

**ICANN**

**VIRTUAL COMMUNITY FORUM**

**70**

**PREP WEEK**



# IDN Program Update



ICANN70 Prep Week  
11 March 2021

# Overview of Session Presentations

---

- ⦿ Internationalized Domain Name (IDN) Program Overview and Progress (Pitinan Kooarmornpatana)
  
- ⦿ Community Updates
  - Greek Generation Panel (Panagiotis Papaspiliopoulos)
  - Korean Generation Panel (Kim Kyongsok)
  - Latin Generation Panel (Bill Jouris)
  - Myanmar Generation Panel (Yin May Oo)
  
- ⦿ Question and Answer

# IDN Program Overview and Progress

Pitinan Kooarmornpatana  
Senior Manager, IDN Programs

ICANN70 Prep Week

# IDN Program Objectives

---

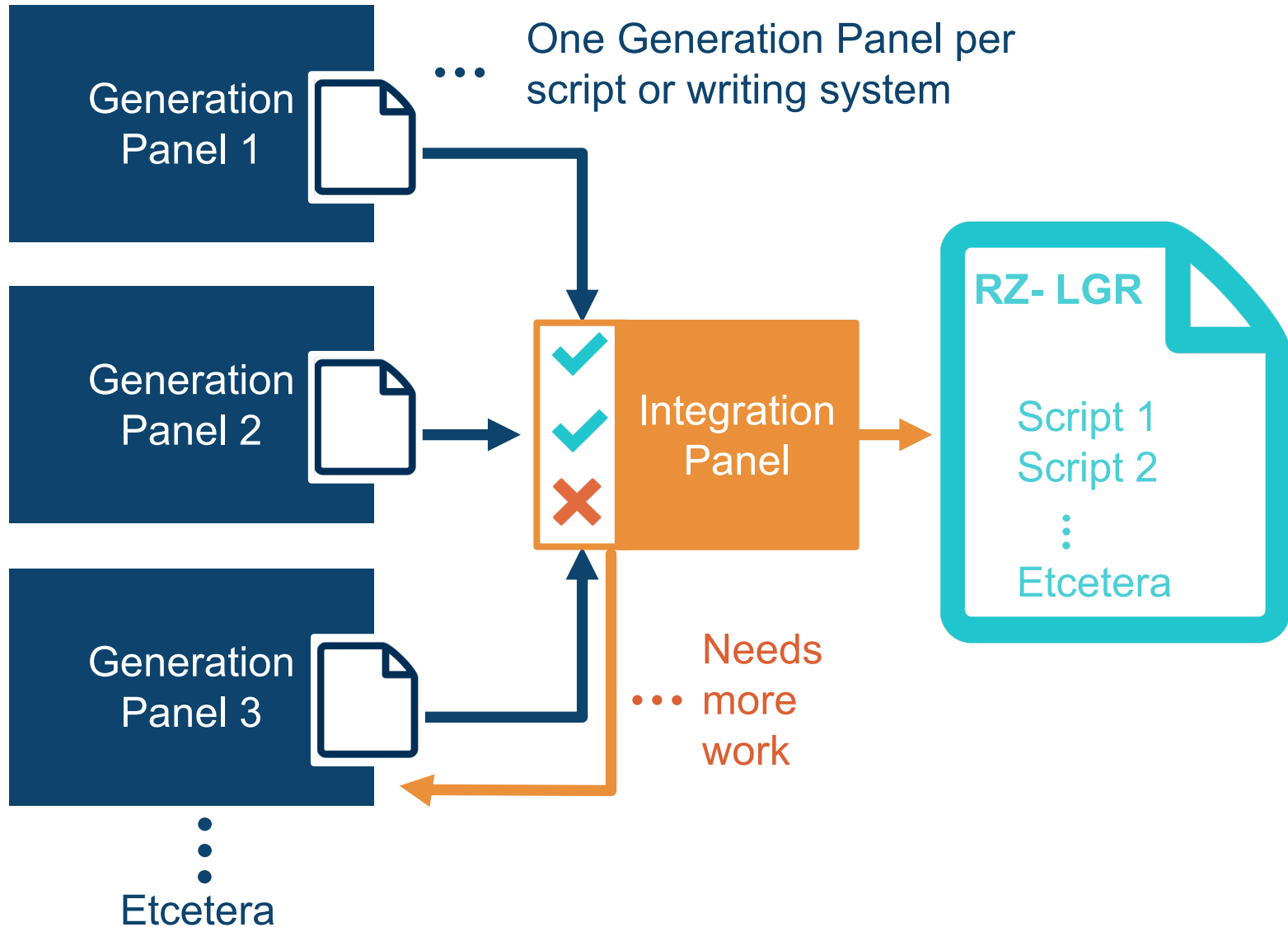
Enable deployment of domain names in the ***local languages and scripts*** of global communities in a ***secure and stable*** manner.

# Overview of IDN Programs

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



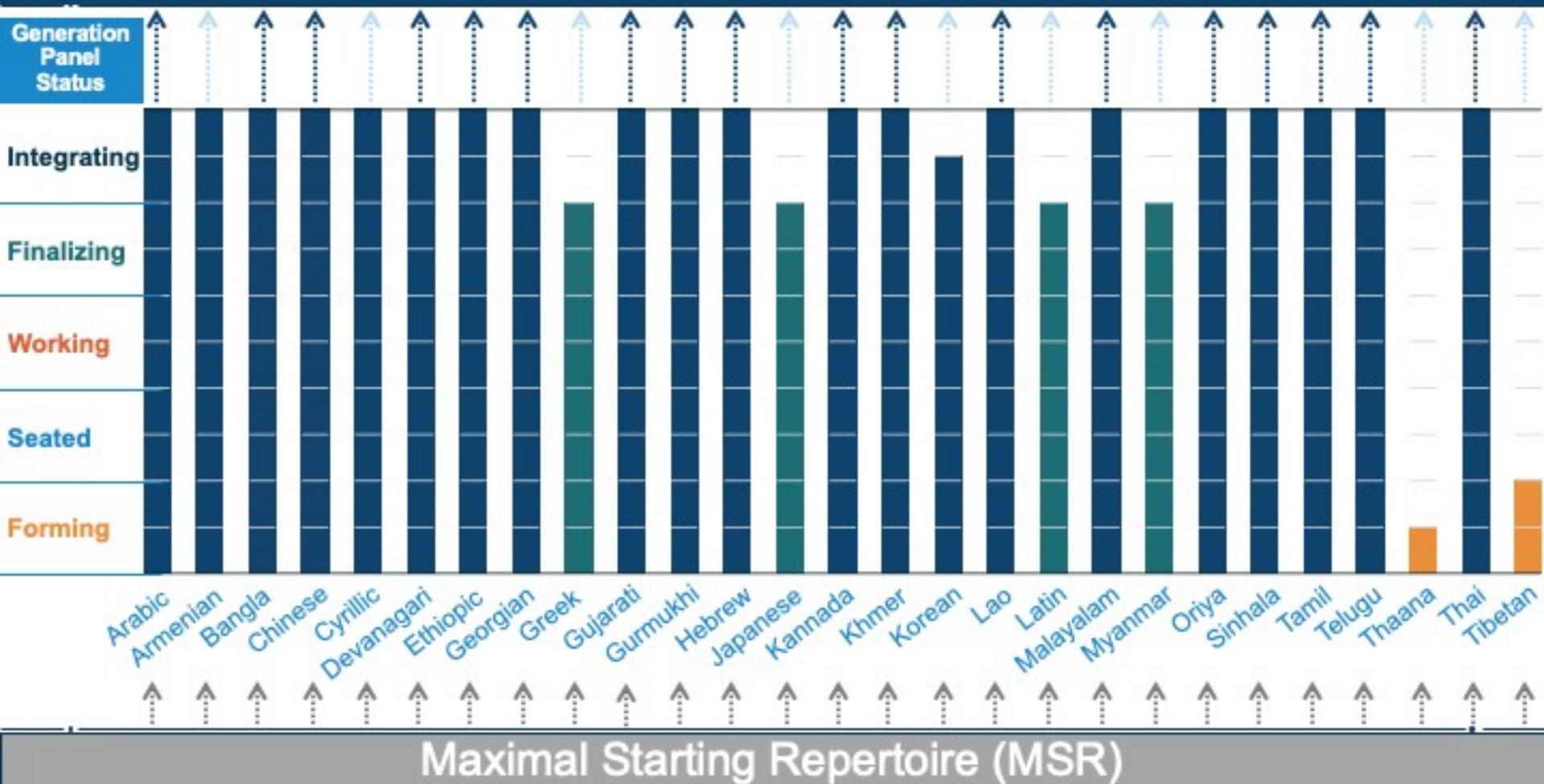
# RZ-LGR Process



# Status of Generation Panels (GPs)

Feb. 2021

## Root Zone Label Generation Rules (RZ-LGR)

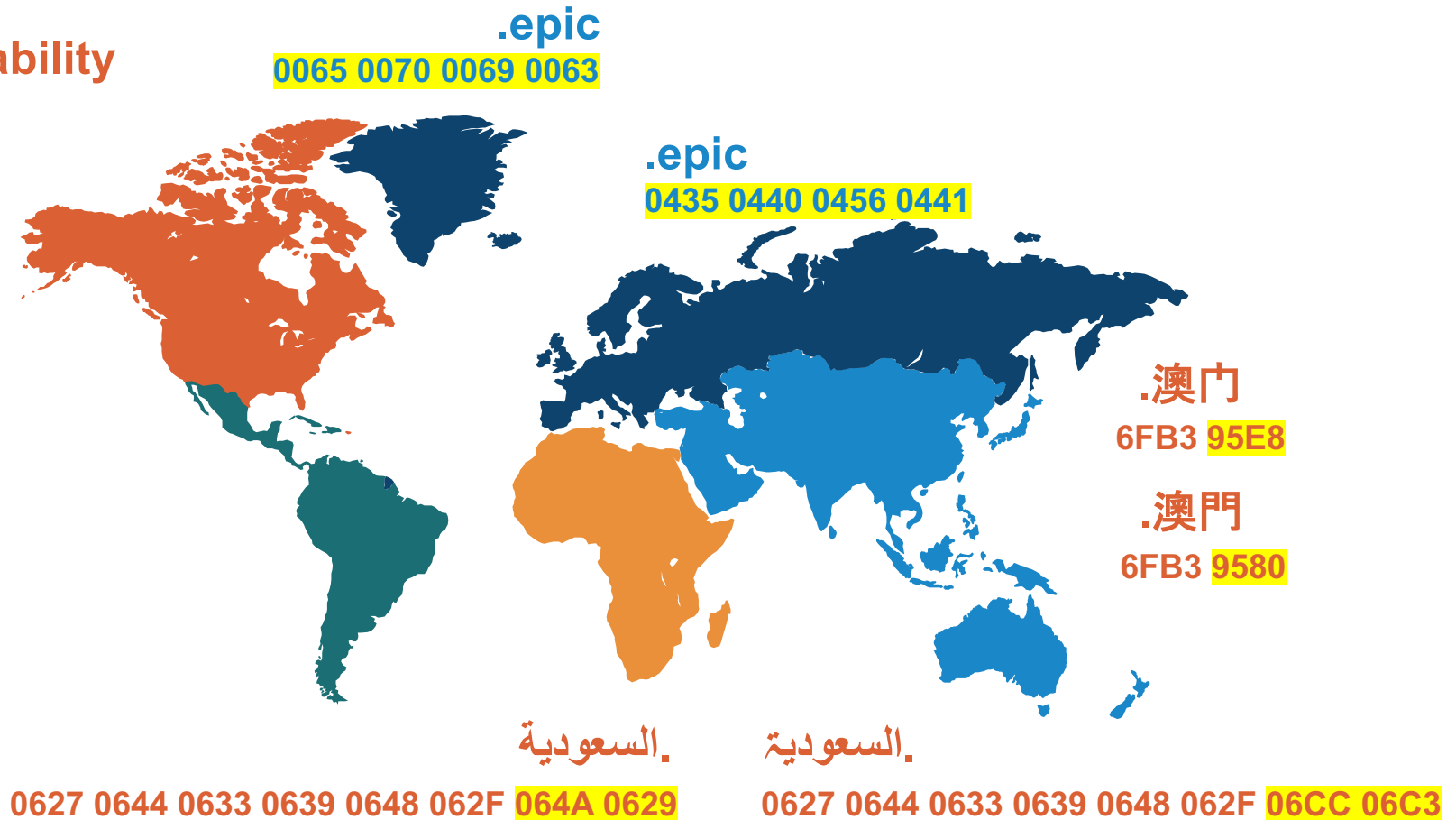




# Understanding IDN Variant TLDs



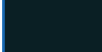
⦿ Security

⦿ Usability



# Which Scripts Have Variant Code Points?

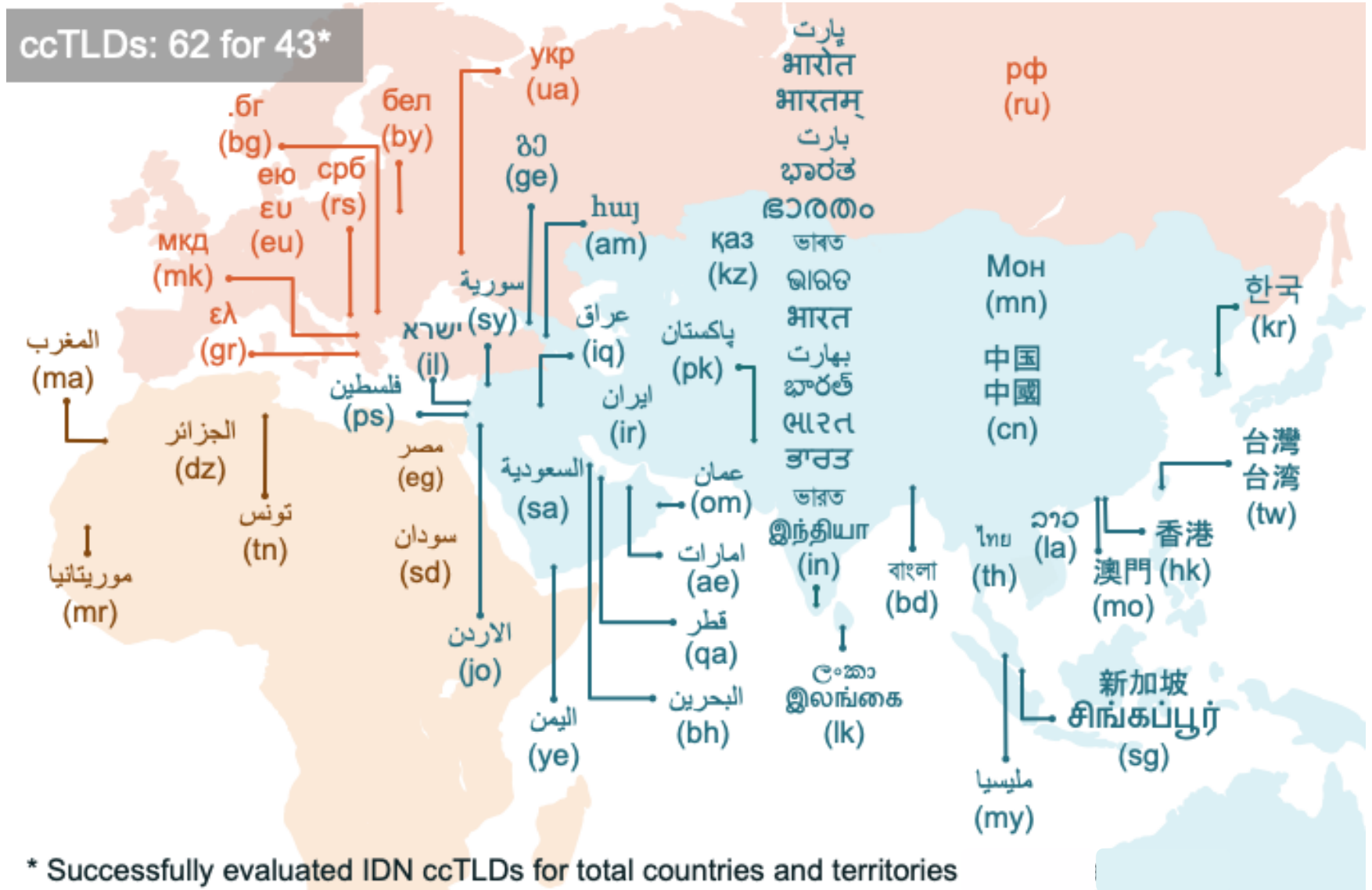
- ◉ **Arabic**
- ◉ **Armenian**
- ◉ **Bangla (Bengali)**
- ◉ **Cyrillic**
- ◉ **Devanagari**
- ◉ **Ethiopic**
- ◉ **Georgian**
- ◉ **Greek**
- ◉ **Gujarati**
- ◉ **Gurmukhi**
- ◉ **Han**
- ◉ **Hebrew**
- ◉ **Japanese**
- ◉ **Kannada**
- ◉ **Khmer**
- ◉ **Korean**
- ◉ **Lao**
- ◉ **Latin**
- ◉ **Malayalam**
- ◉ **Myanmar**
- ◉ **Oriya**
- ◉ **Sinhala**
- ◉ **Tamil**
- ◉ **Telugu**
- ◉ **Thaana**
- ◉ **Tibetan**
- ◉ **Thai**

	Variant code points
	No variant code points
	Work in progress

# 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 2019 after [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.”*
- ⊙ IDN and variant management related 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](#)
    - [IDN EPDP - Drafting Charter](#)

# 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 IDN Guidelines.

# Reference Second Level LGRs (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, 42 reference IDN tables in LGR format are [published](#):
  - Script-based: Bangla, Devanagari, Ethiopic, Georgian, Gujarati, Gurmukhi, Kannada, Khmer, Lao, Malayalam, Oriya, Tamil, Telugu.
  - Language-based: Arabic, Belarusian, Bosnian (Cyrillic), Bosnian (Latin), Bulgarian, Chinese, Danish, English, Finnish, French, German, Hindi, Hungarian, Icelandic, Italian, Korean, Latvian, Lithuanian, Macedonian, Montenegrin, Norwegian, Polish, Portuguese, Russian, Serbian, Spanish, Swedish, Thai, Ukrainian.
- ⦿ Additional reference IDN tables in LGR format are being published for [Public Comment](#) (closing 11 March 2021):
  - Script-based: Arabic, Hebrew, Sinhala
  - Language-based: Hebrew

# LGR (IDN Table) Review Tool

---

- ⦿ The aim of 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 compares an IDN table with the relevant reference LGR.
- ⦿ It will provide the IDN table review report in HTML format as output.
- ⦿ The tool is tentatively planned to be released in April 2021.

# Community Updates

- Greek Generation Panel
- Korean Generation Panel
- Latin Generation Panel
- Myanmar Generation Panel



# Greek Generation Panel Update

Panagiotis Papaspiliopoulos  
Chair, Greek GP

ICANN70 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

---

- ⦿ ISO 15924 Code: Grek
- ⦿ ISO 15924 Key No: 200
- ⦿ ISO 15924 English Name: Greek
- ⦿ Property Value Alias: Greek
- ⦿ Native Name of the script: Ελληνικά
- ⦿ Maximal Starting Repertoire (MSR) version: MSR-4
- ⦿ The Greek script is used to write the Greek language, which is the official language of Greece and Cyprus

# Generation Panel Membership

Name	Role	Organization	Country	Field of Expertise
<b>Panagiotis PAPASPILIOPOULOS</b>	Chair	Hellenic Ministry of Digital Governance	Greece	Technical, Policy Development, GAC Rep
<b>Vaggelis SEGREDAKIS</b>	Vice Chair	Registry of [.gr] & [.ελ] domain names, FORTH-ICS	Greece	Technical (DNS) ccNSO Rep
<b>Konstantinos VALASIS</b>	Member	Hellenic Ministry of Foreign Affairs	Greece	Policy Development
<b>Ioannis KATSARAS</b>	Member	Hellenic Ministry of Foreign Affairs	Greece	Policy Development
<b>Asimina GIANNOPOULOU</b>	Member	Hellenic Telecommunications and Post Commission	Greece	Technical, Regulatory
<b>Giorgos KOLYVAS</b>	Member	Hellenic Telecommunications and Post Commission	Greece	Technical, Regulatory
<b>Michael VASSILAKIS</b>	Member	Registry of [.gr] & [.ελ] domain names, FORTH-ICS	Greece	Technical, ccNSO Rep
<b>Antonis ANTONIADES</b>	Member	Office of Electronic Communications and Postal Regulations	Cyprus	Technical, Policy Development, Regulatory
<b>Constantinos STYLIANOU</b>	Member	Office of Electronic Communications and Postal Regulations	Cyprus	Technical, Policy Development, Regulatory
<b>Alexandros PSYRRIS</b>	Member	Hellenic Body for Standardization	Greece	Technical (Unicode)
<b>Maria GAVRILIDOU</b>	Member	Institute for Language and Speech Processing / Research Centre “Athena”	Greece	Linguistics
<b>Penny LABROPOULOU</b>	Member	Institute for Language and Speech Processing / Research Centre “Athena”	Greece	Linguistics

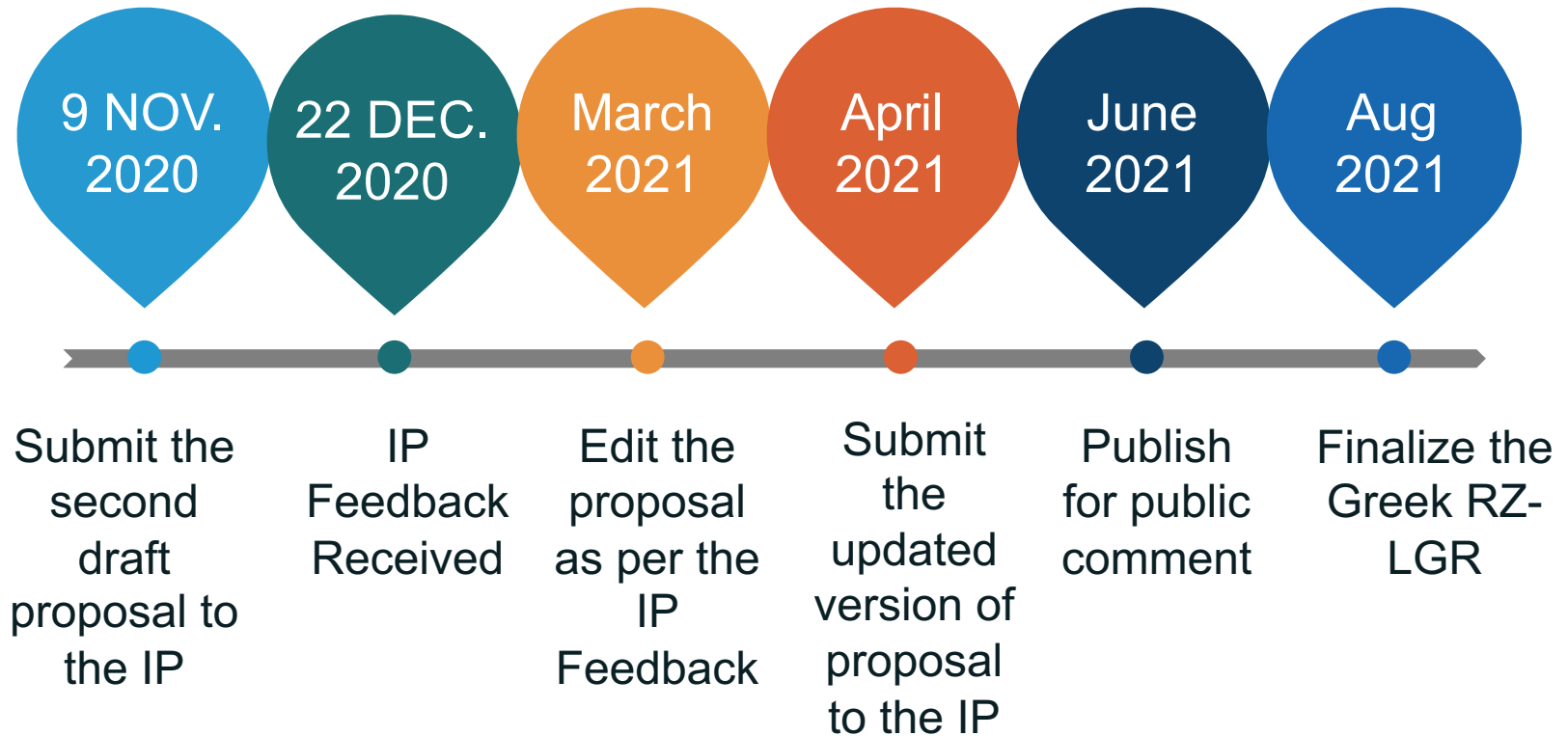
# Progress Summary

---

- ⦿ Greek GP was formed 31 October 2016
- ⦿ Geek GP has developed the LGR proposal 2017 - 2018
- ⦿ Joined Latin GP F2F meeting to discuss cross script variant – Jan 2019
- ⦿ Joint meeting with Latin, Cyrillic, Armenian, Greek GP at ICANN66 – November 2019
- ⦿ Joined Latin GP F2F meeting to discuss imposed in-script variant – Feb 2020

- ⊙ Finalizing Greek LGR proposal for the Root Zone
  - Applying mechanism to limit the allocatable variant labels to two (applied-for label, and the base form). Except that if the labels also contain sigma, the maximal number of allocatable variants per original label is four.
  - Updating the proposal (XML), supporting document, and test labels.

# Plan and Next Steps



# Korean Generation Panel Update

Kim Kyongsok  
Chair, Korean GP

ICANN70 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 spanning 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: Jungsup PARK, Minjung PARK, Boyoung KIM
- ⊙ Registrar: Geonwoo KIM, ChangKi JANG, Myungsoo LEE

# Progress Summary (1): K-LGR v2.3 (2021.02.01.)

- K-LGR v2.1 (2021.02.01.): 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)
  - Total # Variant groups = 283 + 7 – 1 (duplicate\*) = 289
    - \* variant group: [(U+5408 合) (U+95A4 閤) (U+C2B4 슴)]
- 4758 Hanja chars in K-LGR v2.3 (the same as in v1.0)

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

(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.

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, v2.1, v2.2, and v2.3

---

- 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.
- K-LGR v2.2 (2020.12.25.) sent to IP
  - IP feedback RE: K-LGR v2.1 mostly reflected in K-LGR v2.2.
- K-LGR v2.3 (2021.02.01.) sent to IP
  - IP feedback RE: K-LGR v2.2 mostly reflected in K-LGR v2.3.

# Current Work

---

- Waiting for the IP feedback RE: K-LGR v2.3
- K-LGR v2.3 is posting for the second round of Public Comment on 2 March 2021.

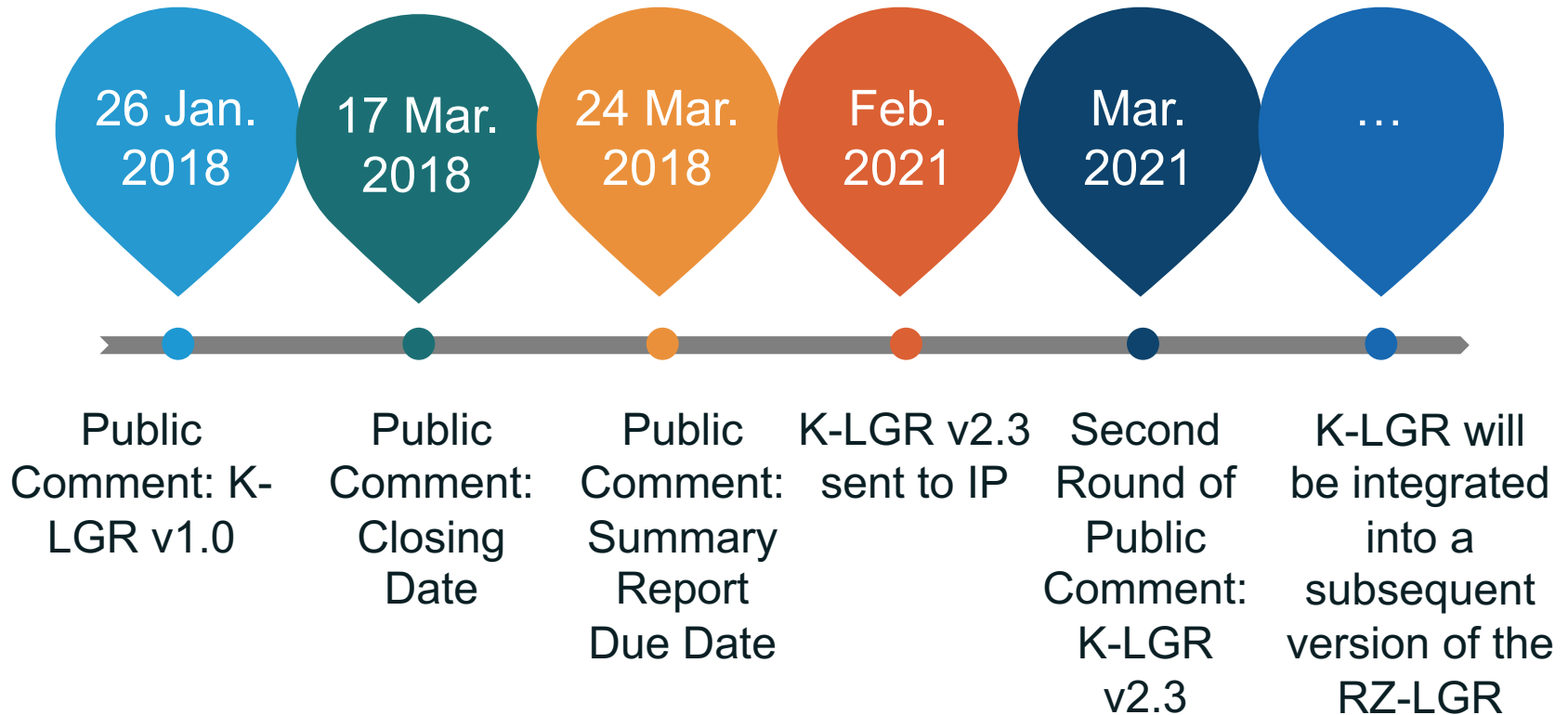


# Brief History of KGP Activities

---

- 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 (ICANN #49 ~ #64).
- Several CJK coordination meetings in Rep. of Korea, China, and Taiwan.

# Plan and Next Steps



# Latin Generation Panel Update

Mirjana Tasić, Bill Jouris  
Chair and Member, Latin GP

ICANN70 Prep Week

# Agenda

---

1

Overview of the  
Script and Language

2

Generation Panel  
Membership

3

Progress Summary

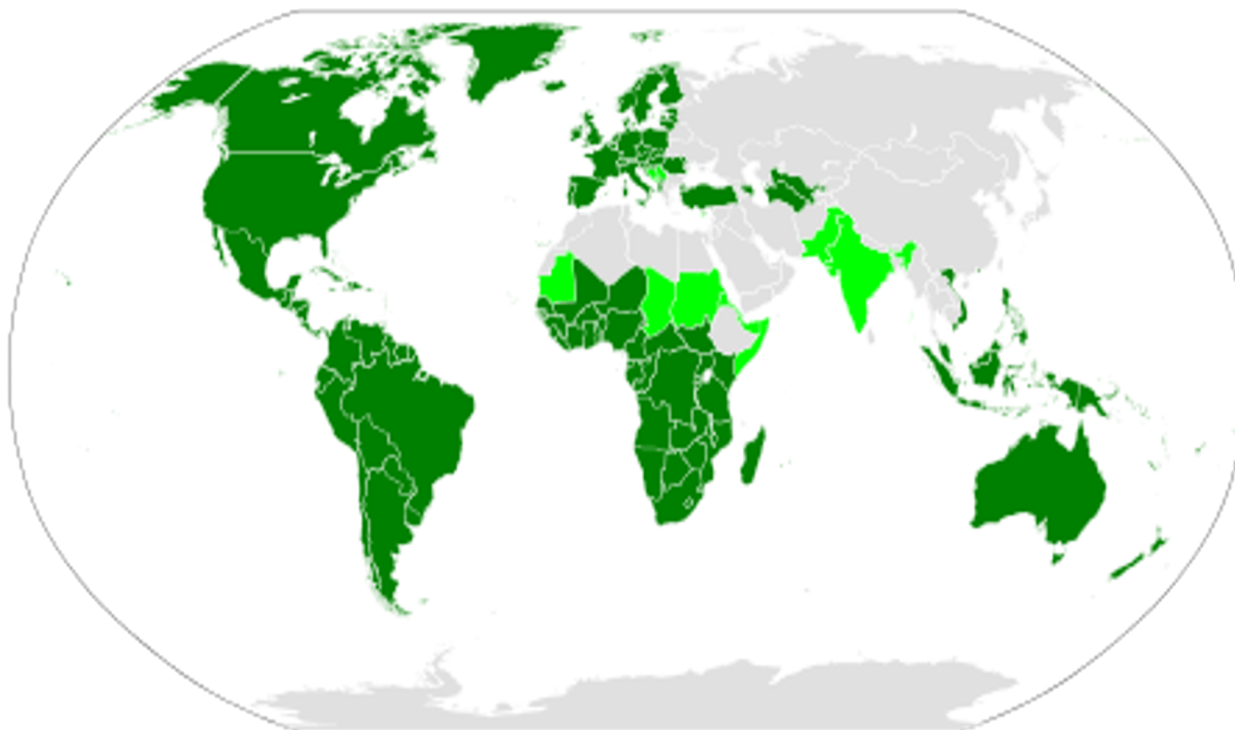
4

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.

A quick calculation shows that about 2.6 billion people (36% of the world population) use the Latin alphabet according to <https://www.worldstandards.eu/other/alphabets>.

# Generation Panel Membership

---

Since 2016, the Generation Panel has seven 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 took approximately two years of work.
  - 181 of 181 EGIDS 1- 4 languages processed.
  - 31 EGIDS 5 languages processed with more than one million speakers.
  - 197 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 took approximately three years of work.
  - 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.

# Progress Summary

---

- ⦿ Submitted the second version of proposal to the IP in May 2018.
- ⦿ Submitted the third version of proposal to the IP in January 2019.
- ⦿ Submitted the fourth version of proposal to the IP in October 2019.
- ⦿ Submitted the fifth version of proposal to the IP on 23 October 2020.
- ⦿ IP comments received on 23 December 2020.

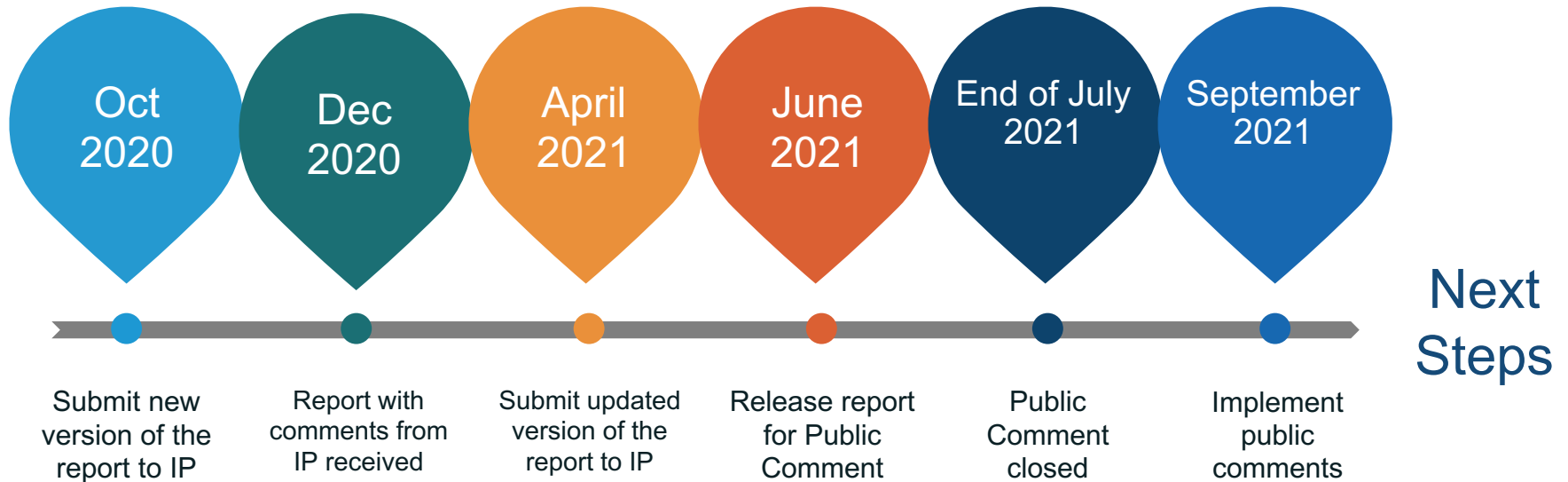


# Current Work

---

- ⊙ GP is now working on IP comments.
- ⊙ Production of test data set.
  - Test data for repertoire completed.
  - Test data for variants in progress.
- ⊙ XML production and check.
- ⊙ Submit new version of report to IP in April 2021.

# 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 September 2021.

# [Myanmar Generation Panel] Update at ICANN70

Yin May Oo  
Myanmar GP Co-Chair

# Agenda

---

1

Overview of the  
Script and Language

2

Summary of the  
Progress

3

Plan and Next Steps

# Myanmar GP – Languages Using Myanmar Script

- Myanmar script have been evolved from the Brahmi script which has flourished in the Indian subcontinent between 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 /lik.táy/	3	3,295,000
Rakhine	[rki]	Myanmar	ရခိုင် /rə.khi/	3	2,020,000
Karen, Sgaw	[ksw]	Myanmar, Thailand	ꨀꨃ /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

# Background of Myanmar-Script Languages

The word “Hello” written in listed languages:

Burmese: /mɪŋ-gə-la-ba/ “မင်္ဂလာပါ”

Shan: /mɪ-suŋ-khə/ “မိၣ်သုၣ်ခါး”

Rakhine: /θa-li-zwə-ba/ “သာလီဝပါ”

Sgaw Karen: /ni:-leə-ə-ye:/ “နီလၢအဂ့ၤ”

Mon: /mŋeə-rə-au/ “မိၣ်ရအံ”

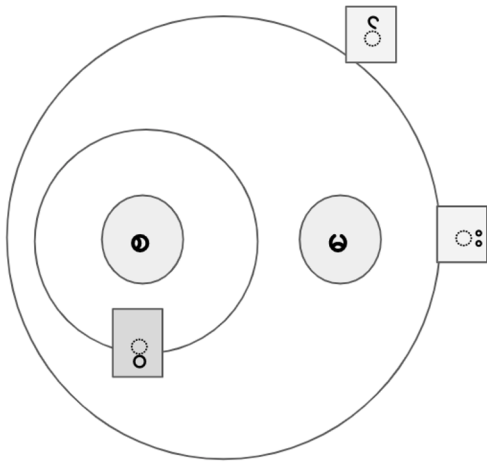
Pa’O Karen: /mɪŋ-gə-lə-dra-wa: ]  
“မင်္ဂလာဒုးဒြာ”

1. Consonant
2. Independent Vowel
3. Diacritic
  1. Dependent Vowel
  2. Medial Signs
  3. Tone marks
  4. Various signs



# Nature of Myanmar Script (LTR)

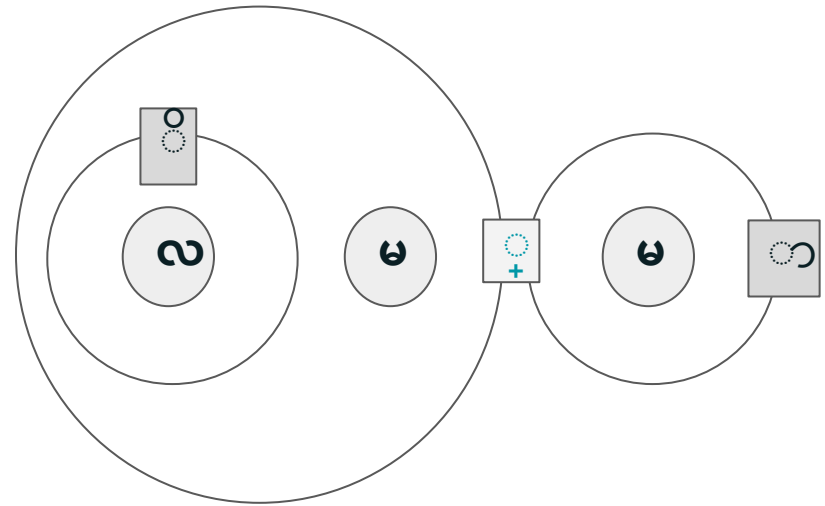
စွမ်း



/စ/ွ/မ/်/း/

/ U+1005 / U+103D / U+1019 / U+103A / U+1038 /

လိမ္မာ



/လ/ိ/မ/ံ/မ/ာ/

/ U+101C / U+102D / U+1019 / U+1039 / U+1019 / U+102C /

# In-Script Variant Analysis

- Myanmar GP defines the following are 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	Detail
1	U+1023 MYANMAR LETTER I	ꨩ	U+1000 U+1039 U+1000 SV1	ꨩ	blocked	Homoglyph
2	U+101F U+103A	ꨵ	U+1015 U+102C U+103A also (U+1015 ST1)	ꨵ	Block	Homoglyph
3	U+1061	ꨱ	U+101B U+103E	ꨱ	Block	Homoglyph
4	U+1004	ꨀ	U+105A	ꨀ	Allocatable	Becomes homoglyph when diacritics added
5	U+1008	ꨀ	U+105B	ꨀ	Allocatable	Semantic, Homoglyph
6	U+1001	ꨀ	U+1076	ꨀ	Allocatable	Homoglyph (for most of the fonts)
7	U+102E	ꨀ	U+1033	ꨀ	Allocatable	Semantic
8	U+102B	ꨀ	U+102C	ꨀ	Allocatable	Semantic



# Cross-Script Variant Analysis

No.	Glyph	Code Point	Character Name
1	ო	U+10D8	GEORGIAN LETTER IN
2	ဂ	U+1002	MYANMAR LETTER GA
3	თ	U+10D7	GEORGIAN LETTER TAN
4	တ	U+1010	MYANMAR LETTER TA
5	ଌ	U+0B47	ORIYA VOWEL SIGN E
6	ꯀ	U+1031	MYANMAR VOWEL SIGN E
7	ଠ	U+0B20	ORIYA LETTER TTHA
8	ഠ	U+0D20	MALAYALAM LETTER TTHA
9	o	U+006F	LATIN SMALL LETTER O
10	ο	U+03BF	GREEK SMALL LETTER OMICRON
11	о	U+043E	CYRILLIC SMALL LETTER O
12	օ	U+0585	ARMENIAN SMALL LETTER OH

# Whole Label Evaluation Rules

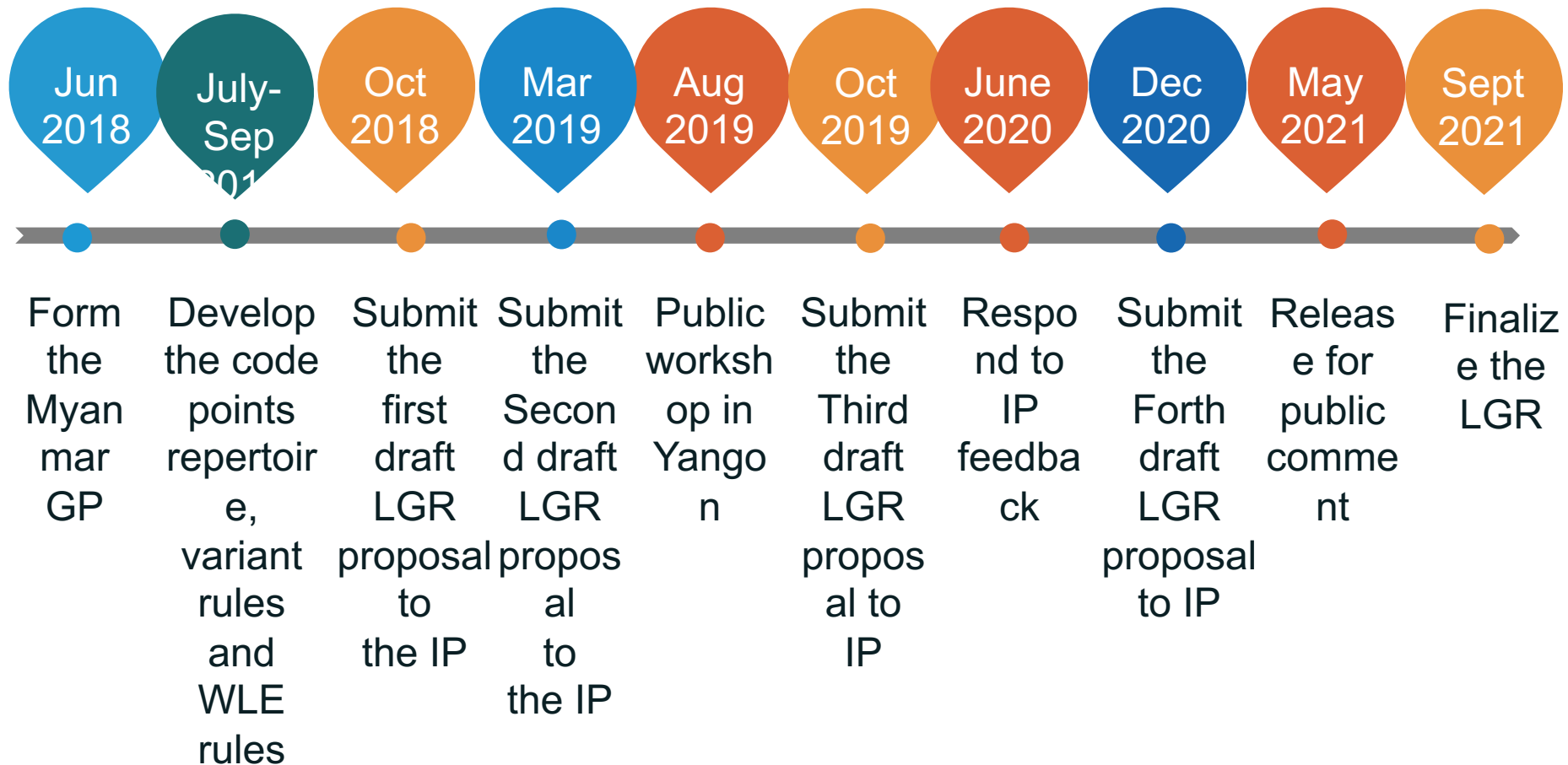
## Classifications - 19

## Rules - 14

### ~~Addition~~ according to In-script variants -

Set#	Unicode Code Point	Glyph	Unicode Code Point	Glyph	Disposition	Detail
4	U+1004	ç	U+105A	Ḅ	Allocatable	Becomes homoglyph when diacritics added
5	U+1008	Ḅ	U+105B	Ḅ̇	Allocatable	Semantic, Homoglyph almost
6	U+1001	ə	U+1076	Ḃ	Allocatable	Homoglyph (for most of the fonts)
7	U+102E	ö	U+1033	Ḅ̇	Allocatable	Semantic

# Timeline and Next Step



# Engage with ICANN and IDN Program



**Thank You and Questions**

Visit us at [icann.org/idn](http://icann.org/idn)

Email: [IDNProgram@icann.org](mailto:IDNProgram@icann.org)



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



[facebook.com/icannorg](https://facebook.com/icannorg)



[youtube.com/icannnews](https://youtube.com/icannnews)



[flickr.com/icann](https://flickr.com/icann)



[linkedin/company/icann](https://linkedin/company/icann)



[slideshare/icannpresentations](https://slideshare/icannpresentations)



[soundcloud/icann](https://soundcloud/icann)