Discussion Paper on Scope of the Similarity Evaluation Panel

This paper is drafted as information for the discussion of the *Scope of the string similarity review on the Request Side* section in the *Public Comment: summary and ANALYSIS version 1*. The paper includes information following the questions.

Q1. As blocked and most allocatable variants from the requested string will never be delegated, how can they impact the confusability of the end

Q2. How the residual mis-connection risk will be addressed by expanding the request side of the base for comparison as suggested?

Q3. What is the impact of inconsistency between ccPDP4 and IDN EPDP?

Q4. What changes needed in the report, to make the policy consistent with the IDN EPDP?

Table: Scope of the string similarity review on the Request Side section in the Public Comment: summary and ANALYSIS version 1

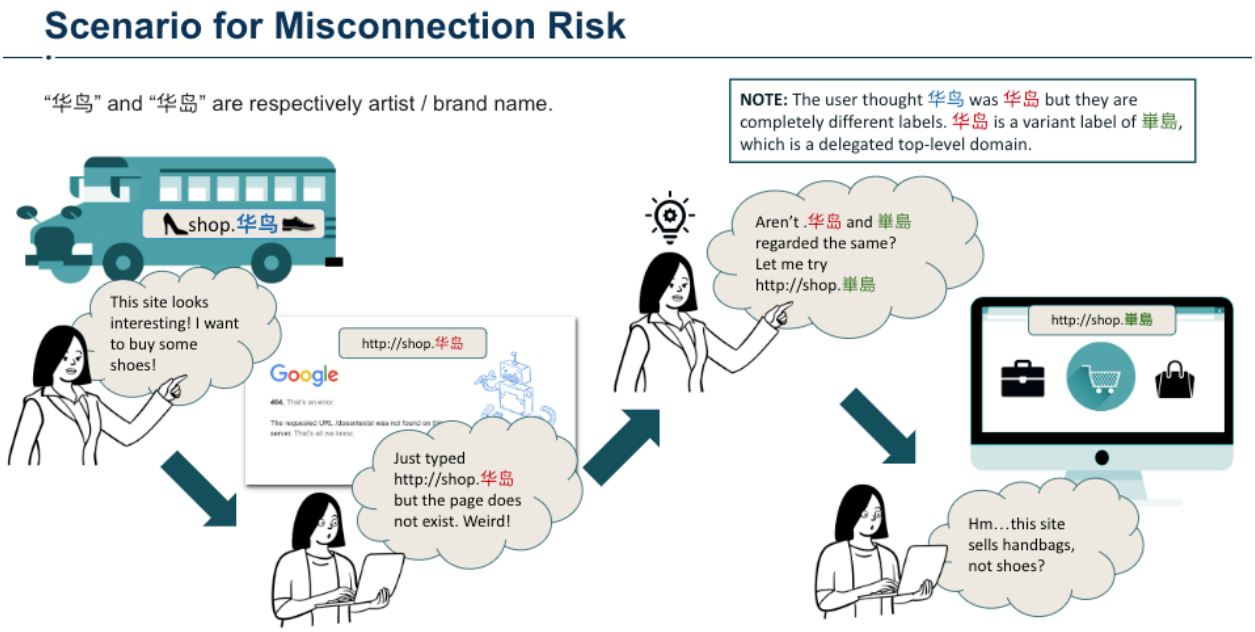
Comment	**WG Analyses**	**Update of Proposed Policy text, if any**
The scope of the string similarity review on the Request Side may not fully address security issues and is not consistent with the GNSO IDN EPDP. ICANN proposes that the Similarity Evaluation Panel " *should determine which additional variants of the basic set of strings should be included in the* ***Request Side,*** *factoring in: The likelihood of misconnection, Scalability, and Unforeseen and/or unwanted side effects. In its report, the Panel must provide its reasoning for its determination, whether to include additional variants of the basic set of strings included in the request side."*	It is noted that only allocatable variant strings that are a meaningful representation of the name of a country in a designated language may be requested as a variant form the selected (or primary string) and hence potentially available a ccTLD string. As stated in the Initial Report of the WG, the WG considered and develop the policy proposals on the SSAC advise in SAC060: when introducing variants, the policy making bodies should consider*, 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 though the similarity review. Therefore, the confusing similarity review is about minimizing the risk i.e., likelihood of misconnection. As blocked and most allocatable variant from the requested string will never be delegated, it is unclear to the WG, which residual mis-connection risk will be addressed by expanding the request side of the base for comparison as suggested. The WG also notes that in some cases variants that meet the criteria may not be requested, or only after	No need to update the proposed policy

	(quite some time) the selected string has been delegated, for example eligible variants of an IDNccTLD string delegated under the Fast Track process.	

Q1: As blocked and most allocatable variants from the requested string will never be delegated, how can they impact the confusability of the end user?

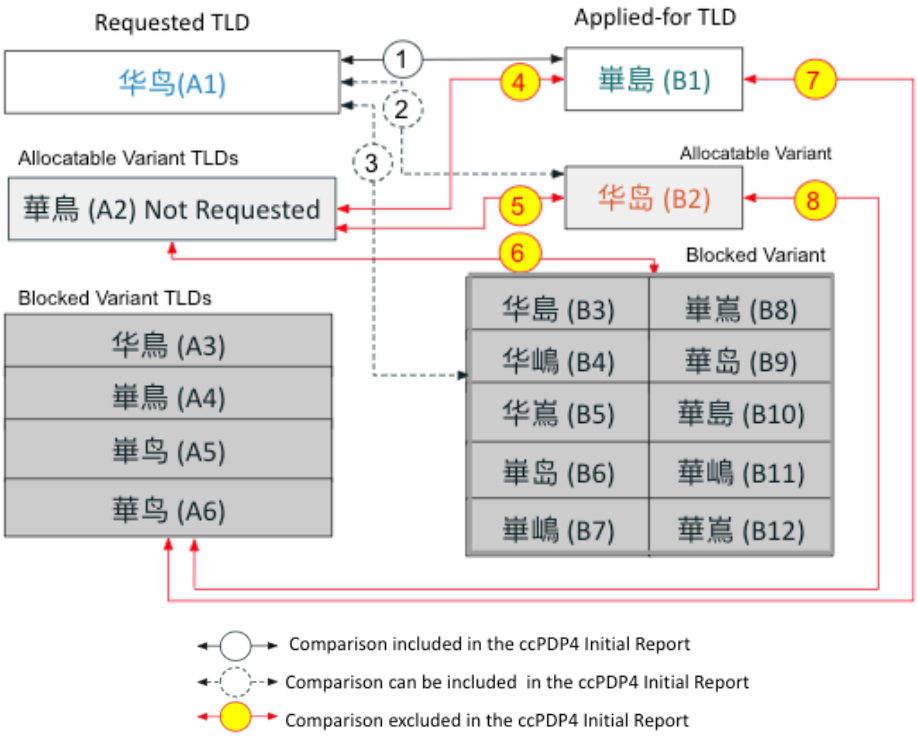
A2: The user can get confused by a variant label, regardless of its delegation status.

- In this example the label 华岛 is not delegated.
- When the user saw the label 华鸟 but thought it was 华岛 and typed shop.华岛, she encountered the no-connection failure mode.
- However, 华岛 and 嶋島 are variant labels, so the user tried again with shop.嶋島. she encountered the mis-connection failure mode.
- End user encounter mis-connection failure mode between 华鸟 and嶋島 due to the non-delegated 华岛.



Q2: How the residual mis-connection risk will be addressed by expanding the request side of the base for comparison as suggested?

A2: Expanding the comparison increase the cases for the String Similarity Review Panel to review. Therefore, it gives a higher possibility to determine the result as similar.



Current Scope	Proposed Expanded Scope
2 华鸟(A1) & 华岛 (B2)	2 华鸟(A1) & 华岛 (B2) 4 華鳥 (A2) & 華島 (B1) 6 華鳥 (A2) & 華嶌 (B8) 7 華鳥 (A4) & 華島 (B1)
Potential outcome: A higher chance that 华鸟(A1) And 華島 (B1) are determined ***non-similar*** and both of them can be delegated. Which can cause **mis-connection** failure mode.	**Potential outcome:** A higher chance that 华鸟(A1) And 華島 (B1) are determined ***similar*** and both of them cannot be delegated. Which can address the **mis-connection** failure mode.

Q3: What is the impact of inconsistency between ccPDP4 and IDN EPDP?

A3: It can create different results in the different timeframes.

- During the next new gTLD Round, both IDN EPDP Recommendations and IDN ccPDP4 Recommendations apply. Based on IDN EPDP Recommendation 4.1-4.3 the comparison will also cover requested ccTLD, its allocatable variant labels, and its blocked variant labels.
- After the next new gTLD Round, only the Recommendations from ccPDP4 apply. Therefore, only requested ccTLD and the required Delegateable ccTLD will be in the Similarity Evaluation scope.
- Using the previous example of 华鸟(A1) and 華島 (B1)
    - During the next new gTLD Round only one label, either A1 or B1, can be proceed and be delegated.
    - After the next new gTLD Round, both can be delegated.

Q4: What changes needed in the report, to make the policy consistent with the IDN EPDP?

A4: There will be two main updates:

- Section 7.2.3.a., add *"It is proposed that the Similarity Evaluation Panel should determine which additional variants of the basic set of strings should be included in the* ***Request Side,…"***
- The result of SEP must apply for all labels in the set, and therefore, section 8.6.4.3.a, 8.6.4.3.b, and 8.6.4.3.c needs to be revised. It is possible that we only need to keep 8.6.4.3.a.