## Work Document section 2.1.2 Criteria for the selection of an IDN ccTLD string

Version 02 – 10 November 2020

The ccNSO Prelimenary Review Team (PRT) conducted an analysis of the proposed overall IDN ccTLD policy as adopted by the ccNSO in 2013 (Basic Document) by comparing the proposed policy with current state of affairs under the Fast Track Process and also looking at other developments. The findings were reported per main section of the proposed overall policy:

- Section in Document. Reference to the specific section in the 2013 Board Report
   (https://ccnso.icann.org/sites/default/files/filefield 41859/idn-ccpdp-board-26sep13en.pdf),
- 2. **Topic.** Description of the topic as included in that Board Report,
- 3. **Comment/Rationale for review/inclusion in list**. The PRT comment and/or rationale for review and inclusion in the topics in the list, and
- 4. **Proposed next step.** The PRT advise to the Council on how to proceed to resolve the issues identified by the RT.

For consideration by the WG, with exception of the section in Board Report on confusingly similarity, the deliberation on Variant Management and the de-selection of IDN ccTLD strings, is to add two additional columns to the tables:

- 5. **Does wording needs to be adjusted?** Response would Y(es) or N(o)
- 6. **Proposed wording by WG.** The wording the WG proposes for the section. The proposed wording will be added to the Basic Document

As noted in the Issue report and Charter of the WG, the topics of confusingly similarity, Variant Management and the de-selection of IDN ccTLD strings are dealt with through sub-groups who are expected to develop their own working method and approach.

## Board report section 2.1.2 CRITERIA for the selection of an IDN ccTLD string

**A.** An IDN country code Top Level Domain must contain at least one (1) non-ASCII character. For example, españa would qualify under these criteria and *italia* would not. españa contains at least one other character other than [-, a-z, 0-9], while still being a valid top-level domain name.

A different way of expressing this is that the selected IDN ccTLD must be a valid U-Label that can also be expressed as an A-label. It cannot be a NR-I DH Label.

For more formal definitions of these terms, see RFC 5890.

B. Eligibility only if the name of territory listed on ISO 3166. To be eligible for a IDN ccTLD string, a country, territory, dependency or other area of particular geopolitical interest (hereafter referred to as: Territory or Territories) must be listed on the 'International Standard ISO 3166, Codes for the representation of names of countries and their subdivisions – Part 1: Country Codes', or, in some exceptional cases a two letter ASCII (letters a-z ) code associated with the Territory already assigned as a ccTLD and listed as an exceptionally reserved ISO 3166-1 code element<sup>1</sup>.

<sup>&</sup>lt;sup>1</sup> In exceptional cases code elements for Territory names may be reserved for which the ISO 3166/MA has decided not to include in ISO 3166 part 1, but for which an interchange requirement exists. See Section 7.5.4 ISO 3166 – 1 : 2006.

12

c. The IDN ccTLD string must be a Meaningful Representation of the name of a Territory. The principle underlying the representation of Territories in two letter (ASCII) code elements is the visual association between the names of Territories (in English or French, or sometimes in another language) and their corresponding code elements<sup>2</sup>.

The principle of association between the IDN country code string and the name of a Territory should be maintained. A selected IDN ccTLD string must be a meaningful representation of the name of the

Territory. A country code string is considered meaningful if it is:

- a) The name of the Territory; or
- b) Part of the name of the Territory that denotes the Territory; or
- c) A short-form designation for the name of the Territory, recognizably denoting the name.

Section in Document	Topic	Comment/Rationale for review/inclusion in list	Proposed next step	Adjust wording	Proposed Wording
2.1.2 C	The IDN ccTLD string must be a Meaningful Representation of the name of a Territory. The principle underlying the representation of Territories in two letter (ASCII) code elements is the visual association between the names of Territories (in English or French, or sometimes in another language) and their corresponding code elements. The principle of association between the IDN		The criteria need to be reviewed in depth in PDP.  Rationale: Proposed criteria have been adopted by the ccNSO Members in 2013.  They are very similar to the criteria used in the Fast Track Process.		

<sup>&</sup>lt;sup>2</sup> See ISO 3166-1: 2006 Section 5.1

1
2
3
4
5
6
7
8
9

Section in	Topic	Comment/Rationale for	Proposed next step	Adjust wording	Proposed
Document		review/inclusion in list			Wording
	country code string and the name	To what extent does the			
	of a Territory should be	selected IDN ccTLD string			
	maintained. A selected IDN ccTLD	need to be (remain?) to be			
	string must be a meaningful	recognised as a ccTLD even			
	representation of the name of the	if you do not know the			
	Territory. A country code string is	language?			
	considered meaningful if it is:				
	a) The name of the Territory; or b)				
	Part of the name of the Territory				
	that denotes the Territory; or c) A				
	short form designation for the				
	name of the Territory, recognizably				
	denoting the name.				

D. A Meaningful Representation of the name of the Territory must be in a Designated Language of the Territory The selected IDN ccTLD string should be a meaningful representation of the name of the territory in a "designated" language of that Territory. For this purpose, a "designated" language is defined as a language that has a legal status in the Territory or that serves as a language of administration <sup>3</sup> (hereafter: Designated Language)<sup>4</sup>.

The definition of Designated Language is based on: "Glossary of Terms for the Standardization of Geographical Names", United Nations Group of Experts on Geographic Names, United Nations, New York, 2002.<sup>5</sup>

<sup>3</sup> 

<sup>&</sup>lt;sup>4</sup> The limitation to Designated Language is recommended as criteria for reasons of stability of the DNS. According to some statistics currently 6909 living languages are identified. See for example: <a href="http://www.ethnologue.com/ethno-docs/distribution.asp?by=area">http://www.ethnologue.com/ethno-docs/distribution.asp?by=area</a>. If one IDN ccTLD would be allowed per territory for every language this would potentially amount to 252\*6909 or approximately 1.7 million IDN ccTLDs.

<sup>&</sup>lt;sup>5</sup> https://unstats.un.org/unsd/ungegn/pubs/documents/Glossary of terms rev.pdf

4 5

> 6 7

> 8

9 10

11 12

13 14

16 17

15

18 19

21 22

20

23 24

25

The language is considered to be a Designated Language if one or more of the following requirements are met:

- The language is listed for the relevant Territory as an ISO 639 language in Part Three of the "Technical Reference Manual for the standardization of Geographical Names", United Nations Group of Experts on Geographical Names (the UNGEGN Manual) (http://unstats.un.org/unsd/geoinfo/default.htm).
- 2. The language is listed as an administrative language for the relevant Territory in ISO 3166-1 standard under column 9 or 10. (Needs to be checked if still correct)
- The relevant public authority in the Territory confirms that the language is used in official communications of the relevant public authority and serves as a language of administration.

Specific requirements regarding documentation of Designated Languages are included in the procedures and documentation recommendations.

If the selected string is not the long or short form of the name of a Territory then evidence of meaningfulness is required. Where the selected string is the long or short form name of the relevant Territory in the Designated Language as listed in the UNGEGN Manual, Part Three column 3 or 4 version 2007, or later versions of that list it is considered to be meaningful.

Where the selected string is not listed in the UNGEGN then meaningfulness must be adequately documented. This is the case when:

- (i) The selected string is not part of the long or short form name of the Territory in the UNGEGN Manual in the Designated Language or
- (ii) An acronym of the name of the Territory in the Designated Language or
- (iii) the Territory or the Designated Language do not appear in the UNGEGN Manual.

If such documentation is required, the documentation needs to clearly establish that:

- The meaning of the selected string in the Designated Language and English and
- That the selected string meets the meaningfulness criteria.

Specific requirements regarding documentation of the Meaningful Representation are included in the procedures and documentation recommendations.

Section in Document	Topic	Comment/Rationale for review/inclusion in list	Proposed next step	Proposed Wording
2.1.2 E	If the selected string is not the	ICANN must make the	The criteria need to be	
	long or short form of the name of	"meaningfulness" criteria	reviewed in depth in PDP.	
	a Territory then evidence of	crystal clear as in the past		
	meaningfulness is required.	ICANN had inconsistent	Rationale: Proposed criteria	
	Where the selected string is the	approaches for the	have been adopted by the	
	long or short form name of the	evaluation of the "adequate	ccNSO Members in 2013.	
	relevant Territory in the	documentation". This	They are very similar to the	
	Designated Language as listed in	applies also to the case	criteria used in the Fast	
	the UNGEGN Manual, Part Three	when one territory has	Track Process.	
	column 3 or 4 version 2007, or	more than one designated		
	later versions of that list it is	language.		
	considered to be meaningful.	Furthermore, the procedure		
	Where the selected string is not	should foresee an appeal		
	listed in the UNGEGN then	step in case the selected		
	meaningfulness must be	string is not accepted		
	adequately documented [].	because of not being		
		"meaningful".		

F. Only one (1) IDN ccTLD string per Designated Language. In the event that there is more than one Designated Language in the Territory, one (1) unique IDN ccTLD for each Designated Language may be selected,

provided the meaningful representation in one Designated Language cannot be confused with an existing IDN ccTLD string for that Territory.

Where a language is expressed in more than one script in a Territory, then it is permissible to have one string per script, although the multiple strings are in the same language.

## **Notes and Comments**

It should be noted that other requirements relating to non-confusability are applicable and should be considered, including the specific procedural rules and conditions for cases when the same manager will operate two or more (IDN) ccTLD's which are considered to be confusingly similar.

Section in Document	Topic	Comment/Rationale for review/inclusion in list	Proposed next step	Adjust wording	Proposed Wording
2.1.2 F	Only one (1) IDN ccTLD string per Designated Language. In the event that there is more than one Designated Language in the Territory, one (1) unique IDN ccTLD for each Designated Language may be selected, provided the meaningful representation in one Designated Language cannot be confused with an existing IDN ccTLD string for that Territory. Where a language is expressed in more than one script in a territory, then it is permissible to have one string per script, although the	It is recommendable that any future IDN ccTLD policy addresses carefully – and with the support of linguist experts – the option of languages that are expressed in more than one script as well as the rules to be produced in case the same registry manages the ccTLD in ASCII and its variant in other script. At present, ICANN approach is not consistent and that may jeopardise the ultimate goal of ensuring the security and	The criteria need to be reviewed in depth in PDP and if deemed appropriate reconfirmed.  Rationale: Proposed criteria have been adopted by the ccNSO Members in 2013. They are very similar to the criteria used in the Fast Track Process.		

Section in Document	Topic	Comment/Rationale for review/inclusion in list	Proposed next step	Adjust wording	Proposed Wording
	multiple strings are in the same language.  Notes and Comments: It should be noted that other requirements relating to non-confusability are applicable and should be considered, including the specific procedural rules and conditions for cases when the same manager will operate two or more (IDN) ccTLDs which are considered to be confusingly similar.	stability of the DNS. Example mentioned is simplified Chinese and Mandarin			

G. The selected IDN ccTLD string should be non-contentious within the territory. The selected IDN ccTLD string must be non-contentious within the territory. This is evidenced by support/endorsement from the Significantly Interested Parties (relevant stakeholders) in the territory.

Concurrent requests for two strings in the same language and for the same territory will be considered competing requests and therefore to be contentious in territory. This needs to be resolved in territory, before any further steps are taken in the selection process.

Section in	Topic	Comment/Rationale for	Proposed next step	Adjust wording	Proposed
Document		review/inclusion in list			Wording
2.1.2 G	The selected IDN ccTLD string	ICANN must make sure	Ensure application of basic		
	should be noncontentious within	there is consistency	principle that IDN ccTLD		
	the territory. The selected IDN	between the delegation of	and ASCII ccTLD should be		
	ccTLD string must be non	an ASCI ccTLD and an IDN	treated similar		
	contentious within the territory.	ccTLD. Therefore,			
	This is evidenced by	contentious requests should			
	support/endorsement from the	be resolved in the territory.			
	Significantly Interested Parties				
	(relevant stakeholders) in the				
	territory. Concurrent requests for				
	two strings in the same language				
	and for the same territory will be				
	considered competing requests				
	and therefore to be contentious in				
	territory. This needs to be resolved				
	in territory, before any further				
	steps are taken in the selection				
	process.				

H. The selected IDN ccTLD string must abide by all Technical Criteria for an IDN TLD string. In addition to the general requirements for all labels (strings), the selected IDN ccTLD string must abide to the normative parts of RFC 5890, RFC 5891, RFC 5892 and RFC 5893.

All applicable technical criteria (general and IDN specific) for IDN ccTLD strings should be documented as part of the implementation plan. For reasons of transparency and accountability they should be made public prior to implementation of the overall policy and endorsed by the ccNSO.

8 9 Validation that a string meets the technical criteria is a process step and shall be conducted by an external, independent panel. The recommended procedure is described in Section 2.1.3, Processes and Documentation.

The method and criteria for the technical validation should be developed as part of the implementation plan and are a critical part of the review process. For reasons of transparency and accountability they should be made public prior to implementation of the overall policy and endorsed by the ccNSO.

Section in Document	Topic	Comment/Rationale for review/inclusion in list	Proposed next step	Adjust wording	Proposed Wording
2.1.2 H	In addition to the general requirements for all labels (strings), the selected IDN ccTLD string must abide to the normative parts of RFC 5890, RFC 5891, RFC 5892 and RFC 5893.  All applicable technical criteria (general and IDN specific) for IDN ccTLD strings should be documented as part of the implementation plan. For reasons of transparency and accountability they should be made public prior to implementation of the overall policy and endorsed by the ccNSO.	It need to be ensured that technical criteria are still valid. It will also need to be reviewed whether the proposed mechanism with respect to including the technical criteria as part of the implementation is appropriate.	The criteria need to be reviewed in depth in PDP and reconfirmed if deemed appropriate.  Rationale: Proposed criteria have been adopted by the ccNSO Members in 2013.		

1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	

15

16 17

- I. Confusing similarity of IDN ccTLD Strings. A selected IDN ccTLD string should not be confusingly similar with:
  - 22 Any combination of two ISO 646 Basic Version (ISO 646-BV) characters<sup>6</sup> (letter [a-z] codes), nor
  - Existing TLDs or Reserved Names as referenced in the new gTLD Applicant Guidebook<sup>7</sup>

The following supplemental rules provide the thresholds to solve any contention issues between the IDN ccTLD selection process and new gTLD process:

- A gTLD application that is approved by the ICANN Board will be considered an existing TLD unless it is withdrawn.
- A validated request for an IDN ccTLD will be considered an existing TLD unless it is withdrawn.

A selected IDN ccTLD string is considered confusingly similar with one or more other string(s) (which must be either Valid-U-labels or any a combination of two or more ISO 646 BV characters) if the appearance of the selected string in common fonts in small sizes at typical screen resolutions is sufficiently close to one or more other strings so that it is probable that a reasonable Internet user who is unfamiliar with the script would perceive the strings to be the same or confuse one for the other<sup>8</sup>.

<sup>&</sup>lt;sup>6</sup> International Organization for Standardization, "Information Technology – ISO 7-bit coded character set for information interchange," ISO Standard 646, 1991

<sup>&</sup>lt;sup>7</sup>Version 2012-06-04, section 2.2.1.2.1 Reserved Names.

<sup>&</sup>lt;sup>8</sup>Based on Unicode Technical Report #36, Section 2: Visual Security Issues

The review of whether or not a selected IDN ccTLD string is confusingly similar is a process step and should be conducted externally and independently. The recommended procedure is described in Section 2.1.3, Processes and Documentation.

The method and criteria to assess confusing similarity should be developed as part of the implementation planning. For reasons of transparency and accountability they should be made public prior to implementation of the overall policy and endorsed by the ccNSO.

The assessment of confusing similarity of strings depends on amongst other things linguistic, technical, and visual perception factors, therefore these elements should be taken into consideration in developing the method and criteria.

Taking into account the overarching principle to preserve and ensure the security, stability and interoperability of the DNS, the method and criteria for the confusing similarity assessment of an IDN ccTLD string should take into account and be guided by the Principles for Unicode Point Inclusion in labels in the DNS Root<sup>9</sup>.

## Notes and Comments

The rule on confusing similarity originates from the IDN WG and Fast Track Implementation Plan and was introduced to minimize the risk of confusion with existing or future two letter country codes in ISO 3166-1 and other TLDs. This is particularly relevant as the ISO 3166 country codes are used for a broad range of applications, for example but not limited to, marking of freight containers, postal use and as a basis for standard currency codes.

<sup>&</sup>lt;sup>9</sup>https://datatracker.ietf.org/doc/draft-iab-dns-zone-codepoint-pples/

The risk of string confusion is not a technical DNS issue, but can have an adverse impact on the security and stability of the domain name system, and as such should be minimized and mitigated.

The method and criteria used for the assessment cannot be determined only on the basis of a linguistic and/or technical method of the string and its component parts, but also needs to take into account and reflect the results of scientific research relating to confusing similarity, for example from cognitive neuropsychology<sup>10</sup>.

		_	
1	r		
	Г		

1

2

3

4

5

Section in Document	Topic	Comment/Rationale for review/inclusion in list	Proposed next step	Adjust wording	Proposed Wording
2.1.2	Confusing similarity of IDN ccTLD Strings.	environment and as domain name end-users/registrants are the same customers all over the internet eco-	The confusing similarity review procedures need to be reviewed in depth. It needs to be done under a ccNSO PDP and if feasible the ccNSO will need to coordinate with the GNSO work in this area.		

Simpson, Ian; Mousikou, Petroula; Montoya, Juan; Defior, Sylvia, Behavior Research Methods; June 2013, Vol. 45 Issue 2, p431

2. Alphabetic letter identification: Effects of perceivability, similarity, and bias. Shane Mueller, Cristoph Weidemann, Acta Psychologica 139, (2012)

<sup>&</sup>lt;sup>10</sup> See for example, M. Finkbeiner and M. Coltheart (eds), Letter Recognition: from Perception to Representation. Special Issue of the Journal *Cognitive Neuropsychology*, 2009. Academic research in the study of letter recognition, neuropsychology and cognition that formed basis for the Extended Process Similarity Review Panel (EPSRP):

<sup>1.</sup> A letter visual-similarity matrix for Latin-based alphabets,

Section in	Topic	Comment/Rationale for	Proposed next step	Adjust wording	Proposed
Document		review/inclusion in list			Wording
	Topic	-	Mechanism to coordinate is through a coordination committee/working party. This model is light weight, without the burden of a Cross-community working group. The results will feed into each of the PDPs. Membership is from the ccNSO PDP and GNSO PDP WG membership with membership open to interested groups.  Before becoming part of the ccNSO PDP WG proposals, the results need to be accepted by the full membership of the ccNSO PDP WG  Once included in the total package it shall be treated in same manner as other		· ·
			package it shall be treated		
			adoption).  Rationale: Confusing similarity review is		

Section in Document	Topic	Comment/Rationale for review/inclusion in list	Proposed next step	Adjust wording	Proposed Wording
			currently part of both of IDN ccTLD and new gTLD processes and policy proposals. Over time the methods, criteria and procedures have evolved differently.		
			The ccNSO Proposals are part of the original recommended policy and need to be updated through a ccNSO PDP.		

To date (March 2013) identifying the issues pertaining to the management of variant TLD's are still under

discussion by the community, in particular the delineation of technical, policy and operational aspects. For this reason, policy recommendations pertaining to the management of variant IDN ccTLDs, if any, are not included,

> Version 01 November 2, 2020

J. Variants PLACEHOLDER

but will be added at a later stage.

Section in Document	Topic	Comment/Rationale for review/inclusion in list	Proposed next step	Adjust wording	Proposed Wording
2.1.2 (J)	Variant Management To date (March 2013) identifying the issues pertaining to the management of variant TLD's are still under discussion by the community, in particular the delineation of technical, policy and operational aspects. For this reason, policy recommendations pertaining to the management of variant IDN ccTLDs, if any, are not included, but will be added at a later stage.		Mechanism to coordinate is through a coordination committee/working party. This model is light weight, without the burden of a Cross-community working group. The results will feed into each of the PDPs. Membership is from the ccNSO PDP and GNSO PDP WG membership with membership open to interested groups.  Before becoming part of the ccNSO PDP WG proposals, the results need to be accepted by the full membership of the ccNSO PDP WG  Once included in the total package it shall be treated in same manner as other proposals (subject to Council and Membership adoption).		