

I C A N N

VIRTUAL ANNUAL GENERAL

69

PREP WEEK

IDN Program Update



ICANN69 Prep Week
6 October 2020

Overview of Session Presentations

- ⦿ Internationalized Domain Name (IDN) Program Overview and Progress (Pitinan Kooarmornpatana)

- ⦿ Community Updates
 - Japanese Generation Panel (Hiro Hotta)
 - Korean Generation Panel (Kim Kyongsok)
 - Latin Generation Panel (Mirjana Tasić)
 - Myanmar Generation Panel (Yin May Oo)

- ⦿ Question and Answer

IDN Program Overview and Progress

Pitinan Kooarmornpatana
Manager, IDN Programs

ICANN69 Prep Week

IDN Program Objectives

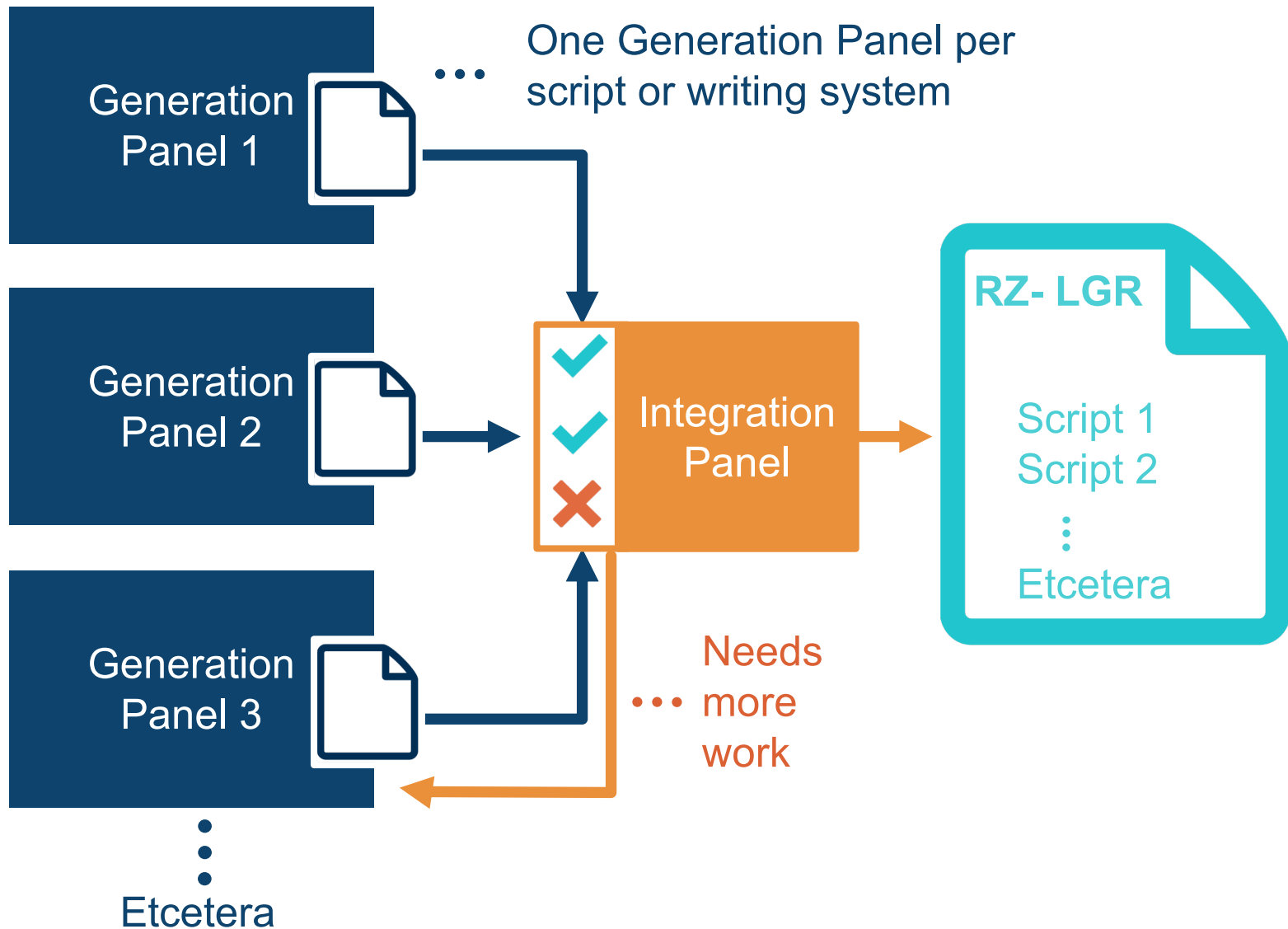
Enable deployment of domain names in the ***local languages and scripts*** of global communities in a ***secure and stable*** manner.

Overview of IDN Programs

- ◉ IDNs at the Top Level
 - Root Zone Label Generation Rules (RZ-LGR)
 - IDN Variant TLD Implementation
 - LGR Toolset
 - IDN ccTLD Fast Track Process
- ◉ IDNs at Second Level for gTLDs
 - IDN Implementation Guidelines
 - Reference IDN Tables
 - IDN Table Review Toolset



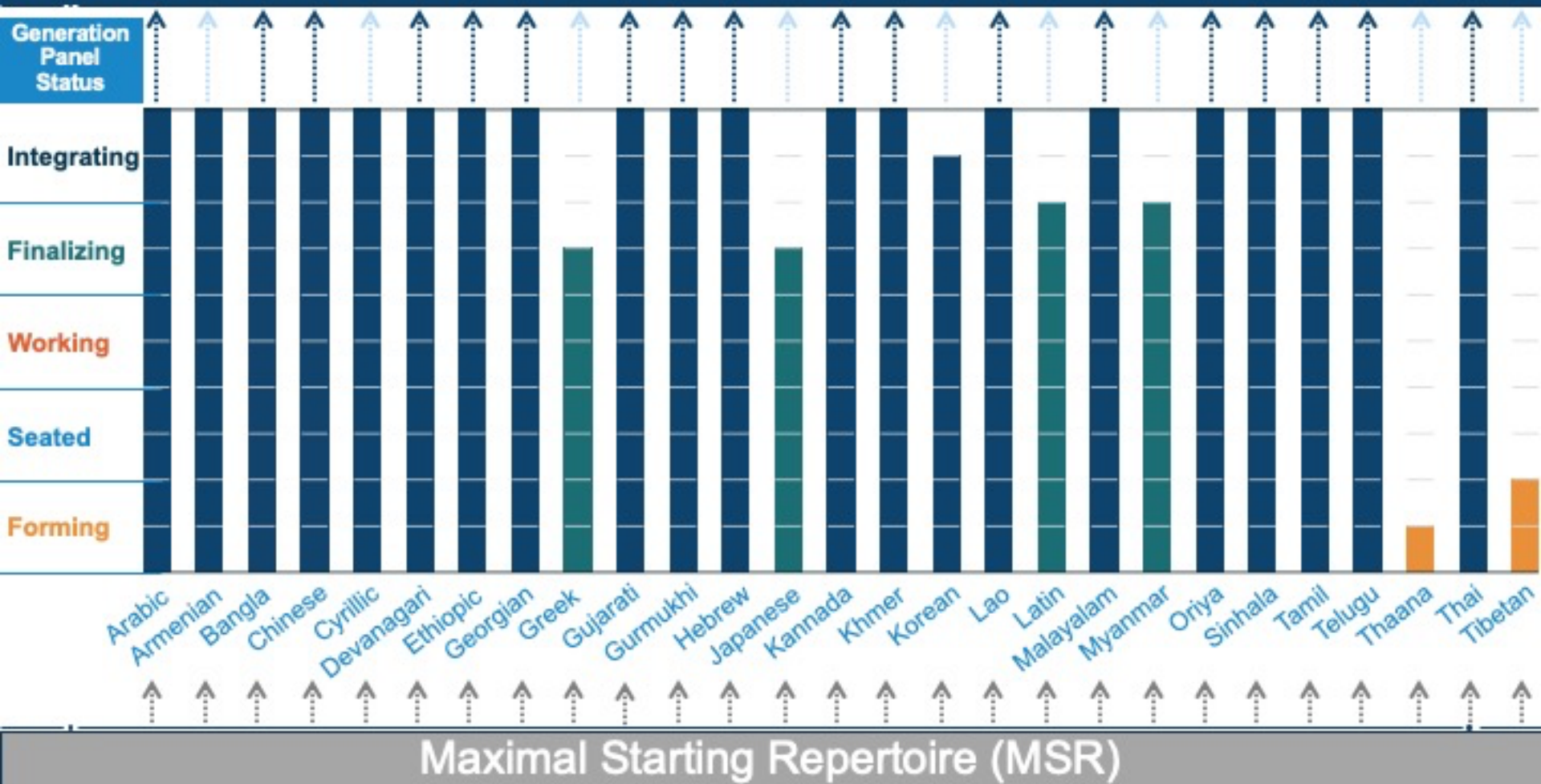
RZ-LGR Process



Status of Generation Panels (GPs)

Sep. 2020

Root Zone Label Generation Rules (RZ-LGR)

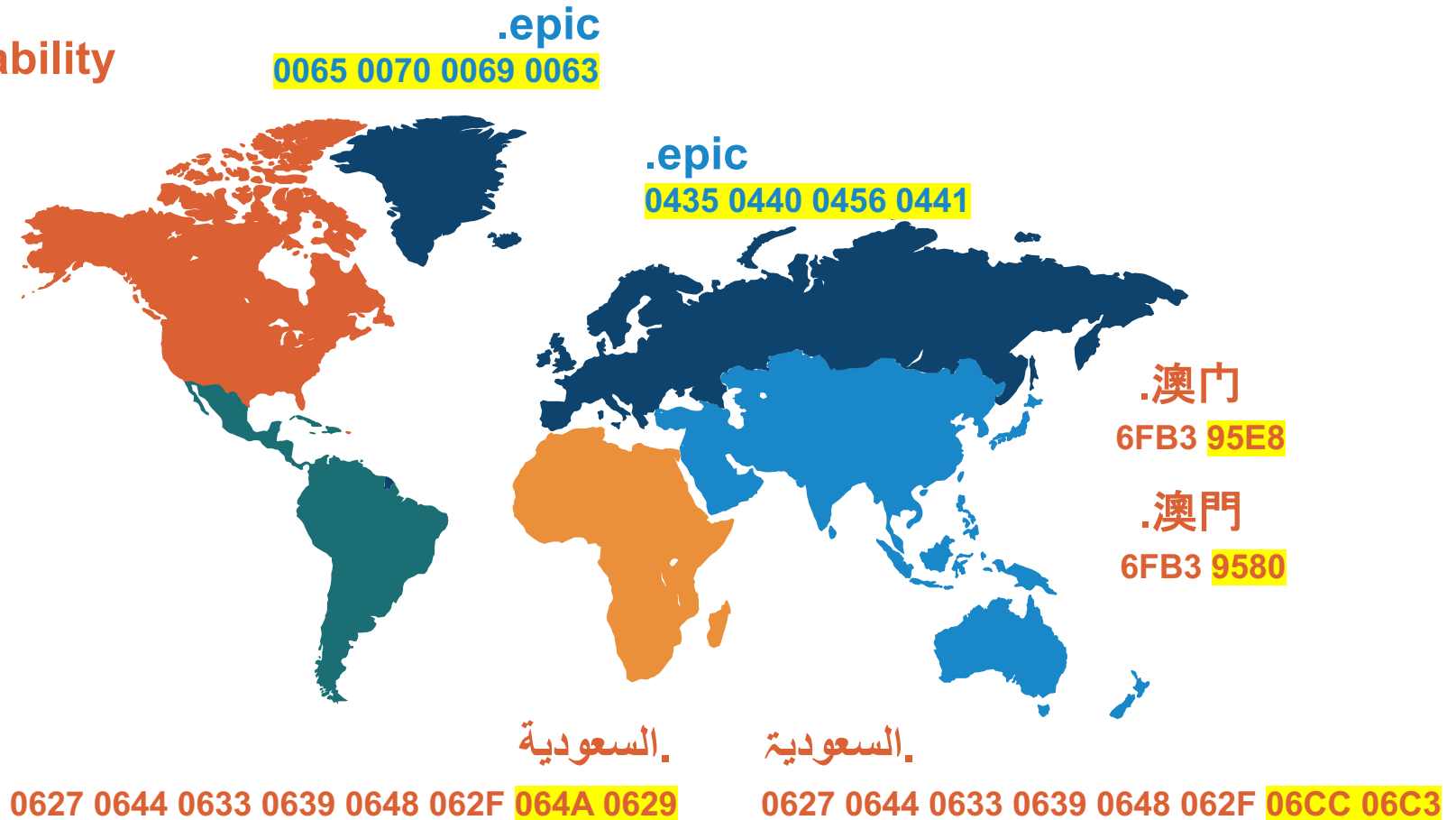


Maximal Starting Repertoire (MSR)

Understanding IDN Variant TLDs

⦿ Security

⦿ Usability



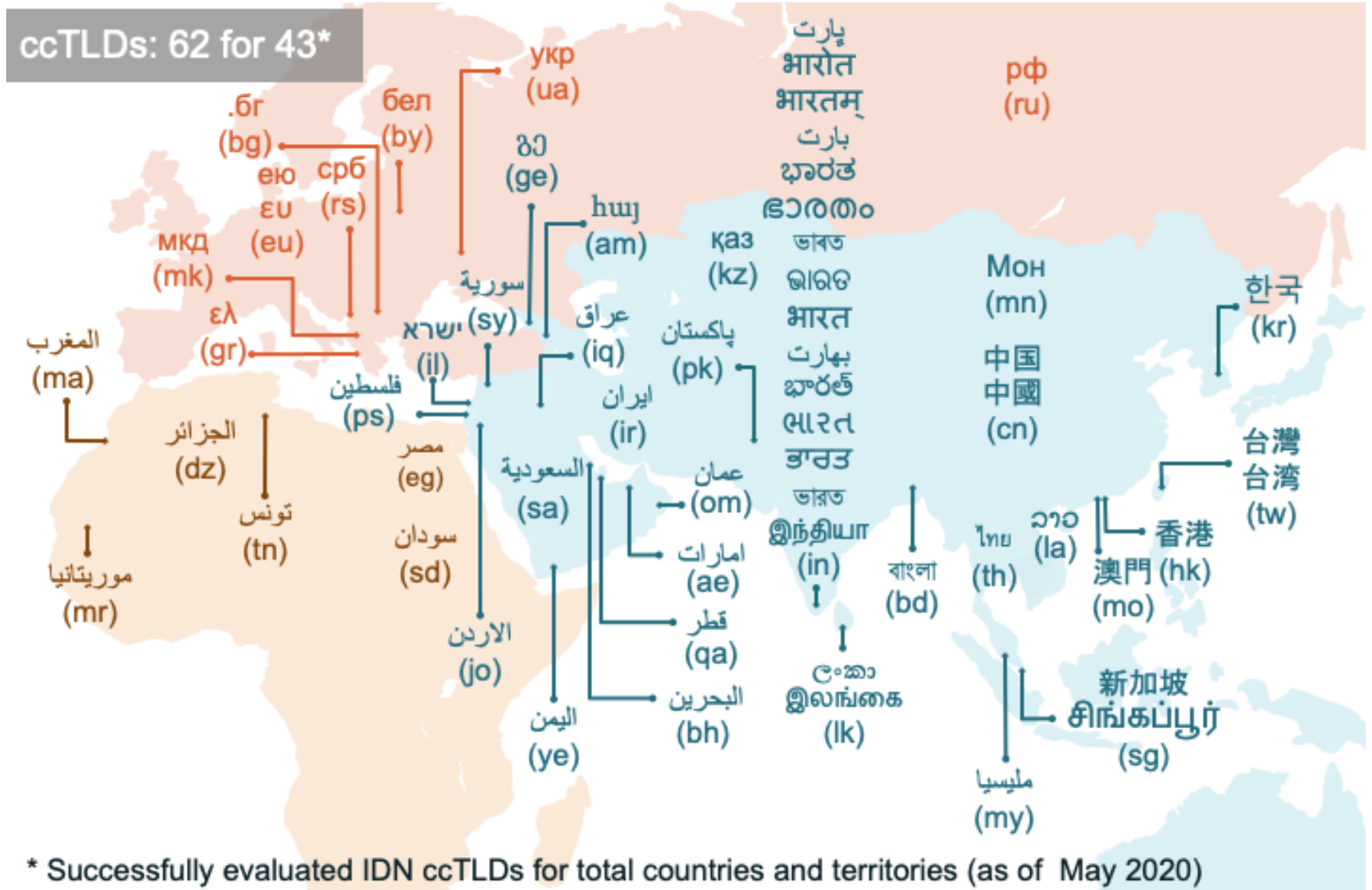
Status of IDN Variant TLDs – Management Mechanism

- ⊙ Solution: ICANN org to work with the community to develop a feasible management mechanism.
 - Recommendations developed by ICANN org and [finalized](#) on 25 January after the [Public Comment](#).
 - ICANN Board [approved](#) 8 March 2019:
“...the Board approves the Variant TLD Recommendations and requests that the ccNSO and GNSO take into account the Variant TLD Recommendations while developing their respective policies to define and manage the IDN variant TLDs for the current TLDs as well as for future TLD applications.”
- ⊙ Policies are being developed by the ccNSO and GNSO.
 - ccNSO: [PDP \(de-\)selection of IDN ccTLD Strings Working Group](#)
 - GNSO: [Draft Final Report on the New gTLD Sub Pro Working Group](#)

LGR Toolset

- ⦿ Label Generation Rulesets (LGRs) used to generate domain name labels as specified in [RFC 7940](#).
- ⦿ LGR Toolset currently allows you to:
 - **Create** single LGR or merge multiple LGRs.
 - **View** LGR in XML form or user-friendly HTML form.
 - **Use** an LGR to validate a label and determine its variant labels.
 - **Manage** LGRs by comparing or combining them.
 - **Review** impact of a new or revised LGR on existing labels.
- ⦿ Online deployment at: <https://lgrtool.icann.org/>
- ⦿ Open source package(s) released with BSD license at GitHub: [picu](#), [lgr-core](#), [lgr-django](#), [munidata](#).
- ⦿ [User guide](#) available for further details.

IDN Country Code Top-Level (ccTLD) Domains



IDN Implementation Guidelines

- ⦿ IDN Guidelines developed to minimize the risk of cybersquatting and consumer confusion.
- ⦿ IDN Guidelines apply to IDN registration policies at the second-level and practices under top-level domains (TLDs).
- ⦿ [Guidelines for the Implementation of IDNs Version 4.0](#) finalized by IDN Guidelines Working Group and [published](#) on 10 May 2018.
- ⦿ On 30 April 2019, the GNSO made a request to the ICANN Board to allow it to study the guidelines before the implementation.
- ⦿ The GNSO is organizing the Operational Track to review the Guidelines.

Reference Second Level IDN Tables

- ⦿ Reference IDN tables in LGR format in [RFC 7940](#).
- ⦿ Registry operators are encouraged to consult the reference IDN tables while they design their own IDN tables.
- ⦿ Reference IDN tables are used as baselines for reviewing IDN tables to promote security and stability of IDNs.
- ⦿ Currently, 26 reference LGRs are published (2016):
 - Belarusian, Bosnian (Cyrillic), Bosnian (Latin), Bulgarian, Chinese, Danish, English, Finnish, French, German, Hungarian, Icelandic, Italian, Korean, Latvian, Lithuanian, Macedonian, Montenegrin, Norwegian, Polish, Portuguese, Russian, Serbian, Spanish, Swedish, Ukrainian
- ⦿ Additional reference LGRs are being published for [Public Comment](#) (closing 15 October 2020):
 - Script-based: Ethiopic, Georgian, Bangla, Devanagari, Gujarati, Gurmukhi, Kannada, Malayalam, Oriya, Tamil, Telugu, Khmer, Lao
 - Language based: Arabic, Chinese, Hindi, Thai.

IDN Table Review Tool

- ⦿ The aim for developing the IDN Table Review Tool is to:
 - Increase efficiency in reviewing the IDN tables.
 - Promote consistency of the reviews.
 - Promote transparency of the reviews.
- ⦿ The tool will be available online for ICANN org and registry operators.
- ⦿ It will take the IDN table as input in three formats:
 - RFC3743, RFC4290, RFC7940 (LGR)
- ⦿ It will provide the IDN table review report in HTML format as output.
- ⦿ The tool is tentatively planned to be released in 2021.

Community Updates

- Japanese Generation Panel
- Korean Generation Panel
- Latin Generation Panel
- Myanmar Generation Panel

Japanese Generation Panel Update

Hiro Hotta
Chair, Korean GP

ICANN69 Prep Week

Agenda

1

Overview of the
Script and Language

2

Generation Panel
Membership

3

Progress Summary

4

Current Work

5

Plan and Next Steps

Overview of the Script and Language

◎ Script and Language

- 3 scripts : Kanji, Hiragana, Katakana
 - Characters can be mingled in any order in a word.
 - Characters defined in JIS (Japanese Industrial Standard) level-1 and level-2 are mostly used in daily life.
 - 6,000+ characters.
- Kanji is used also in the Chinese and Korean languages.

◎ Variants

- Usually, all Japanese characters are regarded as different.
- However, some sets of Kanji characters are regarded as variants of each other in the Chinese language and/or Korean language.

Generation Panel Membership

- ⊙ Hiro Hotta (chair)
 - Policy/business aspects of registry/registrar.
- ⊙ Akinori Maemura (vice chair)
 - Internet governance and domain name in general.
- ⊙ Shigeki Goto
 - Internet in general.
- ⊙ Kazunori Konishi
 - Internet in general.
- ⊙ Tsugizo Kubo
 - Trademarks and domain names.
- ⊙ Yoshitaka Murakami
 - Trademarks and gTLD markets from registry/registrar perspective.
- ⊙ Shuichi Tashiro
 - Character codes.
- ⊙ Yoshiro Yoneya
 - Technical aspects of IDN, LGR.

Progress Summary

- ⊙ Basic framework of Japanese LGR (J-LGR):
 - Repertoire : 6000+ characters in JIS level-1 and level-2
 - Variants : - J-LGR defines no variants that are intrinsic to it.
 - J-LGR accommodates Kanji variants defined in Chinese LGR (C-LGR) and Korean LGR (K-LGR).
 - WLE : - J-LGR has no WLE that is intrinsic to it.
- ⊙ Additional issues have been raised and consideration is required.
 - Reduction of allocatable variants labels.
 - Accommodation of variants from C-LGR and K-LGR causes the generation of large numbers of variant labels for one label.
 - Definition of variants caused by visual identicalness.
 - Identical-looking characters exist between different scripts.



Currently contemplating the employment of field research of human perception.

Additional Work Done – Reduction of Allocatable Labels

- Any combination of characters is allowed in Japanese labels as in the case of Japanese words in daily life.
- The above may result in a large number of variant strings, considering that many variants are imported from C-LGR and K-LGR.
 - e.g., 慶応大学 has 3 variant strings – 慶應大学/慶応大學/慶應大學
- The reduction of the number of allocatable variant labels was required to prevent the explosion of root zone size.
- With IP's suggestion, Japanese GP solved it by limiting allocatable strings by employing the notion that allocatable variant labels must consist of daily-use (Joyo) Kanji.
 - It reduces the maximum number of allocatable labels of an actually registered Japanese label under .JP from 486 to 8.

Current Work - Visual Identicalness

- ⊙ Visual identicalness in Japanese scripts.
 - UNICODE Consortium lists confusable characters between different scripts in <http://www.unicode.org/Public/security/latest/confusables.txt>

- Hiragana Katakana Kanji

へ	へ	
	ニ	ニ
	ハ	ハ
	カ	カ
	ト	ト
	ロ	ロ
	タ	タ
	エ	エ

8 pairs of Japanese characters are pointed as confusable

- Popular words containing above characters

ヘリコプター	コミュニケーション
シャンハイ	ホッカイドウ
インターネット	プロジェクト
コンピューター	ダイエット

- ◎ Field research (under experimentation)
 - a. 8 pairs of single confusable characters + 8 pairs of confusable words
 - b. 9 popular fonts
 - c. 3 font sizes – 18px, 24px, 36px
 - d. 40 examined – among them, 20 read Japanese well, while 20 don't
 - e. Every experiment (=every combination of a. b. c. d.) gives rate 1-5
 - 1 (very similar), 2 (similar), 3 (neutral), 4 (distinct), 5 (very distinct)

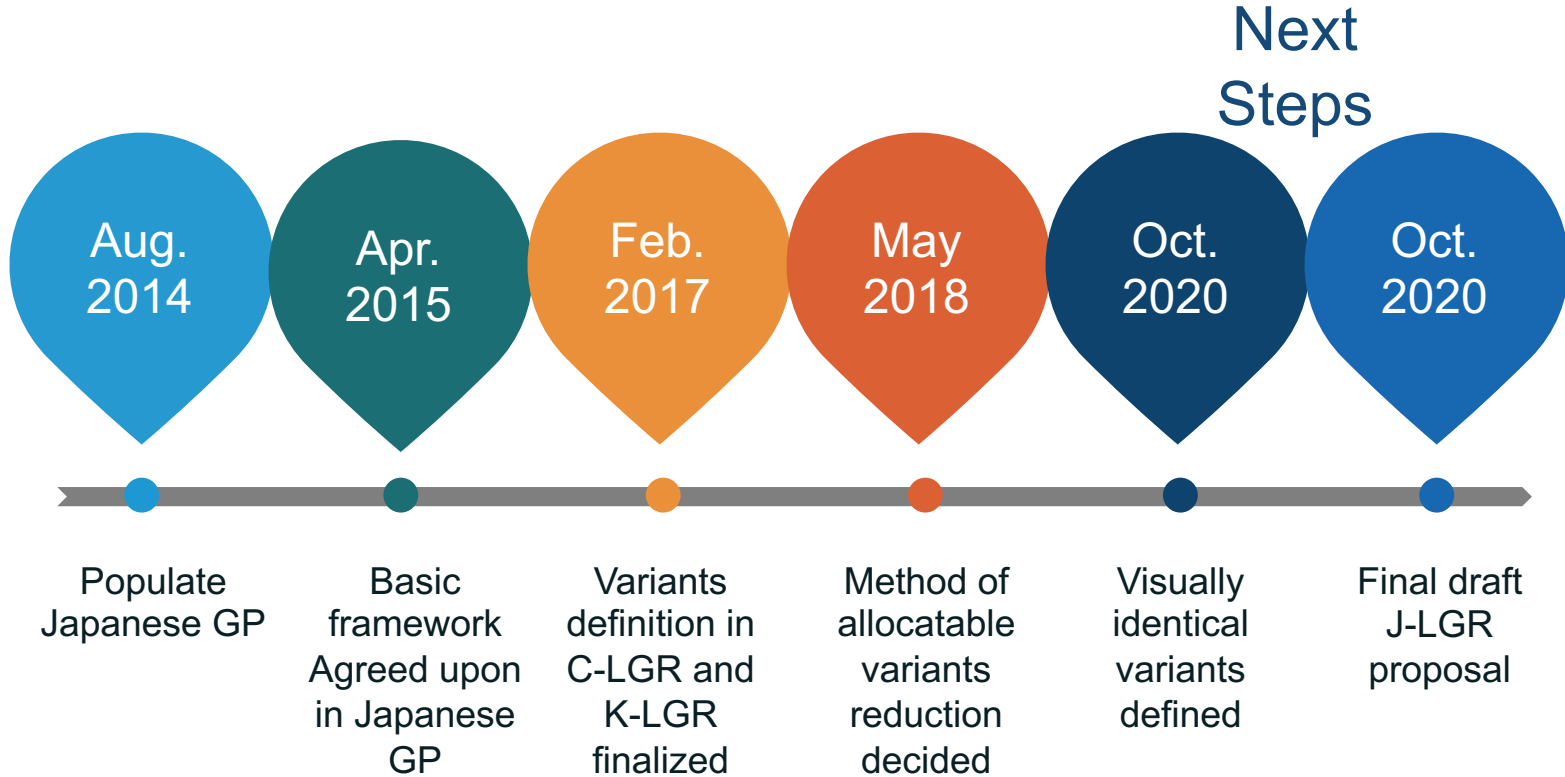
- ◎ Intermediate result

	Those who read Japanese well	Don't read well
Single character	1.4(へへ) – 3.2(ハハ)	1.2 – 3.1
Word	1.1(ヘリコプター) – 3.0(シャンハイ) ヘリコプター シャンハイ	1.1 – 2.3

- ◎ Next step

- Japanese GP will define variant characters based on visual identicalness based on the final result of the field research.

Plan and Next Steps



To Summarize

As there's no to-do item left, Japanese Root Zone LGR is in its final stage (hopefully).

Korean Generation Panel Update

Kim Kyongsok
Chair, Korean GP

ICANN69 Prep Week

Agenda

1

Overview of the
Script and Language

2

Generation Panel
Membership

3

Progress Summary

4

Current Work

5

Plan and Next Steps

Overview of the Script and Language

- K-LGR covers Korean script (= Hangeul + Hanja).
- “Korean script” usually means “Hangeul” or “Hangeul”. However, in the context of the Korean LGR (K-LGR), Korean script is a union of Hangeul (한글) and Hanja (한자).
- Korean language has a long history, more than 2000 years.
- Hangeul was invented in 1443 (about 600 years ago).
- Hanja was used before Hangeul was invented. Hanja is still used in Republic of Korea.
- Korean language is mainly used in Republic of Korea (South Korea) and Democratic People’s Republic of Korea (North Korea).
 - Also used by Korean people living in China, USA, Japan, Europe, Brazil, Russia, Vietnam, and so on.

Korean Generation Panel (KGP) Membership

- ⊙ Technical Experts: Kyongsok KIM (Chair), Dongman LEE
- ⊙ Linguists: Jeongdo CHOI (Hangul), Sanghyun SHIN (Hanja),
Sungduk CHO (Hanja)
- ⊙ Policymakers: Youngeum LEE, Youn Jung PARK
- ⊙ Community: Eungjun JEON, Boknam YUN, Byeongil OH,
Kyuhong BYUN
- ⊙ Registry: Changmin PARK, Minjung PARK, Boyoung KIM
- ⊙ Registrar: Geonwoo KIM, ChangKi JANG, Myungsoo LEE

Progress Summary (1): K-LGR v2.1 (2020.09.01.)

- K-LGR v2.1 (2020.09.01.): repertoire and variant groups
 - Hangul: repertoire – 11172 syllables, no variant groups
 - Hanja: repertoire – 4758 characters, 37 variant groups, 283 variant groups expanded
 - Variant groups composed of Hangul syllables and Hanja chars: 7 (3 Hanja chars: out-of-repertoire variant)
- 4758 Hanja chars in K-LGR v2.1 (the same as v1.0)

Source of Hanja Character Set	# chars
1) KS X 1001 (268 comptb. chars excluded)	4620
2) IICORE - K column marked	4744
K-LGR v1.0 (2017.12.10.): Hanja List (Union of 1) and 2))	4758

Progress Summary (2): Public Comment

(January – March 2018)

- A summary of public comments made:
 - Including Hanja in K-LGR repertoire: positive.
 - Allowing Hangul-Hanja mixed label: several negative comments, some positive comments.
 - Hangul-Hanja variant group: CJK agreement needed.
 - Specific details need be corrected or modified.
- Examples of issues raised by Mr. Kyuhong BYUN:
 - References; quotes; etc.
 - Many Hanja chars allowed for personal names not included in K-LGR.
 - Hangul Jamo not included in K-LGR (actually not in MSR-3).
 - More Hangul-Hanja variant groups need be included.

Progress Summary (3): Public Comment

Reviewed (April 2018 – August 2019)

- Requests by Mr. BYUN for specific details.
 - Reviewed and discussed.
 - Mostly accepted in principle and reflected in K-LGR v2.1.
- Hangeul-only labels, Hanja-only labels, Hangeul-Hanja mixed labels.
 - KGP reconfirmed that there was a general consensus to allow Hangeul-only labels and Hanja-only labels;
 - However, KGP decided not to allow Hangeul-Hanja mixed labels.

Progress Summary (4): K-LGR v2.0 and v2.1

- K-LGR v2.0 (2020.05.01.) sent to IP
 - Public comments RE: K-LGR v1.0 mostly reflected in K-LGR v2.0.
- K-LGR v2.1 (2020.09.01.) sent to IP
 - IP feedback RE: K-LGR v2.0 mostly reflected in K-LGR v2.1.

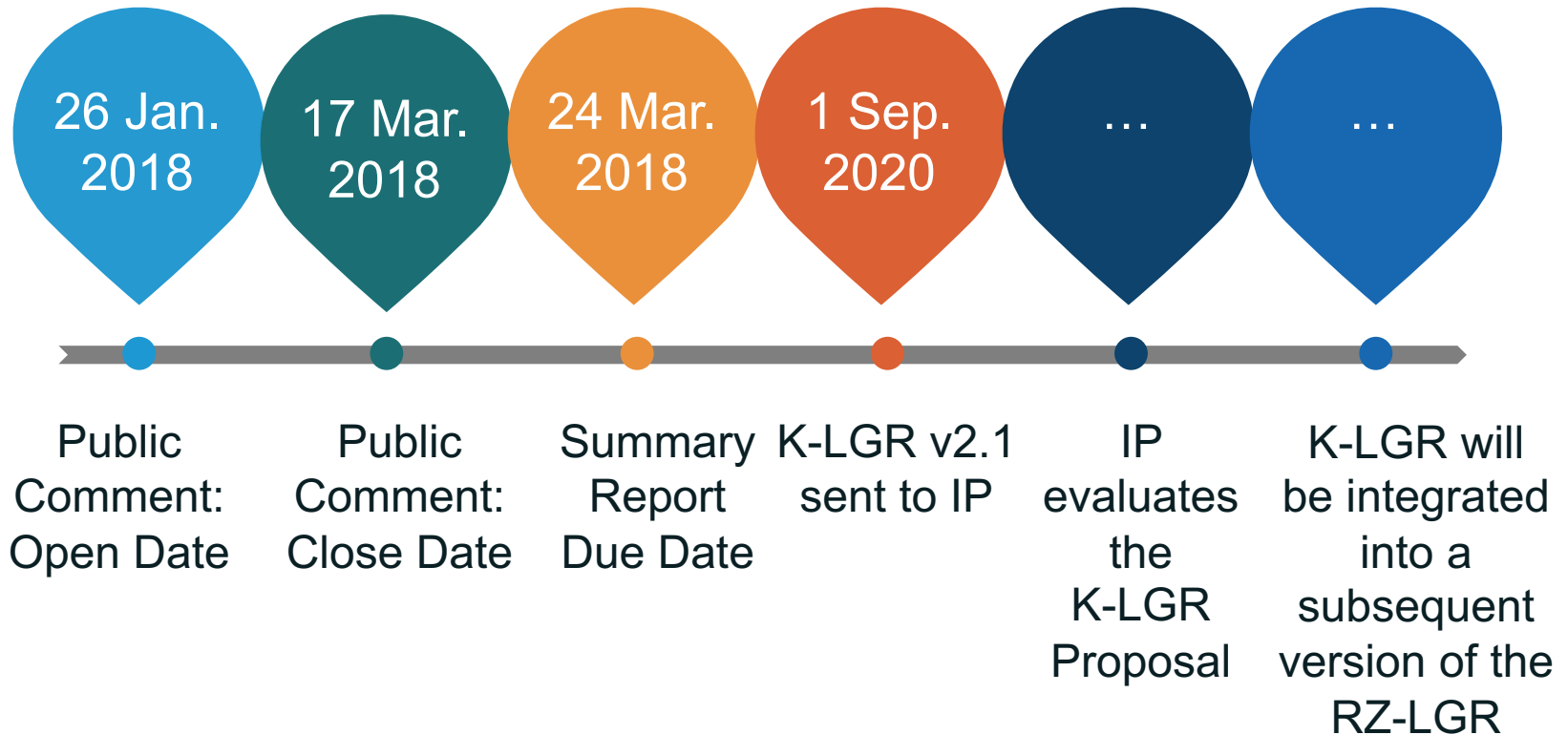
Current Work

- Waiting for the IP feedback RE: K-LGR v2.1

Brief History of KGP Activities

- Dec. 2013: Korean GP (KGP) organized.
- May. 2015: K-LGR v0.1
- Feb. 2016: The Korean community “formally” forms Generation Panel for Developing the Root Zone Label Generation Rules (LGR).
- Dec. 2017: K-LGR v1.0
- January – March 2018: Public Comment for K-LGR v1.0
- March 2018 – August 2019: Public Comment for K-LGR v1.0 reviewed for possible reflection in the next version of K-LGR.
- 39 KGP meetings
- Several CJK coordination meetings during ICANN Public Meetings (49-64).
- Several CJK coordination meetings in Rep. of Korea, China, and Taiwan.

Plan and Next Steps



Latin Generation Panel Update

Mirjana Tasić
Chair, Latin GP

ICANN69 Prep Week

Agenda

1

Overview of the
Script and Language

2

Generation Panel
Membership

3

Progress Summary

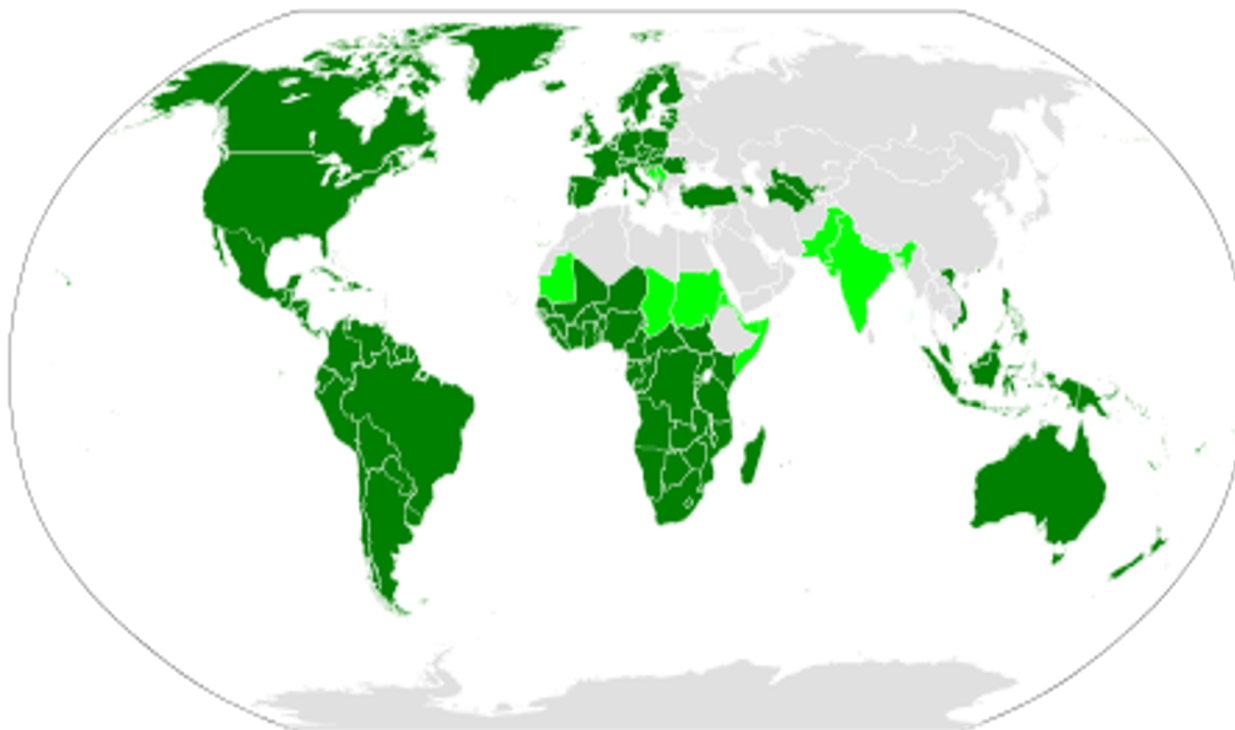
4

Current Work

5

Plan and Next Steps

Overview of the Script and Language



Dark green = Latin script is the sole main script.

Light green = Latin co-exists with other scripts.

Grey = Latin-script alphabets are sometimes used extensively due to the use of unofficial second languages, such as French in Algeria and English in Egypt, and to Latin transliteration of the official script, such as in China or in Japan.

Generation Panel Membership

Since 2016, the generation panel has 7 members who are engaged all the time:

- ⊙ Bill Jouris
- ⊙ Dennis Tan Tanaka
- ⊙ Hazem Hezzah
- ⊙ Mats Dufberg
- ⊙ Meikal Mumin
- ⊙ Michael Bauland
- ⊙ Mirjana Tasić, chair of Latin GP from the year 2016

Progress Summary

- ⦿ Developing Repertoire
 - 181 of 181 EGIDS 1- 4 languages processed.
 - 29 EGIDS 5 languages processed with more than one million speakers.
 - 193 of 279 MSR-2 code points attested.
 - 3 non-MSR-2 code points are included in MSR-3.
 - 3 non-MSR-3 code points are included in MSR-4.
 - 22 Code Point Sequences identified.

- ⦿ Developing Variants
 - In-script variants finalized.
 - Cross-script variants with Armenian script defined.
 - Cross-script variants with Cyrillic script defined.
 - Cross-script variants with Greek script defined.
 - Special HTML Link (underlining) analysis completed.
 - IDNA2003 compatibility analysis completed.
 - Generic glyphs analysis completed.

- ⦿ Submitted the third version of proposal to the IP in May 2018.

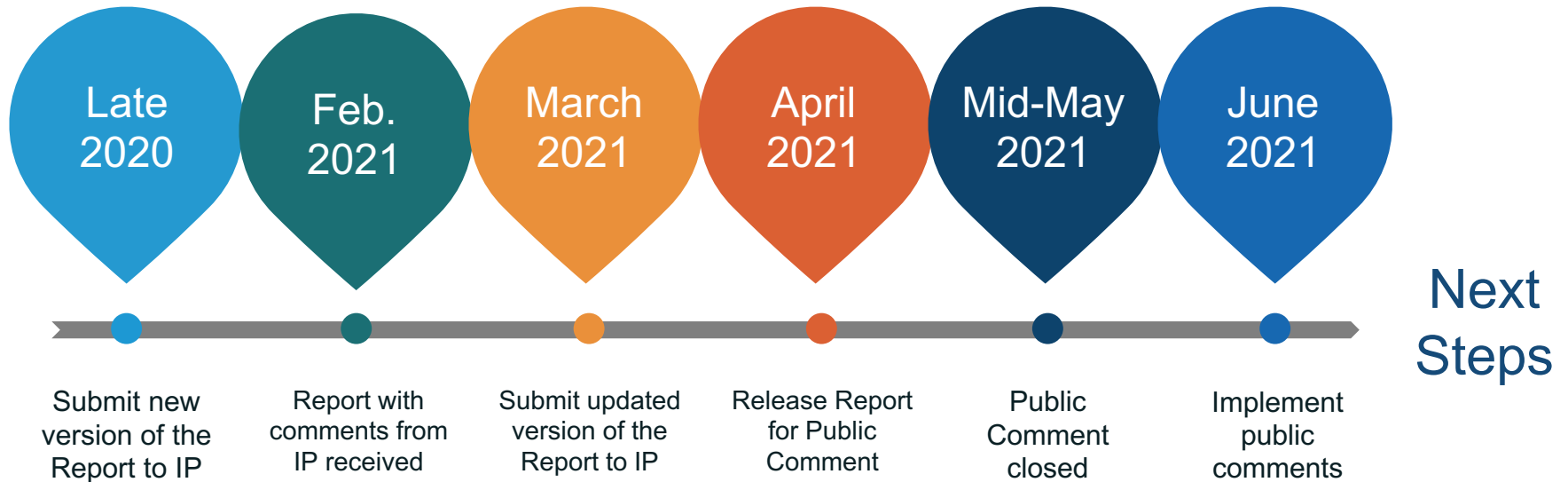
- ⦿ Submitted the fourth version of proposal to the IP in January 2019.

- ⦿ Submitted the fifth version of proposal to the IP in October 2019.

Current Work

- ⦿ Finalization of the next version of the report for submission to IP for comments.
 - Main report has 80 pages.
 - Appendices have more than 100 pages.
- ⦿ Production of test data set
 - Test data for Repertoire completed.
 - Test data for variants in progress.
- ⦿ XML production and check
- ⦿ Consider new findings of Latin glyphs which might be added to Repertoire.
- ⦿ Submission of draft version of the Report to IP for comments in late 2020.

Plan and Next Steps



To Summarize

Although it's not easy to make an exact plan to finish all the Latin GP tasks, it's possible that the Latin GP will finalize its work by the end of June 2021.

Myanmar Generation Panel Update

Yin May Oo
Myanmar GP Co-Chair

ICANN69 Prep Week

Agenda

1

Overview of the
Script and Language

2

Summary of the
Progress

3

Plan and Next Steps

Myanmar GP – Languages Using Myanmar Script

- The Myanmar script evolved from the Brahmi script which flourished in the Indian subcontinent between the 5th Century B.C. and 3rd Century A.D.
- Languages covered by the LGR:

Language	ISO 639-3 Code(s)	Countries	Local Name of the Script	EGIDS Scale	Total Users in All Countries
Burmese	[mya]	Myanmar	မြန်မာ /mja-ma/	1	42,906,490
Shan	[shn]	Myanmar, China, Thailand	လိၵ်ႈတႆး /likʰ.taj/	3	3,295,000
Rakhine	[rki]	Myanmar	ရခိုင် /rə.khi/	3	2,020,000
Karen, Sgaw	[ksw]	Myanmar, Thailand	ꨀꨄ /sɣaʔ/	3	1,560,000
Mon	[mnw]	Myanmar, Thailand	မန် /mun/	5	851,000
Pa'O Karen	[blk]	Myanmar	ပအိုဝ်း /pə.ʔəʊ/	5	560,740

In-Script Variant Analysis [1/2]

- Myanmar GP defines the following as in-script variant code points due to the nearly identical glyph or the character's property of languages.

Set#	Unicode Code Point	Glyph	Unicode Code Point	Glyph	Disposition
1	U+1023 MYANMAR LETTER I	ꠇ	U+1000 U+1039 U+1000 SV1	ꠇ	Block
2	U+101F U+103A Myanmar Letter Ha Asat	ꠏ	U+1015 U+102C U+103A (U+1015 ST1)	ꠏ	Block
3	U+1061 Myanmar Letter Sgaw Karen SHA	ꠑ	U+101B U+103E	ꠑ	Block
4	U+1001 Myanmar Letter KHA	ꠑ	U+1076 Myanmar Letter Shan KHA	ꠑ	Allocatable
5	U+1004 Myanmar Letter NGA	ꠑ	U+105A Myanmar Letter Mon NGA	ꠑ	Allocatable
6	U+1008 Myanmar Letter JHA	ꠑ	U+105B Myanmar Letter Mon JHA	ꠑ	Allocatable

Myanmar Letter NGA “ꠑ” : Mon Letter NGA “ꠑ”
 Burmese “ꠑ” : Mon “ꠑ”
 / U+1004 / U+103A / : / U+105A / U+103A /

In-Script Variant Analysis [2/2]

- Myanmar GP defines the following as in-script variant code points due to the nearly identical glyph or the character's property of languages.

Set#	Unicode Code Point	Glyph	Unicode Code Point	Glyph	Disposition
7	U+102E Myanmar Vowel Sign II	◌်	U+1033 Myanmar Vowel Sign Mon II	◌ံ	Allocatable
8	U+102B Myanmar Vowel Sign TALL AA	◌ါ	U+102C Myanmar Vowel Sign AA	◌ာ	Allocatable

Myanmar Vowel Sign II “◌်” : Mon Vowel Sign II “◌ံ”

Burmese “◌ံ” : Mon “◌ံ”

/U+1000 **U+102E** U+101D **U+102E**/ : /U+1000 **U+1033** U+101D **U+1033**/

Cross-Script Variant Analysis [1/2]

● Myanmar-Georgian

No.	Glyph	Code Point	Myanmar Character Name	Glyph	Code Point	Malayalam Character Name
1	ဂ	U+1002	MYANMAR LETTER GA	ინ	U+10D8	GEORGIAN LETTER IN
2	တ	U+1010	MYANMAR LETTER TA	თ	U+10D7	GEORGIAN LETTER TAN

➤ Georgian word “თთთთ” can be formed using Myanmar Consonants “တ” and “ဂ”

● Myanmar-Oriya

No.	Glyph	Code Point	Myanmar Character Name	Glyph	Code Point	Oriya Character Name
1	ဝ	U+101D	MYANMAR LETTER WA	ଠ	U+0B20	ORIYA LETTER TTHA
2	ေ	U+1031	MYANMAR VOWEL SIGN E	୧	U+0B47	ORIYA VOWEL SIGN E

Cross-Script Variant Analysis [2/2]

● Myanmar-Malayalam

No.	Glyph	Code Point	Myanmar Character Name	Glyph	Code Point	Malayalam Character Name
1	◌	U+101D	MYANMAR LETTER WA	◌	U+0D20	MALAYALAM LETTER TTHA

● Myanmar-Others

No.	Glyph	Code Point	Character Name
1	◌	U+101D	MYANMAR LETTER WA
2	◌	U+006F	LATIN SMALL LETTER O
3	◌	U+03BF	GREEK SMALL LETTER OMICRON
4	◌	U+043E	CYRILLIC SMALL LETTER O
5	◌	U+0585	ARMENIAN SMALL LETTER OH

Confusable Code Point Analysis

- In-script confusable code points

Set#	Unicode Code Point	Glyph	Unicode Code Point	Glyph	Note
1	U+1008	၈	U+1005 U+103B	၈	The sequence U+1005 U+103B is invalid
2	U+1009 U+102C	၉	U+1025 U+102C	၉	The sequence U+1025 U+102C is invalid
3	U+105B	၂	U+1007 U+103B U+103E	၂	The sequence U+1007 U+103B is invalid
...					
20	U+1021	၁	U+1022	၁	Semantic similarity
21	U+102B	၂	U+1083	၂	Semantic similarity
22	U+102C	၂	U+1083	၂	Semantic similarity

Whole Label Evaluation Rules [1/3]

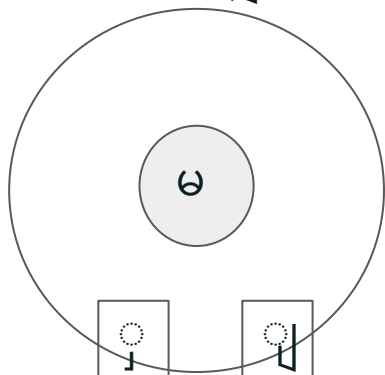
● Classifications

C	→ Consonant (Table 2)	C1	→ 103F သံ
IV	→ Independent Vowel (Section 3.3.2) 1023,1024, 1025, 1026, 1027, 1028, 1029, 102A		(GREAT SA, theoretical combination of two Myanmar Letter Sa)
DV	→ Dependent Vowel Sign 102B, 102C, 102D, 102E, 102F, 1030, 1031, 1032, 1033, 1034, 1036, 1037, 1038, 1062, 1083, 1084, 1086	C2	→ င, ဥ, ည, ထ, န, မ, ယ (1004, 1009, 100A, 100F, 1014, 1019, 101A)
ANUSVARA	→ 1036 ၵ	C3	→ င, ဥ, ည, ထ, န, မ, ဝ, ဝ, ဝ, ဝ, ဝ, ဝ 1004, 1009, 100A, 100F, 1014, 1019, 101D, 1075, 107A, 107C
T_SHORT	→ 1037 ၵ	LV	→ Long Vowel: 102B,102C,102E,1030,1031,1032, 1036
T_LONG	→ 1038 ၵ	SV	→ Short Vowel: 102D, 102F
K	→ Killer or Asat: 103A ၵ	Sh_Tone	→ 1087, 1088, 1089, 108A
VIRAMA	→ 1039 “” (It is always in between two consonants and invisible, “”)	Sh_Vowel	→ 102D, 102E, 102F, 1030, 1031, 1083, 1084, 1086
M	→ Dependent Consonant Sign (Medial) 103B, 103C, 103D, 103E, 105E, 105F, 1060, 1082	Pao_Tone	→ 108F, AA7B
		Sgaw_Tone	→ 1064
		OV	→ Other Various Sign 1035, 1062, 1085

Whole Label Evaluation Rules [2/3]

Implementing WLE Rules for all Medials

Example: မျှ [m̥j̥a]



Unicode Code Point	Glyph	Name
U+1019 U+103B U+103E	မျှ	C_MM2
U+101A U+103B U+103E	ယျှ	C_MM2
U+103B U+103D	ပျှ	MM1
U+103C U+103D	ဥျှ	MM3
U+103D U+103E	ဝျှ	MM5

	Rule: 103B must follow C_103B	Rule: 103C must follow C_103C	Rule: 103D must not follow C_n_103D	Rule: 103E must follow C_103E	Rule: MM1 must follow CMM1	Sequence MM2 (1014, 1019, 101A, 101C, 101E) + 103B	Rule: MM3 must follow CMM3	Sequence MM4 (1004, 1014, 1019) + 103C + 103E	Rule: MM5 must follow CMM5
Consonants (1000-1020)	ျှ U+103B	ဥျှ U+103C	ဝျှ U+103D	ပျှ U+103E	မျှ U+103B+U+103D	ယျှ U+103B+U+103E	ဥျှ U+103C+U+103D	ဝျှ U+103C+U+103E	ပျှ U+103D+U+103E
န U+1014	နျှ	နဥျှ	နဝျှ	နပျှ	နမျှ	နယျှ	နဥျှ	နဝျှ	နပျှ
ဗ U+1015	ဗျှ	ဗဥျှ	ဗဝျှ	ဗပျှ	ဗမျှ	ဗယျှ	ဗဥျှ	ဗဝျှ	ဗပျှ
မ U+1016	မျှ	မဥျှ	မဝျှ	မပျှ	မမျှ	မယျှ	မဥျှ	မဝျှ	မပျှ
ဗ U+1017	ဗျှ	ဗဥျှ	ဗဝျှ	ဗပျှ	ဗမျှ	ဗယျှ	ဗဥျှ	ဗဝျှ	ဗပျှ
ဝ U+1018	ဝျှ	ဝဥျှ	ဝဝျှ	ဝပျှ	ဝမျှ	ဝယျှ	ဝဥျှ	ဝဝျှ	ဝပျှ
မ U+1019	မျှ	မဥျှ	မဝျှ	မပျှ	မမျှ	မယျှ	မဥျှ	မဝျှ	မပျှ

Whole Label Evaluation Rules [3/3]

● Whole Label Evaluation Rules

3. Rules for Single Medial

- 103B must follow consonant C_103B
- 103C must follow consonant C_103C
- 103E must follow consonant C_103E
- 103D must not follow consonant_n103D
- M_mon must follow C_Mon
- 1082 must follow C_shan

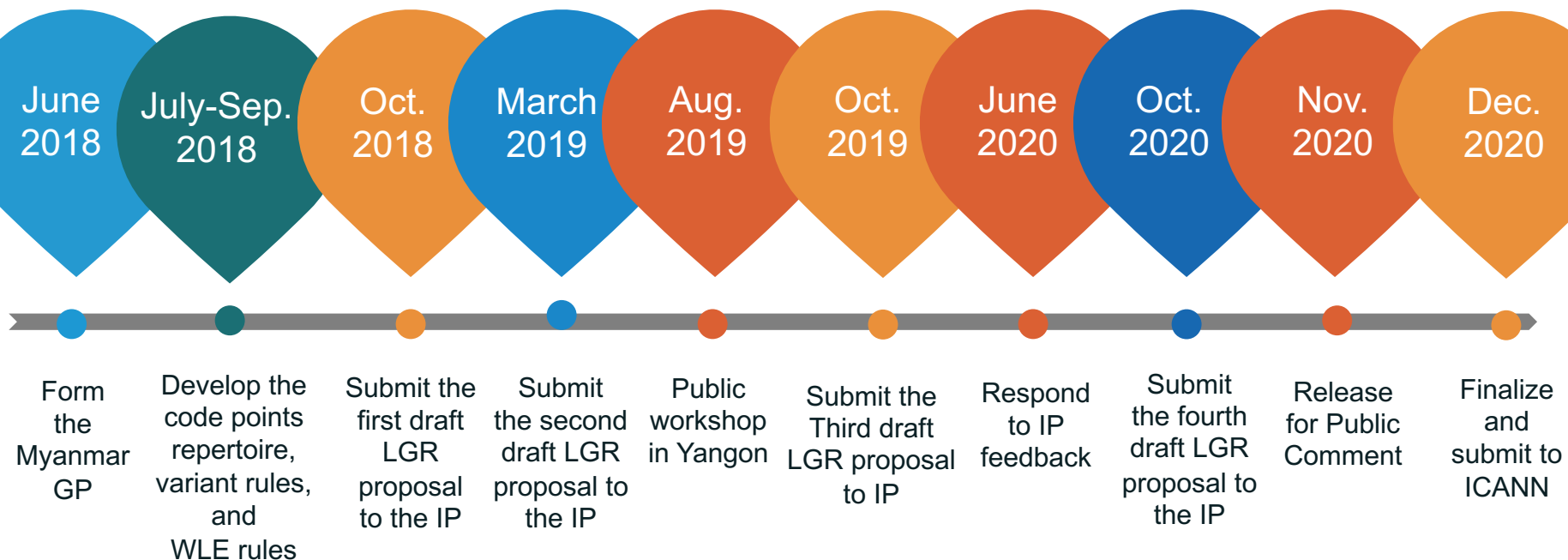
Rules for Combined Medial

- MM1 must follow CMM1
- MM3 must follow CMM3
- MM5 must follow CMM5

8. T_LONG or T_SHORT must follow C or M or LV or S12

10. Pao_Tone must follow DV or M or K

Timeline and Next Steps



Engage with ICANN and IDN Program



Thank You and Questions

Visit us at icann.org/idn

Email: IDNProgram@icann.org



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



facebook.com/icannorg



youtube.com/icannnews



flickr.com/icann



linkedin/company/icann



slideshare/icannpresentations



soundcloud/icann