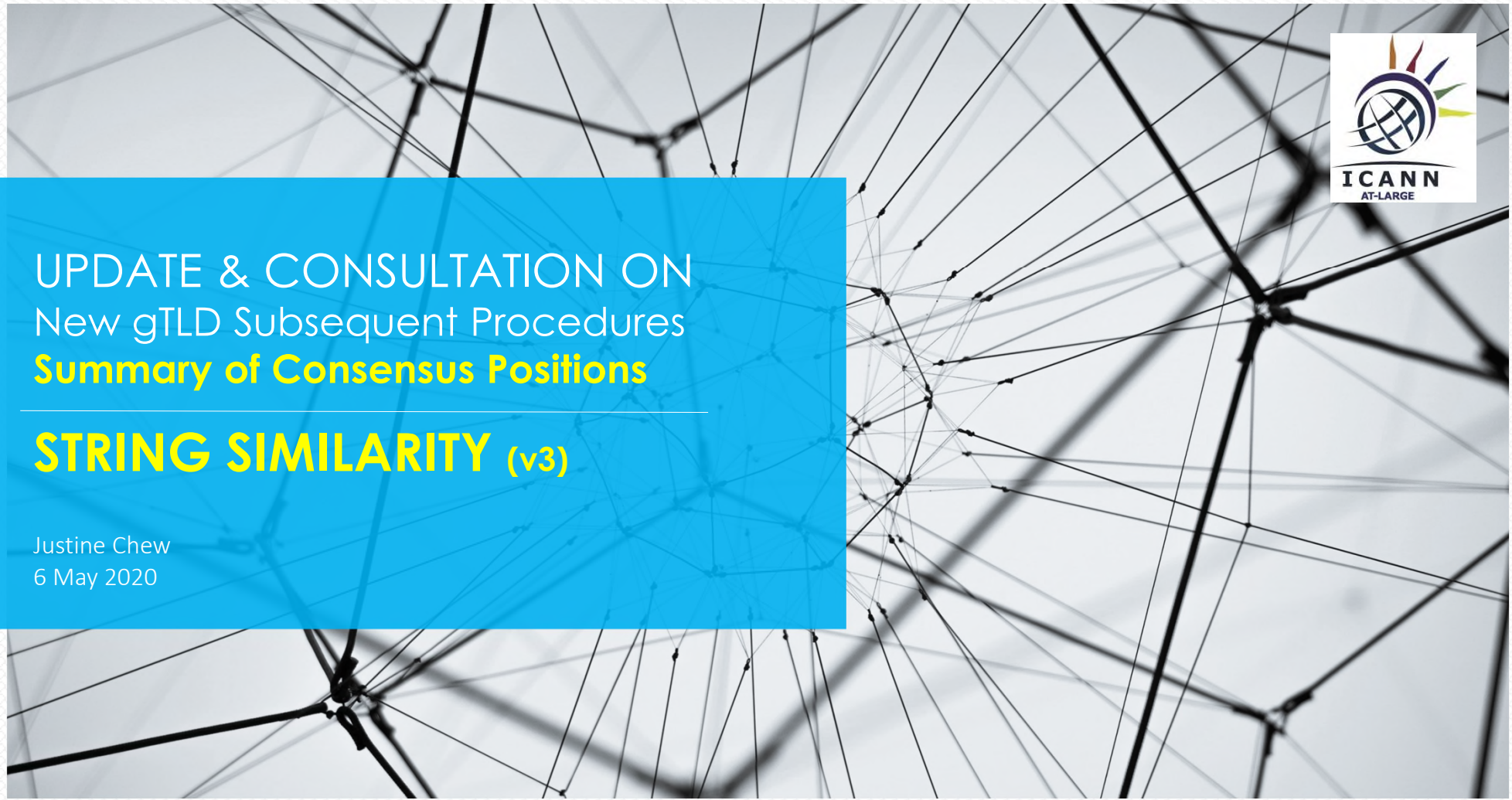

UPDATE & CONSULTATION ON
New gTLD Subsequent Procedures
Summary of Consensus Positions

STRING SIMILARITY (v3)

Justine Chew
6 May 2020

String Similarity in SubPro: Background

- What is the New gTLD Subsequent Procedures ("SubPro")?
  - The set of rules and mechanisms applicable to the next round for New gTLDs, i.e. they DO NOT apply to legacy TLDs, ccTLDs, or delegated new gTLDs or those still unresolved from the 2012 application round
  - "An update" to the 2012 Round rules and mechanisms

- Recap of String Similarity Review in 2012 Round
  - One of 4 string reviews undertaken on applied-for string within the Initial Evaluation Process
  - Undertaken by an independent String Similarity Panel:
    - Identify if an applied-for string is "similar" to any existing TLD, any other applied-for strings, reserved names, and in the case of 2-character IDNs, any single character or any 2-character ASCII string
    - Standard used: visual similarity; where "Similar" means string similarities that would create a probability of user confusion if more than one of the strings is delegated into the root zone
    - Informed by Sword Algorithm – "algorithmic score for visual similarity"
    - If found to be similar, applied-for strings are placed in contention sets
    - Ultimately found 4 strings in contention: .hotels v .hoteis, and .unicom v .unicorn

Summary of Consensus Positions

SubPro PDP WG recommendations

Affirmation #1:

- WG affirms Recommendation 2 from the 2007 policy, "Strings must not be confusingly similar to an existing top-level domain or a Reserved Name."

- Subject to the following recommendation, WG affirms standard used in the String Similarity Review from 2012 to determine **whether an applied-for string is "similar" to any existing TLD**, **any other applied-for string**, **reserved names**, and in the case of 2-char IDNs, any single-char or 2-char ASCII string.
  - ❖ Per s. 2.2.1 of the 2012 AGB, "similar" means "strings so similar that they create a probability of user confusion if more then one of the strings is delegated into the RZ.

- In 2012 round, the **String Similarity Panel was tasked with identifying "visual string similarities that would create a probability of user confusion**." WG **affirms the visual standard** for determining similarity and **recommends that the Panel additionally consider** as part of the standard **whether strings are intended to be used as the singular and plural version of the same word**.

Acceptable. No further intervention needed.

Summary of Consensus Positions

SubPro PDP WG recommendations

- To confirm under "Application Assessed in Rounds" topic re: disallowing application for a string that is still being processed from a previous application opportunity, to avoid creating unintended contention set.
- Consequentially, need a way to terminate any application that has little chance of succeeding and which are not withdrawn in subsequent procedures.

Recommendation #2:

- WG recommends updating the standards of both (a) confusing similarity to an existing top-level domain or a Reserved Name, and (b) similarity for purposes of determining string contention, to address singular and plural versions of the same word, noting that this was an area where there was insufficient clarity in the 2012 round.

- Specifically, WG recommends **prohibiting plurals and singulars of the same word within the same language/script in order to reduce the risk of consumer confusion.**
  - For eg, the TLDs .EXAMPLE and .EXAMPLES may not both be delegated because they are considered confusingly similar.

- This **expands the scope of the String Similarity Review to encompass singulars/plurals of TLDs on a per-language/script basis**.
  - ❑ An application for a single/plural variation of an existing TLD or Reserved Name will not be permitted if the intended use of the applied-for string is the single/plural version of the existing TLD or Reserved Name.
    - For eg, if there is an existing TLD .SPRINGS that is used in connection with elastic objects and a new application for .SPRING that is also intended to be used in connection with elastic objects, .SPRING will not be permitted.
  - ❑ Applications will not automatically be placed in the same contention set because they appear visually to be a single and plural of one another but have different intended uses.
    - For eg, .SPRING and .SPRINGS could both be allowed if one refers to the "season" and the other refers to elastic objects, because they are not singular and plural versions of the same word.
    - However, if both are intended to be used in connection with the elastic objects, then they will be placed into the same contention set.
    - Similarly, if an existing TLD .SPRING is used in connection with the season and a new application for .SPRINGS is intended to be used in connection with elastic objects, the new application will not be automatically disqualified.
    - A mandatory PIC could be a means for a Registry to commit to the use stated in the application and a method for enforcing adherence to this commitment.

- WG recommends using a dictionary to determine the singular and plural version of the string for the specific language.

- Suggest to mention "Use of mandatory PICs" in the recommendation itself, rather than just in the rationale, to give more prominence.

Summary of Consensus Positions

SubPro PDP WG recommendations

Recommendation #3:

- Eliminate the use of the SWORD tool in subsequent procedures.

Acceptable. No further intervention needed immediately. To monitor implementation on feasible replacement tool.

Recommendation #4:

- The deadline for filing a String Confusion Objection must be no less than thirty (30) days after the release of the String Similarity Evaluation results. This recommendation is consistent with PIRR recommendation 2.3.a, “Review the relative timing of the String Similarity evaluation and the Objections process.”

Acceptable, helps ensure that String Confusion Objection period runs for 30 days . No further intervention needed.