

Guidelines for the Implementation of Internationalized Domain Names | **Interim Version for Community Feedback**

This is an interim report drafted to get community feedback on the issues and corresponding recommendations being suggested by the IDN Implementation Guidelines WG.

1 Introduction

~~//Legacy text (to be edited): Internationalized Domain Names (IDNs) Implementation Guidelines (IDN Guidelines or the Guidelines) address the IDN registration policies and practices, designed to minimize the risk of cybersquatting and consumer confusion, and respect the interests of communities using local languages and scripts.~~

~~These guidelines are contractually binding for both registries and registrars offering generic Top Level Domains (gTLDs) with IDNs at the second level and recommended for IDN ccTLDs. For example, it part of the Registry Agreement (Specification 6 Clause 1.4) and 2013 Registrar Accreditation Agreement (Additional Registrar Operation Specification Clause 3) and through the Final Implementation Plan for the IDN ccTLD Fast Track Process.~~

This version of the Internationalized Domain Names (IDNs) Implementation Guidelines (“IDN Guidelines” or the “Guidelines”) reviews version 3.0 of the Guidelines following the expansion of the DNS under the 2012 New gTLD Program. The IDN Guidelines are written for gTLD registries ~~and registrars~~, however the IDN Guidelines are also intended as a support document for other registries establishing IDN policies ~~(e.g. ccTLDs)~~.

This document supersedes [version 3.0](#) of the Guidelines. It was prepared by members of the IDN Guidelines Working Group (IDNGWG) constituted following the [Call for Community Experts](#), and comprises of the following members:

	Name	Supporting Organization/ Advisory Committee
1	Satish Babu	ALAC
2	Wael Nasr	ALAC
3	Mats Dufberg	ccNSO
4	Pablo Rodríguez	ccNSO

	Name	Supporting Organization/ Advisory Committee
5	Edmon Chung	GNSO
6	Christian Dawson	GNSO
7	Chris Dillon	GNSO
8	Kal Feher	GNSO
9	Dennis Tan	GNSO
10	Jian Zhang	GNSO
11	Ram Mohan	SSAC
12	Patrik Fältström (will only review work)	SSAC

2 IDN Guidelines

2.1 Transition

IDNA2008 has been adopted by the registries and registrars offering IDNs at the second level. WG should identify and recommend how to address any residual issues from IDNA2003.

//current recommendation 1: Top-level domain ("TLD") registries supporting Internationalized Domain Names ("IDNs") will do so in strict compliance with the requirements of the IETF protocol for Internationalized Domain Names in Applications. The initial version of this protocol was defined in RFCs 3454, 3490, 3491, and 3492. A revised version is defined in RFCs 5890, 5891, 5892, 5893, and 5894. Both will be in parallel use in applications for an indeterminate transitional period but registries will conform fully with IDNA2008 in the shortest practicable order.

1. Top-level domain ("TLD") registries supporting Internationalized Domain Names ("IDNs") will do so in strict compliance with the requirements of the IETF protocol for Internationalized Domain Names in Applications, as defined in RFCs 5890, 5891, 5892, 5893, and 5894.

- 3.2. *//current recommendation 2: No code point permitted in IDNA2003 but disallowed in IDNA2008 will be accepted for registration regardless of the extent to which such code points appear in names registered prior to the protocol revision. The registrant of a*

domain that is no longer supported by IDNA2008 should be notified that there may be unanticipated consequences for a user attempting to reach it, and such names should be replaced, held, or deleted at registry initiative.

~~6.3.~~ ~~//current recommendation 7:~~ When a preexisting name requires a registry to make transitional exception to any of these Guidelines, the terms of that action will also be made readily available online, including the timeline for the resolution of such transitional matters. The excepted registrations themselves are, however, not part of this documentation. At the end of the transitional period, code points that are prohibited by IDNA2008 will not be permitted even by exception.

~~7.4.~~ ~~//current recommendation 8:~~ No label containing hyphens in the third and fourth positions will be registered unless it is a valid A-label, with reservation for transitional action in accordance with the preceding Guideline. Hyphens in these positions are explicitly reserved to indicate encoding schemes, of which IDNA is only one instantiation. These guidelines are not intended to assist with any other instantiations.

//Also relevant are Appendix A (Comparison of IDNA2003 with IDNA2008) and Appendix B (Additional transitional issues) of [version 3.0](#).

~~//suggested revision:~~

In the case of code points permitted in IDNA2008 but disallowed in IDNA2003, those allowed for use in the Root Zone, typically for scripts added in Unicode versions since 2003, should be allowed in labels at other levels. Code points, however, added to IDNA2008 for other reasons should generally be disallowed in the interests of a good user experience and respecting the Longevity Principle in the Procedure.

Commented [SH1]: Chris Dillon: The word “respecting” is meant to cause discussion. We may want to consider “in the spirit of” or a weaker form of words. This touches on an interesting issue — the Procedure is binding at the Top Level, but what is its relationship to the other levels?

Recommendation:

2.2 Terminology

There has been considerable terminology introduced through the work on Label Generation Rules, relevant RFCs and additional IDN work at ICANN for definition and adoption. These include, but are not limited to, the following: RFC [5564](#), [5890](#), [5891](#), [5892](#), [5893](#), [5894](#), [5895](#), [5992](#), [6912](#). Additional work includes the [Procedure](#) and [additional supporting documents](#) to develop the root zone LGR, the [User Experience Study](#) for IDN variant TLDs, the [Maximal Starting Repertoire \(MSR\)](#) and the [root zone LGR](#).

IDNGWG has identified the relevant terms and documented it in Appendix A.

Recommendation:

2.3 Format of IDN Tables

Based on work by the community, a formal machine readable [specification for representing IDN tables](#) (aka Label Generation Rules or LGR) is now available and being converted to a standards track RFC by IETF. This format should be encouraged for adoption at second level, as it is being done for Root Zone LGR.

//current recommendation 3: A registry will publish one or several lists of Unicode code points that are permitted for registration and will not accept the registration of any name containing an unlisted code point. Each such list will indicate the script or language(s) it is intended to support. If registry policy treats any code point in a list as a variant of any other code point, the nature of that variance and the policies attached to it will be clearly articulated.

Recommendation:

2.4 Consistency of IDN Tables

The content should be made more consistent across registries and across levels for predictable user experience. This could be done by sharing the LGRs across registries, considering reference IDN tables and other relevant work.

//current recommendation 4: All such code point listings will be placed in the IANA Repository for IDN TLD Practices in tabular format together with any rules applied to the registration of names containing those code points, before any such registration may be accepted.

//current recommendation 6: Any information fundamental to the understanding of a registry's IDN policies that is not published by the IANA will be made directly available online by the registry. The registry should also encourage its registrars to call attention to these policies for all prospective IDN registrants. This documentation will include references to the linguistic and orthographic sources used in establishing policies and code point repertoires. If material is provided both via the IANA and other channels the registry must ensure that its substance is concordant across all platforms.

//recommendation 9: TLD registries should collaborate on issues of shared interest, for example, by forming a consortium to coordinate contact with external communities, elicit the assistance of support groups, and establish global fora.

Recommendation:

2.5 IDN Variants

Nomenclature, states of variants and their management process should be made consistent. Relevant policies, e.g. ownership, automatic activation, ceiling value, choice between variants, etc. should be considered and appropriate recommendation should be forwarded.

Recommendation:

2.6 Similarity and Confusability of Labels

The different kinds of confusability of labels at the second level, arising from homoglyphs, cross-script homoglyphs, relevance of upper case, script mixing and other (e.g. semantic) mechanisms should be managed.

*//recommendation 5: All code points in a single label will be taken from the same script as determined by the Unicode Standard Annex #24: Script Names
<http://www.unicode.org/reports/tr24>. Exceptions to this guideline are permissible for languages with established orthographies and conventions that require the commingled use of multiple scripts. Even in the case of 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 is clearly defined.*

Recommendation:

2.7 Registration Data

WG to look into how to represent and manage registration data for IDNs and for variants of IDNs.

Recommendation:

2.8 EPP

WG to look into any recommendations for EPP, as raised by the community in ICANN 55.

Recommendation:

Appendix A: Glossary of Relevant Terms

Term	Acronym	Definition	Additional Notes	Other Related Terms
Writing System				
Whole Label Evaluation Rule	WLE Rule			
Variant Label Disposition				
Variant Label				
Variant Code Point Type				
Variant Code Point				
Variant		"Variant" is an ambiguous term, as it can refer to Variant Code Point or Variant Label , and therefore it should be further qualified whenever it is used.		
U-Label				
Script				
Punycode				

Term	Acronym	Definition	Additional Notes	Other Related Terms
Maximal Starting Repertoire	MSR			
LGR Specification				
Language				
Label: Reserved				
Label: Delegated				
Label: Blocked				
Label: Allocated				
Label: Allocatable				
Label: Activated				
Label Generation Ruleset / Label Generation Rules	LGR			
Label				
Internationalized Domain Names in Applications Protocol 2008	IDNA 2008			
Internationalized Domain Names in Applications Protocol 2003	IDNA 2003			

Term	Acronym	Definition	Additional Notes	Other Related Terms
Internationalized Domain Name Label	IDN Label			
Internationalized Domain Name	IDN	An "internationalized domain name" (IDN) is a domain name that contains at least one A-label or U-label, but that otherwise may contain any mixture of NR-LDH labels, A-labels, or U-labels.	As defined in RFC 5890	A-Label, U-Label
Homoglyph				
Glyph				
Cross-Script Variant Code Points		Variant code points across related scripts, e.g. U+0441 CYRILLIC SMALL LETTER ES 'c' and U+0063 LATIN SMALL LETTER C 'c'		
Code Point Repertoire				
Code Point				
A-Label		An "A-label" is the ASCII-Compatible Encoding form of an IDNA-valid string. It must be a complete label: IDNA is defined for labels, not for parts of them and not for complete domain names. This means, by definition, that every A-label will begin with the	As defined in RFC 5890	U-Label

Term	Acronym	Definition	Additional Notes	Other Related Terms
		IDNA ACE prefix, "xn--", followed by a string that is a valid output of the Punycode algorithm [RFC3492] and hence a maximum of 59 ASCII characters in length. The prefix and string together must conform to all requirements for a label that can be stored in the DNS including conformance to the rules for LDH labels. If and only if a string meeting the above requirements can be decoded into a U-label is it an A-label.		