4.5.1 Internationalized Domain Names and Universal Acceptance

- *4.5.1.1 Explanation of the Subject*

As described by ICANN,[1] Internationalized Domain Names (IDNs) permit the global community to use a domain name in their native language or script. This is enabled by allowing domain names to have characters from different scripts, beyond the letters (a to z), digits (0 to 9) and hyphen (-), as encoded by the Unicode standard[2] and as allowed by relevant IDN protocols (RFC 5890[3], 5891[4], 5892[5], 5893[6], and 5894[7]).

ICANN has instituted the IDN Program to assist and promote the multilingual Internet using IDNs. The program is primarily focused on the planning and implementation of the IDN Top-level Domains (TLDs) that include ccTLDs and gTLDs. The IDN Program also supports and undertakes projects geared towards effective deployment of IDNs at the second-level, as guided by the community.

The IDN Program has been implementing the following projects focused on IDN Top-level Domains.

Top-level Domains

- Root Zone Label Generation Rules (LGR)[8] a community driven project aiming to define conservative mechanism for introducing IDN top-level domains into the Internet's Root Zone in a stable and secure manner.
- LGR Toolset[9] project is being undertaken to make it easier for the community to formally represent, create, use and manage data related to the Label Generation Rules for different languages and scripts. ICANN intends to use the LGR Toolset to assist community in determining the valid Top-Level Domains (TLDs) and their variants (if any) for the different scripts.

Country Code Top-level Domains

- The community has created a special process – the IDN ccTLD Fast Track Process[10] – to evaluate Top-Level Domain labels in different languages and scripts for countries and territories. IDN Program implements various aspects of this process.

IDN Program is implementing the following projects focused on IDNs at the Second-level

---

[1] See https://www.icann.org/resources/pages/idn-2012-02-25-en
[2] See http://www.unicode.org/
[3] See http://www.rfc-editor.org/rfc/rfc5890.txt
[4] See http://www.rfc-editor.org/rfc/rfc5891.txt
[5] See http://www.rfc-editor.org/rfc/rfc5892.txt
[6] See http://www.rfc-editor.org/rfc/rfc5893.txt
[7] See http://www.rfc-editor.org/rfc/rfc5894.txt
[8] See https://www.icann.org/resources/pages/root-zone-lgr-2015-06-21-en
[9] See https://www.icann.org/resources/pages/lgr-toolset-2015-06-21-en
[10] See https://www.icann.org/resources/pages/fast-track-2012-02-25-en

- IDN Implementation Guidelines[11] document the recommended practice for registries implementing IDNs at the second-level through a community led process. These guidelines are designed to promote consistency and minimize the risk of cybersquatting and consumer confusion.
- Second-level LGR Reference[12] are being developed on the request of the community to improve consistency in testing of the IDN tables during Pre-Delegation Testing and Registry Service Evaluation Process.

On the ICANN Universal Acceptance website, ICANN explains the issue as follows: [13]

> *In the earliest days of the Internet, the Domain Name System (DNS) contained a relatively small set of top-level domains (TLDs) such as .com, .net and .org. These were names in the ASCII character set containing three A-Z letters. Those available TLDs were later expanded to include two character Country Code TLDs (ccTLDs). In early 2001 top-level domain space grew to include names with more than three characters. In 2008 top-level domains outside the ASCII character set arrived (Chinese, Cyrillic, Arabic, etc [GC1] .) enabling a multi-lingual Internet. In 2013 the top-level domain name space began growing even more rapidly as new generic top-level domains (gTLDs) were delegated into the root zone.*
>
> *Some internet services and software applications have not sufficiently evolved to properly recognize and consistently handle new gTLDs and Internationalized Domain Names (IDN), thus impeding the added benefits of user choice, user confidence and name space competition to the consumer. Software and service providers have historically been unaware of these problems or had little market or regulatory incentive to invest in solutions that would bring true interoperability to platforms or applications. A coordinated industry effort is underway to ensure a timely, practical, and continuing resolution to these changes.*

ICANN notes: "Universal Acceptance will be considered complete when any person can register and use a domain name in any top-level domain in widely distributed web browsers, email clients, mobile apps, and setting up online accounts for Internet and other services."

Universal Acceptance Roadmap

The universal acceptance roadmap was originally published 11 September 2014. [14] The Roadmap states:

---

[11] See https://www.icann.org/resources/pages/implementation-guidelines-2012-02-25-en
[12] See https://www.icann.org/resources/pages/second-level-lgr-2015-06-21-en
[13] See https://www.icann.org/resources/pages/universal-acceptance-2012-02-25-en#overview
[14] See https://www.icann.org/resources/pages/universal-acceptance-initiative-2014-10-03-en

*The **Universal Acceptance** initiative is an effort to address potential user issues and obstacles observed in the use of new Top Level Domains, issues and obstacles rooted primarily in assumptions based on the TLD. This abridged roadmap, an outcome of the JIG Final Report on Universal Acceptance of IDN TLDs[15] plus other work, presents a proposal, based on community input including public comment[16], as to how ICANN's energy, resources, and actions should be applied as part of the initiative.*

The abridged roadmap emphasizes ICANN's multi-stakeholder model by limiting its scope to ICANN's role and possible actions. Identifying and addressing the issues and obstacles require work and collaboration among many stakeholder groups who have documented their activity independently. ICANN views its primary role as one of active coordination and facilitation, acting as a catalyst in connecting relevant stakeholders with each other and with parties who are in a position to remove these obstacles. The vision includes implementing a 'corporate memory' as a central information depository of progress.

Further, in February 2015, the community created the Universal Acceptance Steering Group (UASG)[17] to lead the effort to promote Universal Acceptance of all valid domain names and email addresses.

Relevant Requirements in the AGB

The AGB states in Part II, Requirements for Internationalized Domain Names:

These requirements apply only to prospective top-level domains that contain non-ASCII characters. Applicants for these internationalized top-level domain labels are expected to be familiar with the Internet Engineering Task Force (IETF) IDNA standards, Unicode standards, and the terminology associated with Internationalized Domain Names.

2.1 The label must be an A-label as defined in IDNA, converted from (and convertible to) a U-label that is consistent with the definition in IDNA, and further restricted by the following, non-exhaustive, list of limitations:

2.1.1 Must be a valid A-label according to IDNA.
2.1.2 The derived property value of all codepoints used in the U-label, as defined by IDNA, must be PVALID or CONTEXT (accompanied by unambiguous contextual rules).
2.1.3 The general category of all codepoints, as defined by IDNA, must be one of (Ll, Lo, Lm, Mn, Mc).

[15] See https://ccnso.icann.org/announcements/announcement-18nov13-en.htm
[16] See https://www.icann.org/public-comments/tld-acceptance-initiative-2014-06-18-en
[17] Universal Acceptance Steering Group (UASG) Wiki: https://community.icann.org/pages/viewpage.action?pageId=47255444

2.1.4 The U-label must be fully compliant with Normalization Form C, as described in Unicode Standard Annex #15: Unicode Normalization Forms. See also examples in http://unicode.org/faq/normalization.html.
2.1.5 The U-label must consist entirely of characters with the same directional property, or fulfill the requirements of the Bidi rule per RFC 5893.

2.2 The label must meet the relevant criteria of the ICANN Guidelines for the Implementation of Internationalised Domain Names. See http://www.icann.org/en/topics/idn/implementation-guidelines.ht.

This includes the following, non-exhaustive, list of limitations:

2.2.1 All code points in a single label must be taken from the same script as determined by the Unicode Standard Annex #24: Unicode Script Property (See http://www.unicode.org/reports/tr24/).

2.2.2 Exceptions to 2.2.1 are permissible for languages with established orthographies and conventions that require the commingled use of multiple scripts.
However, even with this exception, visually confusable characters from different scripts will not be allowed to co-exist in a single set of permissible code points unless a corresponding policy and character table are clearly defined.

IDN Variants

The 2007 Final Report did not provide guidance on IDN variants, but the AGB stated in 1.3.3 that:

A variant TLD string results from the substitution of one or more characters in the applied-for gTLD string with variant characters based on the applicant's top level tables.

Each application contains one applied-for gTLD string. The applicant may also declare any variant strings for the TLD in its application. However, no variant gTLD strings will be delegated through the New gTLD Program until variant management solutions are developed and implemented. Declaring variant strings is informative only and will not imply any right or claim to the declared variant strings.

When a variant delegation process is established, applicants may be required to submit additional information such as implementation details for the variant TLD management mechanism, and may need to participate in a subsequent evaluation process, which could contain additional fees and review steps.

The following scenarios are possible during the gTLD evaluation process:

a) *Applicant declares variant strings to the applied-for gTLD string in its application. If the application is successful, the applied-for gTLD string will be delegated to the applicant. The declared variant strings are noted for future reference. These declared variant strings will not be delegated to the applicant along with the applied-for gTLD string, nor will the applicant have any right or claim to the declared variant strings.*

   Variant strings listed in successful gTLD applications will be tagged to the specific application and added to a "Declared Variants List" that will be available on ICANN's website. A list of pending (i.e., declared) variant strings from the IDN ccTLD Fast Track is available at http://icann.org/en/topics/idn/fast-track/stringevaluation-completion-en.htm.

   ICANN may perform independent analysis on the declared variant strings, and will not necessarily include all strings listed by the applicant on the Declared Variants List.

b) *Multiple applicants apply for strings that are identified by ICANN as variants of one another. These applications will be placed in a contention set and will follow the contention resolution procedures in Module 4.*

c) *Applicant submits an application for a gTLD string and does not indicate variants to the applied-for gTLD string. ICANN will not identify variant strings unless scenario (b) above occurs.*

Each variant string declared in the application must also conform to the string requirements in section 2.2.1.3.2.

Variant strings declared in the application will be reviewed for consistency with the top-level tables submitted in the application. Should any declared variant strings not be based on use of variant characters according to the submitted top-level tables, the applicant will be notified and the declared string will no longer be considered part of the application.

Declaration of variant strings in an application does not provide the applicant any right or reservation to a particular string. Variant strings on the Declared Variants List may be subject to subsequent additional review per a process and criteria to be defined.

It should be noted that while variants for second and lower-level registrations are defined freely by the local communities without any ICANN validation, there may be specific rules and validation criteria specified for variant strings to be allowed at the top level. It is expected that the variant information provided by applicants in the first application round will contribute to a better understanding of the issues and assist in determining appropriate review steps and fee levels going forward.

The IDN Variant TLD Program continues to work on the "creation and maintenance of a label generation ruleset process for the root, which is on the critical path to a variant management process for the root zone. A basic assumption is that no variant TLDs can actually be implemented until the necessary community work on the code point repertoire and label generation rules for the root have been finalized.[18]

- *4.5.1.2 Questions and Concerns Related to Subject*

  ICANN states that the goal of "domain names in a TLD must be useable in applications regardless of the written script, and length or newness of the TLD," which roughly captures the observed issues and obstacles driving the discussion of universal acceptance. The use of names in the IDN TLDs combines all of the challenges mentioned in the goal as well as touching areas of concerns of generic and country-code TLDs. ICANN states, "Registration of names must work, protocols must work, and services/applications impacting the user must work; work also in the sense that domain names and the identifiers built on them are useable in administratively permitted ways. Included in this goal is the usability of internationalized email addresses (RFC 6530)."[19]

  ICANN notes that based on its research of the challenges, the list of issues and obstacles as documented by stakeholders is highly dynamic, diverse, and sometimes overlapping. ICANN's role as part of the initiative is to foster relationships among stakeholders involved with universal acceptance issues. ICANN will also promote internationalized email as a way to enable full functionality of IDN TLDs. In addition, ICANN will "develop a means to accept reports of problems, as well as successes, to pass information amongst stakeholders for resolution [and] engage stakeholders in the effort to exchange information on universal acceptance, whether this is seen as informational or a means to gain insight into issues and obstacles."

  Finally, as noted above, there is currently no variant management process for the root zone and as such, IDN gTLDs currently have no mechanism to delegate IDN variants where they may beneficial.

- *4.5.1.3 Relevant Guidance*

  - Principle B
  - Principle C
  - Recommendation 18

- *4.5.1.4 Rationale for Policy Development*

[18] Information about the IDN Variant TLD Program: https://www.icann.org/resources/pages/variant-tlds-2012-05-08-en#history
[19] Ibid

There are currently community-led initiatives related to Universal Acceptance and IDN Variants. It is not anticipated that a potential PDP-WG on New gTLD Subsequent Procedures would need to produce substantive work on Universal Acceptance as the UASG has been designated by the community to lead this effort. However, there may be a need to consider the outcomes of the IDN Variants Program to determine and develop guidelines for integration into the New gTLD space, so policy development may be needed in that regard.