

**ASIA PACIFIC TOP LEVEL DOMAIN ASSOCIATION (APTLD)  
AD HOC WORKING GROUP ON TRANSLITERATED TLDs**

**MEMBER COMMENTS**

**INTRODUCTION**

APTLD (Asia Pacific Top Level Domain Association) represents the interests of country code Top Level Domain (ccTLD) registries in the Asia Pacific region. APTLD was originally established in 1998 and was legally incorporated in Malaysia in 2003.

APTLD is a forum for information exchange on technological and operational issues between domain name registries in the Asia Pacific region. It also acts as an interface with other international Internet coordinating bodies, fostering the participation of AP ccTLDs in these global fora and acting in the best interest of APTLD members in global Internet policy-making processes.

In July 2012, APTLD established an “*Ad Hoc Working Group on transliterated TLDs*” (“the WG”). The WG’s charter is to study the implications and possible issues arising from the coexistence of TLDs and those which are phonetic transliterations of them and of one another, in different scripts.

The WG’s charter also tasks the group with investigating possible concerns and opportunities inherent in transliterated TLDs with respect to the public policy goal of beneficial development of the Domain Name System in scripts, languages and cultures in the Asia Pacific region.

For the sake of clarity, “transliteration”, for the purposes of this WG, refers to the representation of ASCII TLD strings (either existing or proposed) in non-ASCII scripts, where the representation is achieved via a visual or phonetic mapping of the existing TLD, but may not be a literal *translation* of the string into a meaningful term in the alternate script. For example “.コム” is a transliteration of “.com” in Katakana, pronounced “dotto-ko-mu”, although the character string itself is not immediately semantically meaningful in Japanese.

As an initial priority, the WG determined to develop comments relating to potential issues arising from a number of strings applied for during the first round of new gTLD applications in 2012. This priority was established given the tight timeframes of the new gTLD comment process, and the observations and concerns expressed by a number of APTLD members.

Due to timeframe constraints, this paper does not convey a consensus APTLD position, but rather represents the collected views and comments received from a number of APTLD members and WG participants. The WG also notes that some APTLD members actively expressed no concerns surrounding phonetic similarity.

It is respectfully submitted as a general commentary about the new gTLD process, with the aim of informing the deliberations of ICANN staff and the evaluation panels that are involved in it.

## BACKGROUND

Following an extensive policy development process, ICANN formally invited applications for the establishment and operation of new generic Top Level Domains (gTLDs) on 12 January 2012. The application period closed on 30 May and ICANN announced the full list of 1930 applications it had received on 13 June 2012.

In accordance with ICANN's open invitation for community review and commentary, APTLD members, both individually and collectively, analysed the applications that had been received. Given the script and linguistic diversity of the Asia Pacific region, the substantial number of IDN (non-ASCII) strings applied for, were of particular interest. This analysis led to the determination that a WG should be established to collect and convey the views of APTLD members.

The WG would like to emphasise that the comments and issues raised in this document are not intended as a criticism or objection against a particular new gTLD applicant. Rather, references to particular applications are made as exemplars, to illustrate the issues the WG raises. The WG also notes that, with the exception of preliminary input from Versign Inc. prior to the formal establishment of the WG, the group did not receive clarification of the intent, or awareness of the potential for confusion, from the applications noted in this paper.

## POLICY CONTEXT

The WG notes a number of provisions within the new gTLD applicant guidebook that relate to the prevention of string confusion and loss of user confidence. It is in the context of these provisions that these comments are provided. These provisions include (emphasis added):

### **Section 1.1.2.5:**

*1. String reviews (concerning the applied-for gTLD string). String reviews include a determination that the applied-for gTLD string is **not likely to cause security or stability problems in the DNS, including problems caused by similarity to existing TLDs or reserved names.***

### **Section 2.2: Initial Evaluation**

*The Initial Evaluation consists of two types of review. Each type is composed of several elements.*

*String review: The first review focuses on the applied-for gTLD string to test:*

- *Whether the applied-for gTLD string is **so similar to other strings that it would create a probability of user confusion;***
- *Whether the applied-for gTLD string might **adversely affect DNS security or stability;***

#### **Section 2.2.1.1 String Similarity Review**

*This review involves a preliminary **comparison of each applied-for gTLD string against existing TLDs, Reserved Names (see subsection 2.2.1.2), and other applied-for strings.** The objective of this review is to **prevent user confusion and***

**loss of confidence in the DNS resulting from delegation of many similar strings.**

Note: In this Applicant Guidebook, **“similar” means strings so similar that they create a probability of user confusion** if more than one of the strings is delegated into the root zone.

#### **Section 2.2.1.1.1 Similarity to Existing TLDs or Reserved Names**

This review involves **cross-checking between each applied-for string and the lists of existing TLD strings and Reserved Names to determine whether two strings are so similar to one another that they create a probability of user confusion.**

## ISSUES

### ***Phonetic similarity and possible confusion***

A number of APTLD members noted that some new gTLD applications could potentially generate user confusion – both with existing TLDs and with others that have been applied for during the current application process. This confusion would arise from the phonetic similarity of these current or proposed TLDs, and is an issue that is distinct from the potential for visual confusion that has frequently been discussed and included in guidelines for the new gTLD process. As noted in the Introduction to this paper, this is not a consensus view held by all APTLD members, but is strongly held by some.

For example:

- VeriSign SARL applied for twelve strings that are transliterations of “.com” or “.net”, in a range of scripts.
- Five applications were received for “.site” and six for “.online”. Applications were also received for “сайт” – a Cyrillic transliteration of “site” and “онлайн” a Cyrillic transliteration of “online”. Neither of the Cyrillic-based applications were from an entity that was also an applicant for the ASCII equivalent.
- Better Living Management Company Limited applied for “.THAI”, a direct phonetic transliteration of the current IDN ccTLD “ไทย”.

With regard to the applications received for transliterations of “.com” and “.net”, WG members discussed at length the potential for confusion that could arise from the addition of these strings to the root. For users of a large number of scripts – such as Arabic or Cyrillic – there would be no audible difference between ASCII-based domain names and IDN “equivalents”.

For example, “yahoo.com” and “ياهو.كوم”, would both be pronounced (ja:’hu: (dot) ’ka:m). Similarly, (’gugl (dot) ’ka:m) could be represented as both google.com or гугл.ком.

End users hearing, for example, a radio advertisement featuring one of these domain names, would not be able to distinguish between them, even though no confusion would exist in visually-based media.

Precisely the same issues arise in the case of “.site” and “.online” and the applied-for Cyrillic transliterations, with the potential for phonetic confusion, irrespective of whether the string is ASCII.ASCII, ASCII.IDN or IDN.IDN.

Finally, this is also the case with “.thai” and “ไทย”, illustrating that the issue of possible confusion, based on phonetic similarity, is not limited to a single type of name or string, but exists in all permutations: existing TLD vs. proposed IDN TLD, new gTLD application vs. new gTLD IDN application and new IDN TLD application vs. existing IDN ccTLD.

Some group members also noted, though the Working Group did not discuss at length, the related, reverse issue of transliteration of characters or strings which have meaning in certain scripts into ASCII-based TLD applications. This is particularly relevant in the case of single character IDN strings. Single-character IDN TLDs have been exempted from the first

round of new gTLD applications, though applications denoting ASCII transliterations have been received as part of the current new gTLD process. Although the Working Group did not take a position on this issue, it notes that any decision from ICANN on transliterated TLDs will necessarily also impact on these applications.

### ***Phonetic similarity and security / phishing issues***

In addition to the potential for user confusion, some WG members also noted that the creation of transliterated TLDs, without the development of adequate registration and eligibility policies and procedures, could give rise to an increased risk of phishing and other malicious abuses of the new spaces.

To use the example of “كوم” (Arabic transliteration of .com, though equally valid in other cases), deployment without adequate policy protections could see cyber-squatters register names such as:

- كوم.ياهو - to resemble “yahoo.com”
- كوم.غوجل - to resemble “google.com”
- كوم.فيسبوك - to resemble “facebook.com”
- كوم.ساب - to resemble “sabb.com” (a well know bank in Saudi Arabia)
- كوم.سامبا - to resemble “samba.com” (another Saudi Arabian bank)
- كوم.ارامكو - to resemble “aramco.com” (a well know Oil company)

The WG notes that this potential problem manifests itself at the second level, and is not unique to transliterated TLDs, but would argue that the very nature of these TLDs, and their close similarity to existing TLDs, makes them particularly high-risk targets. Although related to the potential confusion of users, this is a distinct and separate issue, that should be taken into consideration and addressed.

### ***Mechanisms for recourse: Limitations of the string similarity initial evaluation and “Status to object”***

In terms of how this issue can be formally identified, raised and addressed during the new gTLD application process, the WG notes that section 2.2.1.1 “String Similarity Review” of the applicant guidebook provides for an independent String Similarity Panel to make preliminary assessments regarding potential confusion. Applications deemed to be confusingly similar will be grouped as “contending” strings and will be required to resolve confusion issues before a string or strings can be approved. However, the WG notes that element of the Initial Evaluation process is explicitly limited to visual similarity and will not address the issues raised in this paper.

As a subsequent step for recourse, Section 3.2.2 “Standing to object” of the Applicant Guidebook affords existing TLD operators, an IDN fast track requestor or other new gTLD applicants the right to object on the grounds of string confusion. Therefore, the limitations of these provisions also do not expressly allow for community-based interjection based upon the issues raised in this paper.

However, under the provisions of 3.2.2, it could be reasonably expected that the multiple applicants for “.site” and “.online” (which will already be contending) could object to the applications based upon Cyrillic transliterations of these names, and vice-versa. Similarly, the existing operator of “.thai” has recourse to object to the “.thai” application.

However, with regard to the applications for transliterations of “.com” and “.net”, it is unlikely that current mechanisms would generate an objection, as the only valid objector could be VeriSign Inc – the current, US-incorporated .com operator. Some APTLD members questioned whether this represents a “gap” in the current objection guidelines, given that the applicant, VeriSign SARL, is a related Swiss-incorporated entity related to the current operator. As such, it appears that current provisions for both independent review and interested-party objection in the Applicant Guidebook do not provide a clear mechanism for complaint, in this particular instance.

### ***Points for consideration: Sting-bundling***

As a mechanism for resolving the concerns related to both user confusion and the potential for phishing and other security vulnerabilities expressed in this paper, working group members discussed whether the “bundling” of similar strings together would provide a viable solution. By “bundling”, the WG is referring to a possible circumstance where a single registry operates both ASCII and IDN-based versions of a given string and, through policy requirements, bundles the two together and limits registration rights to those with current or future registrations. The term “parallel provisioning” was also used by some WG members. To use a previous example for illustration, Google, as the current registrant of google.com, would be the only entity eligible to register гугл.ком or google.ком. A user entering any of these names would have their query resolved to the same resource.

This concept is not new, and has been discussed extensively during both the policy processes for new gTLDs and IDN fastrack, though typically focussing upon homographs, or visually-confusing strings or characters.

For example, the GNSO Council provided advice to ICANN in July 2010<sup>1</sup> with regard to the potential modification of the Applicant Guidebook to allow for extended review on the grounds of string similarity. This request was specifically based upon possible circumstances where the same Registry Operator (for an existing gTLD or a proposed new gTLD) could apply for a string that is similar to an existing or applied for string in a manner that is “not detrimentally similar from a user point of view”. The intention of the GNSO Council advice was to allow an applicant to apply for both a gTLD with a conventional ASCII label and a corresponding internationalized gTLD (IDN gTLD) that could be deemed to be similar but not cause “detrimental confusion”.

By way of further example, the ICANN Board resolved<sup>2</sup>, in late 2011, to adopt the advice of a sub-group of ccNSO PDP WG1 regarding the handling of cases of confusing similarity in the IDN ccTLD Fast Track process. This resolution clarified the existing rules, to specify that an

---

<sup>1</sup> <http://gns0.icann.org/en/resolutions#201007>

<sup>2</sup> <http://www.icann.org/en/groups/board/documents/minutes-08dec11-en.htm#2.4>

IDN ccTLD application that is confusingly similar to an ASCII ccTLD, should not be rejected on the grounds of confusing similarity, if the application is made by the same TLD operator.

The views of WG members diverged on the issue of whether a string-bundling policy approach would resolve issues of phonetic confusability. Most agreed that it would offer a potential partial solution and would obviate many of the circumstances where confusion could arise.

For example, KRNIC noted that ICANN has received three gTLD applications in Hangeul script and that two of them (.닷컴, .닷넷) are transliterations of existing gTLDs (.com, .net). KRNIC further noted that, given phonetic similarity may cause some user confusion, the scope of this confusion may be decreased significantly if the registry for .닷컴 and .닷넷 provides 1-1 mapping registration services for existing Hangeul.com and Hangeul.net users (and vice-versa).

However, some WG members also noted pertinent limitations to the idea, were it to be applied to the current new gTLD process:

- In the case of the applications for transliterations of .com and .net, VeriSign Inc. helpfully and openly communicated to APTLD that its intent was to “partially”<sup>3</sup> map the new gTLDs to the existing .com and .net registries. Such an approach has utility, as it would avoid a circumstance where, for example, 10 new .com-transliteration gTLDs would be created, each with over 100 million registrations.

This one-by-one approach to mirroring could resolve the issue of exploitation or malicious registration by third parties by limiting rights to existing registrants. However, it would not resolve all potential for confusion.

That is, an end-user who *hears* a given domain name pronounced in Arabic, may not be able to access the site or resource intended if they are unable to ascertain whether the name is ASCII.ASCII, ASCII.IDN, or IDN.IDN. While their query will not be “misdirected” to an unintended site, it will also not be resolved correctly, if the particular domain name has not been applied for in all phonetic representations. Given the prevalence and ever-increasing uptake of IDNs, it would be inappropriate to simply assume that users would necessarily use “.com” in the first instance.

- Once again, while VeriSign Inc. helpfully and openly communicated its intentions to APTLD members prior to the establishment of the WG, it should be noted that the actual applications received by ICANN for transliterations of .com and .net do not expressly state the intention to bundle these transliterations to the existing TLDs. Rather, the applications rather obscurely state (using the example of the Korean / Hangeul transliteration of .net):

### *2.1 Business Goals*

*Our goal is for KOREAN\_TRANSLITERATION\_OF\_.NET to operate as a best-in-class IDN registry. Although the KOREAN\_TRANSLITERATION\_OF\_.NET gTLD **is distinct from** the .net*

---

<sup>3</sup> “Partially”, in the sense that not all registrations in the existing TLDs would be automatically created in the transliterated TLDs, but rather would be activated upon application by the existing registrant.

*gTLD in the DNS, we plan to provide a similar high quality of service that users of .net have come to expect.*

This potential for confusion and/or opacity regarding business goals and intent to string-map should be resolved during the application and assessment process.

- All previous and existing discussions regarding mapping of TLDs revolve around the concept of a single registry operator. Some WG members noted that, legally, the existing operator of .com (VeriSign, Inc) is a different entity to the applicant for transliterated new gTLDs (VeriSign Sarl). Although this is largely a technical / legal point, it is important to note that each one has, or will have, a separate contract/agreement with ICANN, giving rise to the question of whether the registry operator is precisely “the same”. While the WG cannot offer a definitive interpretation of whether this structural separation is, or is not, an issue for major concern, it nonetheless raises it for the consideration of ICANN staff and the new gTLD evaluation panels. Once again, the WG emphasises that the issue raised is equally applicable to all new gTLD applications and the question of legal status regarding VeriSign entities is only presented as an example.
- One-to-one mapping between languages and scripts can become a very onerous and complicated process. The WG notes, for example, that “Google” can be written in Arabic in many forms such as *جوجل، قوغل، قووجل، جوجل، قوغل*. The implementation of bundling algorithms could require experts in both languages and extensive policies and rules to secure the registry, requiring significant time and effort. Even if the most simply technical implementations were proposed, there is no current standard for such a process, potentially resulting in different solutions across registries, exacerbating the potential for user confusion. The WG once again acknowledges that this a universal issue for new gTLD applicants - not just transliterated TLDs. The WG also recognises that second-level registrations and related policies are largely issues for the registry operator and registrants. However, it is raised here because the potential for confusability and resultant damage to trust, security and stability, is greater in the case of transliterated names.
- IDN processes have their own bundling mechanism to overcome the character variants with Unicode tables. With the combination of IDN bundling and language/ bundling as proposed for transliterated TLDs, the potential exists for greater overall complexity and cost in registry operations.
- Separate and aside from the example regarding transliteration of existing TLDs that has been cited above, the potential introduction of mapping or bundling provisions for competing new TLDs could create an impediment to a number of applicants. As noted previously, the applicants for “.online” and “.site” and their Cyrillic transliterations are distinct business entities. No one has applied for both the ASCII and IDN strings. While the applicants could initiate objection processes regarding each other’s applications, this is purely voluntary. Mediation may deliver a solution, but this is not guaranteed. If bundling were determined as an appropriate policy approach to resolve phonetic confusion, there would be significant, potentially unresolvable, dispute between the applicants, each of which would presumably assert an independent and valid right to operate their applied-for registry, distinct from the policy and business goals of the other.



Finally, some WG members raised the issue of cost and the way in which ICANN intends to charge registry fees with regard to entities that operate both new and existing TLDs and intend to implement parallel or bundled provisioning, at the policy level, between them.

For the sake of clarity, this issue is raised not with regard to new gTLD applicants that intend to map entire trees to an existing TLD, but to where a registry takes a proactive policy decision to improve user experience by parallel provisioning only in cases where two labels are so similar that users may reasonably expect them to be simultaneously resolvable and point to the same resource. This practice is already common with regards to IDN variants, where the registrant is often charged a single fee for the provisioning.

Such a registry-level policy decision may be taken to facilitate stability, security and predictability, and yet may also incur significant cost burdens under traditional registry pricing arrangements. Current fee structures could act to discourage registry operators from taking a policy-based mapping decision that would potentially lessen user confusion and the risk of security or phishing-related issues.

While the registry could, alternatively, incur lower fees by activating one label and reserving the other, confusingly similar label, this approach would not improve the end-user experience.

Therefore, it would be useful for ICANN to state whether or not it intends to introduce some form of differentiated pricing structure for new TLDs that take the policy directions outlined above.

### ***Linguistic imposition***

As a final point, a number of WG members expressed their concerns about the effect of adopting strings for common usage that have no semantic meaning in the languages and scripts that they represent.

Some examples of terms arising from the new gTLD application round have already been widely “borrowed” or adopted from English. “.сайт” (.site) and “.онлайн” (.online) are two such examples. However, many others, such as “كوم” have no semantic meaning at all and are not proposed by the local language community. Although research regarding relevance and common-usage may have been undertaken in advance of these strings being applied for, some WG members hold that the introduction of these strings may still represent a negative impact of strings being “forced” onto the targeted language communities.

## **CONCLUSION**

The APTLD *Ad Hoc* Working Group on transliterated TLDs welcomes the opportunity to provide a conduit to express the views and observations of APTLD members, as a contribution to the public commentary period for the new gTLD application process. Although its scope and mission do not limit the WG to developing input to new gTLDs alone, this

process was targeted as a priority, given the immediacy of input deadlines, and the number of applications that provide tangible examples of the WG's concerns.

We hope that ICANN and the broader community involved in the new gTLD process take the comments of various APTLD members into account and that they are used as a basis for further deliberation and decision-making that will result in a process that delivers the best possible outcomes with regard to the stability and security of the DNS and mitigation of the potential for end-user confusion.