

Process and methodology for confusing similarity evaluation

Version 03- 10 October 2022

Introduction

The ccNSO proposed a two-step confusing similarity review in 2013. However over time the Fast Track process evolved further. The IDN Fast Track Process was updated in 2013, following completion of the ccPDP2, to include of the Extended Process Similarity Review Panel. In 2019 the Fast Track was again updated to include of the Risk Mitigation Measures Evaluation. This change was the result of the third review of the Fast Track Process.

For reference are included:

1. ccPDP4 Proposed Policy (page 2-6)
2. References
3. Annex A - Delineation document
4. Annex B - EPSRP Procedures
5. Annex C - Risk Mitigation Measures Evaluation Process.

Basic questions to initiate discussion about the evaluation process

What are pro and cons of each of the 3 steps of the Fast Track Confusing Similarity Review?

Which issue(s) are addressed by each step?

To address this question, suggested to understand and be aware of the pro's and con's of each step and overall process.

Which of the processes need to be included in the proposed policy?

Under assumption that confusing similarity review will be required, should it include both an evaluation and opportunity to suggest risk mitigation?

See:

and has to be conducted by external, independent body(ies), should it be a:

- One (1) step process? Original Fast Track process
- Two (2) step process: CS evaluation and Risk Mitigation or CS evaluation and review of CS evaluation.
- Three (3) step process: as currently under the Fast Track Process

Proposed Process and Method Confusing Similarity Evaluation PDP4

Goal and Standard Confusing Similarity Evaluation

- 1. Goal Confusing similarity review.** The goal of the confusing similarity review is to minimize the risk to the stability and security of the DNS due to user confusion by exploiting potential visual confusing similarity between domain names (eg. .PY in Latin script vs PY in Cyrillic) As such confusing similarity should therefore be minimized and mitigated. The risk of visual confusing similarity is not a technical DNS issue, but can have an adverse impact on the security and stability of the domain name system.

Notes and Observations

The rule on confusing similarity originates from the IDNC 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.

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 neuropsychology¹.

In SAC 060, SSAC advised ICANN (i.e the policy making bodies) that *should they decide to implement safeguards to deal with failing user expectations due to the introduction of variants, a distinction should be made between two types of failure modes: no-connection versus misconnection*"

No-connection may be a nuisance for the user, like a typo, however misconnection may result in the exploitation of the user confusion and this could be avoided through the similarity review.

With the introduction of variants one of the issues in the context of confusing similarity is to delineate the base for comparison, which is defined as the set of

¹

See for example,

- M. Finkbeiner and M. Coltheart (eds), Letter Recognition: from Perception to Representation. Special Issue of the Journal *Cognitive Neuropsychology*, 2009 and:
- Simpson, Ian; Mousikou, Petroula; Montoya, Juan; Defior, Sylvia, A letter visual-similarity matrix for Latin-based alphabets, *Behavior Research Methods*; June 2013, Vol. 45 Issue 2, p431
- Shane Mueller, Christoph Weidemann, Alphabetic letter identification: Effects of perceptability, similarity, and bias, *Acta Psychologica* 139, (2012)

The last two studies were used as basis for the review methodology of the Extended Process Similarity Review.

requested strings (Request Side) that will be compared with the set of potential visual confusingly similar strings (Comparison Side). Delineating the base for comparison is needed for reasons of :

- Scalability
- Avoiding unforeseen and/or unwanted side effects.

2. Standard for evaluation 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 other².

3. Base for comparison Confusing similarity of IDN ccTLD Strings. Under the ccNSO policy a Selected string, and its Requested Delegatable Variants should not be confusingly similar with:

- Any combination of two ISO 646 Basic Version (ISO 646-BV) characters (letter [a-z] codes), nor
- Existing TLDs, which includes the already delegated variants or reserved names.
- Proposed TLDs which are in process of string validation and their requested Delegatable or requested variants (however defined under the ccTLD and gTLD processes)

(From the 2013 policy document) 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.

NOTE; The base for comparison will need to be revisited after competition of CS Process and Methodology

The evaluation 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 **[update section number]**, Validation of IDNccTLD Strings

Stage 2: Validation of IDN ccTLD string

1. General description

The String Validation stage is a set of procedures to ensure all criteria and requirements

² Based on Unicode Technical Report #36, Section 2: Visual Security Issues

regarding the selected IDN ccTLD string (as listed in previous section of the Report) have been met. The actors involved would typically be:

- The IDN ccTLD string requester. This actor initiates the next step of this stage of the process by submitting a request for adoption and associated documentation.
- ICANN staff. ICANN staff will process the submission and coordinate between the different actors involved.
- Independent Panels to review the string (Technical, Similarity & Risk Mitigation Panels).

The activities during this stage would typically involve:

1. Submission of IDN table.
2. Submission of selected string and related documentation.
3. Validation of selected IDN ccTLD string:
 - a. ICANN staff validation of request. This includes
 - i. Completeness of request
 - ii.Completeness and adequacy of Meaningfulness and Designated Language documentation
 - iii. Completeness and adequacy of support from relevant public authority
 - iv. Completeness and adequacy of support from other Significantly Interested Parties
 - b. Independent Evaluations.
 - i. Technical review
 - ii.String Confusion review
4. Publication of selected IDN ccTLD string on ICANN website or notification to requester application was terminated

<snip>

b. Independent Evaluations and Reviews

General description of Technical and string confusion validation

- To validate a selected string is not confusingly similar, ICANN should appoint an external and independent “Similarity Evaluation Panel” to review the selected IDN ccTLD string for confusing similarity.
- To allow for a final validation review relating the confusing similarity, and only if so requested by the requester, ICANN should appoint, an external and independent “Extended Process Similarity Review Panel.”
- To allow for a review of risk mitigation measures if either or both the Similarity Evaluation Panel and EPSRP have found the requested string to be confusingly similar.

The goal of the validations is to provide external and independent advice to the ICANN Board whether a selected string and/or its requested delegatable variant(s) meet(s) the required technical criteria and is/are not considered to be confusingly similar.

If according to the definite outcome of the validation a selected string does not meet one or more of the technical criteria and/or is considered confusingly similar to another string, the requested IDNccTLD string is not eligible under this policy.

It is recommended that ICANN appoint the following external and independent Panels:

- To validate the technical requirements ICANN should appoint a “Technical Panel³” to conduct a technical review of the selected IDN ccTLD string.
- To evaluate a string for string similarity, an external and independent “Similarity Evaluation Panel” conducts a review of the requested IDN ccTLD string.
- To review the decision of the “Similarity Evaluation Panel” by using a different framework, an external and independent “Extended Process Similarity Review Panel” (hereafter: EPSRP) conducts a review of the requested IDN ccTLD string, only if so requested by the requester. Due to the specific nature of the confusing similarity and its subjective elements the “Extended Process Similarity Review:” is considered a specific review mechanism, not to be confused with the general ccTLD Review Mechanism.
- To allow for a review of risk mitigation measures if either or both the Similarity Evaluation Panel and EPSRP have found the requested string to be confusingly similar appoint a Risk Mitigation Evaluation Panel

Notes and observations

The details of the roles and responsibilities of the various panels and membership requirements and the details of the methods, procedures for evaluations and reviews by the respective panels should be developed as part of the implementation planning. It is noted that these details have been developed and tested under the IDNccTLD Fast Track Process and could be used as an example. The various details of EPSR and Risk Mitigation review are included in Annex B (EPSR) and Annex C (Risk Mitigation Evaluation).

Note that under the Fast Track Process the “Technical Panel” and “Similarity Evaluation Panel” were combined under the function of the DNS Stability Panel. Whether in future, under the ccPDP4 policy, the two Panels will be combined is a matter of implementation.

A. Process for Technical Validation

1. After completion of the ICANN staff validation of the request, ICANN staff will submit the selected IDN ccTLD string to the “Technical Panel” for the technical review.

³ Or any other name ICANN would prefer.

2. The Technical Panel conducts a technical string evaluation of the string submitted for evaluation. If needed, the Panel may ask questions for clarifications through ICANN staff.
3. The findings of the evaluation will be reported to ICANN staff. In its report the Panel shall include the names of the Panelists and document its findings, and the rationale for the decision.

Usually the Panel will conduct its review and send its report to ICANN staff within 30 days after receiving the IDN ccTLD string to be evaluated. In the event the Panel expects it will need more time, ICANN staff will be informed. ICANN staff shall inform the requester accordingly.

- 4 If according to the technical review the string meets all the technical criteria the string is technically validated. If the selected string does not meet all the technical criteria the string is not-valid. ICANN staff shall inform and notify the requester accordingly.

B. Process for confusing similarity validation

B.1 Similarity Evaluation Process.

B.1.1 After completion of the Technical Validation ICANN staff will submit the selected IDN ccTLD string to the String Similarity Panel for the confusing similarity string evaluation.

B.1.2 The Panel shall conduct a confusability string evaluation of the string submitted for evaluation. The Panel may ask questions for clarification through ICANN staff.

B.1.3 The findings of the evaluation will be reported to ICANN staff. In the report the Panel will include the names of the Panelists, document the decision and provide the rationale for the decision. Where the string is considered to be confusingly similar the report shall at a minimum include a reference to the string(s) to which the confusing similarity relates and examples (in fonts) where the panel observed the similarity.

ICANN staff shall inform and notify the requester accordingly.

Usually the Panel will conduct its review and send its report to ICANN staff within 30 days after receiving the IDN ccTLD string to be evaluated. In the event the Panel expects it will need more time, ICANN staff will be informed. ICANN staff shall inform the requester accordingly.

B.1.4 If according to the review, the Panel does not consider the string to be confusingly similar, the selected IDN ccTLD is validated.

B.2. Requested IDNccTLD string and confusing similar string refer to same Territory. If according to the review the selected IDN ccTLD string presents a risk of string confusion with a ccTLD string (see Base for Comparison above) and this (variant) ccTLD string is associated with the same Territory as represented by the selected IDNccTLD or requested delegatable variant IDNccTLD string(s), this should be noted in the report. ICANN staff shall inform the requester accordingly.

If, within 3 months of receiving the report the requester shall confirm that:

- (i) The intended manager and intended registry operator for the IDN ccTLD and the ccTLD manager for the confusingly similar country code are one and the same entity; and
- (ii) The intended manager of the IDN ccTLD shall be the entity that requests the delegation of the IDN ccTLD string; and
- (iii) The requester, intended manager and registry operator and, if necessary, the relevant public authority, accept and document that the IDN ccTLD and the ccTLD with which it is confusingly similar will be and will remain operated by one and the same manager, and
- (iv) The requester, intended manager and registry operator and, if necessary, the relevant public authority agree to specific and pre-arranged other conditions with the goal to mitigate the risk of user confusion as of the moment the IDN ccTLD becomes operational;

then the IDN ccTLD string is deemed to be valid.

If either the requester, intended manager or the relevant public authority do not accept the pre-arranged conditions within 3 months after notification or at a later stage refutes the acceptance, the IDN ccTLD shall not be validated.

Alternatively, the requester may defer from this mechanism and use the procedure as described under C or D.

C. Similarity Review Process

C.1 If according to the review the selected IDN ccTLD string is found to present a risk of string confusion, ICANN staff shall inform the requester in accordance with paragraph 3 above. The requester may call for an Extended Process Similarity Review and provide additional documentation and clarification referring to aspects in the report of the Panel. The requester should notify ICANN within three (3) calendar months after the date of notification by ICANN, and include the additional documentation. After receiving the notification from the requester, ICANN staff shall call on the Extended Process Similarity Review Panel (EPSRP).

C.2 The EPSRP conducts its evaluation of the string, based on the standard and methodology and criteria developed for it, and, taking into account, but not limited to, all the related documentation from the requester, including submitted additional documentation, IDN tables available, and the finding of the Similarity Review Panel. The EPSRP may ask questions for clarification through ICANN staff.

C.3 The findings of the EPSRP shall be reported to ICANN staff and will be publicly announced on the ICANN website. This report shall include and document the findings of the EPSRP, including the rationale for the final decision, and in case of the risk of confusion a reference to the strings that are considered confusingly similar and examples where the panel observed this similarity.

If according to the Extended Process Similarity Review, the EPSRP does not consider the string to be confusingly similar the selected IDN ccTLD is valid.

D. Risk Mitigation Evaluation Process.

- D.1 In the event that the SEP and/or EPSRP determines that a requested (delegatable variant) IDNccTLD string is confusingly similar the relevant Panel shall document this finding in its report to ICANN (Findings Report).
- D.2 If, at the time of the request or within three months after receiving the notification of the Findings Report, the requester, and, if considered necessary by ICANN, the relevant public authority, provide(s) a clarification that demonstrates and documents to ICANN that:
- D.3 The intended manager shall propose, agree upon and implement adequate pre-arranged risk mitigation measures with the goal to reduce the potential risk of user confusion as of the moment the IDNccTLD becomes operational, including specific consideration of confusability from the perspective that any domain name may be displayed in any case (lower- or upper-case), depending on the software application and regardless of the user's familiarity with the language or script.
- D.4 If these measures are agreed upon by the time the delegation request of the IDN ccTLD string is submitted then the requested string is deemed to have passed the string evaluation.
- D.5 If the intended IDN ccTLD manager does not propose mitigation measures or does not implement the agreed upon risk mitigation measures sufficiently within the timeline described above, the Termination Procedure will be initiated.

To determine whether the proposed risk mitigation measures are adequate ICANN will consult experts in the area of relevant Risk Mitigation measures and the IDN ccTLD string requestor. The proposed measures are to be evaluated together with the finding of the confusability evaluation.

References & Background material

- Guideline EPSRP: <https://www.icann.org/en/system/files/files/epsrp-guidelines-04dec13-en.pdf>
- Guideline Risk Mitigation Measures Evaluation:
<https://www.icann.org/en/system/files/files/guideline-risk-mitigation-measures-evaluation-28mar19-en.pdf>
- EPSRP and Risk Mitigation Reports for IDN ccTLD Applications:
<https://www.icann.org/resources/pages/epsrp-reports-2014-10-14-en>
- Joint ccNSO SSAC Response to ICANN Board (on introduction of Risk Mitigation)
<https://ccnso.icann.org/sites/default/files/field-attached/epsrp-final-response-17aug17-en.pdf>

Annex A - Delineating Confusing Similarity

Introduction

At its last meeting, the CS sub-group discussed the scope of the base for comparison for the confusing similarity review. The discussion focused on which variants, if any, to include in the comparison to assess possible confusing similarity of requested strings. Basicly the group started with assumption that the review should be based on:

- On the submission/ request side:
 - the requested label (level 1) and all allocatable variants (level 2).
- On the other side it would include:
 - 1 Any combination of two ISO 646 Basic Version (ISO 646-BV) characters⁴ (letter [a-z] codes),
 - 2 Existing TLDs or reserved names, their allocatable (level 2) and blocked variants (level 3), and
 - 3 Proposed TLDs which are in process of string validation, their allocatable (level 2) and blocked variants (level 3)

The scope will need to be revisited again. Starting point of this discussion is the goal of the confusing similarity review.

Goal Confusing similarity review

The agreed upon goal of the confusing similarity review is to minimize **the risk to the stability and security of the DNS due to user confusion by exploiting potential visual confusing similarity between domain names (eg. .PY in Latin script vs PY in Cyrillic)** As such confusing similarity should therefore be minimized and mitigated. The risk of visual confusing similarity is not a technical DNS issue, but can have an adverse impact on the security and stability of the domain name system.

In SAC 060, SSAC advised ICANN (i.e the policy making bodies) that *should they decide to implement safeguards to deal with failing user expectations due to the introduction of variants, a distinction should be made between two types of failure modes: (no-connection) versus misconnection.*

- **No-Connection (Denial of Service):** *the user attempts to visit http://example.Y, reading it as being the same Uniform Resource Identifier (URI) as the http://example.X that, for example, he or she saw in an advertisement, but the connection does not work (lookup fails) because Y is either blocked, withheld, or X has no variant at all, and example.Y is not registered.*
- **Misconnection:** *the user attempts to visit http://example.Y, reading it as being the same URI as the http://example.X that, for example, he or she saw in an advertisement, but arrives at a site controlled by a registrant different to that of example.X.*

⁴ International Organization for Standardization, "Information Technology – ISO 7-bit coded character set for information interchange," ISO Standard 646, 1991

In case of no-connection, the user is frustrated and may conclude that “the Internet does not work,” but no serious harm has arisen.

From a risk perspective: although there is a possibility (p_1) of confusion (C), there is no harm (H) nor potential (p_2) harm. The overall estimated impact of the risk is therefore zero [$p_1*C*p_2*0(=H)$].

The second case is problematic even if this effect is not the result of malicious work on the part of Y’s operator or example.Y registrant. Misconnections to a perfectly legitimate site operating at example.Y present issues of possible credential compromise or other accidental disclosure of information in addition to user confusion and frustration.

From a risk perspective: there is not only a possibility (p_1) of confusion (C), there is also a potential (p_2) harm (H) to be associated with the confusing similarity. The overall estimated impact of the risk is therefore not zero [p_1*C*p_2*H , whereby $H>0$] and should be avoided.

Under the evolution of the Fast Track Process a joint ccNSO-SSAC working group⁵ noted that in dealing with risks associated with confusing similarity *there is no general hard and fast rule with respect to the mitigation measures that should be implemented or with respect to the acceptable level of risk. It all depends very much on the circumstances, context and interplay of proposed measures and current and future risks associated with the confusing similarity of proposed strings.*

Linking these two risk categories to the goal of the confusing similarity review

Visual similarity is relevant for those situations where as a result of visual similarity a user does not connect or misconnects. In line with the SAC060 distinction between No-connection and Misconnection a distinction should be made whereby as a result of visual confusion no-connection or a misconnection is established.

No-connection may be a nuisance for the user, like a typo, however misconnection may result in the exploitation of the user confusion and this could be avoided through the similarity review.

Scope of comparison

Taking into account the goal of the confusing similarity review, **minimize the risk to the stability and security of the DNS due to user confusion by exploiting potential visual confusing similarity between domain names (eg. .PY in Latin script vs PY in Cyrillic)** the confusing similarity review is limited to avoid misconnection resulting from visual similarity of strings.

With the introduction of variants one of the issues in the context of confusing similarity is to delineate the base for comparison, which is defined as the set of requested strings (Request

⁵ <https://ccnso.icann.org/sites/default/files/field-attached/epsrp-final-response-17aug17-en.pdf>

Side) that will be compared with the set of potential visual confusingly similar strings (Comparison Side)

As a result of the introduction of variants, the potential scope of the Base for Comparison will expand exponentially. For example, as part of the confusing similarity review a selected IDNccTLD string needs to be compared with the string “Pakistan” in the Arabic script. As a result of introducing the comparison could expand to over 1200 strings (including all allocatable and blocked variants of “Pakistan” in the Arabic script). Therefore delineating the base for comparison is needed for reasons of :

- Scalability:
 - Be able to scale the review appropriately. It is expected that for the upcoming years, confusing similarity reviews have to be done manually.
 - Without proper limitation, the review may become too resource intensive and/or long in duration, which may additional issues, for example around predictability.
- Avoiding unforeseen and/or unwanted side effects.
 - If the full set of blocked variants of a would be included in the Comparison Side, a requested selected IDNccTLD could be “invalid” and further processing terminated although the variant string included in the Compare Side is from another script, and co-mingling of scripts is not allowed. In other words, the comparison may include strings/labels, which are not allowed under policy.
 - If a string includes or is comprised of or contains blocked variants it will never be delegated.

Comparison Side. To assess confusing similarity of strings the requested strings needs to be compared with and should not be visual similar to other strings (Comparison Side) that would include visual comparable strings from the following set:

- Any combination of two ISO 646 Basic Version (ISO 646-BV) characters⁶ (letter [a-z] codes), nor
- Existing TLDs or reserved names.
- Proposed TLDs which are in process of string validation.

Delineating Scope of Request Side

The primary question to determine the scope of the Request Side Question:

Which set of variants should be taken into request side of the base for comparison?

1. Only the selected string and the requested delegatable variants?
2. The selected string and **all delegatable variants**?
3. The selected string and **all allocatable variants** of the selected string, or
4. The selected string and **all variants (allocatable and blocked)**.?

Proposed Request Side. The proposed policy the request side for the Base for Comparison is comprised of the:

⁶ International Organization for Standardization, "Information Technology – ISO 7-bit coded character set for information interchange," ISO Standard 646, 1991

- Selected string, and
- Requested delegatable variants (only those allocatable variants, which are a meaningful representation of the name of the territory in the designated language and related script and requested at the time of submission of the request)

Rationale

1. The IDN selection process is open and ongoing. Variants may be requested any time as long as they meet all criteria, including meaningfulness.
2. The focus should be minimizing the risk of Misconnection to minimize and/or mitigate harm.
Abstracting from variants, if the selected string “X X” is considered confusingly similar with the string “xx”, which belongs to the pool of:
 - Any combination of two ISO 646 Basic Version (ISO 646-BV) characters⁷ (letter [a-z] codes),
 - Existing TLDs or reserved names.
 - Proposed TLDs which are in process of string validation
 The potential misconnection results from this confusing similarity between “X X” and “xx” and for that reason “X X” is deemed to be invalid and processing under the policy will end.
3. From a technical point of view the selected sting “X X” and its delegatable variants should be viewed as separate TLDs. Therefore each of the requested strings should be reviewed on confusing similarity.
4. As IDNccTLD process is open and at a later stage additional variant strings may be requested (for example variants of already delegated IDNccTLD under the Fast Track process). Each of these requested variants of an already delegated selected string, should be reviewed at its own merits with respect to confusing similarity.

Delineating Scope of Comparison Side.

Re-iterating, the goal of the confusing similarity review is to minimize **the risk to the stability and security of the DNS due to user confusion by exploiting potential visual confusing similarity between domain names** or to paraphrase in terms of SAC 060 (*Examining the User Experience Implications of Active Variant TLDs*) the goal is to minimize the risk of Misconnection due to visual confusability of two strings.

The minimum level of the Comparison Side, before the introduction of variants, includes:

- 4 Any combination of two ISO 646 Basic Version (ISO 646-BV) characters⁸ (letter [a-z] codes), nor
- 5 Existing TLDs or reserved names.
- 6 Proposed TLDs which are in process of string validation.

⁷ International Organization for Standardization, "Information Technology – ISO 7-bit coded character set for information interchange," ISO Standard 646, 1991

⁸ International Organization for Standardization, "Information Technology – ISO 7-bit coded character set for information interchange," ISO Standard 646, 1991

After the introduction of the variants, the minimum set of strings in the Comparison Side, could be defined as:

- 7 Any combination of two ISO 646 Basic Version (ISO 646-BV) characters⁹ (letter [a-z] codes), nor
- 8 Existing TLDs, which includes the already delegated variants or reserved names.
- 9 Proposed TLDs which are in process of string validation and their requested delegatable or requested variants (however defined under the ccTLD and gTLD processes)

In other words, all strings that:

1. Should never be delegated under any existing policy (the reserved names),
2. Should always be delegatable because of other existing policy (ASCII two-letter country-code TLDs, RFC 1591)),
3. Have been delegated (existing TLDs and their delegated variants), and
4. Are in the process of validation at the time the request for the selected IDNccTLD and its requested delegatable variants was submitted. This would include the variants of the selected IDNccTLD strings and new gTLD labels and their requested variants.

Secondly, all allocatable variants could be included of all already delegated TLDs, and those which are in process.

Although, by definition allocatable variants may be requested at a later stage. The allocatable variants will need to be reviewed against all criteria, including confusing similarity and meaningfulness if they are to be delegated. By including all allocatable variants in the comparison side, the confusing similarity review could become a reservation system. Allocatable variants, which have not been requested and may never be requested could block the introduction and delegation of a selected IDNccTLD.

And again, the goal of the confusing similarity review is to minimize risk of misconnection, and therefore avoid that a requested string is potentially delegated. The goal is not to minimize or avoid Denial of Service or Non-Connection.

With respect to including the blocked variants. The arguments to exclude all allocatable variants apply even in a stronger sense.

In summary: Under the ccNSO policy a Selected string, and its Requested Delegatable variants should not be confusingly similar with:

- 10 Any combination of two ISO 646 Basic Version (ISO 646-BV) characters (letter [a-z] codes), nor
- 11 Existing TLDs, which includes the already delegated variants or reserved names.
- 12 Proposed TLDs which are in process of string validation and their requested delegatable or requested variants (however defined under the ccTLD and gTLD processes)

⁹ International Organization for Standardization, "Information Technology – ISO 7-bit coded character set for information interchange," ISO Standard 646, 1991

Annex B – Extended Process Similarity Review

Extended Process Similarity Review Panel

Introduction

As part of the DNS Stability Evaluation external and independent advice to the ICANN Board is provided whether a selected string is not confusingly similar to other existing or applied for TLDs. If according to the DNS Stability Evaluation the selected string is considered confusingly similar to another string, the request for the IDN ccTLD with that particular selected string is not eligible under the Fast Track Process.

To evaluate potential similarity, the DNS Stability Evaluation includes the following evaluation Panel:

- To evaluate a string for string similarity, an external and independent “Similarity Review Panel” conducts a review of the requested IDN ccTLD string.
- To evaluate a string for string similarity If a selected string is found to be confusingly similar by the “Similarity Review Panel”, an external and independent “Extended Process Similarity Review Panel” (hereafter: EPSRP) conducts a review of the requested IDN ccTLD string second panel, using a different framework, and, only if so requested by the requester.

The EPSRP shall review the requested string(s) on the basis of the framework described below, with a clear focus on the overarching principle to preserve and ensure the security, stability and interoperability of the DNS.

Extended Proces Similarity Review Procedure

The EPSRP can be requested to conduct a second and final confusing similarity assessment of the requested IDN ccTLD string if:

- 1) The DNS Stability Panel, in performing its string similarity review, deems the string to be invalid; and
- 2) If the requester seeks review by the EPSRP within three (3) months of ICANN’s notification of the DNS Stability Panel’s determination.

Transitional arrangement: If an IDN ccTLD string request submitted under the Fast Track Process is still in process or has been terminated due to non-validation of the string per confusing similarity criteria, the requester has the option to request a second and final validation review by the Extended Process Similarity Review Panel. This option is available to the requester within three (3) calendar months of the date when the EPSRP is appointed. Requesters who fall in this category will be notified by ICANN staff of their eligibility for this process when the panel has been seated.

If ICANN is not notified within three (3) calendar months after the date of notification by ICANN of DNS Stability Panel findings, or under the transitional arrangement within three (3) months of the date the EPRSP is appointed, the Fast Track Termination Process will be initiated (See section 5.4. of the Implementation Plan).

The requester may call for the second and final Extended Process Similarity Review by sending a request to <insert address> (INCLUDE SAMPLE?). Additional documentation and clarification, if any, referring to aspects in the report of the DNS Stability Panel may be also provided. The additional material should be send to: <insert address>.

After receiving the notification from the requester, ICANN shall call on the EPSRP.

Within one (1) month after receiving the notification from ICANN staff, the EPSRP will request the external research team to measure similarity and confusability of the selected IDN ccTLD string to similar and dissimilar comparison letter strings, taking into account the documentation provided. The request will include at a minimum the font and font size conditions.

The EPSRP conducts its evaluation of the string based on the methodology and criteria described below, and, taking into account, but not limited to:

- All the related documentation from the requester, including submitted additional documentation,
- IDN tables and
- The findings of the DNS Stability Panel.

During the evaluation process the EPSRP may seek further clarification from the requester through ICANN staff, if deemed necessary.

The findings of the EPSRP shall be reported to ICANN and will be publicly announced on the ICANN website. This report shall include and document the findings of the EPSRP, including:

- The final decision
- The rationale for the final decision.

In case the string is deemed to be invalid the report shall also include:

- A reference to the strings that are considered confusingly similar and
- Examples where this similarity was noted.
- Report of the external research team.

If according to the EPSRP the selected IDN ccTLD string is valid on string similarity grounds, the requester is notified by ICANN staff that the DNS Stability Evaluation has successfully been completed and that the requested string(s) will be queued for public posting.

Methodology and criteria

A selected IDN ccTLD string should not be confusingly similar with:

- Any combination of two ISO 646 Basic Version (ISO 646-BV) characters¹⁰ (letter [a-z] codes), nor
- Existing TLDs or reserved names.

As stated in the proposed IDN ccTLD policy, the rule for confusing similarity is that if the appearance of the selected string, in both upper and lower case, in common fonts in small

¹⁰ International Organization for Standardization, "Information Technology – ISO 7-bit coded character set for information interchange," ISO Standard 646, 1991

sizes at typical screen resolutions, is sufficiently close to one or more other strings, it is probable that a reasonable Internet user who is unfamiliar with the script perceives the strings to be the same or confuses one for the other¹¹.

In order to determine whether this is the case in particular for the two letter codes, under the Fast Track Process, the EPSRP will establish whether a selected IDN ccTLD string is too similar to another to recommend acceptance, based on a behavioral metric that objectively measure the visual similarity of a candidate IDN ccTLD strings to other letter strings, and in particular the reserved 2-letter ISO3166-1 country codes. The behavioral metric provides quantitative and statistical evidence about the likelihood of confusing two possible IDN ccTLDs and its methods are open and repeatable to enable replication by third parties¹².

An external and independent research team will provide the behavioral metric relating to the selected IDN ccTLD string under evaluation by the EPSRP. The metric itself is a combined metric derived from three (3) different measuring methods to assess similarity:

- Subjective Rating Task: Participants judge on a multi-point scale the visual similarity of two letter strings. Although this is necessarily a subjective measure, the outcomes from such ratings can be very reliable within and between raters, and this can easily be treated as a numerical scale.
- Delayed Match to Sample / 2-AFC: Participants in the test are shown a stimulus, which later must be selected from a set of options. When only two options are given, this is sometimes referred to as a two-alternative forced choice (2-AFC) task.
- Visual Search Task: Participants search for and identify a stimulus either by matching a target or miss-matching the rest of the stimuli in a field of text strings.

Panelists Extended Process Similarity Review Panel (Initially include a placeholder)

Research Team

¹¹ Based on Unicode Technical Report #36, Section 2: Visual Security Issues

¹² This takes into account the latest literature in study of letter recognition, neuropsychology and cognition for example:

A letter visual-similarity matrix for Latin-based alphabets,
Simpson, Ian; Mousikou, Petroula; Montoya, Juan; Defior, Sylvia,
Behavior Research Methods; June 2013, Vol. 45 Issue 2, p431

Alphabetic letter identification: Effects of perceptability, similarity, and bias.
Shane Muleler, Cristoph Weidemann, **Acta Psychologica** 139, (2012)

Annex C - Risk Mitigation Evaluation Procedure

1. Introduction

As per proposed policy, a requested IDN ccTLD string should not be confusingly similar with (i) any Reserved Name, existing TLDs (both ccTLDs and gTLDs) or potential future TLDs to avoid risk associated with “misconnection” (see Annex A above).

To evaluate possible confusing similarity, ICANN has appointed the following two panels:

- **Similarity Evaluation Panel (SEP).** The DSP conducts the initial DNS Stability Evaluation, which includes a string similarity review of the requested IDN ccTLD string.
- **Extended Process Similarity Review Panel (EPSRP).** The EPSRP conducts a review of the requested IDN ccTLD string for contention cases identified by DSP upon the request of the requester, using the same criteria but with a different methodology from DSP¹³.

The process description includes the evaluation of mitigation measures to reduce risks associated with confusingly similarity of TLD strings. This describes the process on how to propose and review mitigation measures.

2. High level overview Risk Treatment Appraisal Process

At the request of the requester of an IDN ccTLD string and under the eligibility conditions of this guideline, the Risk Treatment Appraisal Process Panel (RTAP Panel) will need to be satisfied that the requester has followed an appropriate risk management process and adequate, related risk mitigation measures.

Should the RTAP Panel have concerns as to the adequacy of the proposed risk management process or the proposed mitigation measures, the RTAP Panel will communicate with ICANN and the requester during the process to understand the objective and the Risk Mitigation Proposal (RMP), and the requester may provide additional information and clarification.

3. Conditions for Application of these Guidelines

In accordance with the proposed ccPDP4 procedure and under the following limited set of conditions, a requester is eligible to propose measures to mitigate the risk associated with confusing similarity:

- If the DSP or EPSRP evaluation have determined that the requested string is confusingly similar in uppercase only.
- The requester has filed a request for a review of its proposed mitigation measures within three months from the date the results from the DRP and/or EPSRP have been communicated to the requester or, if at a later date, within 3 months after the date at which this guideline becomes effective.

¹³. Following the methodology in its guidelines, for the scripts which are bicameral the EPSRP provides separate recommendations for uppercase and lowercase versions of the applied-for IDN ccTLD strings given that from a visual similarity point of view, uppercase and lowercase characters of the same letter are distinct entities (see for example: <https://www.icann.org/en/system/files/files/epsrp-greece-30sep14-en.pdf>)

- In the request for a review of proposed mitigation measures, the requester has included - at a minimum – a reference to the proposed, internationally recognized and appropriate risk management and mitigation process the requester intends to use, and the related, proposed mitigation measures.
- The requester commits to implement the proposed and agreed upon mitigation measures as of the moment the IDN ccTLD becomes operational.

If the above conditions are met, the review and evaluation of the proposed methodology and related mitigation measures shall be undertaken by an independent panel (the ‘RTAP Panel’), appointed by ICANN.

The RTAP Panel shall evaluate the proposed risk management process and related risk mitigation measures to assess whether the risks of confusing similarity identified through the evaluation or review has been mitigated.

4. Objective and Criteria for Review of Risk Mitigation Measures

The mitigation measures proposed in the RMP should meet the objective of Risk Mitigation Measures and the criteria for review of Risk Mitigation Proposal.

The requester should make clear how the risk management process and proposed mitigation measures contained in the RMP meet the objective and criteria and should be evaluated together with the confusability findings.

The residual level of risk, if any, due to the confusability of domain names is expected to be in the same range as which would occur by adding another IDN ccTLD which has not been found similar to existing or reserved TLD.

4.1 The Objective of the Review of Risk Mitigation Measures

The objective is to determine if the risk will be effectively mitigated, as per the statement below:

If a requested string has been found to be confusingly similar with the upper case version of other strings, the proposed mitigation measures should reduce the risks associated with the confusing similarity to an acceptable level or threshold. The proposed mitigation measures should be evaluated in relation to the strings identified by the relevant panel (DSP or EPSRP) as confusingly similar to the applied-for string. In accordance with the IDN ccTLD Implementation Plan, the RTAP Panel should consider the likelihood of confusing similarity with specific consideration of confusability from the perspective that any domain name may be displayed in either upper- or lower-case, depending on the software application and regardless of the user’s familiarity with the language or script.

4.2 The Criteria for assessing the risk mitigation measures

1. **Proportionate:** The mitigation measures will be in proportion to risks identified. The higher the risks, the greater the mitigation measures will be required; conversely, lower mitigation measures will be a proportionate response to risks that are identified as low severity or low likelihood,
2. **Adequate:** For each of the case(s), the measures should reduce the risk of user confusion arising from the potential use of the applied-for TLD to an acceptable level.

The residual level of risk, if any, due to the confusability of domain names is expected to be in the same range as which would occur by adding another IDN ccTLD which has not been found similar to existing or reserved TLD.

3. **Self-contained:** The proposed mitigation measures can only apply to the registration policies of the applied-for TLD and do not assume any restrictions on the availability or registration policies of other current or future TLD labels.
4. Global Impact: The proposed mitigation measures must have global applicability, and not apply to confusability within the intended user community only.

5. Risk Treatment Appraisal Process Panel (RTAP Panel)

Effective risk analysis and mitigation require expertise in the area of risk management and risk management processes and procedures. To guide the discussion and coordinate the assessment work and given the paramount nature of this kind of expertise, at least one person on the panel should be a recognized expert in this area. The RTAP Panel members shall appoint one of their members to be the chair of the RTAP Panel.

The team doing the risk analysis should also include persons who are 1. considered experts in the area of internationalized domain names and how related registration policies are implemented by the registries (to review the practicality of implementing the RMP), 2. how IDNs may be confusing, to what extent such confusion can cause harm and how such confusion and harm could be prevented.

Therefore, the RTAP Panel will have three (3) to five (5) members, ensuring all the following requirements/skill sets are represented:

- Expertise in and understanding of various risk mitigating processes and standards and risk mitigation practices.
- Expertise on IDN implementation by registries, good understanding of the implementation opportunities and challenges for different IDN policies at the second and other levels, and knowledge of the relevant security and technical standards relating to IDNs.
- Expertise in brand protection, trade mark law and domain name disputes pertaining to the use of domain names as instruments for phishing and other sorts of abusive use, their impact and measures to address them.
- Expertise in the relevant language(s)/scripts.

ICANN organization convenes the RTAP Panel to review the anticipated RMP.

The names of the members of the RTAP Panel will be listed on the ICANN Website as soon as possible following their appointment, and included in the report.

6. Risk Treatment Appraisal Procedure

1. Requester submits the RMP within three (3) months after receiving the communication of the string similarity review decision¹⁴

¹⁴ For applications in the process before the implementation of these guidelines, this period will start from the date of publishing of the announcement that these guidelines are applicable.

2. ICANN organization convenes the RTAP Panel, and forwards RMP to RTAP Panel within one (1) week of the formation of the RTAP Panel
3. The RTAP Panel creates a review plan within three (3) weeks for the completion of the work, which includes at a minimum:
 - a. Tentative work plan and timeline
 - b. Request, if any, for additional information which may be needed or helpful
4. ICANN organization reviews the RTAP Panel's evaluation plan, and informs the requester of the timeline and any additional information needed.
5. Requester considers the review plan and shares any feedback, and additional information requested with respect to the RMP, and any other information considered necessary and /or relevant as soon as possible and confirms whether to proceed with the RTA. If the confirmation is not received within eight (8) weeks of receiving the review plan, the application is closed
6. ICANN organization forwards the updates with respect to the RMP, if any, to RTAP Panel, within one (1) week of receiving it.
7. RTAP Panel undertakes analysis of the RMP. ICANN organization coordinates any additional interaction between RTAP Panel and requester with respect to any clarifying question RTAP Panel may have or additional information the requestor intends to provide with respect to the RMP.
8. The RTAP Panel creates and hands over to ICANN organization a first RTA-Interim Report within eight (8) weeks of receiving the requester's confirmation to proceed with the RTAP,
9. ICANN organization passes RTA-Interim Report to the requester within one (1 week) of receiving it.
10. Requester submits its response and any additional information it considers relevant on the RTA-Interim Report and updated RMP (if at all) to ICANN organization within four (4) weeks of receiving the RTA-Interim Report.
11. ICANN organization sends the response and updates of the RMP (if any) to RTAP from the requester. If requester has not submitted a response within four (4) weeks after receiving the Interim Report, ICANN will inform the RTAP Panel that they may continue to next steps.
12. The RTAP Panel creates the RTA-Final Report and sends it to ICANN organization within (4) weeks of receiving the requester response on the RTA-Interim Report, or if no response is received within four (4) weeks of the expiry of the deadline for filing a response. ICANN organization coordinates any clarifying questions between RTAP Panel and the requester.
13. ICANN organization sends the RTA-Final Report to the requester and publishes it one (1) week after sending it to the requester

7. Closure of procedure

The end result of the review procedure is either:

- A documented and consolidated recommendation from the RTAP Panel, following consultations with the requester, confirming that:

- The requester has adopted an appropriate risk management methodology and framework;
- The mitigation measures are proportionate and adequate to treat the risk(s) identified by the DSP or EPSRP (as the case may be);
- The requester/ IDN ccTLD operator has committed to implement the mitigation measures prior to or on launch of the IDN ccTLD string(s);
or
- A documented and consolidated recommendation confirming the risk is not adequately treated, given the list of mitigation measures being proposed by the requester.
- The end result of the review, will be made public.

8. Risk Treatment Appraisal (RTA) Reports

There are two kind of reports generated by the panel. There is *RTA-Interim Report* which identifies gap(s) and (possibly) recommends any additional controls and solutions to mitigate risks identified. The second, the *RTA-Final Report* provides the final consolidated recommendation after evaluating the RMP by the requester. These reports would contain at least the following:

RTA-Interim Report

1. *Objective and scope of the risk management process.*
2. *Summary of the external and internal context and how it relates to the system being assessed.*
3. *Summary of the methodology used for various stages of risk management.*
4. *Assessment of risk and breakdown of overall risk into its itemized component risks, with description of each component risk, the gap it causes, the end-user communities it impacts, and its evaluation.*
5. *Summary of the initial RMP by the requester, its break down into constituent controls, and how applicable constituent controls address each component risk.*
6. *Analysis of the degree (and description) of residual risk for each component risk after applying the proposed constituent controls.*
7. *For each component risk and in accordance with the objective and criteria set out in these guidelines, a detailed evaluation if the residual risk is still at significant level. Why? Why not?*
8. *Any suggestions, if available, for effectively addressing any of the residual risks which is still considered significant.*
9. *Based on the RMP, the residual risk for each component risk, what is the interim consolidated recommendation: is the cumulative risk effectively mitigated based on the RTA objective? Why? Why not?*

RTA-Final Report

1. *Objective and scope of the risk management process.*
2. *Summary of the external and internal context and how it relates to the system being assessed.*
3. *Summary of the methodology used for various stages of risk management.*

4. *Assessment of risk and breakdown of overall risk into its itemized component risks, with description of each component risk, the gap it causes, the end-user communities it impacts, and its evaluation.*
5. *Summary of the initial RMP, and any response or changes to the mitigation measures proposed by the requester in response to the RTA-Interim report,*
6. *Summary of the final RMP, its break down into constituent controls, and how applicable constituent controls address each component risk.*
7. *Analysis of the degree (and description) of residual risk for each component risk after applying the proposed constituent controls.*
8. *For each component risk, and in accordance with the objective and criteria set out in this guideline, a detailed evaluation if the residual risk is still at significant level. Why? Why not?*
9. *Based on the RMP, the residual risk for each component risk, what is the final consolidated recommendation: is the cumulative risk effectively mitigated based on the RTA objective? Why? Why not?*

Glossary

- Risk Mitigation Proposal, by the requester – RMP. The RMP should include at a minimum the proposed internationally recognized and appropriate risk management and mitigation process the requester has used and intends to use, and the proposed mitigation measures.
- Risk Treatment Appraisal Process- RTAP
- Risk Treatment Appraisal Process Panel – RTAP Panel (none DRP EPSPR or ICANN employees or contractors)

To be updated

Process Flow Diagram

