**Working Document**

Topic E: Adjustments to objection process, string similarity review, string contention resolution, reserved strings, and other policies and procedures

| **CHARTER QUESTIONS** |
| --- |

| **e1)** In considering the conclusion(s) with respect to question **b4a)**, what role, if any, do TLD labels “withheld for possible allocation” or “withheld for the same entity” play vis-a-vis:   * objection process; and * string similarity review process? |
| --- |

**High-Level Notes:**

* *As e1(the part related to “withheld-same-entity” labels) and e3 are interrelated, the discussion for these charter questions was combined, with the deliberations captured under e3 below.*

| **e2)** Under the rules of the most recent gTLD application round, there are four criteria for objections to a string (see *gTLD Applicant Guidebook*, version 2012-06-04, section 3.2.1).[[1]](#footnote-0) The SubPro PDP has also affirmed the continuation of these four criteria for objections to a string, while proposing recommendations and implementation guidance to enhance/adjust these criteria.[[2]](#footnote-1)  The WG and the SubPro IRT to coordinate to ensure consistency in the implementation of the **objection** process for the variant label applications of existing and future TLDs. |
| --- |

| **e3)** In the Initial Evaluation for new gTLD applications, a proposed applied-for TLD is checked against several criteria as part of the string similarity review process (see *gTLD Applicant Guidebook*, version 2012-06-04, section 2.2.1.1.1).[[3]](#footnote-2) The SubPro PDP affirmed these standards, while proposing recommendations and implementation guidance to enhance the process.[[4]](#footnote-3)  The WG and the SubPro IRT to coordinate to ensure consistency in the implementation of the **string similarity review** procedure for variant label applications of existing and future gTLDs.[[5]](#footnote-4) |
| --- |

**High-Level Notes:**

* The current scope of EPDP discussion of this question was limited to future new gTLDs going forward.
* One member suggested the following: the string similarity review should only cover the variant labels requested by the applicants. The primary applied-for gTLD should be evaluated first, and the requested variant labels should be evaluated next.
  + If the primary string is confusingly similar to existing TLDs, requested ccTLDs, or Reserved Names, the application is rejected.
  + If the primary string is confusingly similar to another applied-for gTLD in the same round, they go forward with contention resolution.
  + If the primary string passes evaluation, but the requested variant is confusingly similar to existing TLDs, string requested as ccTLDs, or Reserved Names, the applicant should have an opportunity to withdraw its request for the variant label.
  + If the primary string passes evaluation, but the requested variant is confusingly similar to another applied-for gTLD in the same round, the applicant should have an opportunity to either withdraw its request for the variant, or go forward with contention resolution.
  + If the applicant would like to request a variant that it didn’t previously apply for, it needs to undergo a new string similarity review process, as circumstances may have changed with respect to potential string contention, and this will need to be taken into account in the evaluation process.
* Another member suggested that the applicant might also be given the chance to “switch” its primary applied-for gTLD label to one of the variants.
* In the development of the Staff Paper, three levels of analysis were conducted regarding the coverage of String Similarity Review:
  + Level 1: Applied for strings (including requested variants) compared against other applied for, delegated strings, and reserved strings;
  + Level 2: Applied for strings (including requested variants) compared against other delegated strings, reserved strings, applied-for strings and their allocatable variants;
  + Level 3: Applied for strings (including requested variants) compared against other delegated strings, reserved strings, applied-for strings and all of their variants (including blocked ones).
* Staff Paper ultimately recommended the maximally conversative approach (Level 3) as there was little knowledge of how variants would be operated and it seems safer to err on the side of caution.
* There was support that evaluating all valid variants is overly restrictive and conversative. For certain scripts, allocatable variants could be tens of labels, and blocked variants can be very large for not only Arabic, but also Latin script (e.g., hundreds or thousands). Conservative approach could be costly and the EPDP Team needs to consider adding parameters to make the process easier and less costly.
* One member suggested the full set of the reserved strings should be evaluated against, as a certain reserved string could become unreserved in the future.
* Another member commented that if the language community didn’t want a certain blocked string, shouldn’t it have also blocked the string that is confusingly similar to the blocked string? If they didn’t block it, maybe they are okay with it being delegated. A variant of a blocked string should only be rejected IF it is confusingly similar from a visual perspective.
* The leadership/staff has been tasked to develop a matrix to clearly lay out the different levels of comparison and identify which elements of each set are compared to which elements of the other set, as well as the respective pros and cons of each level of comparison.
* Staff developed a [comparison matrix](https://community.icann.org/download/attachments/192217195/EPDP%20Team%20Meeting%20%2330%20Slides%20-%20E3%2C%20E1%2C%20E3a.pdf?version=1&modificationDate=1649976844000&api=v2) to show the three levels of string similarity review and their respective “pros” and “cons” as conversation starter: Level 1 - Primary + Only Requested Allocatable Variants; Level 2 - Primary + All Allocatable Variants; Level 3 - Primary + All Variants
* One member believes that all three levels guarantee that there will be no confusability, and no level provides better protection than another level in this regard.
* Discussion of Level 1 - Primary + Only Requested Allocatable Variants:
  + An RO has no automatic “ownership rights” to allocatable variants. It only has rights to a variant that is requested and delegated. It should be acceptable if an applicant does not request an allocatable variant at first but is then denied when it later requests it, because at that point in time that variant is confusingly similar to an already delegated string.
  + Registry agreement has the following clause regarding “Ownership Rights”: “Nothing contained in this Agreement shall be construed as (a) establishing or granting to Registry Operator any property ownership rights or interests of Registry Operator in the TLD or the letters, words, symbols or other characters making up the TLD string, or (b) affecting any existing intellectual property or ownership rights of Registry Operator.”
  + Some allocatable variant labels will never be applied for, so they should not be taken into account in the string similarity review.
  + If activation of variant labels is possible between application rounds, an additional “con” would be the string similarity review would have to be done at any point in time. However, the EPDP Team has not agreed on whether activation requests between rounds are allowed.
  + When there are multiple rounds and if there is overlap in timing of evaluation processes, the string similarity review of variant labels may not be that simple. A consideration is needed regarding the specific rules applied across rounds that govern how activation of variants by ROs is prioritized. SubPro also has addressed the prioritization of applications where there may be overlapping rounds.
  + Comments supporting this level include:
    - Support this option if variants can only be requested inside application rounds.
    - This level is sufficient because it is a safeguard against allocating variants that would be confusable to other strings allocated. No need to check any future potential confusability of labels as that would only unnecessarily restrict the DNS space.
    - Level 1 entails the least similarity work and the lowest cost. The only drawback is that if an entity does not apply for an allocatable variant initially it might not be able to get it later. If the applicant is aware of this potential consequence, it can apply for all the allocatable variants it wants at the outset.
* Discussion of Level 2 - Primary + All Allocatable Variants:
  + To set a policy that gives a right of first refusal to an existing RO to any variants goes against the notion that the RO doesn’t have ownership rights to those variant strings. Essentially it is reserving the right for the RO for the string should they ever want it, and over anyone else who applies for a string that might be confusingly similar.
  + To review the allocatable but not requested variants is not to give IP rights to the applicant. In the 2012 round, the self identified variants were incorporated in the string similarity review but the applicant did not have any IP rights over those strings.
  + An applicant is not an RO bound to the registry agreement or the “Ownership Rights” clause.
  + Comments supporting this level include:
    - Support this option if variants can be requested outside of application rounds. We need to ensure from the start that allocation of all allocatable variants will be possible.
    - It narrows the chances of confusability in the same round, which improves predictability. This creates predictability because all entities have their original string and allocatable variants being compared against each other. No other TLD in the future could apply for those allocatable variants.
    - It presents a good balance from a cost perspective.
    - It provides flexibility and opportunity while ensuring a degree of protection of confusability.
    - This level depends on other elements such as costs associated with each string applied for and when a string could be applied for. If the cost is relatively low for applied-for string, ALAC believes Level 2 is more appropriate.
* Discussion of Level 3 - Primary + All Variants:
  + The maximally conservative approach is the safest but might be too much.
  + Similarity is considered at the perception layer and it cannot be done in an automatic fashion. In the 2012 round there was a SWORD tool, but it did not perform adequately and SubPro recommends eliminating such tool.
  + When a string is formally requested there is a need to notify the community, provide an opportunity for objections, and complete related processes. SubPro did not favor requesting applications on a rolling basis because it is difficult to make this predictable. If the activation of variants is made possible between rounds, this will go against the principle of predictability that SubPro spent a lot of time working on.
  + Level 3 is more predictable, as all variants are reviewed upfront.
  + Giving right of first refusal on every variant label to the existing registry operator is predictable, but it goes against many other principles put forward by SubPro.
  + Why do blocked strings need to be taken into consideration? GPs may have multiple motivations for blocking a string. The main motivation was that the RZ-LGR procedures suggested that allocatable variants should be minimized and blocked variants should be maximized. The panel may look at it from a usability perspective. If there are two characters that are nearly visually identical and there is no usability argument, then they would make it a blocked variant. If there is a usability argument, they would make it an allocatable variable.
  + In the case of Level 3, the first-come-first-serve rule would not be an issue, but it would be an extra consideration for Level 1 and Level 2.
* There was divergence regarding which of the three levels is the most appropriate. There was a discussion about whether Level 1 or Level 3 is more consistent with the SSAC conservative principle. From one perspective, Level 3 is a broader evaluation and therefore more conservative. From another perspective, Level 3 creates rights to a wider range of variants, and therefore Level 1 is more conservative because it treats the pool of variants in a more limited way.
* There was a suggestion regarding activation of labels between rounds: What if there was a rule that if an applicant is applying for an IDN TLD, they know what the variant set looks like and they decide what they want to apply for? There is one opportunity to apply for the primary and the variant set you are interested in applying for. If the requested variants are still available in the future, there is still an opportunity.
* Need to examine the issue from end user perspective to make the space less confusing; that’s the underlying motivation of the string similarity review. Irrespective to whether the variant is allocatable or blocked, there can still be confusability for end users when there are two similar strings in the DNS space.
* Example: Simplified and traditional Chinese – in most cases the two are visually distinct. If applicant A applies for a simplified Chinese version, which has a traditional version variant that looks different but is considered to be the same in meaning and has not been applied for. If applicant B applies for a traditional Chinese character that is visually very similar to traditional variants of the simplified character applied for by applicant A. If the comparison is done only for applied-for strings, the applied-for strings for applicant A and b will both be delegated, the end users can get confused by seeing the traditional version is the same as the simplified version. This will create confusion for all registrations under those TLDs. Level 1 will not prevent this scenario.
* One member supported the model proposed in the Staff Paper, where the applied-for primary label + requested variants are compared against other primary TLDs + all allocatable and blocked variants.
* Confusing similarity has always been only about visual similarity. Applying that principle, the maximum conservative approach is not consistent. How would this work in the case of variants that are not visually similar and why would additional protections be provided to them? Hard to compare these levels without concrete examples. The EPDP Team agreed to use a small group mechanism to develop examples related to Levels 1, 2, and 3 to assist further deliberation.
* Even if the EPDP Team reaches agreement on when variants can be applied (inside application rounds or outside), the appropriate level of string similarity review still can be identified without further deliberation.

| **e3a)**  After a requested variant string is rejected as a result of a string similarity review, should the other variant strings in the same variant set remain allocatable? Should individual labels be allowed to have different outcomes/actions (e.g., some labels be blocked and some be allowed to continue with an application process)?[[6]](#footnote-5) |
| --- |

**High-Level Notes:**

* The current scope of EPDP discussion of this question was limited to future new gTLDs going forward.
* The team discussed two possible scenarios related to the outcome of an applied-for variant string getting rejected as a result of a string similarity review: 1) only the applied-for variant string is rejected while the other allocatable variant labels continue to remain allocatable; 2) the entire variant set including the applied-for variant string is rejected. The team had split views.
* One member emphasized the need to think through the idea of atomicity of the variant set (i.e., the whole set of variant labels is inseparable). If the application of the set is treated as an atomic whole, the evaluation of the set should be treated the same way. Once the set is split up, we may run into potential issues about how the variants are conceptualized.
* Another member emphasized that the string similarity review is a visual test. How can you justify rejecting a string if it’s not visually confusingly similar to another string? In that logic, where the applied for string is rejected, the other allocatable variants should still be eligible for activation unless they are also confusingly similar.

| **e4)** Under current procedures, resolution of string contention for applied for gTLD strings may include components such as a settlement between the parties, a community priority evaluation (if a community-based applicant in a contention set elects this option), and an auction. SubProp PDP affirmed these components while proposing recommendations and implementation guidance to enhance the mechanisms for string contention resolution.[[7]](#footnote-6)  The WG and the SubPro IRT to coordinate to ensure consistency in the implementation of the **string contention resolution** mechanism for variant label applications of existing and future new gTLDs.[[8]](#footnote-7) |
| --- |

| **e5)** The WG and the SubPro IRT to coordinate and consider the following questions in order to develop a consistent solution: should the reserved strings ineligible for delegation for existing and future gTLDs be updated to include any possible variant labels? Consider this question by taking into account the data to be collected in the “Data and Metric Requirements” section of this charter. |
| --- |

**High-Level Notes:**

* Staff [presented](https://community.icann.org/download/attachments/192217197/EPDP%20Team%20Meeting%20%2331%20Slides%20-%20ccPDP4%20update%2C%20E5.pdf?version=1&modificationDate=1650567274965&api=v2)on the background on the Reserved Names and Strings Ineligible for Delegation, including their recent development and implications for future rounds.
* Staff may need to investigate other additional categories of TLDs, including numeric, single and two character ASCII, and single character IDN strings mentioned in the 2012 AGB.
* The IDN “[test](https://www.iana.org/domains/reserved)” strings are currently not delegated but remain reserved.
* All Reserved Names, except for the IDN “test” string, are ASCII strings that don’t have allocatable variants.
* The IGO/INGO identifiers are currently not available for registration at the second level, and in the subsequent rounds they will also be ineligible for delegation at the top level.
* For Strings Ineligible for Delegation, going forward, there will be an exception procedure that would allow a party to apply for their own strings. These strings ineligible for delegation would not be included in the string similarity process. It is protection for the precise terms on the list and only preventing delegation of those precise strings.

Discussion of “Reserved Names”:

* There is some support for checking the variants (including the blocked variants) of the Reserved Names in the String Similarity Review, as well as creating a clear definable set of variant labels of the Reserved Names. The blocked variants of Reserved Names may come into play if the EPDP Team recommends that string similarity review should include blocked variants.
* The variant labels of the Reserved Names can be generated, but the list could be in the thousands or tens of thousands. Usually if an ASCII label has certain vowels and about 5-6 characters, it can generate a large number of blocked variants.
* Some members believe the number of blocked variants is not an issue of concern. If they are blocked, no one would want those variants and there is no negative impact to the market. If the atomicity of variants can be preserved without hurting the gTLD market, a more conservative approach should be the way to go. If the EPDP Team recommends that all variants of Reserved Names are blocked, the evaluation panel just needs to run the LGR tool to check if there is a variant relation with the Reserved Names for the applied-for label. The LGR tool does not care whether the number of variants is 2 or 20,000 as long as those variants are blocked.
* Leadership team and staff to consider next steps on data collection regarding variants of Reserve Names, to guide the EPDP Team discussion of how security and stability benefits weigh against the complexity of implementing.
* ICANN org staff presented on the [data](https://docs.google.com/spreadsheets/d/11OkBT_1-kABdUgy7kbrf9bd8PHFPtP9A/edit?usp=sharing&ouid=101698682360672018983&rtpof=true&sd=true) related to the variant labels of Reserved Names by running them through the RZ-LGR. Some labels have hundreds of thousands of variant labels (all of which are blocked).
* Some members believed that if the Reserved Names list is to be expanded, only the variant strings that could actually be applied for should be included (e.g., mixed script variant strings are not allowed). However, the team recognized that if some variants of Reserved Names are added to the Reserved Names list, they need to be checked against during the String Similarity Review; it means that every applied-for string will be compared against a much larger pool of Reserved Names.
* Members discussed an example of a string that is a variant of a Reserved Name but not be considered visually similar, and its potential impact. Example: ßac and ssac are variants; if BAC and ßac are considered visually similar, should it be possible to apply for BAC? From one perspective, it should be possible and we shouldn’t block BAC just because ßac is a variant of ssac.
* The Staff report suggested a coherent analysis looking at all the possible TLD strings across all categories rather than offering one approach for an applied-for string and different rules for Reserved Names.
* The group converged on the idea that the Reserved Names list should stay as is and no variants of Reserved Names should be added to the list. To the extent there is an interest from certain groups to extend the Reserved Names list, they can follow the specific process.
* Rationale for this idea includes that the reserved strings are reserved for a purpose. If an application looks like the reserved name, this will be addressed in the string similarity review. If there is a variant that looks different from the Reserved Name, it should not be taken into account in the string similarity review. Adding to the Reserved Names list should not be taken lightly. A lot of thought went into making this list. The trend over time has been to limit the Reserved Names.
* The group also converged on the idea that all variants of the Reserved Names will be automatically blocked. For example, ßac (variant of ssac) should be blocked even if it would not have been caught in the string similarity review because it is not visually similar. There may be old browsers that might format the string ßac into ssac which may cause a real problem.

Discussion of “Strings Ineligible for Delegation”:

* Would extending preventative protections for variants
  + Circumvent the careful work of the IGOs PDP?
  + Extend rights beyond those that are expressly identified in relevant treaties?
  + In other words, are variants for these strings in scope for the IDN-EPDP?
* It is important to make the distinction between reserved names and strings ineligible for delegation and consider that different treatment may be appropriate for the two groups. Strings ineligible for delegation are limited to exact match, so the EPDP Team may not be in a position to extend these rights. Given the sensitivities around this topic and the finite nature of this list, is there an argument/rationale for protecting variants of such strings?
* One member believes that the variant labels of the Strings Ineligible for Delegation need to be blocked as well. If the variants are not protected, the first-come first-serve rule may not create a fair situation and the IGOs potentially won’t be able to access their variants.
* Some members commented that there is an existing policy/procedure to extend the list of strings ineligible for delegation. To the extent there is an interest from those IGOs and INGOs to extend the list, they can follow the specified process.
* However, some members believe that there needs to be a consistent way to treat variants across all strings and focus more on the issue in a principled manner, and less on the number of variants of those protected strings.
* One of the fundamental principles is that a primary label and its variants are allocated to the same entity. If protection is not extended to the variant labels, there is a possibility that a variant is allocated to an entity that is different from the IGO or INGO and that entity may take the original label away from that organization, which breaks that principle and is a significant change.
* Therefore, some members suggest blocking all variant labels of the strings ineligible for delegation. And to avoid a long list, the list only contains the original strings ineligible for delegation, and link to a resource for calculating the variants that should be blocked.
* Some other members pointed out that the names that are ineligible for delegation are not reserved for the organization. If ineligible for delegation, it’s ineligible for all. They suggest that the work on the IGO-INGO names that took years to complete should not be modified. The GNSO Council should weigh in on whether this is in scope for the EPDP.

| **e6)** The WG and the SubPro IRT to coordinate and consider the following questions in order to develop a consistent solution: is there any reason to permit the registration of gTLDs consisting of decorated two-character Latin labels which are not variant labels of any two-letter ASCII labels?[[9]](#footnote-8) If so, rationale must be clearly stated. |
| --- |

| **e7)** Besides the objection process, string similarity review, and string contention resolution, what other ICANN policies and procedures should be updated to enforce the “same entity” rule and the use of RZ-LGR as the sole source to calculate the variant Labels and disposition values?[[10]](#footnote-9) See the list of ICANN Consensus Policies here: <https://www.icann.org/resources/pages/registrars/consensus-policies-en> |
| --- |

* Evaluation elements for variant labels of TLDs with restrictions (e.g., community, Geo, Category 1, Brand)

1. The four criteria are: String Confusion Objection; Legal Rights Objection; Limited Public Interest Objection; and Community Objection. [↑](#footnote-ref-0)
2. See “Topic 31: Objections” in the SubPro PDP Final Report, pp.145-154: <https://gnso.icann.org/sites/default/files/file/field-file-attach/final-report-newgtld-subsequent-procedures-pdp-02feb21-en.pdf#page=145> [↑](#footnote-ref-1)
3. These criteria are: existing TLDs and reserved names; other applied-for strings; strings requested as IDN ccTLDs; and applied-for 2-character IDN gTLD strings against every other single character and any other 2-character ASCII string. [↑](#footnote-ref-2)
4. See “Topic 24: String Similarity Evaluations” in the SubPro PDP Final Report, pp.108-114: <https://gnso.icann.org/sites/default/files/file/field-file-attach/final-report-newgtld-subsequent-procedures-pdp-02feb21-en.pdf#page=108> [↑](#footnote-ref-3)
5. The Staff Paper recommends that the string similarity process to compare strings under consideration not just against all allocated or applied-for strings, but also all variants of those strings (including allocatable, withheld-same-entity, and blocked). For example, if a string is merely withheld-same-entity and a second string is visually similar, then allocating the second string undermines the predictability of the outcome of variant processing from the RZ-LGR. Similarly, if a string is blocked under the RZ-LGR, but a visually similar string is allocatable, then the second (visually similar) string might become a “work around” for the blocked string. This approach is maximally conservative. It is nevertheless worth noting that this expands considerably the number of strings that might need to be considered; the entire similarity review process will consequently probably become more expensive to operate. See Section 3.8 Adjustments in String Similarity Process in the Staff Paper, pp.18-19: <https://www.icann.org/en/system/files/files/idn-variant-tld-recommendations-analysis-25jan19-en.pdf#page=18>

   Staff Paper further recommends that in the event that two or more applied-for variant strings are visually similar, they may only be allocated if they are associated with the same variant set and are being requested by the same entity. In case of such conflicts across variants, the entire IDL set gets processed as one contention set; if one of the labels is already allocated, the contention is resolved in favor of the current operator. The Staff Paper recommends that it is necessary to perform the visual similarity checks for every requested-to-be-allocated variant in any given set against all the possible variants in every other set. This is because such an available variant could be requested at any time in the future. See Section 3.8.1 in the Staff Paper, pp.20-21: <https://www.icann.org/en/system/files/files/idn-variant-tld-recommendations-analysis-25jan19-en.pdf#page=20> [↑](#footnote-ref-4)
6. The Staff Paper recommends that the following outcomes may be considered: 1) only the variant string requested for delegation is rejected. For example, the requested variant t1v2 of top-level label t1 will get rejected while t1v1 and t1v3 from the same variant set continue to remain allocatable; or 2) the entire variant set is rejected. For example, the requested variant t1v2 of top-level label t1 will get rejected including t1v1 and t1v3 from the same variant set as t1v2. This outcome appears to be difficult to justify, though an applicant could decide that, if it cannot receive t1v2 then it does not wish to proceed with the application. See Section 3.8.2 in the Staff Paper, pp.21: <https://www.icann.org/en/system/files/files/idn-variant-tld-recommendations-analysis-25jan19-en.pdf#page=21> [↑](#footnote-ref-5)
7. See “Topic 35” in the SubPro PDP Final Report, pp. 173-182: <https://gnso.icann.org/sites/default/files/file/field-file-attach/final-report-newgtld-subsequent-procedures-pdp-02feb21-en.pdf#page=173> [↑](#footnote-ref-6)
8. For contention issues that involve the same entity, the Staff Paper suggests that the following resolution options may be considered, with a preference to the second option: 1) When the requested variant strings are placed in a contention set for later evaluation, the applicant is notified of the contention set and has the opportunity to establish that both applications are from the same entity. 2) It may be more efficient to establish early on in the string similarity review that the variant strings are being requested by the same entity prior to reaching the contention phase. See Section 3.8.2 in the Staff Paper, p. 21: <https://www.icann.org/en/system/files/files/idn-variant-tld-recommendations-analysis-25jan19-en.pdf#page=21> [↑](#footnote-ref-7)
9. The ccTLD labels in the root depend on an external registry (ISO 3166) that allocates alphabetic codes to countries. In order to ensure that no conflicts with future assignments by ISO can happen, ICANN has traditionally also maintained a restriction against the use of two-letter TLDs for all Latin script letters; no variants should be generated for ccTLDs based on the ISO3166 codes. This principle is also reaffirmed by the SubPro PDP. See Recommendation 21.6 in the SubPro Final Report, p.95: <https://gnso.icann.org/sites/default/files/file/field-file-attach/final-report-newgtld-subsequent-procedures-pdp-02feb21-en.pdf#page=95> [↑](#footnote-ref-8)
10. IDN Variant TLD Implementation Staff Paper: <https://www.icann.org/en/system/files/files/idn-variant-tld-recommendations-analysis-25jul18-en.pdf> [↑](#footnote-ref-9)