

Introduction to Internationalized Domain Names (IDNs)

Pitinan Kooarmornpatana
IDN Program Senior Manager, ICANN

AFRALO Webinar
17 November 2022

ASCII Domain Name Label

www.cafe-123.com



2

Forming ASCII Labels

Use LDH

- Letters [A..Z a..z]
- Digits [0..9]
- Hyphen [-]

Label length = 63

Other constraints (e.g., on hyphen)

1

Forming ASCII Labels

Use only Letters

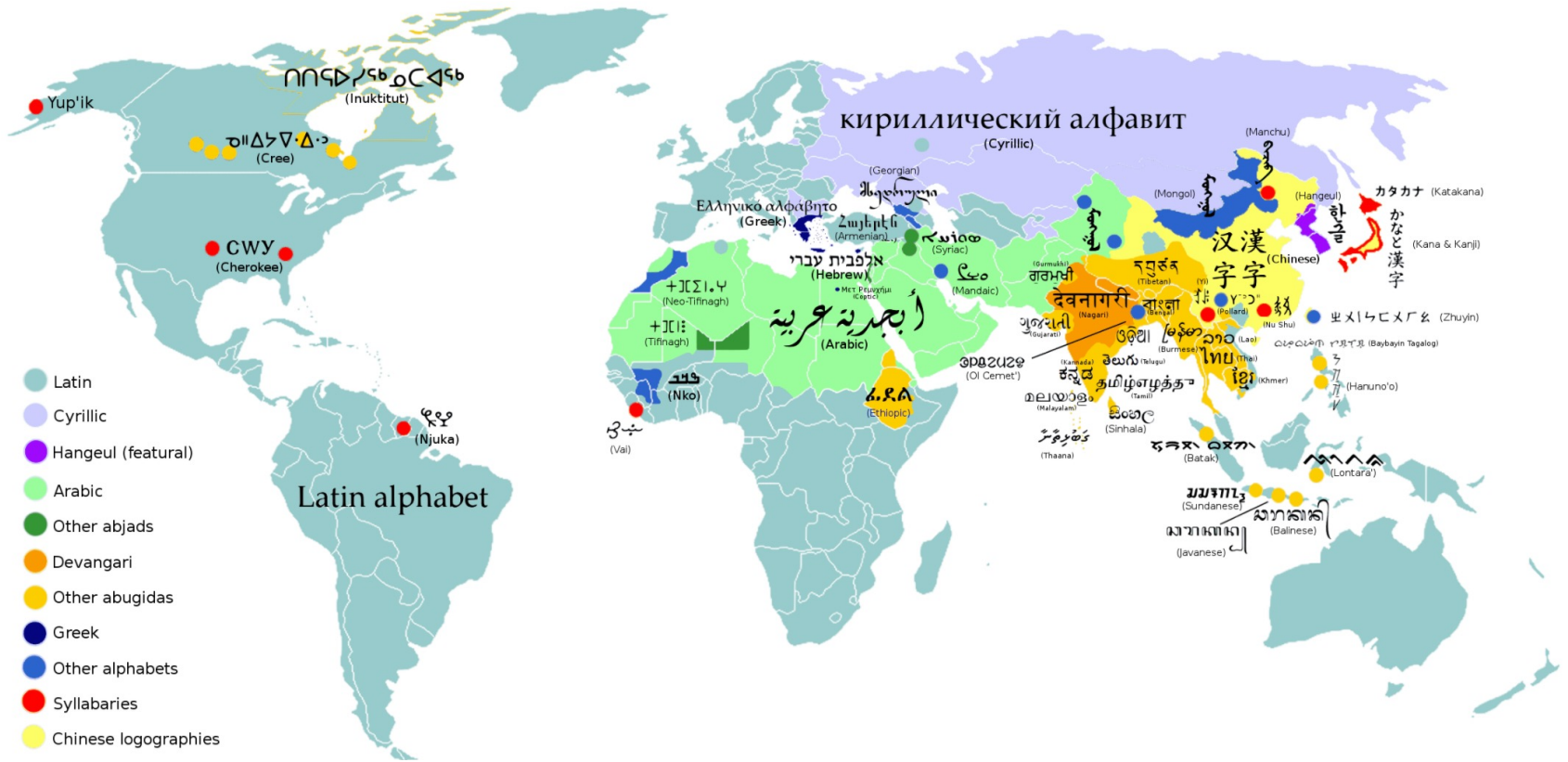
- Letters [A..Z a..z]

Label length = 63

Why Do We Need Internationalized Domain Names (IDNs)?

“The Web does not just connect machines, it connects people.”
-- Tim Berners-Lee

“Language is just as important to building human connections online as it is offline.”
-- [The Digital Language Divide](#)



Scripts Used in the Africa Region

ASCII

lubakasai
northernsotho swahili
afrikaans farefare
malagasy chokwe swazi
bemba kinyarwanda
nyungwe southernndebele
xhosa shona
english seychelloiscreole
sotho luvale zulu ndebele
swati kituba kikongo
capeverdeancreole
umbundu ganda lingala
mampruli

Latin Script

chichewa dazaga
hassaniyya bulu oromo
wolof fang guinea oshiwambo
tamazight mandinka
portuguese venda tonga
tswana bissau creole jula krio
kimbundu gen tem
hausa igbo lozi bari maba yoruba
pular zande
kanembu equatorialguineanpidgin
bagirmi tsonga

Ethiopic script

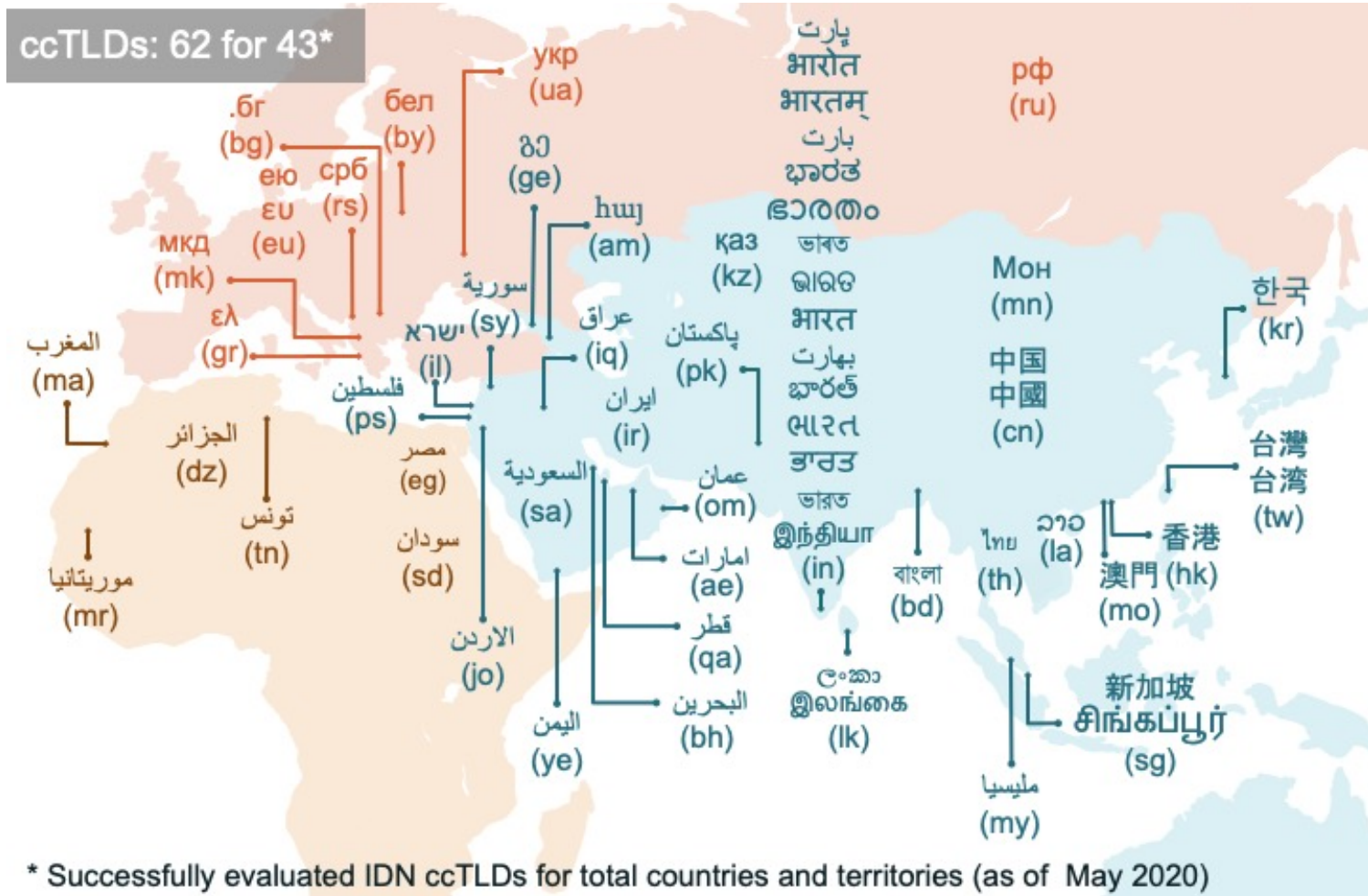
tigrina
tigriya
amharic

Arabic script

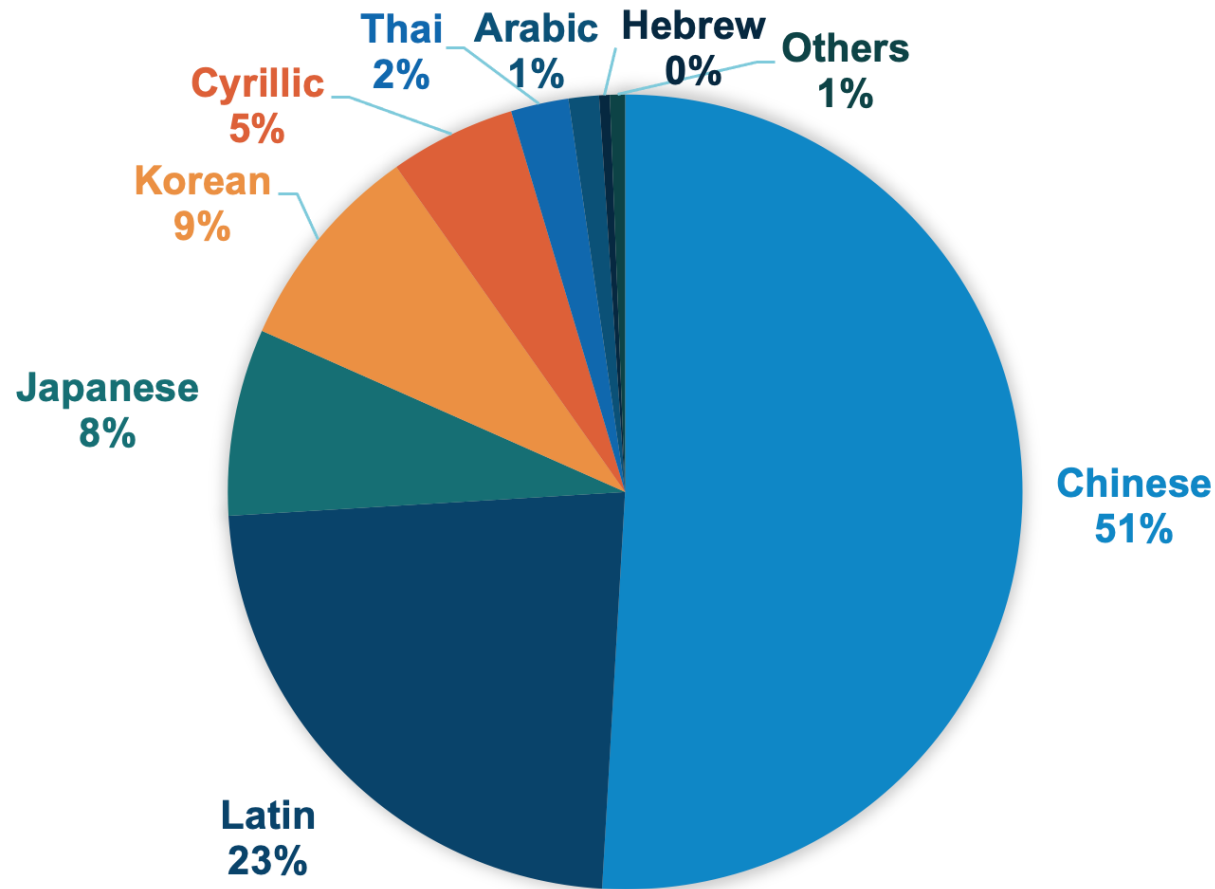
arabic
arabicegyptian

IDN Country Code Top-Level Domains (ccTLDs)

- Implement the string evaluation as part of the IDN ccTLD Fast Track process

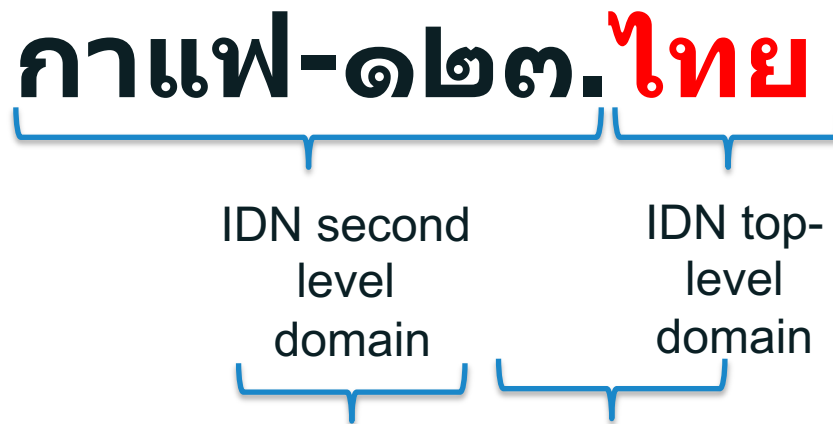


IDN Registrations by Script Under gTLDs (April 2021)



Based on 1.6 million registrations

Internationalized Domain Names (IDNs)



2

Syntax of IDN Labels

Valid U-Label: Unicode code points as constrained by **the “LDH” scheme** within IDNA 2008 – sets baseline limiting to letters, marks and digits.

Syntax of IDN Labels

Valid U-Label, further constrained by **the “letter” principle** for TLDs.

1

Domain Names in Different Scripts

Arabic

Armenian

Georgian

Korean

Malayalam

Oriya

1. համընդհանուր-ընկալում-թեստ. հայ
2. മുൻ‌കരുതൽ-വകുപ്പുകൾ-ടെസ്റ്റ്. തിരുവ
3. უნივერსალური-თავსობადობის-ტესტი. გე
4. 다국어도메인이용환경테스트. 한국
5. സാർവത്രിക-സ്വീകാര്യതാ-പരിശോധന. ഭാരതം
6. تجربة-القبول-الشامل. موريتانيا

Domain Names in Different Scripts

Armenian

Georgian

Korean

Malayalam

Oriya

1. համընդհանուր-ընկալում-թեստ. հայ
2. മുൻ‌കരുതൽ-വകുപ്പുകൾ-ടെസ്റ്റ്. തിരുവ
3. უნივერსალური-თავსობადობის-ტესტი. გე
4. 다국어도메인이용환경테스트. 한국
5. സാർവത്രിക-സ്വീകാര്യതാ-പരിശോധന. ഭാരതം
6. تجربة-القبول-الشامل. موريتانيا

Arabic

Domain Names in Different Scripts

- | | Korean | Malayalam | Georgian | Oriya | Armenian | Arabic |
|----|--------------------------------------|-----------|----------|-------|----------|--------|
| 1. | համընդհանուր-ընկալում-թեստ. հայ | | | | | |
| 2. | മുൻ‌കരുതൽ-വകുപ്പ്-ടെസ്റ്റ്. മലയാളം | | | | | |
| 3. | უნივერსალური-თავსობადობის-ტესტი. გე | | | | | |
| 4. | 다국어도메인이용환경테스트. 한국 | | | | | |
| 5. | സാർവത്രിക-സ്വീകാര്യതാ-പരിശോധന. ഭാരതം | | | | | |
| 6. | تجربة-القبول-الشامل. موريتانيا | | | | | |

Domain Names in Different Scripts

Korean	Malayalam	Oriya
1. համընդհանուր-ընկալում-թեստ.hայ		Armenian
2. മുൻ‌കരുമാൽ-എഴുത്തുക-രേഖ.ଭାରତ		
3. უნივერსალური-თავსობადობის-ტესტი.გე		Georgian
4. 다국어도메인이용환경테스트.한국		
5. സാർവത്രിക-സ്വീകാര്യതാ-പരിശോധന.ഭാരതം		
6. تجربة-القبول-الشامل.موريتانيا		Arabic

Domain Names in Different Scripts

Malayalam

Oriya

1. համընդհանուր-ընկալում-թեստ.hայ

Armenian

2. ୟୁନିଭରସାଳ-ଏକସେପ୍ଟ-ଟେଷ୍ଟ.ଭାରତ

3. უნივერსალური-თავსობადობის-ტესტი.გე

Georgian

4. 다국어도메인이용환경테스트.한국

Korean

5. സാർവത്രിക-സ്വീകാര്യതാ-പരിശോധന.ഭാരതം

6. تجربة-القبول-الشامل.موريتانيا

Arabic

Domain Names in Different Scripts

- | | |
|--|-----------|
| | Oriya |
| 1. համընդհանուր-ընկալում-թեստ.հայ | Armenian |
| 2. ୟୁନିଭରସାଳ-ଏକସେପ୍ଟ-ଟେଷ୍ଟ.ଭାରତ | |
| 3. უნივერსალური-თავსობადობის-ტესტი.გე | Georgian |
| 4. 다국어도메인이용환경테스트.한국 | Korean |
| 5. സാർവത്രിക-സ്വീകാര്യതാ-പരിശോധന.ഭാരതം | Malayalam |
| 6. تجربة-القبول-الشامل.موريتانيا | Arabic |

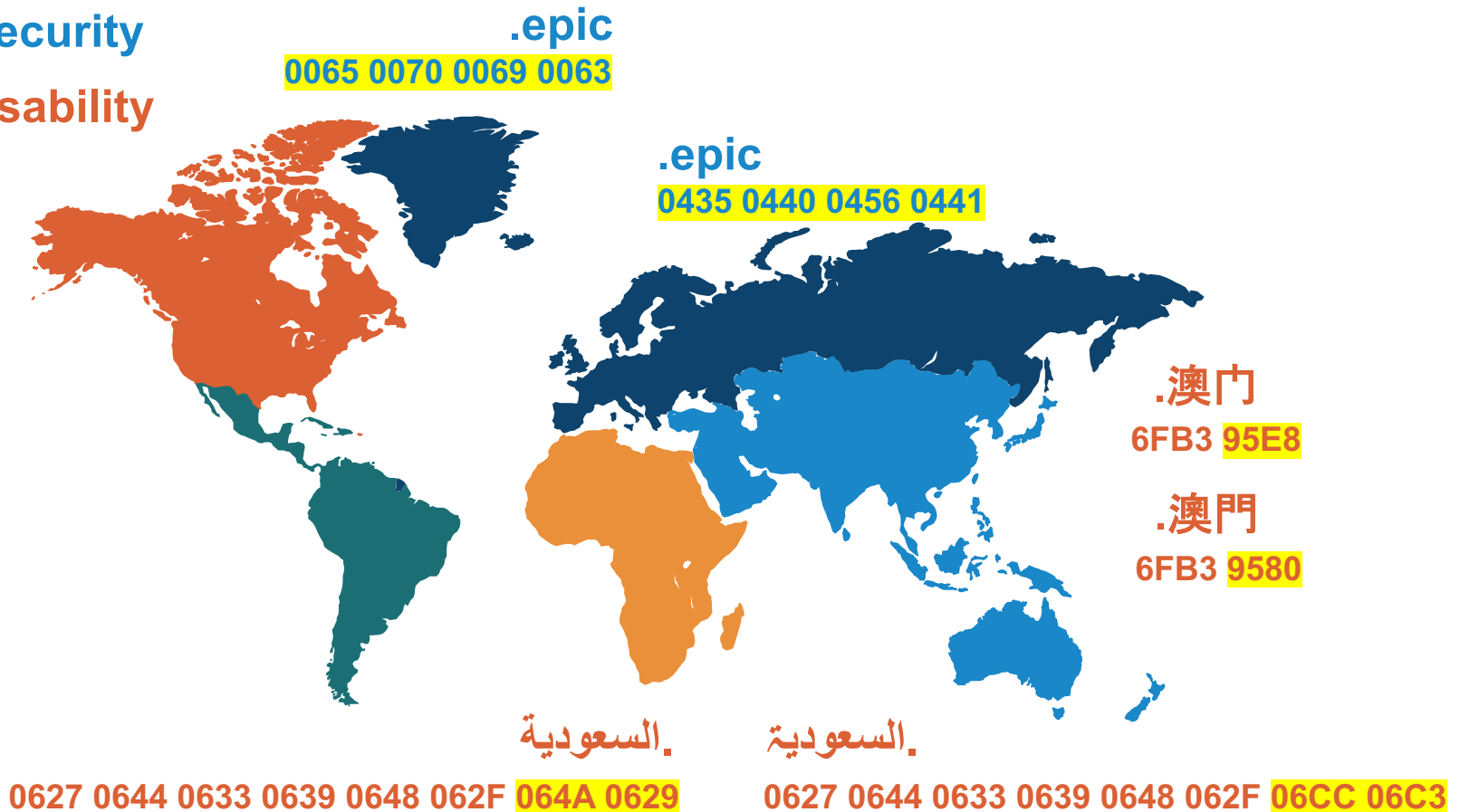
Domain Names in Different Scripts

1. համընդհանուր-ընկալում-թեստ.հայ Armenian
2. ୟନିଭରସାଳ-ଏକସେପ୍ଟ-ଟେଷ୍ଟ.ଭାରତ Oriya
3. უნივერსალური-თავსობადობის-ტესტი.გე Georgian
4. 다국어도메인이용환경테스트.한국 Korean
5. സാർവത്രിക-സ്വീകാര്യതാ-പരിശോധന.ഭാരതം Malayalam
6. تجربة-القبول-الشامل.موريتانيا Arabic

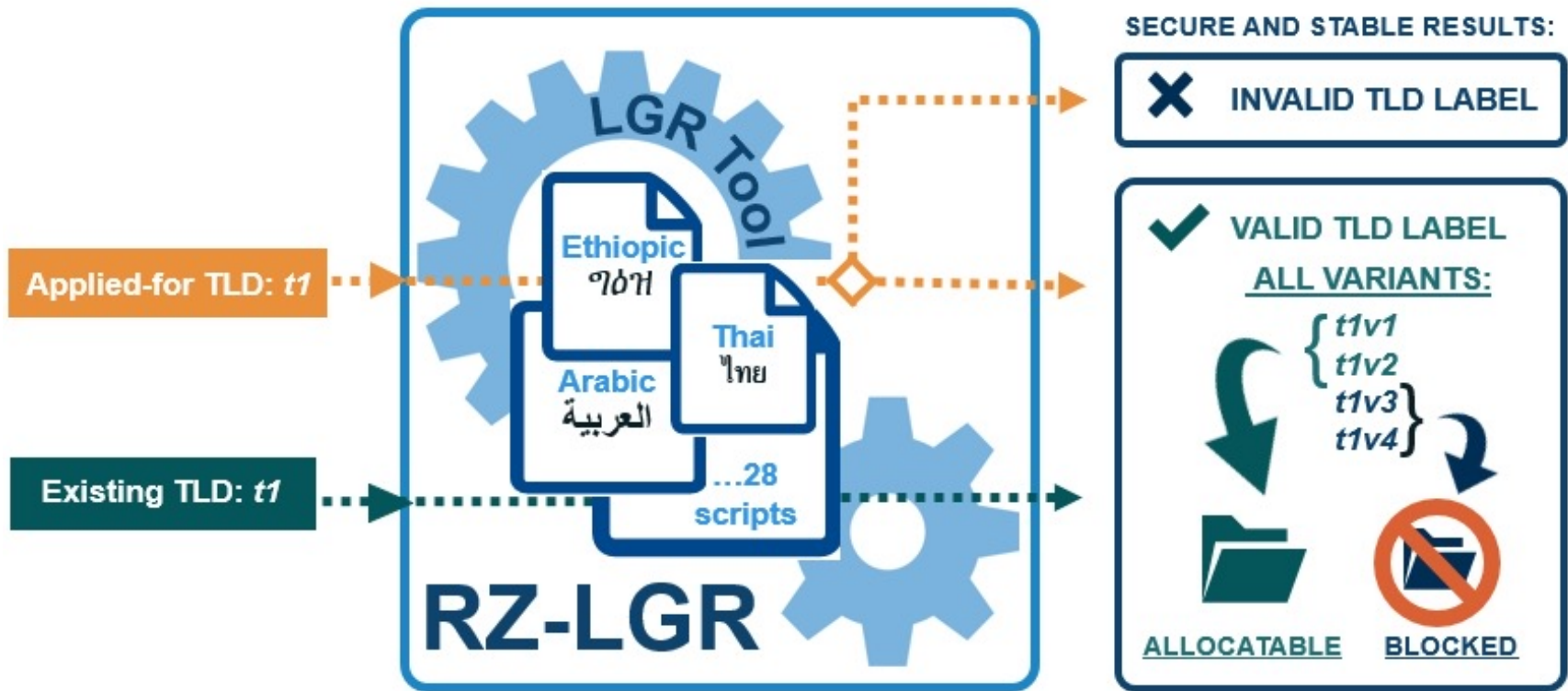
What are Variant Labels?

Variant labels are labels considered the “same” in some measure by a given community of Internet users.

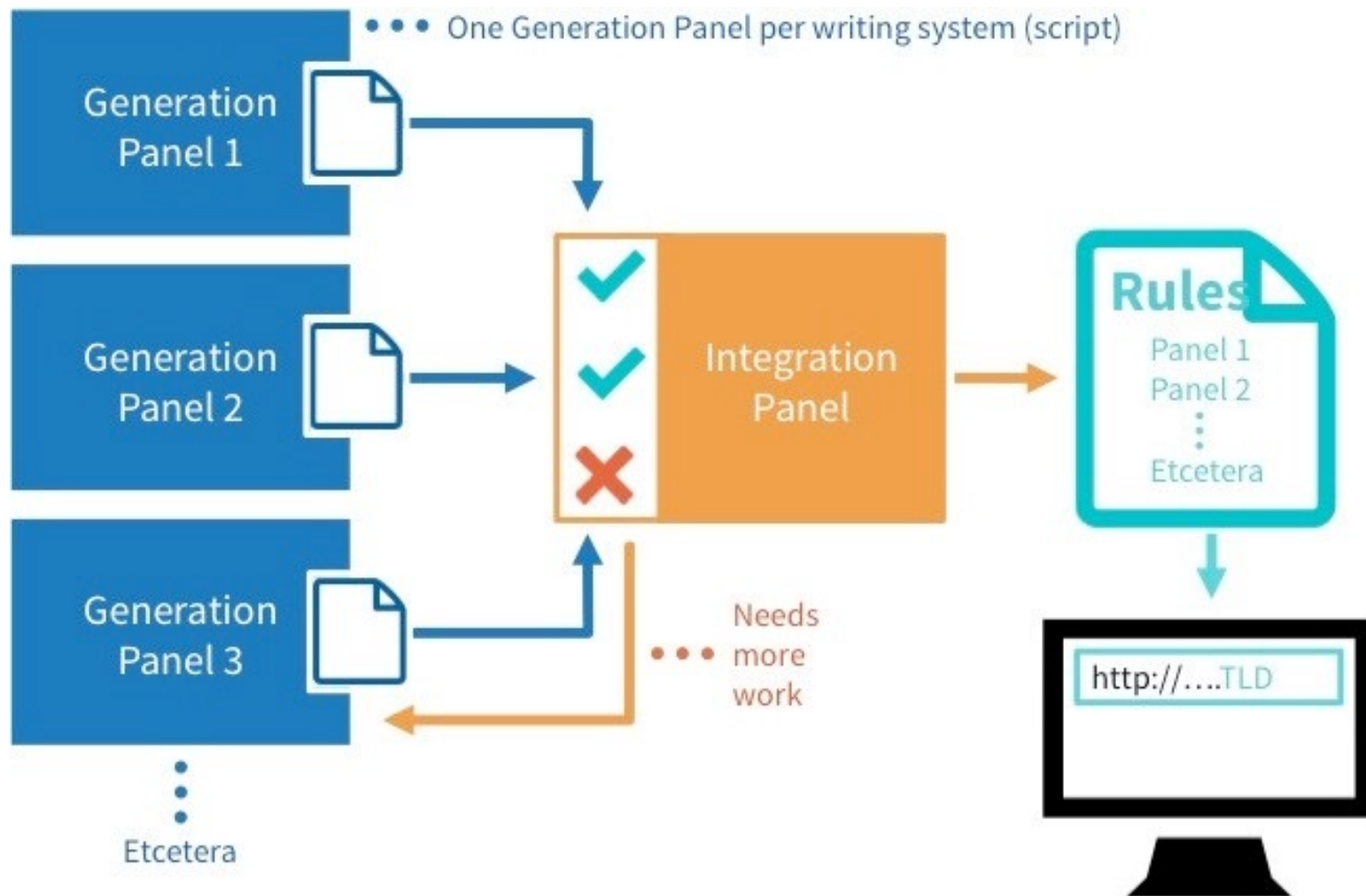
- Security
- Usability



How Does Root Zone Label Generation Rules Work?



RZ-LGR Proposal Development Process



Community Contribution to IDNs

44

countries
represented
by GP members

10,000+

volunteer hours

270+

script community
volunteers

9

years of work

386+

languages supported by
Generation Panel work

26

unique scripts covered by
LGR proposals

30+

Public Comment
proceedings

17

Generation
Panels



RZ-LGR-5

- ⦿ In 2019, the ICANN Board [resolved](#) that the GNSO and ccNSO take into account the Recommendations for Managing the IDN Variant TLDs - which integrated the use of RZ-LGR - in their policy development processes.
- ⦿ In 2020, the ICANN Board [resolved](#) that the GNSO and ccNSO take into account the Recommendations for the Technical Utilization of the Root Zone Label Generation in their policy development processes.
- ⦿ In 2021, the GNSO published its [Report on New gTLD Subsequent Procedures](#) which incorporates the use of RZ-LGR for the next round.
- ⦿ RZ-LGR-5 [published](#) on 26 May 2022 and integrated 26 scripts: **Arabic**, Armenian, Bangla, Chinese (Han), Cyrillic, Devanagari, **Ethiopic**, Georgian, Greek, Gujarati, Gurmukhi, Hebrew, Japanese (Hiragana, Katakana, and Kanji [Han]), Kannada, Khmer, Korean (Hangul and Hanja [Han]), Lao, **Latin**, Malayalam, Myanmar, Oriya, Sinhala, Tamil, Telugu, and Thai.

What's Next? Reference LGRs for the Second Level

- How are they developed?

Reference LGRs have been developed using [guidelines](#), followed by a consultation with the script community. They are then finalized after Public Comment.

- Currently, 46 reference LGRs are [published](#):

30 language-based LGRs: Arabic, Belarusian, Bosnian (Cyrillic), Bosnian (Latin), Bulgarian, Chinese, Danish, English, Finnish, French, German, Hebrew, Hindi, Hungarian, Icelandic, Italian, Korean, Latvian, Lithuanian, Macedonian, Montenegrin, Norwegian, Polish, Portuguese, Russian, Serbian, Spanish, Swedish, Thai, and Ukrainian.

16 script-based LGRs: Arabic, Bangla (Bengali), Devanagari, Ethiopic, Georgian, Gujarati, Gurmukhi, Hebrew, Kannada, Khmer, Lao, Malayalam, Oriya, Sinhala, Tamil, and Telugu.

- What's coming next?

Additional reference LGRs for the following seven scripts: Armenian, Cyrillic, Greek, Japanese, Korean, Latin, and Myanmar.

Thank you

Engage with ICANN – Thank You and Questions



One World, One Internet

Visit us at icann.org



[@icann](https://twitter.com/icann)



linkedin/company/icann



facebook.com/icannorg



slideshare/icannpresentations



youtube.com/icannnews



soundcloud/icann



flickr.com/icann



instagram.com/icannorg