ccNSO
Emoji Study Group
Outline

TABLE OF CONTENTS

Commented [BT1]: Should we just call this SSAC Advice?

Commented [BT2]: Just wondering should we not use IN vs AS given you could have valid IDNA2008 characters mixed in with emoji's?

Commented [BT3R2]:

Commented [BT4]: Will include this in my next version.

INTRODUCTION AND SUMMARY

There has been community interest in the use of emoji in domain names and some country code Top Level Domains (ccTLDs) allow domain names with emoji to be registered at the second level, and several registrars accept labels with emoji for registrations in other TLDs. The SSAC has analyzed the use of emoji for domain names and published the findings in the SAC 095 advisory, which was published on 25 May 2017. Based on their findings, the SSAC recommends not allowing the use of emoji in TLDs, and discourages their use in a domain name in any of its labels. The SSAC also advises registrants of domain names with emoji that such domains may not function consistently or may not be universally accessible.

After considering the SAC095 Report the ICANN Board passed the following resolutions regarding the use of emoji's in domain names[1]:

> *Resolved (2017.11.02.09), the Board hereby directs that conformance to IDNA2008 and its successor will continue to be a necessary condition to determine valid IDN TLD labels.*
>
> *Resolved (2017.11.02.10), the Board requests that the Country Code Names Supporting Organization (ccNSO) and the Generic Names Supporting Organization (GNSO) engage with the SSAC to more fully understand the risks and consequences of using a domain name that includes emoji in any of its labels, and inform their respective communities about these risks.*
>
> *Resolved (2017.11.02.11), the Board requests that the ccNSO and GNSO integrate conformance with IDNA2008 and its successor into their relevant policies so as to safeguard security, stability, resiliency and interoperability of domain names.*
>
> *Resolved (2017.11.02.12), the Board directs the ICANN CEO, or his designee(s), to engage with gTLD and ccTLD communities on the findings and recommendations in SAC095.*

The study group was established by the ccNSO Council in response to the resolutions by the ICANN Board and to provide to ccTLD community and the ccNSO Council a comprehensive overview of the issues and practices associated with the use of Emoji in second level domains by ccTLD managers which allow this. If considered appropriate by the Study Group it may advise on a course of further actions.

The Study Group requested information from ccTLDs who could be allowing the registration of domains which include Emoji's (TBC)…….

[1] https://www.icann.org/resources/board-material/resolutions-2017-11-02-en#1.e

Emojis and their use: SAC 095

Commented [MOU5]: Agree

Commented [BT6]: No fond of this title.

Summary of SAC95

The ICANN Security and Stability Advisory Committee (SSAC) considered the use of Emoji's in domain names in its SAC095 Advisory published 25 May 2017. The core of report can be best summarized by listing its Findings and Recommendations:

Findings

Finding 1: Emoji are disallowed by the IDNA standard; domain names with emoji will not be accepted or processed consistently by applications.

Finding 2: Emoji are not required by design, standard, or convention to be visually uniform (one code point displayed the same way in all circumstances) or visually distinguishable (different code points displayed in ways that permit them to be disambiguated regardless of context). As a result, a user will be exposed to problems of confusability and accessibility. Different code points that are rendered the same or one code point that renders differently to different users will lead to inconsistent results depending on the display or rendering technology used.

Finding 3: Emoji modifiers and "glue" arrangements allow for a potentially much larger set of composed multi-codepoint symbols with even greater rendering variation and potential for ambiguous interpretation.

Finding 4: A fundamental property of the DNS is that it is an exact-match lookup service. For a given query, either there is a single name that matches or there is no match. When two domain names are identical in appearance except for ordinary typographic style variations (which, at present, have no equivalent for emoji), but have different underlying code points, they identify two different DNS domains.

Finding 5: It is unrealistic to expect that just because a code point is included in Unicode, it should be used as part of a domain name.

While Unicode is used in the DNS, such usage should be considered secondary and for several reasons outlined in this report the repertoire must be limited.

Unicode, as an encoding system, is intended to accommodate a range of requirements for the encoding of natural language text (including printing). Natural language text has a number of flexibilities, including the assumption that readers will spot typographical and similar errors and be able to deduce what was intended from the context.

Identifiers such as domain names do not generally have such associated context. This is particularly true for TLD labels, where labels in the root zone do not have any linguistic context.

Finally, domain names are used by end users. In ordinary circumstances, they need to be constructed in such a way as to allow easy and accurate transcription by the end user from one context to another.

These factors constrain the number and the classes of Unicode characters suitable to be used as part of a domain name.

Recommendations

Recommendation 1: Because the risks identified in this Advisory cannot be adequately mitigated without significant changes to Unicode or IDNA (or both), the SSAC recommends that the ICANN Board reject any TLD (root zone label) that includes emoji.

Recommendation 2: Because the risks identified in this Advisory cannot be adequately mitigated without significant changes to Unicode or IDNA (or both), the SSAC strongly discourages the registration of any domain name that includes emoji in any of its labels. The SSAC also advises registrants of domain names with emoji that such domains may not function consistently or may not be universally accessible as expected.

Emojis as SLDs: ccTLD experience

One of the tasks of the Study Group is to provide an overview of the need for and current practice by ccTLD managers to allow Emojis as second level domains To achieve this goal the SG first had to identify on more than an anecdotal basis those ccTLDs that accept Emojis as second level domains. As a second step the SG invited these ccTLDs to provide would any presentations, memoranda or any other material deemed relevant by the ccTLD, with the intent to understand the different approaches and incorporate as many perspectives as possible in the overview. The letter inviting the ccTLDs is included in Annex A.

The SG noticed that some ccTLDs have considered accepting Emojis as second level domains, however to date have decided against it. To provide a complete picture, views of some of these ccTLDs are included as well.

Identifying ccTLD accepting Emojis as second level domain names

To avoid an anecdotal approach of the subject, based on articles/ columns in the relevant media, 3 different, easy to use methods were used to identify ccTLDs that allow Emojis. Two of the methods are using a simple script that gets the root zone, picks out the ccTLDs, and looks for emoji domain names that seem likely to be registered if a ccTLD allows emojis.

The first method used the following Emojis:

xn--228h: U+1F618, Face throwing a kiss
xn--ls8h: U+1F4A9, Pile of poo
xn--y8h: U+26BD, Football

The second method extended this search with following Emojis:

unified id	unified character name	emoji shorthand name	A-Label
1F602	FACE WITH TEARS OF JOY	:joy:	xn--g28h
2764	HEAVY BLACK HEART	:heart:	xn--qei
267B	BLACK UNIVERSAL RECYCLING SYMBOL	:recycle:	xn--26h
1F60D	SMILING FACE WITH HEART-SHAPED EYES	:heart eyes:	xn--r28h
2665	BLACK HEART SUIT	:hearts:	xn--g6h
1F62D	LOUDLY CRYING FACE	:sob:	xn--o38h
1F60A	SMILING FACE WITH SMILING EYES	:blush:	xn--o28h
1F612	UNAMUSED FACE	:unamused:	xn--w28h
1F495	TWO HEARTS	:two_hearts:	xn--0r8h
1F618	FACE THROWING A KISS	:kissing_heart:	xn--228h
1F4A9	PILE OF POO	:poop:	xn--ls8h

The 3rd method looked at a third party listing of ccTLDs accepting registrations and check whether their websites.

The SG is aware the methodology used is neither exhaustive with respect to the Emojis registered, nor with respect to the ccTLD allowing Emojis. The SG believes that priority should be given over documenting diversity in practices and experiences within a reasonable time frame over 100% coverage.

Based on the combined methods, the following ccTLDs were identified as allwoing oen or more of the Emojis listed above, and have been approached:

.cf.
.ga.
.gg.
.gq.
.je.
.la.
.ml.
.mp.
.st.
.tk.
.to.
.vu.
.ws.

Experience ccTLDs accepting Emoji as SLDs

.WS Experience

Considerations of ccTLD, which considered accepting registration

(.RS experience)

From Barcelona meeting: (.RS) we had some proposals to introduce emoji at the 2nd level. We have been pushed hard to offer them. We are not ready for introduction, issue of potential misuse is real.

Patrik F: Outreach to TLDs that currently do not offer emoji – based on a conscious evalutation - might be a good idea, asking others to share the results of their evaluation, such as .rs and their decision not to offer emojis. Mirjana Tasic will be happy to send such documentation explaining the evaluation.

Observations Study Group

Annex A : letter to ccTLDs

(ccTLD Manager/operator)

Subject: CCNSO Study on the use of Emoji's at the second level in ccTLDs.

On 26 February 2018 the ccNSO Council constituted the Emoji Study Group (ESG) to provide it with a comprehensive overview of the issues associated with the use of Emoji in second level domains as well as any current practices by ccTLDs which accept such registrations. For more details please refer to <https://ccnso.icann.org/sites/default/files/field-attached/emoji-sld-purpose-scope-activities-26feb18-en.pdf>
You are receiving this communication as the result of some initial ad-hoc work by the ESG which potentially identified your registry as one which accepts the registration of Emojis as or in second level domains. If you do not accept such registrations we apologize for any inconvenience and would appreciate you advising us of this.
If you do accept such registrations we are seeking your assistance with our study. I would greatly appreciate you forwarding to us via Kimberly.carlson@ICANN.org pointers to relevant public information related to this practice in your registry. This might be a list of accepted Emoji's, any technical specifications or requirements.
The ESG would also welcome any presentations, memoranda or any other material you may deem relevant, and you would wish to share on this subject. The intent is to understand the different approaches and incorporate as many perspectives as possible into the Study Group report.
The ESG's current plan is to complete the first draft of its report for the ccNSO Council by ICANN 64 to be held in Kobe Japan and as such if you wish to provide any information, I kindly request you to do so as soon as possible.
For further information on the study group, please refer to
https://ccnso.icann.org/en/workinggroups/emoji-sld.htm
Thanking you for your collaboration

Peter Koch
Chair ccNSO-ESG

Annex B : Detailed explanation of method used to identify ccTLDs

Paul Hoffman
Search the whole ccTLD space. I wrote a simple Python script that gets the root zone, picks out the ccTLDs, and looks for three simple single-emoji domain names that seem likely to be registered if the ccTLD allows emojis. The results are below.
Please let me know if you have any questions.

There are 1537 TLDs
There are 247 ccTLDs
Testing:
xn--228h: U+1F618, Face throwing a kiss
xn--ls8h: U+1F4A9, Pile of poo
xn--y8h: U+26BD, Football

In cf: xn--228h.cf. xn--ls8h.cf. xn--y8h.cf.
In ga: xn--ls8h.ga. xn--y8h.ga.
In gg: xn--y8h.gg.
In gq: xn--228h.gq. xn--ls8h.gq. xn--y8h.gq.
In je: xn--y8h.je.
In la: xn--ls8h.la.
In ml: xn--228h.ml. xn--y8h.ml.
In mp: xn--ls8h.mp.
In st: xn--228h.st. xn--ls8h.st. xn--y8h.st.
In tk: xn--228h.tk. xn--ls8h.tk. xn--y8h.tk.
In to: xn--228h.to. xn--ls8h.to. xn--y8h.to.
In vu: xn--ls8h.vu.
In ws: xn--228h.ws. xn--ls8h.ws. xn--y8h.ws.

Peter Koch
Sample with a slightly different set of code points.
It started with the 10 most popular emojis on
<https://urldefense.proofpoint.com/v2/url?u=http-3A__www.emojitracker.com&d=DwICAg&c=FmY1u3PJp6wrcrwll3mSVzgfkbPSS6sJms7xcl4I5cM&r=nC4gfbyyEQpMRtwgTZRqwi_KEPkorJJN6BPIwtsjBWQ&m=LtGzoolCKwI-29r209wOwgXBoyM_IGQGfgdNZAWAZt4&s=tnyGyOPtXBHE1oWll2J5bQrkRpl3d0FoV89QEKemlYs&e=>
and tested for the infamous 'xn--ls8h' as well as for the DNS wildcard '*' explicitly. Where a TLD is marked as supporting '*', the column will flag any other label only if that's delegated. Due to the wildcard, any label will match a DNS query, of course.

GA is missing in the sample due to connectivity issues during taking
the sample from my particular vantage point.

unified id	unified character name	emoji shorthand name	A-Label
1F602	FACE WITH TEARS OF JOY	:joy:	xn--g28h
2764	HEAVY BLACK HEART	:heart:	xn--qei
267B	BLACK UNIVERSAL RECYCLING SYMBOL	:recycle:	xn--26h
1F60D	SMILING FACE WITH HEART-SHAPED EYES	:heart_eyes:	xn--r28h
2665	BLACK HEART SUIT	:hearts:	xn--g6h
1F62D	LOUDLY CRYING FACE	:sob:	xn--o38h

```
1F60A           SMILING FACE WITH SMILING EYES          :blush:                 xn--o28h
1F612           UNAMUSED FACE                           :unamused:              xn--w28h
1F495           TWO HEARTS                              :two_hearts:            xn--0r8h
1F618           FACE THROWING A KISS                    :kissing_heart:         xn--228h

1F4A9           PILE OF POO                             :poop:                  xn--ls8h
```

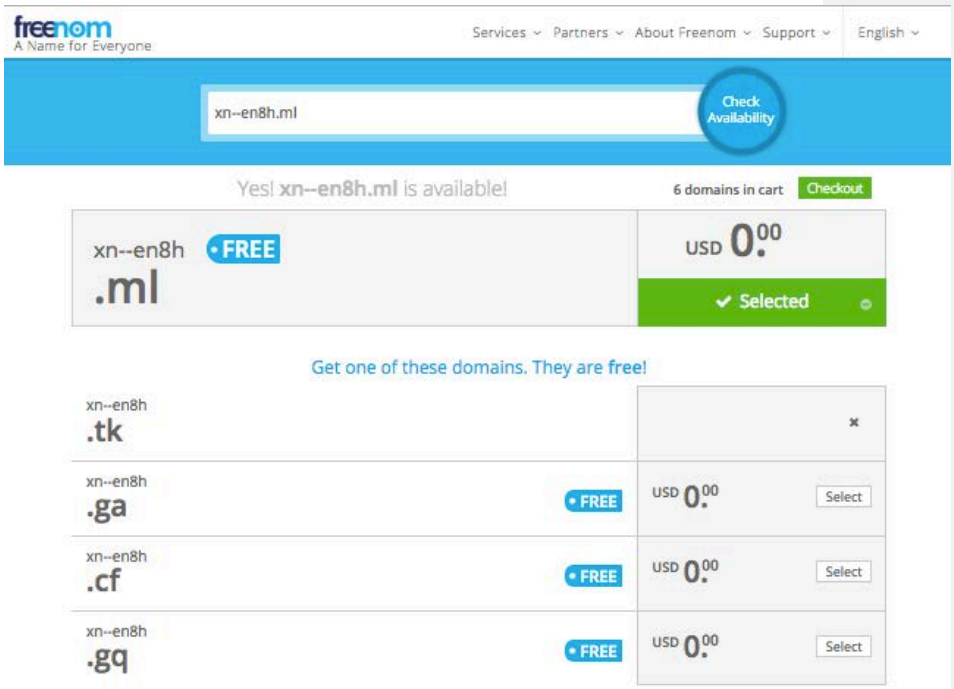
*	FM.	LA.	PH.	VG.	WS.
xn--g28h	CF.	GQ.	ML.	ST. TK.	UZ. VG.	
xn--qei	CF.	GG. GQ. JE.	ML.	ST. TK. TO.	UZ. VG.	WS.
xn--26h			ML.	TK. TO.	VG.	WS.
xn--r28h		LA.		ST. TK.	UZ. VG.	WS.
xn--g6h	CF.	GQ.	LA. ML.	TK. TO. TV.	UZ. VG.	WS.
xn--o38h	CF.	GQ.	ML.	TK. TO.	VG.	WS.
xn--o28h	CF.	GQ.	ML.	TK. TO.	UZ. VG.	WS.
xn--w28h	CF.	GQ.		TK.	VG.	WS.
xn--0r8h	CF. FM.	GQ.	ML.	TK. TO.	VG.	WS.
xn--228h	CF.	GQ.	ML.	ST. TK. TO.	UZ. VG.	WS.
xn--ls8h	CF.	GQ.	LA. ML. MP.	ST. TK. TO.	VG. VU.	WS.

```
This is mostly consistent with what Paul and Alejandra have found already.
We should keep in mind that the DNS wildcard or any special 'tricks' applied
in responses might taint the result for some TLDs. Also, so far we've probably
only looked for single 'character' labels to avoid combinatorics kicking in,
so our samples do not claim completeness.  To that extent, the 'tests' are
a little less reliable than those applied by the wildcard study group.
```

Policy review: Alejandra Reynoso

.cf - none

1. .ga - none (seems to have the same website provider as .cf, same message error: Not Found. The requested URL /registration/registerpaid was not found on this server.)
2. .gg & .je - none
3. .gq - none (seems to have the same website provider as .cf & ga, same message error: Not Found. The requested URL /registration/registerpaid was not found on this server.)
4. .fm: https://dot.fm/emoji/ [dot.fm] - policy: https://dot.fm/policy.cfm [dot.fm] (Appendix for the EmojiDomain Names (Emoji).)
5. .la - none
6. .ml - none (seems to have the same website provider as .cf, ga & gq, but no error message this time). Maybe this are free domains? I searched for white flag xn--en8h, because the common ones were already taken.



1.
7. .mp - none
8. .st - none
9. .tk - none (similar interface as .ml)
10. .to - none
11. .vu - none
12. .ws - none in the official samoanic.ws [samoanic.ws] - but here https://xn--i-7iq.ws/ [i%e2%9d%a4.ws] there is a lot of emoji publicity and even a timeline of the evolution of emoji domain names.

Annex C. Terms of Reference

Annex D. Membership