For the delegation of “.中国”:

Whereas, the applicant has undertaken to operate the two top-level domains in a manner which does not cause confusion to the Internet user community, as documented in their implementation plan published online at <http://www.cnnic.cn/html/Dir/2010/06/12/5852.htm>.

For the delegation of “.台灣”:

Whereas, the applicant has undertaken to operate the two top-level domains in a manner which does not cause confusion to the Internet user community, as documented in their implementation plan published online at http://www.twnic.net/english/dn/dn_07a.htm.

¹ Section 3.3 of the IDN Implementation Working Team Final Report for reference:

3.3 Proposal for management at the top-level of both Type 1 and Type 2 Variants

1. In addition to the Base IDN TLD Label, a TLD applicant may identify a number of Variant Labels.
2. The applicant must classify each Variant Labels as 'Desired' or 'undesired'.
3. Desired Variant Labels will be allocated, and may also be delegated, to the applicant.
4. Undesired Variant Labels will neither be allocated nor delegated, but will be blocked.
5. Desired Variant Labels may be delegated to the (same) applicant provided the following conditions are met:
 - A. The applicant requests the Variant Label.
 - B. The domains directly below the variant TLD are to be mapped exactly to the corresponding domains under the Base TLD.
 - C. The applicant must demonstrate how this mapping is to be achieved, preferably by presenting earlier work on a testbed.
 - D. The applicant must provide information on how it intends to inform its registrars/resellers/registrants about the correct set-up of name servers, and all other application servers that process domain names directly, so that end-user confusion can ideally be prevented, and at least held to an acceptable minimum. (This is discussed in further detail in Annex 3.)
 - E. ICANN shall conduct periodic checks to ensure that the said mapping is in place.
 - F. The delegation document for the Base IDN TLD Label (whatever form it takes: contract, MoU, accountability framework, ...) will commit the registry to the points above.
6. It is understood that that technical or other considerations may necessitate changes in the items under 5 above.

At its meeting on September 25, 2010, in Trondheim, Norway, the ICANN Board resolved to request the CEO to develop (in consultation with the board ES-WG) an issues report (<http://www.icann.org/en/minutes/resolutions-25sep10-en.htm#2.5>), and noted that “The recent delegation of Chinese-language ccTLDs does not yet provide a generally workable approach for gTLDs; there are serious limits to extending this approach at this time.”

Subsequently, on February 21, 2011, a Draft Proposal for the Study of Issues Related to the Delegation of IDN Variant TLDs was published for public comments (<http://www.icann.org/en/announcements/announcement-5-21feb11-en.htm>). The announcement further explains that “*The proposed project work plan includes a series of case studies to investigate issues associated with delegating and using IDN variant TLDs in different communities, a staffing plan involving community members and a resource allocation plan. Importantly, this work will be coordinated with the Joint IDN Working Group and other community work on this topic.*”

Besides the IDN-Implementation Working Team Final Report, the GNSO Final Report on the Introduction of New gTLDs, and the Final Report of IDNC Working Group, and the Board Resolutions described above, in conducting its work, the JIG is also guided by the following:

- The overarching requirement to preserve the security and stability of the DNS;
- Compliance with the IDNA protocols and ICANN IDN Guidelines;
- Input and advice from the technical community in respect to the implementation of IDNs;
- Draft New gTLD Applicant Guidebook (DAG – <http://www.icann.org/en/topics/new-gtlds/draft-rfp-clean-28may10-en.pdf>) and subsequent versions as they become available, along with corresponding comments received; and,
- IDN ccTLD Fast Track Final Implementation Plan (<http://www.icann.org/en/topics/idn/fast-track/idn-cctld-implementation-plan-16nov09-en.pdf>) and relevant subsequent updates.

Furthermore, the JIG refers to the ongoing IDNccPDP, the Policy Development Process (ccPDP) for the long term implementation of IDN ccTLDs, and the two working groups formed: ccNSO IDN PDP Working Group 1 (<http://ccnso.icann.org/workinggroups/idn-pdp-wg1-charter.pdf>), to report on and identify a feasible policy for the selection and delegation of IDN ccTLDs associated with the territories listed in the ISO 3166-1; and, ccNSO IDN PDP Working Group 2 (<http://ccnso.icann.org/workinggroups/idn-pdp-wg2-charter-23mar10-en.pdf>), to report on changes to Article IX and relevant Annexes in the ICANN Bylaws to include IDN ccTLD's as full members in the ccNSO on equal footing as the current members.

Besides the IDNccPDP, the current discussion at the IETF DNSEXT WG is also of relevance to the IDN Variant discussion at the JIG: <http://www.ietf.org/id/draft-yao-dnsext-identical-resolution-01.txt>.

II. Glossary / Nomenclature of Terms Used

[Probably should have a session explaining the nomenclature used for this document/discussion.]

Base IDN: The basis for which an IDN Variant is generated from (Similar to Primary IDN)

Primary IDN: The IDN TLD String applied for

IDN Variant: A string different from the Primary IDN but considered an IDN Variant of the Primary IDN (see section IV.) based on the corresponding IDN Language Policy

Possible additional references: RFC3743 (JET Guidelines for IDN) and RFC4713 (Registration and Administration Recommendations for Chinese Domain Names)

III. Policy Aspects of IDN Variants

In considering the scope of the work at the JIG, the group is cognizant that there are relevant ongoing discussions on the subject of IDN Variants at various different forums and groups. Following the charter of the JIG, the scope and focus of this group would be on policy aspects relevant to issues of common interest. More specifically, for IDN Variants, the JIG will focus on policy operatives for implementation by ICANN in its technical coordination of the DNS root zone.

The idea of a “variant” for a domain string has been discussed in various forums to potentially mean various different things. In the context of this document nevertheless, it is an important fundamental understanding that in this discussion, an “IDN Variant” is considered to be a domain that is a technically different string from the Primary IDN, and that the Primary IDN along with any delegated IDN Variant should be considered by an appropriate IDN Language Policy as one domain. The specific IDN Language Policy for the determination of such IDN Variants is not within the scope of this discussion, except for the resulting “types” of IDN Variants by their allocation and delegation nature.

The JIG has identified the following aspects that should be considered for the development of an IDN Variant policy that could be consistently applied by ICANN for all IDN TLDs in the root zone:

1. The requirements for a string to be considered an IDN Variant (of its Primary IDN) and a framework of attributes constituting an IDN Language Policy for producing IDN Variants
2. The types of IDN Variants with respect to their allocation and delegation properties
3. Policy operatives corresponding to the types of IDN Variants
4. Requirements for zones directly managed by a TLD operator of an IDN Variant TLD
5. Adding IDN Variant TLDs subsequent to initial delegation of a Primary IDN TLD

[The above list is a first draft for discussion so far.]

IV. Preliminary Viewpoints & Possible Approaches

[This section will include viewpoints collected in the workgroup discussion regarding the aspects mentioned above. This draft contains discussions collected so far.]

Aspect 1. The requirements for a string to be considered an IDN Variant (of its Primary IDN) and a framework of attributes constituting an IDN Language Policy for producing IDN Variants

[viewpoints collected so far]

While this is a policy aspect identified by the JIG, upon the commencement of the work proposed for the Study of Issues Related to the Delegation of IDN Variant TLDs (<http://www.icann.org/en/topics/new-gtlds/idn-variant-tlds-delegation-21feb11-en.pdf>), specific policy aspects for the definition and the determination of the appropriateness of an IDN Variant with respect to its Primary IDN would generally reference that process.

[NEVERTHELESS, I wonder if the framework below should/should not be kept for the Initial Report with special description that this will be further discussed in the Studies as mentioned above]

In order to be considered an IDN Variant, a string should be:

- Consistently produced based on a Primary (or Base) IDN along with an appropriate IDN Language Policy (described further below)
- Is technically (based on the DNS standards) a different string than the Primary IDN

An appropriate IDN Language Policy should include the following components

- A set of character tables (IDN Language Tables)
- If IDN Variants are to be produced, at least one of the character tables being a character variant mapping table
- A set of rules describing how IDN Variants are to be generated from a given Primary IDN
- A set of rules describing how IDN Variants are to be categorized (into the types identified in Aspect 2. below)
- A set of rules describing any allowed or disallowed combination of characters

An appropriate IDN Language Policy should demonstrate the following properties

- Same set of tables and rules to be applied for second level registrations (or any level the registry provides registrations for)
- Due consideration for user cultural and linguistic considerations

Aspect 2. The types of IDN Variants with respect to their allocation and delegation properties

The IDN Implementation Working Team Final Report identified two types of IDN Variants:

Variant Types: Variants may be classified in two types to differentiate between those that are graphemically identical (visually indistinguishable), and variants that are of dissimilar appearance but have the same meaning (semantically equivalent):

- *Type 1 Variants: Variants that are visually identical to the base label.*
- *Type 2 Variants: Variants that are not visually identical but are considered equivalent to the base label in some orthographic sense.*

However, subsequent discussion in the Report does not distinguish between the operative policies between Type 1 and Type 2. But rather, the operative principles are based on the concepts of “Desired Variants” and “Undesired Variants”:

An applicant who requests a Base Label and either Type 1 or Type 2 Variant Labels must indicate each variant as either a Desired Variant or an Undesired Variant. Desired variants are intended for use interchangeably with the base IDN TLD label. All other variants are not desired.

The JIG will expand on this classification rather than the “Type 1” and “Type 2” differentiation.

This report will adapt the terminologies used by RFC 3743:

- Base IDN / Primary IDN
- Preferred IDN Variants
- Other IDN Variants
- IDN Zone Variants

[viewpoints collected so far]

- Preferred IDN Variants – allocated and delegated together with the Primary IDN
- Other IDN Variants

- Reserved IDN Variants – allocated, not delegated, but reserved for use by (i.e. delegation to) Primary IDN only
- Blocked IDN Variants – unallocated, and should not be allocated or available for further application

Another type of IDN Variants that can be identified as:

- IDN Zone Variants – In the case where some of the reserved variants are also delegated (“activated”) into a zone along with the Primary IDN and the Preferred Variants (i.e. Preferred Variants + delegated Reserved Variants)

3. Policy operatives corresponding to the types of IDN Variants

In considering policy operatives for IDN Variants as a matter of ICANN’s coordination of the root zone the following 2 questions should be addressed:

- A. Whether or not an IDN Variant TLD should be:
 - a. Allocated
 - b. Delegated
 - c. Reserved
- B. What policies should regulate the delegation and ongoing operations of an IDN Variant TLD

The IDN Implementation Working Team Final Report explained and clarified the concept and distinction between TLD “allocation” and “delegation”:

TLD Allocation: *A Base IDN TLD Label and its Desired Variant Label(s) must be indicated by the applicant. Once a Base IDN TLD Label is assigned to an applicant, the Desired Variant Label(s) will be reserved solely for the use of the applicant. This process is called “Allocation”.*

TLD Delegation: *The addition of the Allocated TLD Label (Base IDN TLD Label or Desired Variant Label(s)) to the DNS root zone.*

Based on the classifications explained in Aspect 2 above:
[viewpoints collected so far]

The JIG discussed a two step process to implement the policy operatives described for IDN Variants:

Step 1: IDN Variant Legitimacy (Allocation Evaluation)

- Should be based on an approach as described in Aspect 1 above
- Should be consistent with IDN Variant allocation and delegation policies at the 2nd level (or lower levels for which the TLD offers registrations)
- Should demonstrate broad acceptance from relevant language community
- Should incorporate findings of the Study of Issues Related to the Delegation of IDN Variant TLDs

Step 2: Business Considerations (Delegation Evaluation)

- Delegation of Primary IDN TLD
- Consideration of Preferred IDN Variant TLDs (i.e. those for which must be delegated together to avoid user confusion and cultural sensitivity, e.g. Simplified and Traditional Chinese IDNs)

- Consideration of Blocked IDN Variant TLDs (i.e. which IDN Variant TLDs should not be applicable for delegation)
- Consideration of Reserved IDN Variant TLDs (i.e. IDN Variant TLDs that may be further applied for delegation)

The Allocation Evaluation (Step 1) could be incorporated into the String Evaluation (and Contention resolution) processes as described for new gTLDs in the draft Applicant Guidebook, and for IDN ccTLDs in the IDN ccTLD Fast Track process, and the IDN ccPDP. It is important that the Allocation Evaluation include IDN Variant TLDs to also identify contention sets caused by overlapping IDN Variant TLDs (even when the Primary IDN TLDs may not be themselves considered direct contention, i.e. indirect contention)

The Delegation Evaluation could take place after the evaluation and contention processes and during the Transition to Delegation process.

In the development of Step 1, Aspect 1 may need to be considered and an evaluation on the submitted and/or adopted IDN Language Policy may need to be analyzed for its appropriate for use as per IDN Variant TLDs managed in the ICANN root zone.

The “activation” (or application) of a Reserved IDN Variant TLD is further discussed on Aspect 5 below.

The IDN Implementation Working Team Final Report described the concept and distinction between a “Base Label” and a “Variant Label” as follows:

Base Label: *A Base IDN TLD Label is the primary form used in an IDN TLD request, in a specified language.*

Variant Label: *A Variant Label is resulting from the substitution of one or more characters in a Base IDN TLD Label with variant characters.*

The JIG will build on this identification and include the concept for a Primary IDN as distinguished from a Base IDN:

- Primary IDN: is the form used in an IDN TLD request
- Base IDN: is the form used to generate IDN Variants

In most cases, Primary IDN and Base IDN would be the same. However, there are cases for which IDN Variants should be generated from a Base IDN, while an applicant may wish to use a different Primary IDN for its users. An example could be found for Arabic IDNs where:

- Primary IDN may be a string with some select diacritics included
- Base IDN may be a string with all diacritics included

4. Requirements for zones directly managed by a TLD operator of an IDN Variant TLD

The IDN Implementation Working Team Final Report recommended the following rule as a requirement for zones directly managed by a TLD operator of an IDN Variant TLD:

(Section 3.3, 5.B.) The domains directly below the variant TLD are to be mapped exactly to the corresponding domains under the Base TLD.

[viewpoints collected so far]

- Before a technical standard for mapping is generally available, zones directly below, and managed by the TLD operator of the IDN TLD should essentially serve the same data for the Primary IDN TLD as for all delegated IDN Variant TLDs
- Should an appropriate technical standard for mapping become available along with ICANN policies for migrating towards such standard, the TLD operator must implement such standards accordingly
- IDN TLD Variant zones should produce consistent responses for all domain labels for which the registry operator offers registration (i.e. delegation) for
- The same IDN Language Policy for producing the IDN TLD Variant must be used for IDN registrations for all domain name labels for which the registry operator offers registration for

5. Adding IDN Variant TLDs subsequent to initial delegation of a Primary IDN TLD

[viewpoints collected so far]

The JIG has identified the following scenarios for which an IDN Variant TLD may be added (“activated”) subsequent to the initial delegation of a Primary IDN TLD (along with its Preferred IDN Variant TLDs):

- A. Activation of a Reserved IDN Variant TLD by the operator of a Primary IDN TLD
- B. Update/Change of IDN Language Policy

For A., there are two general approaches that may be used:

- i. Treat each addition (“activation”) as a separate TLD application
- ii. Utilize a special streamlined process

For a Reserved IDN Variant TLD, because it can be assumed, based on policies suggested above, that it would be managed by the same TLD operator, and that the IDN Language Policies and selected strings would have already undergone extensive evaluation, it may be beneficial for the community to have a more streamlined process to manage such applications.

It is assumed that B. should be an exceptionally rare case, especially should an IDN Language Policy be constructed properly. On this point, the advice from the IESG Statement on IDN already explained that *“It is suggested that a registry act conservatively when starting to accept IDNA-based domain names. Equivalences are very hard (if not impossible) to define after registration has started. Assume that the labels “x” and “y” at first are different, but later the tables for the registry are changed so that “x” and “y” are then treated as being the same. If x.example.com and y.example.com both were already registered to different registrants, it is unclear which of them has to withdraw the registration, how that selection process done, and so on. Thus, having complete, publicly-stated policies before accepting registration will lead to a much more stable registration process.”*

However, in such rare cases where the evolution of an IDN Language Policy requires for additional variants, policies should be included to handle such processes. The elements for such process may include the following:

- Whether such additional IDN Variant would conflict with another then existing IDN TLD
- Whether such additional IDN Variant would overlap with IDN Variants of another then existing IDN TLD
- Whether such addition would create excessive user confusion

V. Considerations for a Holding Pattern

The JIG understands that there are ongoing discussions at the technical community for new technical solutions to implement IDN Variants. However, before such technologies can be widely deployed, the following holding pattern policy should be considered by ICANN:

- Allocate and Delegate Preferred Variants only to the same set of NS delegations (to the Primary IDN) to a successful applicant
- Require that the TLD operator adhere to the IDN standards and ICANN IDN Guidelines, as well as to develop a comprehensive implementation plan to undertake to operate the IDN Variant TLDs in a manner which does not cause confusion to the Internet user community
- Maintain Reserved Variants as “reserved” (unallocated and un-delegated) to the successful applicant (of the corresponding Primary IDN) for the time being until a special allocation and delegation process can be developed or the applicant applies for it as a separate application.
- For Preferred and Reserved Variants, they should be taken into consideration when considering contention sets (i.e. two applications for different Primary IDNs with an overlap of one or more IDN Variants should be identified as a contention set)
- For Blocked Variants, do not allow overlap (i.e. two applications for different Primary IDNs with an overlap of one or more Blocked Variants cannot be allowed to proceed separately)