

**Guidelines for the Implementation of Internationalized Domain
Names
Final Draft Version 2
Pending ICANN Board Approval**

A [draft Version 2 of the IDN Guidelines](#) was published on 20 September 2005. It reflected the experiences of the IDN TLD registries in the implementation of [Version 1.0 of the Guidelines](#). A wide range of remarks and suggestions on the initial draft of Version 2 were submitted to a [forum for public commentary](#) that was open through 23 October 2005. On the basis of that material, this final draft Version 2 of the Guidelines was prepared and submitted to the ICANN Board for endorsement.

That text appears below and was prepared by:

- * gTLD Registry Constituency Representatives:
 - o Cary Karp, MuseDoma
 - o Pat Kane, VeriSign
 - o Ram Mohan, Afiliis
- * ccNSO Representatives:
 - o Hiro Hotta, JPRS
 - o Mohammed EL Bashir, .sd Registry
- * ICANN Staff:
 - o Tina Dam

with the grateful acknowledgement of the all the people whose names appear on the public forum, and those who provided further expert assistance.

Introduction

The initial [Version 1.0 of the Guidelines for the Implementation of Internationalized Domain Names](#), was published on 20 June 2003, coinciding with the initiation of IDN deployment in accordance with the IETF Proposed Standard for Internationalized Domain Names in Applications as stated in RFCs [3454](#), [3490](#), [3491](#), and [3492](#). The implementation approach set forth in the Version 1.0 Guidelines was endorsed by the ICANN Board on [27 March 2003](#). That document stated the conditions under which a TLD registry requiring ICANN's authorization to accept IDN registration could begin doing so. The Guidelines were further intended as a support document for other registries establishing IDN policies.

The experience gathered in actual registry practice would then serve as a basis for the revision of the Guidelines whenever such need was apparent. During the course of the review preceding the present revision, and as indicated in the comments received on the resulting draft, the initial version of the Guidelines required extensive modification. The requisite changes could not readily be made by simple incremental changes to the initial text. However, given the urgent nature of some IDN concerns and the corresponding need for rapid response, the working group assigned to the task decided to produce a revised version of the Guidelines retaining their initial format as rapidly as possible, and then proceed with an alternate instrument with which to replace them altogether.

The text presented below does not address all of the concerns that currently attach to IDN. (A list of such issues has been extracted from the public comments on the draft text, and will be posted separately.) The next intended editorial action is to reframe the Guidelines in a manner that is amenable to further development as a Best Current Practices (BCP) document, for which formal IETF status will also be sought.

The Guidelines as presented below have no direct conformance implications with respect to the IDN standards referenced below. The term "will" is not to be read as it would be in a formal normative instrument. Although the Guidelines apply directly to the gTLD registries, they are intended to be suitable for implementation in other registries on the second and lower levels. Any residual lack of clarity that may be inherent in the present wording will be dealt with in the successor BCP.

Guidelines

1. Top-level domain registries that implement internationalized domain name capabilities will do so in strict compliance with the technical requirements described in RFCs [3454](#), [3490](#), [3491](#), and [3492](#) (collectively, the "IDN standards").
2. In implementing the IDN standards, top-level domain registries will employ an "inclusion-based" approach (meaning that code points which are not explicitly permitted by the registry are prohibited) for identifying permissible sets of code points from among the full Unicode repertoire, as described below.
3. (a) In implementing the IDN standards, top-level domain registries will associate each label in a registered internationalized domain name, as it appears in their registry with a single script. This restriction is intended to limit the set of permitted characters within a label. If greater specificity is needed, the association may be made by combining descriptors for both language and script. Alternatively, a label may be associated with a set of languages, or with more than one designator under the conditions described below. (b) A 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. Any such table will be designated in a manner that indicates the script(s) and/or language(s) it is intended to support. (c) All code points in a single label will be taken from the same script as determined by the Unicode Standard Annex #24: Script Names at <http://www.unicode.org/reports/tr24>. Exception to this is permissible for languages with established orthographies and conventions that require the commingled use of multiple scripts. In such cases, visually confusable characters from different scripts will not be allowed to co-exist in a single set of permissible codepoints unless a corresponding policy and character table is clearly defined. (d) All registry policies based on these considerations will be documented and publicly available, including a character table for each permissible set of code points, before the registration of any IDN associated with such an aggregate may be accepted.
4. Permissible code points will not include: (a) line symbol-drawing characters (as those in the Unicode Box Drawing block), (b) symbols and icons that are neither alphanumeric nor ideographic language characters, such as typographic and

pictographic dingbats, (c) characters with well-established functions as protocol elements, (d) punctuation marks used solely to indicate the structure of sentences. (e) Punctuation marks that are used within words may only be permitted if they are not excluded by any of the preceding points, are essential to the language of the IDN registration, and are associated with explicit prescriptive rules about the context in which they may be used. (f) Under corresponding conditions, a single specified character may be used as a separator within a label, either by allowing the hyphen-minus to appear together with non-Latin scripts, or by designating a functionally equivalent punctuation mark from within the script.

When a pre-existing registered name requires a registry to make transitional exception to any of these rules, the terms of that action will be made readily available online. A registry may not even by exception permit code points that are prohibited by the IDN standards.

5. A registry will define an IDN registration in terms of both its Unicode and ASCII-encoded representations. The availability of a given Unicode sequence is currently determined by its encodability into the scheme defined in RFC 3491, and changes to that component of the IDN standard can have disruptive consequences for the operability of a Unicode name. Since the appearance of hyphens in the third and fourth positions of a label indicates an encoding scheme, the registration of any label containing hyphens in these positions must not be permitted unless the hyphens follow a two-letter designator for a sanctioned scheme and the label conforms to the corresponding specifications.

6. Top-level domain registries will work collaboratively with relevant stakeholders to develop IDN-specific registration policies, with the objective of achieving consistent approaches to IDN implementation for the benefit of DNS users worldwide. Top-level domain registries will work collaboratively with each other to address common issues, for example by forming or appointing a consortium to coordinate contact with external communities, elicit the assistance of support groups, and establish global fora.

7. Top-level domain registries will make definitions of what constitutes an IDN registration and associated registration rules available to the IANA Registry for IDN Tables. If material fundamental to the understanding of a registry's IDN policies is not published by the IANA, it will otherwise be made readily available online by the registry, which should also ensure that its registrars call the attention of prospective holders of IDN names to it.

8. The top-level domain registries should provide resources containing information about the sources and references that were used in the formation of the corresponding IDN registration policies for all languages and scripts in which they offer IDN registrations.

Additional remarks

The deceptive use of visually confusable characters from different scripts is discussed in detail in the Unicode Technical Report #36 on 'Unicode Security Conditions' at

<http://www.unicode.org/reports/tr36/> and in a draft Unicode Technical Report #39 at <http://www.unicode.org/reports/tr39/>. Limitations to the character repertoire available for IDNs are suggested in UTR#36 in tables presented under the heading “Data files”.

The current restriction of top-level labels to the 26-letter basic Latin alphabet makes it necessary to determine the language attributes of an IDN without consideration of the top-level label. The discussion that is in progress about permitting a more extensive character repertoire in top-level labels may change this, as well as raise need for guidelines specific to the new condition.