## ICANN|57 HYDERABAD





# IDN Program | ICANN 57 | 6 November 2016

### **Overview of Session Presentations**

- ⊙ IDN Program Overview and Progress
- ⊙ Update by Integration Panel
- IDN Implementation Guidelines
- ⊙ Community Updates
  - Ethiopic GP
  - Georgian GP
  - Neo-Brahmi GP
  - Thai GP
- ⊙ Q/A

- Sarmad Hussain
- Asmus Freytag
- Edmon Chung

- Dessalegn Yehuala
- Sophio Elizbarashvili
- Akshat Joshi
- Wanawit Ahkuputra



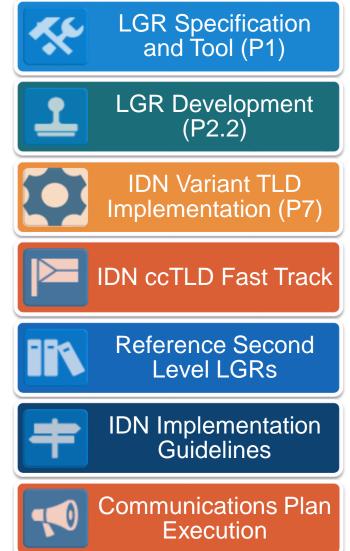
### **IDN Program Overview and Progress**

Sarmad Hussain Director, IDN Program

### **Overview of Presentation**

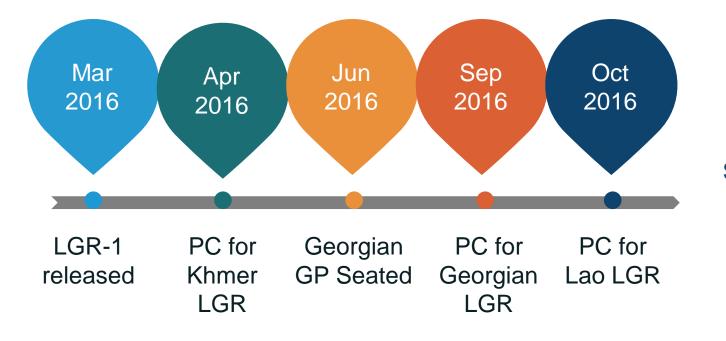
#### ⊙ IDNs at Top Level

- IDN TLD Program
  - Label Generation Rules (LGR)
  - LGR Toolset
  - IDN Variant TLD Implementation
- IDN ccTLD Fast Track Process
- $\odot\,$  IDNs at Second Level for gTLDs
  - IDN Implementation Guidelines
  - Reference Second Level LGRs
- Community Outreach and Involvement





### Root Zone Label Generation Rules (LGR)

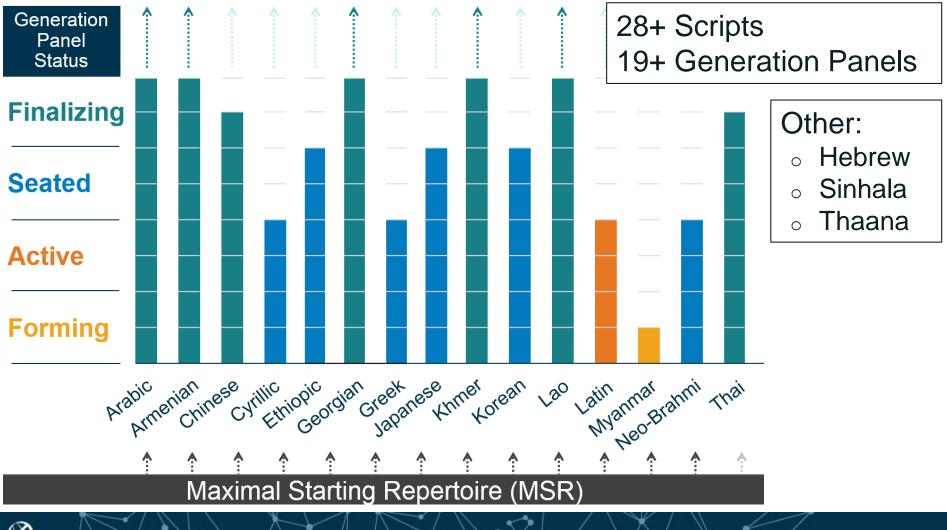


Additional scripts will be incorporated in future versions of LGR



### Status of LGR Development

#### Label Generation Rules (LGR)



October 2016

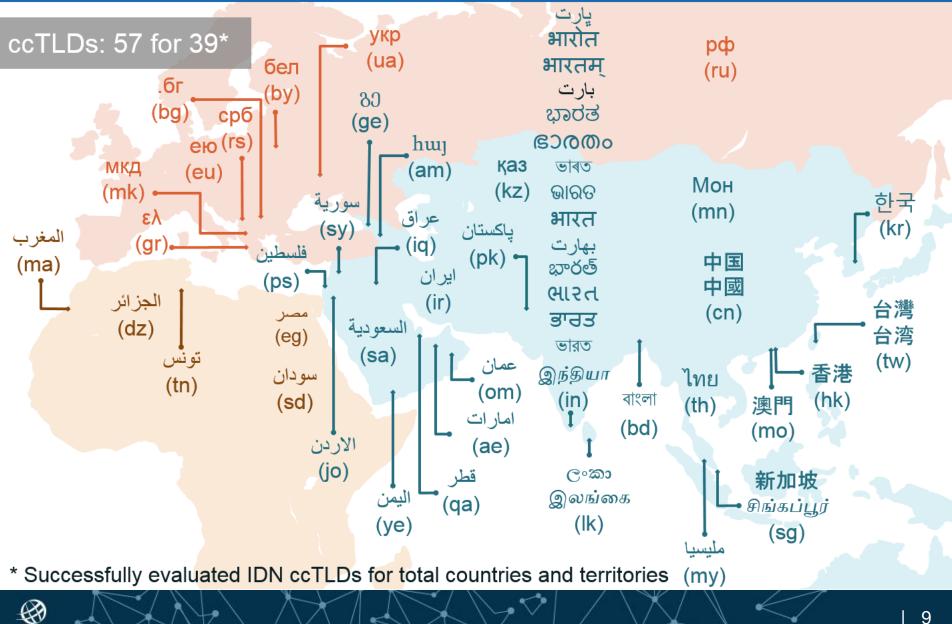
### LGR Toolset (beta)

- Label Generation Rulesets (LGRs) used to generate domain name labels, as specified in <u>RFC 7940</u>
- LGR Toolset allows for the following:
  - Create a LGR
  - Use a LGR to validate a label and determine its variants
  - Manage LGRs, by comparing or combining them
  - Review possible impact of a new or a revised LGR on existing labels
- Online beta deployment
  - Visit <u>https://lgrtool.icann.org/</u>
- Open source package(s) release with BSD license
  - Released at github: <u>lgr-core</u>, <u>lgr-django</u>, <u>munidata</u>
- <u>User guide</u> available for further details



### IDN Country Code Top-Level Domains

ICANN



### IDN ccTLD Fast Track Process

#### ⊙ Launched in late 2009

- 57 IDN ccTLDs evaluated representing 39 countries/territories
- 47 IDN ccTLDs delegated representing 37 countries/territories
- Requests cover 21 scripts for 35 languages
- ⊙ Currently under review
  - <u>Public comment</u> in Jan 2015 raised issues with second similarity review process (EPSRP)
  - <u>Board resolution</u> in June 2015 for ccNSO to review EPSRP (with input from GAC and SSAC)
  - ccNSO formed EPSRP Working Group (WG)
  - Public comment in July 2016 by WG on revised EPSRP



### Reference Second Level LGRs

- To assist applicants and registry operators in the Pre-Delegation Testing (PDT) and Registry Services Evaluation Policy (RSEP) process, based on the <u>Guidelines</u>
  - Use the reference LGRs or derive custom LGRs
- Public comment on LGRs and evaluation process in June 2016
- ⊙ 27 final LGRs <u>published</u> in Oct 2016 based on <u>RFC 7940</u>
  - Latin: Bosnian, Danish, English, Finnish, French, German, Hungarian, Icelandic, Italian, Latvian, Lithuanian, Norwegian, Polish, Portuguese, Spanish, Swedish
  - Cyrillic: Belarusian, Bosnian, Bulgarian, Macedonian, Montenegrin, Russian, Serbian, Ukrainian
  - Mixed scripts: Korean
  - Others: Chinese, Hebrew



### **Communication and Outreach Efforts**

- ⊙ Updated IDN web pages at <u>icann.org/idn</u>
- ⊙ IDN Program sessions at ICANN meetings
- ⊙ IDN Program updates to SOs/ACs at ICANN meetings
- Direct outreach

| 21-23 July 2016  | Ethiopic Generation Panel Training  | Addis Ababa, Ethiopia |
|------------------|-------------------------------------|-----------------------|
| 27-29 July 2016  | APRIGF                              | Taipei, Taiwan        |
| 17-18 Aug. 2016  | SDNOG                               | Khartoum, Sudan       |
| 26-28 Sept. 2016 | <b>TF-AIDN Face to Face Meeting</b> | Istanbul, Turkey      |
| 29-30 Sept. 2016 | CJK GPs Face to Face Meeting        | Taipei, Taiwan        |

- ⊙ Blogs
  - Blogs: <u>Coffee Beans and the Ethiopic Script</u>; <u>LGR-1</u>
- IDN community wiki pages
- ⊙ IDN mailing lists
  - o {vip, Igr, ArabicGP, ArmenianGP, ChineseGP, ...}@icann.org



### Thank You

- For information on IDN Program projects, please visit: <u>www.icann.org/idn</u>
- For queries regarding the IDN Program, please email: <u>IDNProgram@icann.org</u>



## **Update by the Integration Panel**

Asmus Freytag Integration Panel



### Agenda

- $\odot$  IP activities since ICANN 55
- $\odot$  Considerations
  - Southeast Asian
  - $\circ$  CJK
  - $\circ$  Others



### IP activities since ICANN 55 (Marrakech)

#### $\odot$ Detailed review for (near) final LGRs

- ⊙ Khmer, went to public comment
- ⊙ Georgian, went to public comment
- ⊙ Lao
- ⊙ Thai
- ⊙ Interim review
  - $\odot$  Chinese
- ⊙ Initial feedback
  - ⊙ Korean
  - ⊙ Ethiopic
- Communications
  - ⊙ Japanese
  - ⊙ Neo Brahmi (governing principles)
- ⊙ IP F2F session in June



- ⊙ Rate of progress is very encouraging
- IP gaining useful experience with complex scripts
- ⊙ "Complex scripts"
  - depend on layout engines to support an internal structure of each syllable for correct rendering.
  - require restricted placement of certain code points to contexts expected by the layout engines.
  - some limits on letter combinations are customary, not structural
    - ⊙ "spelling rules" do not have to be enforced in LGR



⊙ Encouraging: Seeing more CJK progress

- ⊙ GPs have shared preliminary LGR drafts
- $\odot$  Allows IP to give detailed feedback
- ⊙ Progress on variants, but more work to do
- Challenge: Need to reduce the number of allocatable variants
- Need to document the source and use case for all variants



### **Considerations – Other**

- GP sharing of preliminary LGR drafts with IP for early feedback: enables faster convergence
- Some scripts (Ethiopic) do not seem to have an agreed-upon spelling, which makes LGR design challenging
- Encourage communities, especially Latin, Greek, Cyrillic, to progress
- LGR formal specification is now IETF Standard Track (RFC 7940).



## Update on IDN Implementation Guidelines

Edmon Chung IDN Guidelines WG Co-Chair



### **Background and Purpose**

- ⊙ Purpose
  - Guidelines for second level IDN registration policies and practices
  - Designed to address end-user concerns, e.g. user confusion
- $\odot$  Relevance
  - For gTLD registries and registrars offering IDNs at the second level
  - For IDN ccTLDs
- Status
  - GNSO community requested for updating the Guidelines
    - Previous version (<u>3.0</u>) updated in 2011
  - Currently being reviewed and updated by IDN Guidelines Working Group



### IDN Guidelines WG Members

|    | Name                                | Organization    | SO/AC |
|----|-------------------------------------|-----------------|-------|
| 1  | Satish Babu                         | ISOC-TRV        | ALAC  |
| 2  | Wael Nasr                           | TLDVILLA LLC    | ALAC  |
| 3  | Mats Dufberg                        | IIS             | ccNSO |
| 4  | Pablo Rodríguez                     | Puerto Rico TLD | ccNSO |
| 5  | Edmon Chung                         |                 | GNSO  |
| 6  | Christian Dawson                    | i2Coalition     | GNSO  |
| 7  | Chris Dillon                        |                 | GNSO  |
| 8  | Kal Feher                           | Neustar         | GNSO  |
| 9  | Dennis Tan                          | Verisign        | GNSO  |
| 10 | Jian Zhang                          | KNET            | GNSO  |
| 11 | Ram Mohan                           | Afilias         | SSAC  |
| 12 | Patrik Fältström (will only review) |                 | SSAC  |



### Summary of Items Covered by the Guidelines

- 1. Transition from IDNA2003 to IDNA2008
- 2. Terminology
- 3. Format of IDN Tables
- 4. Consistency of IDN Tables
- 5. User Acceptance
- 6. IDN Variant Labels Cont.
- 7. Similarity and Confusability of IDN Labels TBD
- 8. Registration Data TBD
- 9. EPP TBD

Detailed text of the recommendations available at IDNGWG Wiki page: <a href="https://community.icann.org/display/IDN/IDN+implementation+Guidelines">https://community.icann.org/display/IDN/IDN+implementation+Guidelines</a>



○ Complete the draft guidelines for remaining topics

- ⊙ IDN Variant Labels cont.
- $\odot$  Similarity and Confusability of Labels
- Registration Data
- ⊙ EPP
- Release proposed guidelines for public comment
- $\odot$  Finalize guidelines and publish for adoption



 $\odot$  For detailed guidelines on the topics, please visit:

 IDN Guidelines WG Wiki page: <u>https://community.icann.org/display/IDN/IDN+implementat</u> <u>ion+Guidelines</u>

• For feedback, email at:

⊙ <u>idngwg@icann.org</u> or <u>IDNProgram@icann.org</u>



## Update by the Ethiopic GP

Dessalegn Yehuala Ethiopic GP Chair



- $\odot$  Introduction to Ethiopic script
- ⊙ Overview of GP members
- ⊙ Challenges
- $\odot$  Progress and timeline for remaining activities



### Introduction to Ethiopic Script





## Used for the National Languages of Ethiopia and Eritrea



## Languages or Writing Systems Using the Language:

Several languages use the script in their writing systems. However, only eight are considered for the LGR proposal:

 Amharic, Argobba, Awngi, Harari, Hamt'agna, Silt'e, Tigrè, Tigrigna



### Introduction to Ethiopic Script



#### **Users Community**

Eritrea, Djibouti, Egypt, South Africa, South Sudan, Kenya, United States, Canada, Israel, Europe, Australia



## Each Letter of the "Syllabary" has both a consonant and a vowel

 $\circ h = ke \qquad \circ h = ki \\ \circ h = ku \qquad \circ h = ka$ 



### **Code Point Redundancy or Phonemic Decay**

There are about 69 code points with sound and meaning alike in Amharic language writing system

- ሐልዮ Song
- /ነልዮ Theory



### **GP** Members

- Dessalegn Mequanint Yehuala (Chair), Lecturer and researcher in computer science
- Mulugeta Seyoum (Secretary), Assistant Professor in linguistics
- ⊙ Gezahegn Tadesse, Director IT projects, Awash International Bank
- ⊙ Balcha Reba, Director standards at MCIT
- ⊙ Assefa Kore, IT expert at Ethio-telecom
- ⊙ Kinfe Michael Yilma Desta. Ph.D. student in Law
- ⊙ Tigabu Dagne Akal, Lecturer at AIT in computer engineering
- Halefom Hailu Abraha, expert in cyber law, INSA
- ⊙ Teshome Yehualashet, Assistant Professor in linguistics
- ⊙ Kassahun Lemlemu, Journalist at MK
- Destaye Alemayehu, M.Sc. student in computer science



- ⊙ Languages are under-resourced
  - Getting pertinent information about the languages both in paper and digital forms has been a challenge
- Deviation from the original timeline
  - GP members being immersed in their primary duty and lack of time slot for other things



### Progress and Timeline for Remaining Activities



#### **To Summarize**

Draft Proposal for the Ethiopic Script LGR for the Root Zone has been submitted. We are expecting to receive the IP's feedback in the immediate future. Afterwards it will be a case of mapping the LGR document into machine readable XML using the online tool available. We are planning to conclude the project in two month's time.



### Thank You



## **Update by the Georgian GP**

Sophio Elizbarashvili Georgian GP Chair



- $\odot$  Background on script and principal languages using it
- GP members
- LGR overall development process and methodology
- Challenges behind the process
- ⊙ Current progress and timeline



### Introduction to Georgian Script - მხედრული

| <ul> <li>Three writing systems:</li> <li>Asomtavruli</li> <li>Nuskhuri</li> <li>Mkhedruli</li> </ul> | ISO 15924: Geor<br>ISO 15924 N°: 240                          | The oldest<br>Mkhedruli<br>inscription found is<br>dated back to 982<br>AD            |
|--|---|---|
| Written horizontally<br>following the<br>standard left-to-<br>right, with spaces<br>between words    | Spoken by about<br>4.1 million<br>people mainly in<br>Georgia | Georgian script also<br>used to write:<br>• Mingrelian<br>• Laz<br>• Svan<br>• Abkhaz |



### Georgian Script Generation Panel Members

| Name                    | Role                           | Organization                                       |
|-------------------------|--------------------------------|--|
| Sophio Elizbarashvili   | Generation Panel<br>Chair      | Information Technologies<br>Development Centre     |
| Konstantine Karosanidze | DNS / IDNA /<br>Unicode Expert | Georgian National<br>Communications Commission     |
| Mzia Gogilashvili       | Policy Expert                  | Bank of Georgia                                    |
| David Birman            | Community<br>Representative    | Information Technologies<br>Development Centre     |
| Malcolm Taylor          | XML Code<br>Developer          | Information Technologies<br>Development Centre     |
| Marine Beridze          | Linguistic Expert              | Arnold Chikobava State Institute<br>of Linguistics |
| la Feradze              | Registry Expert                | Caucasus Online                                    |



### LGR Development Process and Methodology

- Define 33 from 37 code points from MSR Version 2, needed for the root zone
- No confusing code points or label constraints in the use of the selected code points, as there is no confusing similarities with any other scripts
- ICANN assisted Georgian Script Generation Panel in the development of the XML file, based on the solution proposed by the GP. The GP reviewed and finalized it



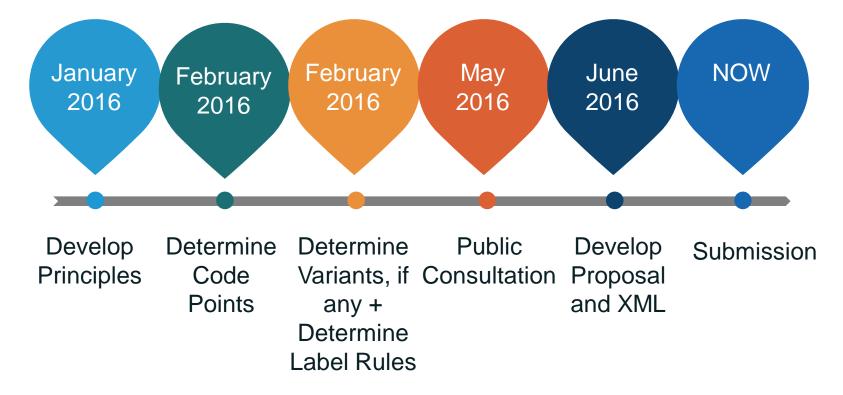
### **Challenges Behind the Process**

| ა                   | გ                  | გ                   | Q                   | 9                    | 3                    | ъ                     | က                     | 0                          |
|---------------------|--------------------|---------------------|---------------------|----------------------|----------------------|-----------------------|-----------------------|----------------------------|
| ანი<br>ani<br>1     | ბანი<br>bani<br>2  | განი<br>gani<br>3   | დონი<br>doni<br>4   | ენი<br>eni<br>5      | ვიწი<br>vini<br>6    | ზენი<br>zeni<br>7     | თანი<br>tani<br>9     | обо<br>ini<br>10           |
| 3                   | ლ                  | 9                   | б                   | ო                    | 3                    | ป                     | რ                     | Ն                          |
| კაწი<br>k'ani<br>20 | ლასი<br>lasi<br>30 | მაწი<br>mani<br>40  | წარი<br>nari<br>50  | ოწი<br>oni<br>70     | პარი<br>p'ari<br>80  | ჟაწი<br>zhani<br>90   | რაე<br>rae<br>100     | სანი<br>sani<br>200        |
| ð                   | უ                  | ფ                   | ց                   | ღ                    | g                    | в                     | в                     | в                          |
|                     | $\mathbf{U}$       | 0                   | 0                   |                      | 0                    |                       |                       | 0                          |
| ტარი<br>ťari<br>300 | უნი<br>uni         | ფარი<br>pari<br>500 | ეანი<br>kani<br>600 | ღაწი<br>ghani<br>700 | ტარი<br>q'ari<br>800 | ටීoඊo<br>shini<br>900 | ჩიწი<br>chini<br>1000 | ი<br>ცანი<br>tsani<br>2000 |
| ťari                | •                  | ფარი<br>pari        | ქანი<br>kani        | ღაწი<br>ghani        | ყარი<br>q'ari        | shini                 | chini                 | ცანი<br>tsani              |

Due to the unique characters of Mkhedruli, the panel unanimously adopted the decision that there are no variants or Whole Label Evaluation (WLE) rules in the use of the selected code points, as there are no cross-script variants with any other scripts.



### **Current Progress and Timeline**



#### **To Summarize**

Collect public feedback, summarize and make final corrections for submission



# Thank You

### ⊙ Sophio Elizbarashvili

- o sopho@itdc.ge
- +995555230500
- ITDC Information Technologies Development Centre



# Update by the Neo-Brahmi GP

Akshat Joshi Neo-Brahmi GP



- ⊙ Introduction to scripts languages being covered
- ⊙ Introduction to GP members
- Outreach efforts
- ⊙ Current progress
- ⊙ Challenges in developing Neo-brahmi LGR
- $\odot$  Timelines for completion



| Script     | Language                        | Devanāgarī  | Nepali – नेपाली                    |
|------------|---------------------------------|-------------|------------------------------------|
| Bengali    | Assamese – অসমীয়া              | Devanāgarī  | Sanskrit – संस्कृतम्, संस्कृतावाक् |
| Bengali    | Bengali – বাংলা                 |             |                                    |
| Bengali    | Manipuri – মনিপুরি              | Devanāgarī  | Santali/Santhali – संथाली          |
| Devanāgarī | Bodo – बडो                      | Devanāgarī  | Sindhi – सिंधी                     |
|            |                                 | Gujarati    | Gujarati – ગુજરાતી                 |
| Devanāgarī | Dogri – डोगरी                   | Gurumukhi   |                                    |
| Devanāgarī | Hindi – हिन्दी                  |             | Punjabi – ਪੰਜਾਬੀ                   |
| Devanāgarī |                                 | Kannada     | Kannada – ಕನ್ನಡ                    |
|            | Kashmiri – कॉशुर, Kāšur, Koshur | Malayalam   | Malayalam – മലയാളം                 |
| Devanāgarī | Konkani – कोंकणी,               | Oriya(Odia) | Odia – ଓଡ଼ିଆ                       |
| Devanāgarī | Maithili – मैथिली, भৈथिली       | Tamil       | Tamil – தமிழ், Tamizh              |
| Devanāgarī | Marathi – मराठी                 | Telugu      | Telugu – తెలుగు                    |



### **GP** Members

- (Chair) Udaya Narayana Singh -Bangla, Maithili, Hindi, English
- Anupam Agrawal Hindi, Bangla
- Akshat Joshi Hindi, Marathi
- Abhijit Dutta Bengali, Hindi
- Mahesh Kulkarni Marathi, Hindi
- Neha Gupta Hindi
- Nishit Jain Hindi
- Prabhakar Pandey Hindi
- Raiomond Doctor English, Hindi, Marathi, Gujarati
- ⊙ N. DeivaSundaram Tamil
- Shantaram Walawalikar Konkani
- ⊙ Bal Krishna Bal Nepali
- Ganesh Murmu Santali
- Balaram Prasain Nepali

- Rajib Chakraborty Bangla
- Gurpreet Singh Lehal Panjabi
- ⊙ Saroja Bhate Sanskrit
- ⊙ Shambhu Kumar Singh Maithili
- ⊙ SwarnaPrabha Chainary Bodo
- Ghanashyam Nepal Nepali
- Kalyan Vasudeo Kale Marathi
- Shashi Pathania Dogri
- Santhosh Thottingal Malayalam, Sourashtra, Tamil
- Uma Maheshwar G Telugu
- ⊙ Girish Chandra Mishra Odia
- ⊙ K. C. Tikayat ray Odia
- DEBAJIT SHARMA Assamese
- BASANTA KUMAR PANDA Odia
- Arvind Bhandari Gujarati
- ⊙ Harish Chowdhary Hindi



### Outreach efforts

- Conducted a workshop in AprIGF-2014 for awareness and call for participation in LGR procedure
  - Topic: "Bringing diverse linguistic communities together for a unified IDN ruleset"
  - The panel discussion touched upon the various aspects of creation of the LGR for the Neo-Brahmi scripts
  - <u>http://2014.rigf.asia/agenda/workshop-</u> proposals/workshop-proposal-13/
- Participation and presentation in 49th ICANN Public Meeting at Singapore
- Participation and presentation in 50th ICANN Public Meeting at London



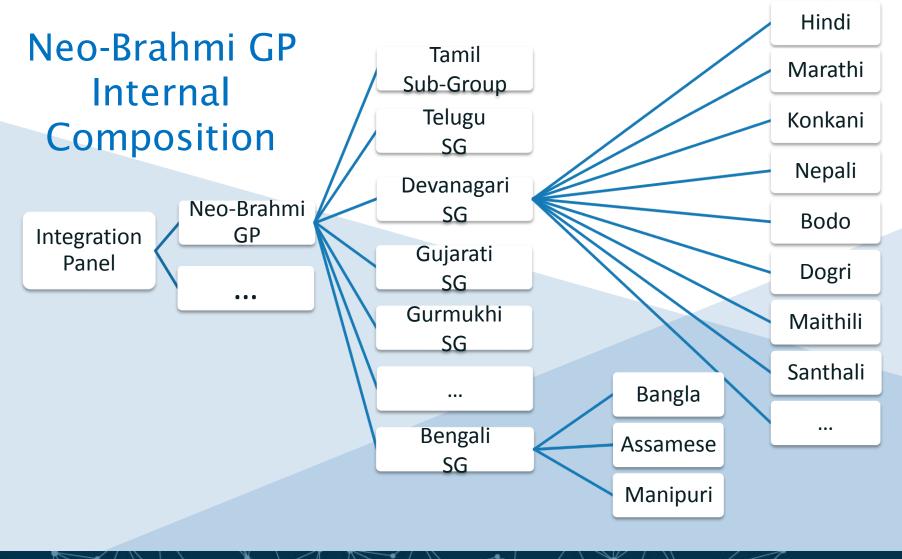
### **Current Progress**

- First face-to-face meeting of the GP was held in July 2015 at Pune
- Following were the major decisions taken during the first F2F:
  - Inviting the identified additional experts
  - Revisiting the timelines due to need for additional experts
  - Meeting over fortnightly/monthly call
  - Second face-to-face meeting for new members and remote participation via Adobe Connect for the existing members
- Code-point repertoires are in the process of being finalized
  - They are almost final for the following languages:
    - Hindi, Marathi, Bodo, Dogri, Konkani, Maithili, Nepali, Sanskrit, Santali, Sindhi, Bangla, Assamese and Panjabi

⊙ The Whole Label Evaluation rules are under construction



### Challenges in Developing Neo-Brahmi LGR





### Challenges in Developing Neo-Brahmi LGR



### **Diverse and large** 21 languages written in 9 different scripts



#### New set of conditions than existing policy

The existing exhaustive policy for Indian language domain names under ".in" framework needs revisiting to a considerable extent

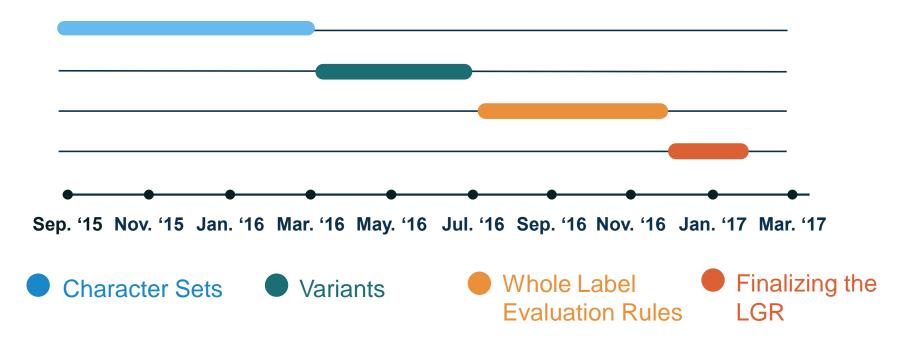


### **Core team challenges**

Due to some inevitable changes in commitments of core team of volunteers managing the activity, work is slowly progressing



### **Timeline for Completion**



#### Timeline proposed at the end of first F2F meeting in Pune:

As the decision of addition of new members was taken during the first F2F in Pune, a new set of timelines was proposed.







# Update by the Thai GP

Wanawit Ahkuputra Thai GP Chair



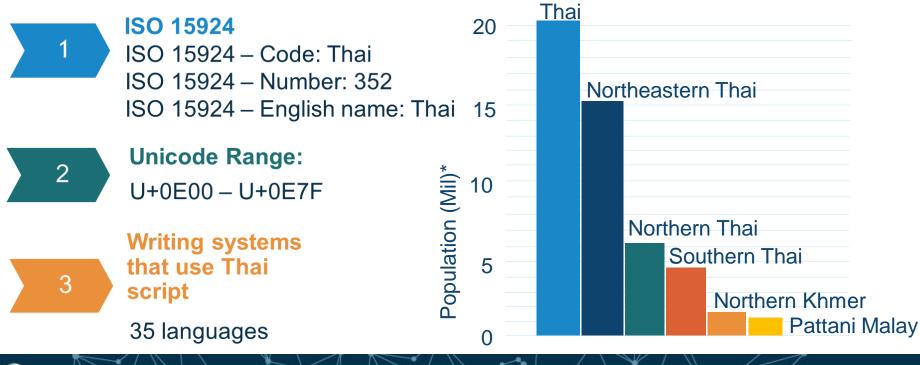


- $\odot$  Background on script and principal languages using it
- $\odot$  Code point repertoire
- ⊙ Variants
- ⊙ Cross script homoglyphs
- $\odot$  Whole Label Evaluation (WLE) rules



# **Background on Script and Principal Languages**

- Thai script is an abugida script, written left-to-right, without spaces between words
- ⊙ No notion of uppercase and lowercase characters
- ⊙ Some vowels are written before and after the main consonant
- Certain vowels, tone marks, and diacritics are written above and below the main character
   Selected Languages written in Thai script



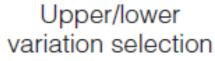


# Background on Script and Principal Languages

### $\odot$ The unique characteristic of script

- Allow both upper and lower position, The different input order of upper and lower marks
- Stored codes are different
- ⊙ Problematic issue
  - Fail string matching
  - Confuse the output method

### $\odot$ Solve this problem





- The WTT 2.0 specification defines the canonical order of Thai character strings. WTT 2.0 compliance requires that certain inputsequence rules must be met, and most (if not all) input syntactic errors are eliminated at the time of data entry
- The WTT2.0 later become the national standard "TIS 1566 Thai Input/ Output Methods for Computers"



### **Code Point Repertoire**

- The Thai GP takes code points shortlisted in MSR-2 as a starting point for Thai Script analysis for Root Zone Label Generation Rules
- The Thai GP refers to Thai script writing system from Royal Institute of Thailand and refers to various standards such as:
  - $\circ~$  TIS 620 series Standard for Thai Character Codes for Computers
  - TIS 820 series Layout of Thai Character Keys on Computer Keyboard
  - TIS 1566 Thai Input/ Output Methods for Computers



### **Code Point Repertoire**

⊙ Starting from the 71 code points in MSR-2 for the Thai script

- 3 code points to be excluded,
- 2 additional combined code points will be included
- In total, the repertoire includes 70 code points

| # | Unicode Code Point | Glyph | Unicode Code Point Name    |
|---|--------------------|-------|----------------------------|
| 1 | 0E45               | ſ     | THAI CHARACTER LAKKHANGYAO |
| 2 | 0E46               | ണ്    | THAI CHARACTER MAIYAMOK    |
| 3 | 0E4E               | ँ     | THAI CHARACTER YAMAKKAN    |

Code point repertoire excluded

#### Code point repertoire included 2 additional code points

| #  | Unicode Code Point | Glyph | Unicode Code Point Name                           |
|----|--------------------|-------|---|
| 69 | 0E24 + 0E45        | ฤๅ    | THAI CHARACTER RU +<br>THAI CHARACTER LAKKHANGYAO |
| 70 | 0E26 + 0E45        | ฦๅ    | THAI CHARACTER LU THAI +<br>CHARACTER LAKKHANGYAO |



### Variants

- Point of consideration:
- 1. U+0E40 (THAI CHARACTER SARA E, l) and U+0E40 (THAI CHARACTER SARA AE, ll)  $\rightarrow$  Handle by WLE rule
- 2. U+0E32 (THAI CHARACTER SARA AA,  $\mathbf{1}$ ) and U+0E45 (THAI CHARACTER LAKKHANGYAO,  $\mathbf{1}$ )  $\rightarrow$  Handle by additional code points
- U+0E33 (THAI CHARACTER SARA AM, <sup>Č</sup>**1**)
   excluded from IDNA2008 → Out of scope of this proposal
- ⊙ In conclusion, there is no blocked variant proposed within Thai script



## Whole Label Evaluation (WLE) Rules

- 1. No leading combining mark (default WLE)
- 2. Every leading vowel must precede a consonant
  - A leading-vowel must be followed by a consonant
- 3. Code points which must follow a consonant
  - A below-vowel, an above-vowel, a below diacritic and a U+0E47(MAITAIKHU) must be after a consonant
- 4. Context of U+0E31 (MAI HAN-AKAT)
  - MAI HAN-AKAT must be between a consonant and either tone or consonant
- 5. Context of U+0E30 (SARA-A)
  - A following-vowel-SARA-A cannot be after following-vowel-SARA-A



# Whole Label Evaluation (WLE) Rules

- 6. Context of U+0E32 (SARA-AA)
  - A following-vowel-SARA-AA cannot be after following-vowel-SARA-AA
- 7. Context of tone mark
  - A tone-mark cannot be after (a below-diacritic or an above-diacritic or a tone-mark or start)
- 8. Context of diacritic
  - An above-diacritic cannot be after another above-diacritic
  - An above-diacritic-MAITAIKHU must be after a consonant
  - An above-diacritic-NIKHAHIT can follow tone or consonant



# Engage with ICANN and IDN Program



### **Thank You and Questions**

Reach us at: IDNProgram@icann.org Website: icann.org/idn



twitter.com/icann







linkedin.com/company/icann



youtube.com/user/icannnews



#### gplus.to/icann



weibo.com/ICANNorg



flickr.com/photos/icann



slideshare.net/icannpresentations

