



78 | ANNUAL
GENERAL
MEETING



Internationalized Domain Name (IDN) Implementation Roundtable



Pitinan Kooarmornpatana
IDN Program Senior Manager

ICANN78
24 October 2023

Agenda

- IDN Implementation Guidelines Version 4.1 (Pitinan Kooarmornpatana)
- Additional Reference LGR: Indonesian Scripts (Yudho Giri Sucahyo, Cokorda Pramatha)
- Policy Implementation Related to New gTLD Program Next Round
 - Single Character TLDs (Kenny Huang, Hiro Hotta, Kim Kyongsok)
 - Planning for String Similarity Review (Pitinan Kooarmornpatana)
- Q&A

IDN Implementation Guidelines Version 4.1

Pitinan Kooarmornpatana, IDN Program Senior Manager

Background

- The IDN Implementation Guidelines were developed by the community and designed to minimize the risk of cybersquatting and consumer confusion under TLDs. The notice of implementation for version 3.0 was published in 2011.
 - Contractually binding through the gTLD Registry Agreement.
 - Recommended through the IDN ccTLD Fast Track Process.
- Version 4.0 was [published](#) on 10 May 2018. The GNSO Council requested the ICANN Board to defer the adoption.
- GNSO IDN Expedited Policy Development Process (IDN EPDP) was formed to develop IDN-related policy overlapping with some guidelines.
- After subsequent discussion between the ICANN Board and GNSO Council, the ICANN Board deferred implementation of guidelines 6a, 11, 12, 13, 18 in version 4.0 as they overlapped with ongoing work in the IDN EPDP.
- In 2022, the ICANN Board directed ICANN org to publish the remaining guidelines as version 4.1 and implement the remaining guidelines.

Summary of Obligations in v4.1 vs. v.3

- **Guidelines 1-3 (update):** TLD registries that support IDNs must do so in strict compliance with IDNA2008 (already in the Base Registry Agreement since 2013).
 - Follow RFCs 5890, 5891, 5892 and 5893 or any RFC that replaces or updates the listed RFCs.
 - Example: ‘ ▪ **com**’ , the code point ▪ (U+30FB, KATAKANA MIDDLE DOT) requires at least one character in the label to be in Hiragana, Katakana, or Han scripts.
 - Not allow exceptions for registrations that are IDNA2003 compliant but not IDNA2008 compliant.
 - Example: ‘√’ (U+221A), ‘♻’ (U+267A), ‘⦿’ (U+2706), these were PVALID in IDNA2003 but DISALLOWED in [IDNA2008](#)
 - When a pre-existing domain name requires transitional exception, the terms of that action must also be made readily available online.

Summary of Obligations in v4.1 vs. v.3

- **Guidelines 4 (no change):** No label containing hyphens in both the third and fourth positions may be registered unless it is a valid A-label
 - Example: “0-~~-----~~0”, “a-~~---~~b”, “ok-~~---~~computer”
- **Guidelines 5 (update terminology):** A TLD registry must publish one or several repertoires of Unicode code points that are permitted for registration and must not accept the registration of any domain name containing an unlisted code point. Each such list must indicate the script or language(s) it is intended to support.
- **Guidelines 6 (no change):** IDN Tables must be placed in the IANA Repository for IDN Practices.
- **Guidelines 6(c) (update terminology):** The IDN Table must include the complete repertoire of code points, any IDN variant code points and any applicable contextual rules which the TLD registry uses to determine if an IDN label is acceptable for registration.

Summary of Obligations in v4.1 vs. v.3

- **Guidelines 8 (new and aligned with the current practice):** TLD registries may use [Second Level Reference Label Generation Rules \(LGRs\)](#) as is or as a reference. IDN Tables may deviate from Reference Second Level LGRs. Notwithstanding the foregoing, registries seeking to implement IDN Tables (i.e. new or modifications of existing ones) that pose any security and/or stability issues must not be implemented.
 - Notes:
 - This guideline only affects new IDN Tables or the updated IDN Tables of the TLD.
 - The previously approved IDN Tables for the TLD will not be affected by the implementation of this guideline. Any security issues found in the previously approved IDN Table of a TLD have been communicated separately and TLD registries are encouraged to address the issues.

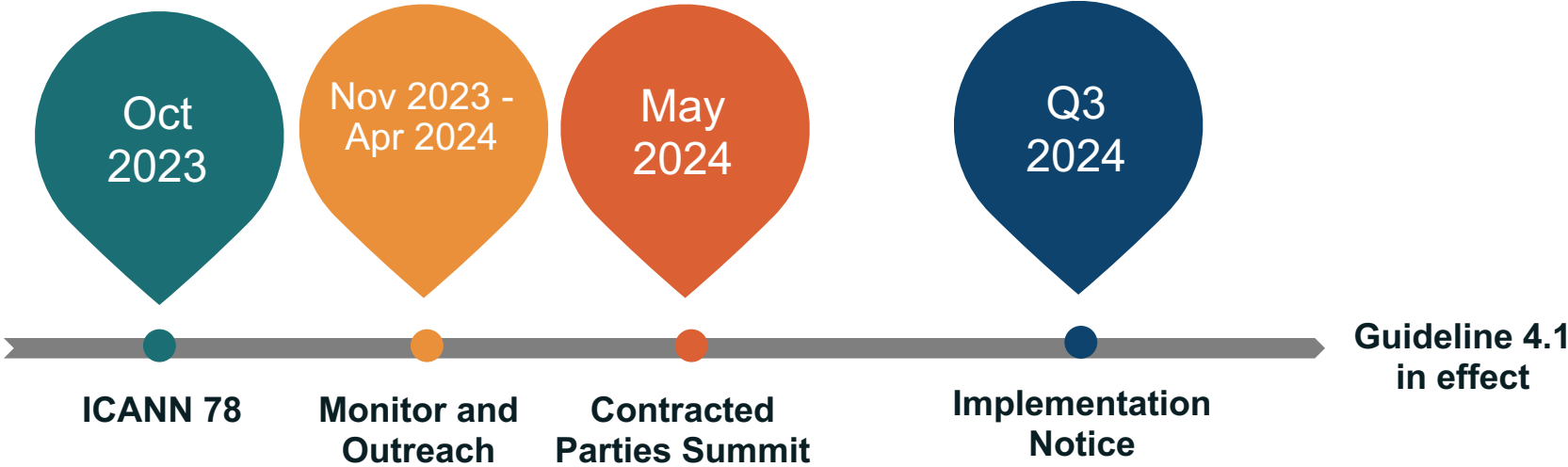
Summary of Obligations in v4.1 vs. v.3

- **Guidelines 10 (no change):** Any information fundamental to the understanding of a TLD registry's IDN policies that is not published by IANA must be made directly available online by the TLD registry.
- **Guidelines 15 (no change):** All code points in a single IDN label must be taken from the same Unicode script as determined by the Unicode Standard. Exceptions are permissible for languages with established orthographies and conventions that require the commingled use of multiple Unicode scripts.
 - Example of violation: 'pay^{pal}' (U+0070 U+0061 U+0079 U+0440 U+0061 U+006C), the mix of Cyrillic and Latin scripts in a label is not allowed.
- **Guidelines 16 (no change):** In the case of any exceptions made allowing mixing of Unicode scripts, visually confusable characters from different scripts must not be allowed to co-exist in a single set of permissible code points unless a corresponding IDN policy and IDN Table is clearly defined to minimize confusion between domain names.

Implementation Approach

- **Guidelines 1–4, 15:** The Guidelines related to the existing registrations and existing IDN Tables.
 - Automated checks for IDNA2008 non-compliant registrations and IDN Tables.
 - Automated checks for the invalid mixed-script labels.
 - If any issue is found, the affected registry operators will receive an email for more details about steps to take (November 2023 - April 2024).
 - Work out the timeline with registry operator for IDNA non-compliance.
- **Guidelines 5, 6, 6(c), 8, 10, 16:** The guidelines related to new or updated IDN Tables.
 - No change in the current process. Continue using the existing IDN Table Review Process under the IDN Services and RST Process.
 - Continue to use reference LGRs and IDN Table tool for transparency and consistency as requested by the RySG.
- Other guidelines are non-binding and contracted parties are encouraged to adopt.
- All guidelines will be promoted through the communication plan.

Timeline



Preparation Phase	● Oct 2023: share implementation approach at ICANN78. ● Nov 2023 - Apr 2024: - Run the automated checks on existing registrations and existing IDN Tables. - Communicate any issues found to relevant registry operators. - Collaborate with registry operators to address the issues. ● May 2024: - Communicate the implementation of IDN Guidelines v4.1 to ROs. - Share updates and progress at CP Summit 2024.
Implementation Notice	● Q3 2024: deliver legal notice on IDN Guidelines Version 4.1 being in effect.
IDN Guidelines v4.1 in Effect	● Regularly run automated check for registrations. ● Continue the IDN Table Review Process under IDN Services and RST Process.

Reference LGRs and IDN Table Review Tool

- [The reference LGRs](#), developed using input from the relevant script communities, are generally consistent with the analysis done by the community for the [RZ-LGR](#).
- All reference LGRs include consultations with the script communities and are finalized after Public Comment proceedings.
- TLD registry operators can use reference LGRs when they develop rules for IDN registration under their TLDs.
- The online [IDN Table Review Tool](#) is available for reviewing the designed IDN Table with a reference LGR or RFCs' based requirements. The IANA table is being used for the IDNA2008 check. The latest version is based on Unicode 12.0.0.
- ICANN org uses reference LGRs and the LGR Tool to review the registration rules under gTLDs.
- This contributes to consistency and transparency of the review as requested by the Registries Stakeholder Group (RySG).

- Additional 7 script-based LGRs and related updates will be published which will bring the number of reference LGRs to 54.
- 31 language-based LGRs
 - Arabic, Belarusian, Bosnian (Cyrillic), Bosnian (Latin), Bulgarian, Chinese, Danish, English, Finnish, French, German, Hebrew, Hindi, Hungarian, Icelandic, Italian, Japanese, Korean, Latvian, Lithuanian, Macedonian, Montenegrin, Norwegian, Polish, Portuguese, Russian, Serbian, Spanish, Swedish, Thai, and Ukrainian.
- 23 script-based LGRs
 - Arabic, Armenian, Bangla (Bengali), Cyrillic, Devanagari, Ethiopic, Georgian, Greek, Gujarati, Gurmukhi, Hebrew, Japanese (Hiragana, Katakana, Kanji[Han]), Kannada, Korean (Hangul, Hanja [Han]), Khmer, Lao, Latin, Malayalam, Myanmar, Oriya, Sinhala, Tamil, and Telugu.

Additional Reference LGR: Indonesian Scripts

Prof. Yudho Giri Sucahyo, Ph.D
Cokorda Rai Adi Pramatha, Ph.D

Background of Indonesian Scripts

- More than **17.000** islands, **600** ethnic groups, **719** mother-tongue languages (**13 have vanished**), **37 using non-Latin scripts**.
- In digital world, when a language is not present → not exist.
- 9** scripts in UNICODE:
 - Limited**: Balinese, Batak, Javanese, Makasar, Sundanese.
 - Excluded**: Buginese, Kawi, Rejang
 - Recommended**: Pegon
- Since **2019**, PANDI has been actively **increasing public awareness** on the importance of indigenous language **digitization** and **digitalization**.

Unicode 15.1 Character Code Charts

[SCRIPTS](#) | [SYMBOLS & PUNCTUATION](#) | [NAME INDEX](#)

Find chart by hex code: [Help](#) [Conventions](#) [Terms of Use](#)

11 Bahasa Daerah Punah, 25 Terancam

Sebanyak 11 bahasa daerah dinyatakan punah oleh Badan Pengembangan dan Pembinaan Bahasa Kementerian Pendidikan dan Kebudayaan. Enam lainnya dinyatakan kritis dan 25 terancam punah.

Punah

- Bahasa Tandia Papua Barat
- Bahasa Mawas Papua
- Bahasa Kayeli Kayeli Maluku
- Bahasa Piru Maluku
- Bahasa Mokela Maluku
- Bahasa Palumata Maluku
- Bahasa Ternate Maluku Utara
- Bahasa Hukuma Maluku
- Bahasa Hoti Maluku
- Bahasa Milla Maluku
- Bahasa Sersa Maluku

Terancam punah

- Bahasa Hulung Maluku
- Bahasa Samasuru Maluku
- Bahasa Mander Papua
- Bahasa Namia Maluku
- Bahasa Uku Papua
- Bahasa Maklew/Maklew Papua
- Bahasa Bku Papua
- Bahasa Mansim Borai Papua
- Bahasa Ponsokan Sulawesi Utara
- Bahasa Koriw Sulawesi Selatan
- Bahasa Bajau Tungkal Satu Jambi
- Bahasa Lematang Sumatera Selatan
- Bahasa Dubu Keerom Papua
- Bahasa Irarutu Fakfak Papua
- Bahasa Podena Sarmi Papua
- Bahasa Sangihe Talaud Minahasa Sulawesi Utara
- Bahasa Minahasa Gorontalo
- Bahasa Nidebang NTT
- Bahasa Adang Gorontalo
- Bahasa Adang NTT
- Bahasa Benggaun Sulawesi Barat
- Bahasa Bobat Maluku
- Bahasa Namia Papua
- Bahasa Arguni (Taver) Papua Barat
- Bahasa Kalabra Papua Barat

Kritis

- Bahasa Rota Papua NTT
- Bahasa Sapori Papua
- Bahasa Ibo Maluku
- Bahasa Meher Maluku Tenggara Barat
- Bahasa Letti Maluku
- Bahasa Sapori Papua

Infografis: Adhian Yuhai | Naskah: Febyan Ramayanti | Sumber: Badan Pengembangan dan Pembinaan Bahasa Kementerian Pendidikan dan Kebudayaan

CN

Indonesian & Philippine Scripts

Balinese

Batak

Buginese

Buhid

Hanunoo

Javanese

Kawi

Makasar

Rejang

Sundanese

Sundanese Supplement

Tagalog

Tagbanwa

10 October at 01:32 · 🌐



10 October at 01:32 · 🌐

Together with [#ICANN](#), find out what you can do to support IDNs. <https://go.icann.org/3PNfFlk>



1. Javanese: 84,3 M
2. Sundanese: 34 M
3. Buginese: 3,5 M
4. Balinese: 3,3 M

Background of Indonesian Scripts

Some of PANDI's efforts that have been made towards the IDN goal:

- Collaborated with **UNESCO** and worked with more than **30 cultural institutions** in Indonesia, consisting of elements of universities, government, and the script and cultural communities.
- Organized more than **40 activities** (competitions, symposiums, seminars, webinars, congresses) to encourage the **standardization** of the archipelago's script and its **massive** use in the **digital realm**.
- Worked together with the **National Standardization Agency** (BSN) to obtain: Establishment of SNI FONT and KEYBOARD for **Javanese, Sundanese, Balinese, Kawi**, and **Pegon** Script.
- Improving the status of the Indonesian scripts in UNICODE from **Limited Use** to **Recommended** by optimizing the digitization of scripts.
- Working together with script **experts** especially in **Balinese, Javanese**, and **Pegon** which directly directed by ICANN to make a **Label Generation Rules** (LGR) proposal for the **second level domain** (SLD).

Balinese Script: Significance

- The need for using Balinese script for the younger generation.
- Bali Governor Regulation **number 80 of 2018** concerning the Protection and Use of Balinese Language, Script and Literature and the Implementation of Balinese Language Month
- Bali Province Regional Regulation **number 4 of 2020** concerning Strengthening and Advancing Balinese Culture



Balinese Script: Categorization

- Vowel script (*Aksara Suara*)

- Dependent vowel (14)
- Independent vowel (14)

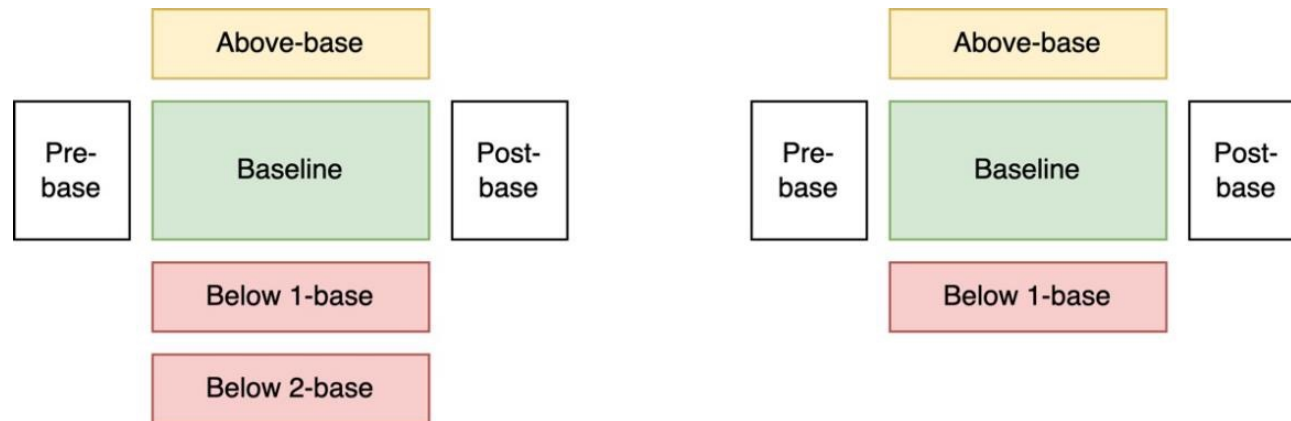
◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌
◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌

- Consonants

- *Wreastra* (18)
- *Swalalita* (15)

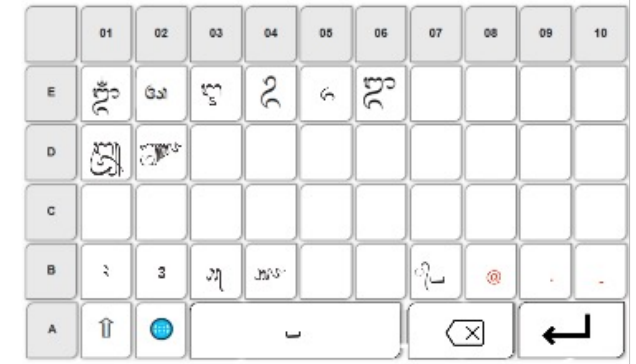
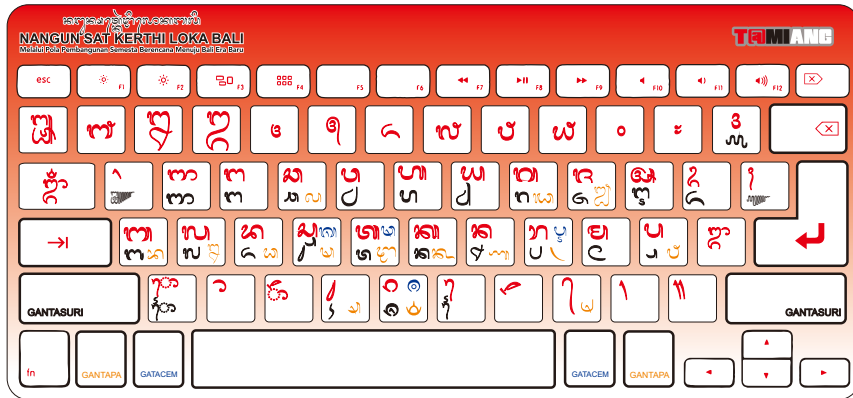
◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌
◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌

- Signs (7) ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌ ◌◌



“◌◌◌◌◌◌◌◌” > “◌◌◌◌◌◌◌◌”

Balinese Script: Supporting Ecosystem/tools



1. = Vimala
2. = Pustaka Bali
3. = Kadiri
4. = Noto Sans Balinese
5. = Natya
6. = Bali Banat
7. = Bali Raung
8. = Mangu sastra
9. = Narada
10. = Sapuh Luih
11. = Sudamala
12. = Tutud Bali
13. = Urdhasastra
14. = Singhambara



KEPUTUSAN KEPALA BADAN STANDARDISASI NASIONAL
NOMOR 545/KEP/BSN/11/2021
TENTANG
PENETAPAN SNI 9047:2021 FON (FONT) AKSARA NUSANTARA

KEPALA BADAN STANDARDISASI NASIONAL,



KEPUTUSAN KEPALA BADAN STANDARDISASI NASIONAL
NOMOR 546/KEP/BSN/11/2021
TENTANG
PENETAPAN SNI 9048:2021 TATA LETAK PAPAN TOMBOL AKSARA
NUSANTARA

KEPALA BADAN STANDARDISASI NASIONAL,

Balinese Script: Repertoire

No.	Code Point	Glyph	Name
1	1B02	ᮀ	BALINESE SIGN CECEK = anusvara
2	1B03	ᮁ	BALINESE SIGN SURANG = final r • also used for repha in transliteration of Kawi → A982 ᮘjavanese sign layar
3	1B04	ᮂ	BALINESE SIGN BISAH = visarga
4	1B05	ᮃ	BALINESE LETTER AKARA = a
5	1B06	ᮄ	BALINESE LETTER AKARA TEDUNG = aa ≡ 1B05 ᮃ 1B35 ᮀ

No.	Code Point	Glyph	Name
63	1B41	ᮓᮀ	BALINESE VOWEL SIGN TALING REPA TEDUNG = au ≡ 1B3F ᮓᮀ 1B35 ᮀ
64	1B42	ᮓᮁ	BALINESE VOWEL SIGN PEPET = ae
65	1B43	ᮓᮂ	BALINESE VOWEL SIGN PEPET TEDUNG = oe ≡ 1B42 ᮓᮁ 1B35 ᮀ
66	1B44	ᮓᮃ	BALINESE ADEG ADEG = virama

Balinese Script: Variants

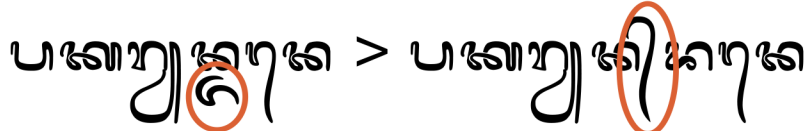
- Different shape but can be interchanged

Unicode Code Point	Glyph	Unicode Code Point	Glyph	Sound
1B14	ꦏꦲ	1B13	ꦏꦲ	Ka
1B16	ꦒꦲ	1B15	ꦒꦲ	Ga
1B19	ꦕꦲ	1B18	ꦕꦲ	Ca
1B1B	ꦗꦲ	1B1A	ꦗꦲ	Ja
1B1D	ꦠꦲ	1B22	ꦠꦲ	Ta
1B1E	ꦠꦲ	1B22	ꦠꦲ	Ta
1B1F	ꦢꦲ	1B24	ꦢꦲ	Da
1B20	ꦢꦲ	1B24	ꦢꦲ	Da

Unicode Code Point	Glyph	Unicode Code Point	Glyph	Sound
1B21	ꦚꦲ	1B26	ꦚꦲ	Na
1B23	ꦠꦲ	1B22	ꦠꦲ	Ta
1B25	ꦢꦲ	1B24	ꦢꦲ	Da
1B28	ꦥꦲ	1B27	ꦥꦲ	Pa
1B2A	ꦧꦲ	1B29	ꦧꦲ	Ba
1B30	ꦱꦲ	1B32	ꦱꦲ	Sa
1B31	ꦱꦲ	1B32	ꦱꦲ	Sa

Balinese Script: Rules

- Rules to write the Balinese Script:

- *Pasang jajar*: without spaces between word *scripto continua*
- *Pasang palas*: with space between word
 - Pakaryan dane  pakaryan dane

- Permissible Entities (Hypnen)

- www.dumogi-becik.id :  $\text{www.dumogi-becik.id}$

- Not Permissible (Code-Page Mixing)

-  $\text{www.dumogi-becik.id}$ (not allowed)

- Not Permissible (Digits and Punctuation)

-  $\text{www.dumogi-becik.id}$

Policy Implementation Related to New gTLD Program Next Round

Single Character TLDs

Update from the Chinese Script Generation Panel (CGP)

- CJK meeting
 - 05/23; 06/27; 07/19; 07/27
- CGP meeting
 - 03/22; 05/31
- CGP considerations
 - Single character domain policy has been reviewed in CNNIC/TWNIC.
 - The purpose of establishing CJK GP is to develop CJK Generation Rules, not for the creation of single character CJK TLDs. Evaluating the technical eligibility of single character CJK TLDs falls outside the scope of CJK GP.
- CGP statement
 - CGP reckons that the “Final Report on the new gTLD Subsequent Procedures Policy Development Process”, as well as the “Proposal for a Chinese Scripts Root Zone LGR”, are suffice to address any lingering concerns surrounding the application and delegation of Chinese single character TLD in the upcoming sub pro round, thus no need to provide any additional rules for IDN EPDPD WG for that matter

Update from the Japanese Script Generation Panel (JGP)

- JGP (Japanese Generation Panel) was dissolved (March 2022).
 - After that, ex-Chair Hiro Hotta has been the contact point to the Japanese IDN community in place of JGP
- CGP, JGP, and KGP were asked by GNSO and ICANN about how the rules and/or guidelines of single Han character TLDs should be (May 2022).
- Positions to that question from Japanese script community is :
 - From the Japanese script community perspective only, **any Han character that RZ-LGR allows as a TLD string should be allowed** as a single Han character TLD.
 - However, **if CGP and/or KGP consider there should be restriction imposed on single Han character TLDs, they must be respected and adopted.**

Update from the Korean Script Generation Panel (KGP)

- Most of other scripts' variants: ONLY shape variant groups (SVG)
- Current variant definitions for Korean script (Hangul and Hanja (Han) scripts) in the RZ-LGR:
 - 1) Within Hangul script: NO SVG, NO meaning variant group (MVG)
 - 2) Within Hanja (Han) script: ONLY MVG (meaning variant group)
[(機 6A5F), (机 673A)] – different shapes, the same meaning
 - 3) Between Hangeul and Hanja (Han) script: NO MVG, 7 (seven) SVG
[(𐮎 C2A5), (今 4ECA)], [(슴 C2B4), (合 5408)]: similar shapes
- For single character Hanja (Han) gTLD:
 - Question: Any additional SVG for Hanja (Han) needed?
SVG examples: [書 畫 畫], [慄 慄], [梁 梁], [壺 壺], [刺 刺], [壘 壘] – in an address window
- Currently KGP is trying to collect opinions from the Korean community.

Policy Implementation Related to New gTLD Program Next Round

Planning for String Similarity Review

String Similarity Review at ICANN78

- **Planning for String Similarity Review in the Next gTLD Round**
- Thursday, 26 October • 09:00 - 10:00, Hall 1 or join remotely [here](#).
- Agenda:
 - Scope, Methodology and Outcomes of String Similarity Review
 - String Similarity Review Guidelines

Engage with ICANN – Thank You and Questions



One World, One Internet

Visit us at **icann.org**



@icann



facebook.com/icannorg



youtube.com/icannnews



flickr.com/icann



linkedin/company/icann



soundcloud/icann



instagram.com/icannorg