Introduction to ICANN
The I* Organizations

- Internet Corporation for Assigned Names and Numbers (ICANN)
- Internet Society (ISOC)
- Internet Engineering Task Force (IETF)
- Internet Architecture Board (IAB)
- Regional Internet Registries (RIRs)
- Regional TLD Organizations (RTLDOs)
- Internet Governance Forum (IGF)
ICANN develops and implements policies for Internet Names and Numbers in a bottom-up, consensus-driven, multi-stakeholder model

- Anyone interested is invited to be part of this process

- Website at http://icann.org

- Online Learning Platform at http://learn.icann.org
ICANN Offices

- Main Location: Los Angeles, California
- Hub Offices
  - Singapore
  - Istanbul
- Other Offices and Engagement Centers:
  - Beijing
  - Brussels
  - Geneva
  - Montevideo
  - Seoul
  - Washington DC
IDN Program at ICANN
Internationalized Domain Name (IDN) Label

https://youtu.be/wnauGpYh96c
ASCII Domain Name Label

www.cafe.com

Top Level Domains (TLDs)
- Country Code TLDs (ccTLDs)
  - .sg, .cn, .kh, .la, .mm, .th, .ca, ...
  - Two letter [a..z] codes, reserved for countries and territories by ISO 3166 standard
- Generic TLDs (gTLDs)
  - .com, .org, .net, .edu, … - organizations
  - New gTLDs – 1930 applications in 2012

Domain Stakeholders
- ICANN
- Registry
- Registrar
- Reseller
- Registrant
- End-User
ASCII Domain Name Label

www.cafe.com

Third Level Domain  Second Level Domain  Top Level Domain (TLD)

Forming ASCII Labels
Use LDH
• Letters [a-z]
• Digits [0-9]
• Hyphen (LDH)
Label length = 63
Other constraints (e.g. on hyphen)

Forming ASCII Labels
Use only Letters
• Letters [a-z]
Label length = 63
Internationalized Domain Name (IDN) Labels

Syntax of IDN Labels

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

Valid A-Label

Syntax of IDN Labels

Valid U-Label: Unicode code points as constrained by IDNA2008

Valid A-Label - “xn--” followed by punycode of U-Label of length 59
Overview of IDN Program at ICANN

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

Report and documentation of all completed projects available at:
https://www.icann.org/resources/pages/reports-2013-04-03-en

Community agreed to define a Label Generation Rules (LGR)

Projects:
P1 LGR XML Specification
P2.1 LGR Process for the Root Zone
P6 User Experience Study for TLD Variants

Projects:
P1 LGR Specs. and Toolset
P2.2 LGR Development
P7 LGR Implementation
Label Generation Rules for the Root Zone

- For the Root Zone, single “table” containing data for all scripts
  - As it is a shared resource, must be conservative
  - Must be stable and secure
  - Must be based on inclusion based analysis

- For each script or writing system:
  - Which code points are valid for use?
  - Are any of these code points variants of each other?
  - Are the any additional constraints on the labels?
LGR for the Root Zone

Unicode

<table>
<thead>
<tr>
<th>Code</th>
<th>Character</th>
</tr>
</thead>
<tbody>
<tr>
<td>0000</td>
<td>@</td>
</tr>
<tr>
<td>0001</td>
<td>P</td>
</tr>
<tr>
<td>0002</td>
<td>`</td>
</tr>
<tr>
<td>0003</td>
<td>p</td>
</tr>
<tr>
<td>0004</td>
<td>!</td>
</tr>
<tr>
<td>0005</td>
<td>1</td>
</tr>
<tr>
<td>0006</td>
<td>A</td>
</tr>
<tr>
<td>0007</td>
<td>Q</td>
</tr>
<tr>
<td>0008</td>
<td>a</td>
</tr>
<tr>
<td>0009</td>
<td>q</td>
</tr>
<tr>
<td>0010</td>
<td>#</td>
</tr>
<tr>
<td>0011</td>
<td>3</td>
</tr>
<tr>
<td>0012</td>
<td>C</td>
</tr>
<tr>
<td>0013</td>
<td>S</td>
</tr>
<tr>
<td>0014</td>
<td>s</td>
</tr>
<tr>
<td>0015</td>
<td>$</td>
</tr>
<tr>
<td>0016</td>
<td>4</td>
</tr>
<tr>
<td>0017</td>
<td>D</td>
</tr>
<tr>
<td>0018</td>
<td>T</td>
</tr>
<tr>
<td>0019</td>
<td>d</td>
</tr>
<tr>
<td>0020</td>
<td>&amp;</td>
</tr>
<tr>
<td>0021</td>
<td>6</td>
</tr>
<tr>
<td>0022</td>
<td>F</td>
</tr>
<tr>
<td>0023</td>
<td>v</td>
</tr>
<tr>
<td>0024</td>
<td>(</td>
</tr>
<tr>
<td>0025</td>
<td>H</td>
</tr>
<tr>
<td>0026</td>
<td>X</td>
</tr>
<tr>
<td>0027</td>
<td>x</td>
</tr>
<tr>
<td>0028</td>
<td>)</td>
</tr>
<tr>
<td>0029</td>
<td>I</td>
</tr>
<tr>
<td>0030</td>
<td>Y</td>
</tr>
<tr>
<td>0031</td>
<td>i</td>
</tr>
<tr>
<td>0032</td>
<td>y</td>
</tr>
<tr>
<td>0033</td>
<td>:</td>
</tr>
<tr>
<td>0034</td>
<td>J</td>
</tr>
<tr>
<td>0035</td>
<td>Z</td>
</tr>
<tr>
<td>0036</td>
<td>j</td>
</tr>
<tr>
<td>0037</td>
<td>+</td>
</tr>
<tr>
<td>0038</td>
<td>K</td>
</tr>
<tr>
<td>0039</td>
<td>l</td>
</tr>
<tr>
<td>0040</td>
<td>,</td>
</tr>
<tr>
<td>0041</td>
<td>L</td>
</tr>
<tr>
<td>0042</td>
<td>\</td>
</tr>
<tr>
<td>0043</td>
<td>I</td>
</tr>
<tr>
<td>0044</td>
<td>=</td>
</tr>
<tr>
<td>0045</td>
<td>M</td>
</tr>
<tr>
<td>0046</td>
<td>]</td>
</tr>
<tr>
<td>0047</td>
<td>m</td>
</tr>
<tr>
<td>0048</td>
<td>.</td>
</tr>
<tr>
<td>0049</td>
<td>N</td>
</tr>
<tr>
<td>0050</td>
<td>^</td>
</tr>
<tr>
<td>0051</td>
<td>n</td>
</tr>
<tr>
<td>0052</td>
<td>~</td>
</tr>
<tr>
<td>0053</td>
<td></td>
</tr>
<tr>
<td>0054</td>
<td>/</td>
</tr>
<tr>
<td>0055</td>
<td>?</td>
</tr>
<tr>
<td>0056</td>
<td>O</td>
</tr>
<tr>
<td>0057</td>
<td>o</td>
</tr>
</tbody>
</table>

---

...
LGR for the Root Zone

Unicode

IDNA2008 – by IETF
LGR for the Root Zone

Unicode

IDNA2008

Maximal Starting Repertoire – by Integration Panel of ICANN
LGR for the Root Zone

LGR Proposal: Select code points allowed for top-level labels – by Generation Panel of Script Community

<table>
<thead>
<tr>
<th>Unicode</th>
<th>IDNA2008</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Maximal Starting Repertoire (MSR)</td>
</tr>
</tbody>
</table>

- X
- X
- X
- X
- X
- X
- X
- X
- X
- X
Label Generation Rules (LGR)

- Valid code points
- Variants code points

Label constraints
- Cannot mix ک and گ in a label

Example:
- کلاکت
  - ✅
- کلاکت
  - ✅
- کلاکت
  - ✗
- کلاکت
  - ✗
Cross-Script Variants – Example from Armenian

Armenian Letter
- U+0572 Armenian small letter GHAD
- U+0582 Armenian small letter YIWN
- U+0585 Armenian small letter OH

Greek Letter
- U+03B7 Greek small letter ETA
- U+03B9 Greek small letter IOTA
- U+03BF Greek small letter OMICRON

Armenian Letter
- g U+0581 Armenian small letter CO
- ղ U+0570 Armenian small letter HO
- h U+0578 Armenian small letter VO
- օ U+0585 Armenian small letter OH
- ք U+0566 Armenian small letter ZA
- ւ U+057D Armenian small letter SEH
- տ U+0582 Armenian small letter YIWN

Latin Letter
- g U+0067 Latin small letter G
- h U+0068 Latin small letter H
- n U+006E Latin small letter N
- o U+006F Latin small letter O
- q U+0071 Latin small letter Q
- u U+0075 Latin small letter U
- ɩ U+0269 Latin small letter IOTA
1. Which code points from Greek script must be included in the Root Zone
   - Are exclusions from MSR (pink) correct?
   - What must be included in LGR?
     - “everyday, general purpose [use ...] in a stable and widespread manner”

2. Are there any variant code points for Greek
   - Two code points form confusingly similar labels
     - Within-script for Greek
     - Cross-script between Greek and other scripts

3. Are there any label-level constraints in Greek
   - Well-formedness of a cluster?
   - Constraints on initial or final position in a label?
   - Other?
Label Generation Rulesets (LGRs) used to generate domain name labels, as specified in RFC 7940.

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 https://lgrtool.icann.org/

Open source package(s) release with BSD license
- Released at github: lgr-core, lgr-django, munidata

User guide available for further details
LGR Specification and Toolset (beta)

Domain or label to test

آزمايشي

Code Point Rules
Variant Rules
WLE Rules

LGR Tool

Machine-readable Label Generation Ruleset

System returns a list of variants based on the rules
Root Zone LGR Procedure

Generation Panels
- Generate proposals for script specific LGRs, based on community expertise and requirements

Integration Panel
- Integrates them into common Root Zone LGR while minimizing the risk to Root Zone as shared resource

Label Generation Rules (LGR)
- Which labels are permissible
- Which variant labels exist
- Which variant labels may be allocated
Root Zone LGR Procedure

- One Generation Panel per writing system (script)

Generation Panel 1
Generation Panel 2
Generation Panel 3

Integration Panel
Rules
Panel 1
Panel 2
Etcetera

Needs more work
http://.....TLD

Etcetera
Status of Root Zone LGR Development

Label Generation Rules (LGR)

28+ Scripts
19+ Generation Panels

Generation Panel Status

Finalizing

Seated

Active

Forming

Arabic
Armenian
Chinese
Cyrillic
Ethiopic
Georgian
Greek
Japanese
Khmer
Korean
Lao
Latin
Myanmar
Neo-Brahmi
Thai

Maximal Starting Repertoire (MSR)

Other:
- Hebrew
- Sinhala
- Thaana
Get Involved!
Speak up for your language
Engage with ICANN

Thank You and Questions
Email: IDNProgram@icann.org
Website: icann.org/IDN

twitter.com/icann
gplus.to/icann
facebook.com/icannorg
weibo.com/ICANNorg
linkedin.com/company/icann
flickr.com/photos/icann
youtube.com/user/icannnews
slideshare.net/icannpresentations