TLDs and Domain Names

Fahd A. Batayneh | pkSIG 2015 | October 7, 2015

Topics of Discussion

- Top Level Domains (TLDs)
- ccTLDs
- gTLDs
- New gTLDs
- DNS Root Servers
- Registry/Registrar/Registrant
- Dispute Resolution
- Business Opportunities in the Domain Name Industry

What is the DNS?

http://www.youtube.com/watch?v=72snZctFFtA

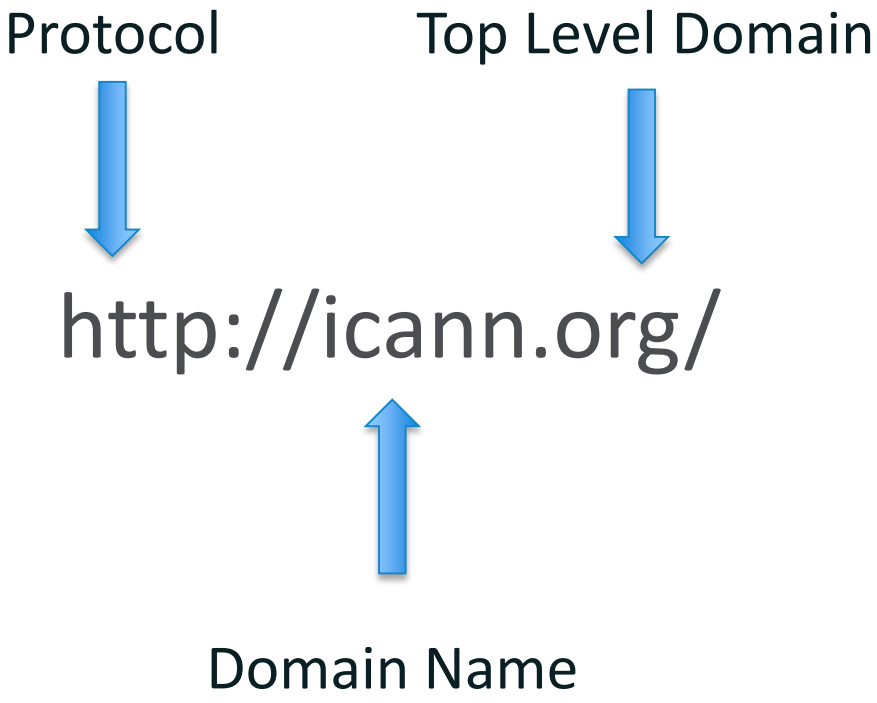
Top Level Domains (TLDs)

- Exists at the highest level of the DNS hierarchy
- Is the entry installed into the root-zone
- Consists of various groups:
  - ccTLDs (.pk, .kw, .jo, .eg, .tn)
  - gTLDs
    - Sponsored (.asia, .cat)
    - Unsponsored (.com, .net, .org)
    - New (.apps, .shop, .موقع, .شبكة)
  - IDN TLDs (.تونس, .بازار, .ابوظبي)

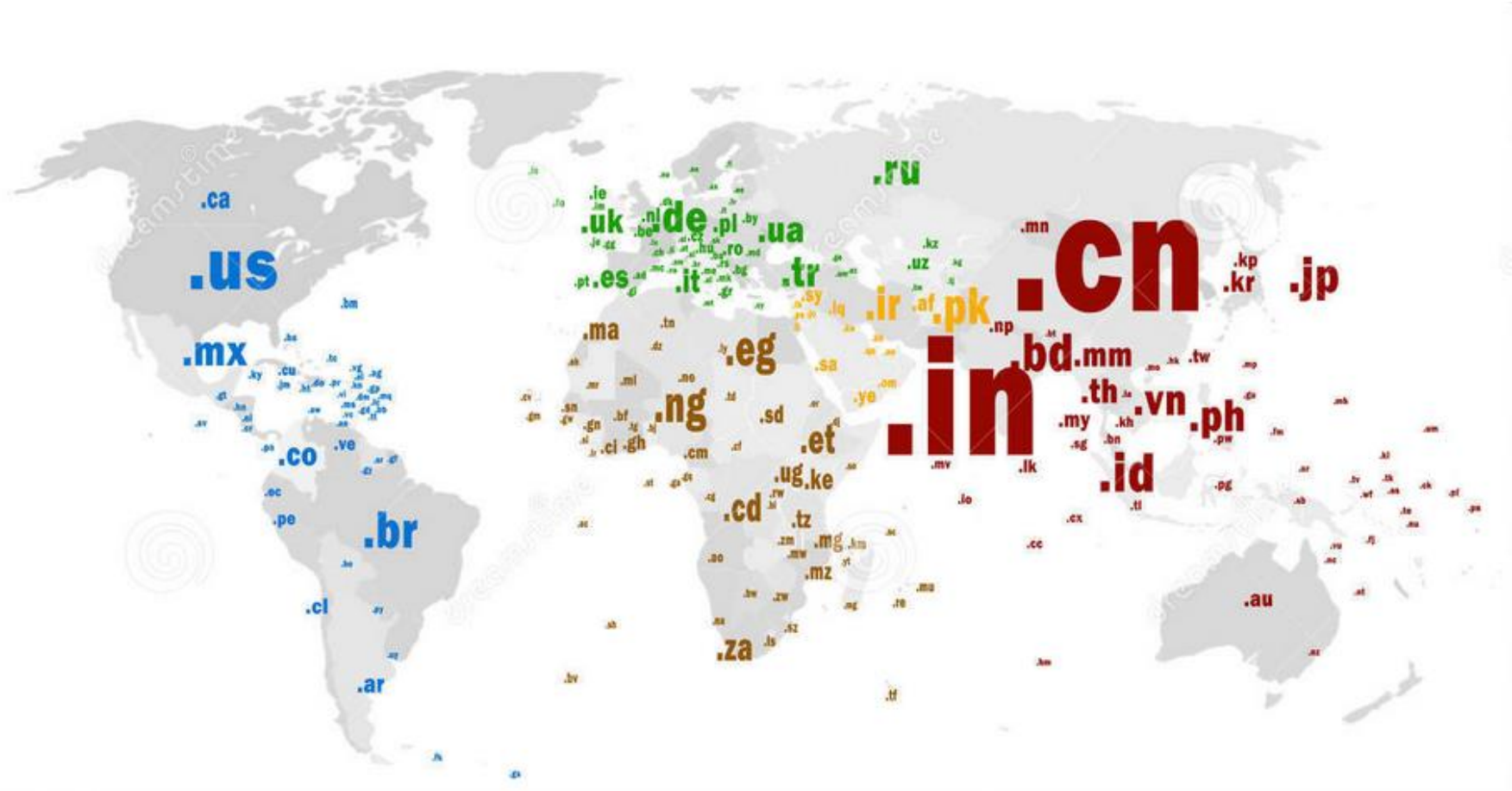
What is a Domain Name?

http://www.youtube.com/watch?v=2ZUQ2Szu-JI

Parts of a Domain Name?



Country Code TLDs (ccTLD)

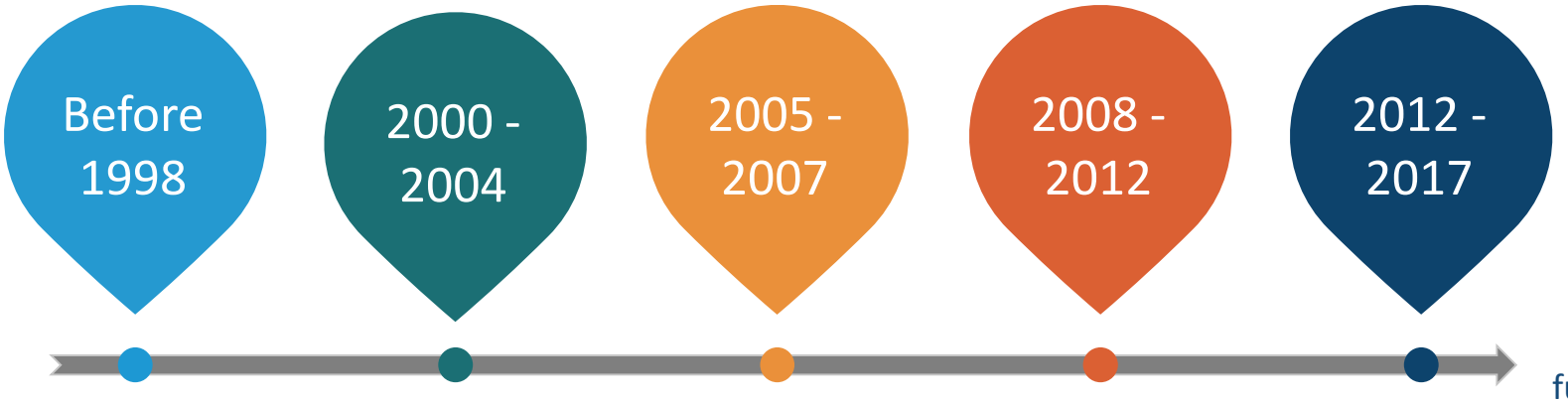


Source at http://dreamstime.com/

Where do ccTLDs get Their Code From?

- The ISO-3166 provides 2-letter codes and 3-letter codes of countries or a grouping of countries
- 249 assigned 2-letter codes
  - According to the UN, the world has 196 countries
- More at https://www.iso.org/obp/ui/#search

History of TLD



Before 1998	2000 - 2004	2005 - 2007	2008 - 2012	2012 - 2017
8 gTLDs pre-date ICANN's creation	2 previous rounds of new gTLDs	Policy Development Process	Policy Implementation (Design)	Policy Implementation (Deploy, Support)
.com .edu .gov .int .mil .net .org .arpa	(2000) .aero .biz .coop .info .museum .name .pro	(2004) .asia .cat .jobs .mobi .post .tel .xxx .travel	Introduction of new gTLDs: Principles, Recommendations and Implementation Guidelines	Development of the New gTLD Applicant Guidebook	Evaluation, Delegation and Launch of new gTLDs

The 2012 New gTLDs Program

- Considered the 3rd round of New TLDs
- More than 5 years of community collaboration to come out with the “TLD Applicant Guidebook”
- A one-time application fee of USD 185,000, and an annual fee of USD 25,000
- Allowed the application of non-Latin names
- 1,930 applications received for 1,420 unique strings
  - 116 non-Latin strings

… cont. (The 2012 New gTLDs Program)

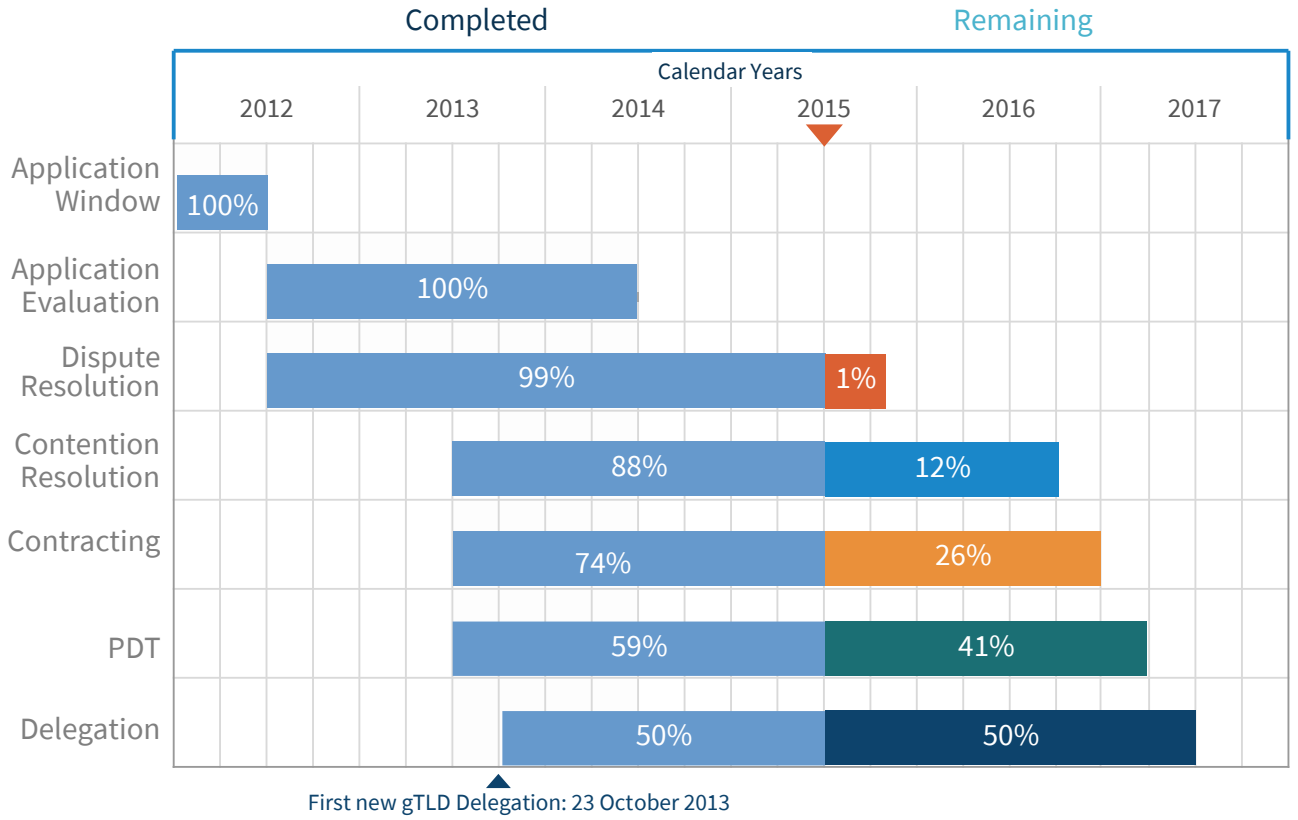
- A rigors application and evaluation process
  - To ensure that TLDs remain operational and that registrants do not loose their domain names due to the instability of a TLD
- The New gTLD microsite at http://newgtlds.icann.org/
- More statistics on the program can be found at http://newgtlds.icann.org/en/program-status/statistics

Understanding New gTLDs

New gTLDs Application Breakdown by Region

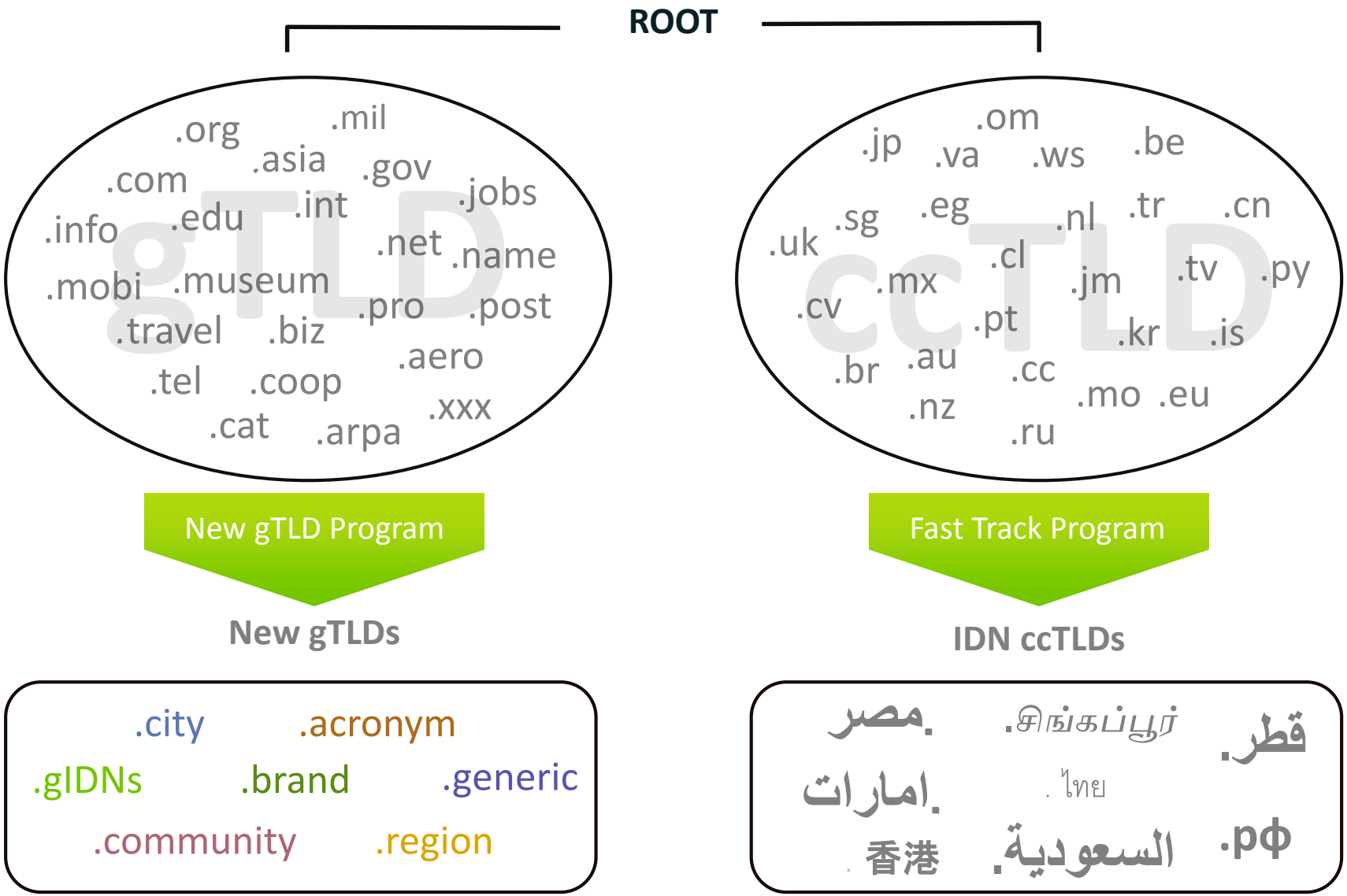
Where do Things Stand At?

- 1930 applications
- 1300+ potential TLDs delegated by 2017
- 750 new gTLDs delegated (as of 18 September 2015)
- More than 50% of the IDN gTLDs have been delegated (103 applications)



The New Landscape

How does the Name Space Look Nowadays?



Root Servers

- Publish the root zone file to other DNS servers and clients on the Internet
- The root zone file describes where the authoritative servers for the DNS TLDs are located
- The root name server operators publish the root zone file as received from the IANA

Root Server Operators

1. A - VeriSign Global Registry Services
2. B - University of Southern California - Information Sciences Institute
3. C - Cogent Communications
4. D - University of Maryland
5. E - NASA Ames Research Center
6. F - Internet Systems Consortium, Inc.
7. G - U.S. DOD Network Information Center
8. H - U.S. Army Research Lab
9. I - Autonomica/NORDUnet
10. J - VeriSign Global Registry Services
11. K - RIPE NCC
12. L – ICANN
13. M - WIDE Project

13 root-servers

12 Operators

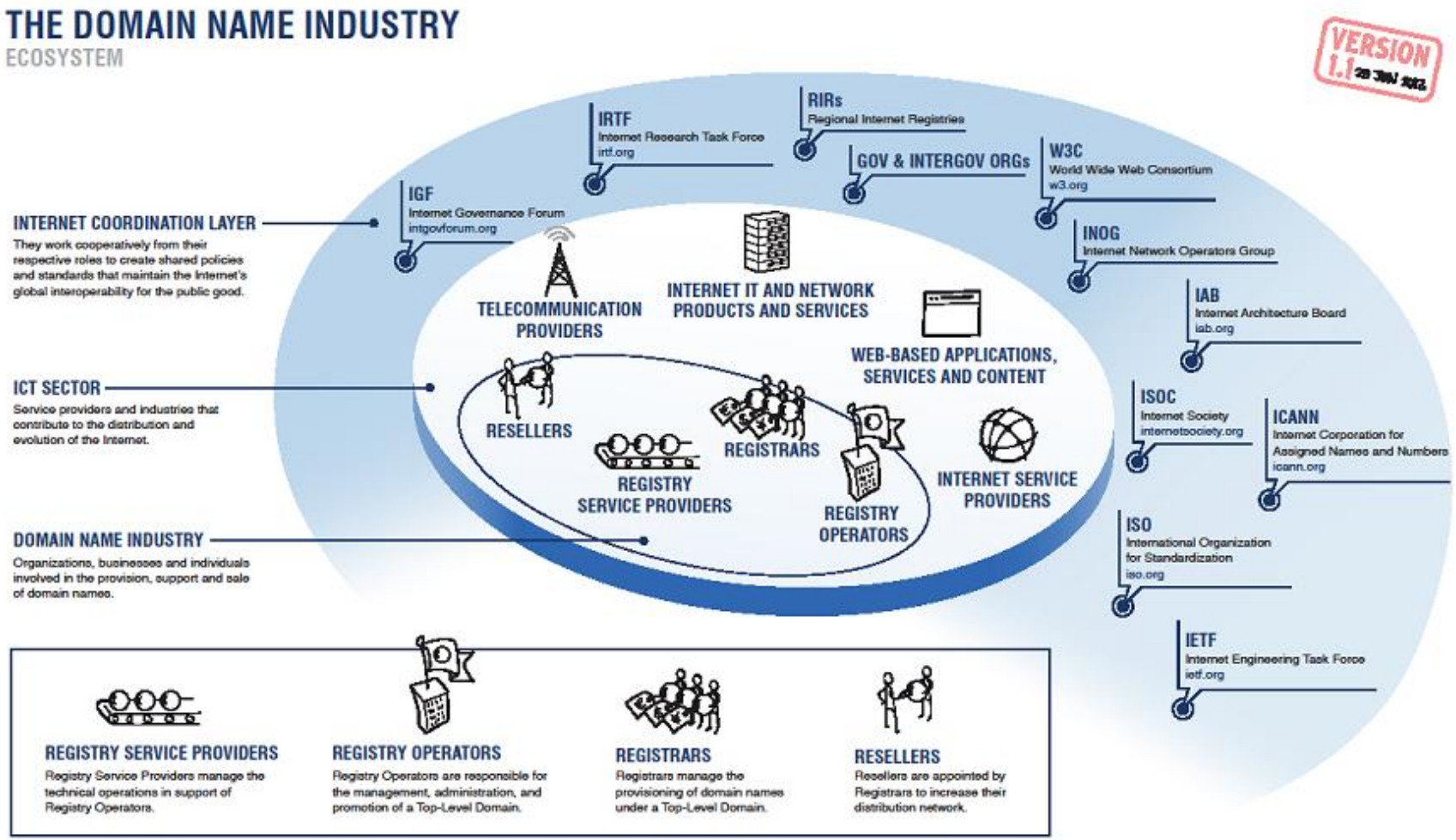
Root Servers Around the World (~480 instances)

Source at http://root-servers.org/

Registry/Registrar/Registrant

- **Registry** – The entity that operates a TLD
- **Registrar** – A middle-agent that sells domain names to registrants on behalf of a TLD Registry
  - ICANN Accredited Registrar vs. Resellers
  - ccTLDs vs. gTLDs
- **Registrant** – Entity registering a domain name

The DNS Industry Ecosystem



Dispute Resolution

- Cases could arise where a registrant registers a domain name that could infringe the naming rights of others. This is also called “Cybersquatting”
- When ICANN was formed in 1998, one of the first things it worked on was the “**Uniform Domain Name Dispute Resolution Process (UDRP)**”
  - 5 accredited providers for gTLDs to date
  - ccTLDs can customize this to fit their needs and local laws
  - An expensive and sometimes long process
- All 5 providers listed at http://bit.ly/1j6Rehz

New gTLDs and Dispute Resolution

- With the introduction of 100s of New gTLDs, cybersquatting is on the high
- Several protection and DRP mechanisms devised
  - Trademark Clearinghouse (TMCH)
  - Uniform Rapid Suspension (URS)
  - Post Delegation Dispute Resolution Procedure (PDDRP)
- UDRP would not be the best path for clear cut cases that need urgent attention
  - TMCH is good for initial protection, while URS comes in for a faster and cheaper path for resolving disputes

Trademark Clearinghouse (TMCH)

http://www.youtube.com/watch?v=ZXP2dieeZes

Business Opportunities in the DNS Industry

- Registry Services
  - 4 ccTLDs in the region have state-of-the-art registries. Many others are moving to a shared system
  - Having a Registry/Registrar model is the way forward
- New gTLDs
  - Providing more options in the TLD space away from the traditional .com/.net/.org and the ccTLDs, and providing multilingual TLDs
- Accredited Registrars
  - More than 1000 registrars
  - More at http://www.icann.org/en/resources/registrars/accreditation

… continue (Business Opportunities)

- Dispute Resolution Services
  - Five providers in total the last of which joined in 2013
- Online Intellectual Property Protection Services
  - Domain names are digital assets online
- Domain Name Aftermarket
  - insurance.com sold for USD 35.6 million in 2010
  - 360.com sold for USD 17 million in 2015
- Development of Domain Name Tools and Software

Domain Name Ecosystem in Pakistan

- ccTLD is .pk, and is run by PKNIC
- IDN ccTLD applied for is باکستان. in Urdu by the MoITT
- 1 ICANN Accredited Registrar
- 5 root-server instances (2 in Karachi | 2 in Islamabad | 1 in Lahore)
- Membership from MoITT and PTA in the GAC
- ISOC Islamabad Chapter is an ALS
- Many Pakistani stakeholders attend ICANN meetings; mainly through the ICANN fellowship program

Questions?!

Fahd Batayneh
Manager, Stakeholder Engagement, Middle East
E: fahd.batayneh@icann.org
W: http://icann.org/

twitter.com/icann
twitter.com/fahdbatayneh

gplus.to/icann

facebook.com/icannorg

weibo.com/ICANNorg

linkedin.com/company/icann

flickr.com/photos/icann

youtube.com/user/icannnews

slideshare.net/icannpresentations

Internationalized Domain Names (IDNs)

Sarmad Hussain | pkSIG 2015 | October 7, 2015

Internationalized Domain Names (IDNs)

http://www.youtube.com/watch?v=wnauGpYh96c

Overview of Presentation

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

ASCII Domain Name Label

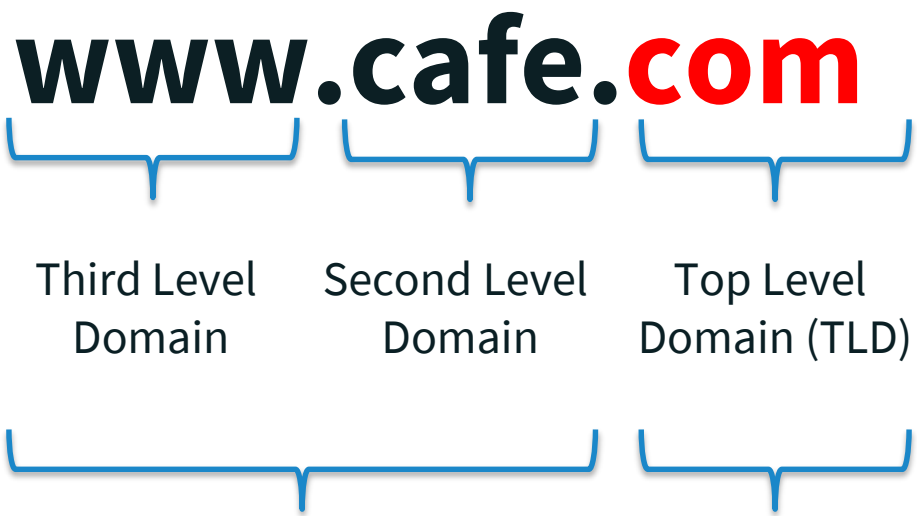
Top Level Domains (TLDs)

- 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



Forming ASCII Labels

Use LDH

- Letters [a-z]
- Digits [0-9]
- Hyphen (LDH)

Label length = 63

Other constraints (e.g. on hyphen)

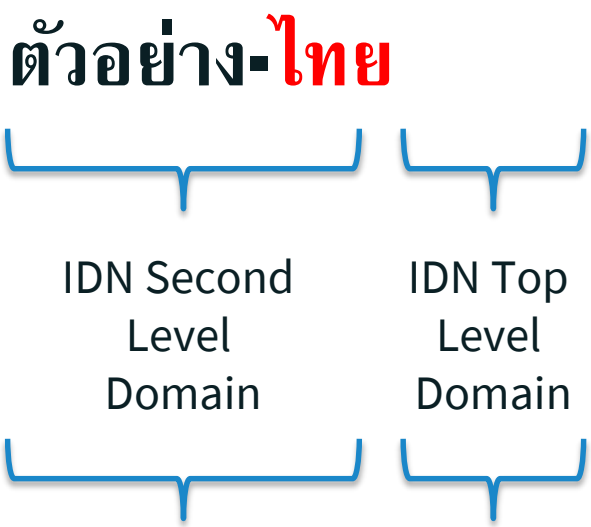
Forming ASCII Labels

Use only Letters

- Letters [a-z]

Label length = 63

Internationalized Domain Name (IDN) Labels



বাংলা
Бел
الجزائر
հայ
中国
ಭಾರತ್
한국
ලංකා

Syntax of IDN Labels

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

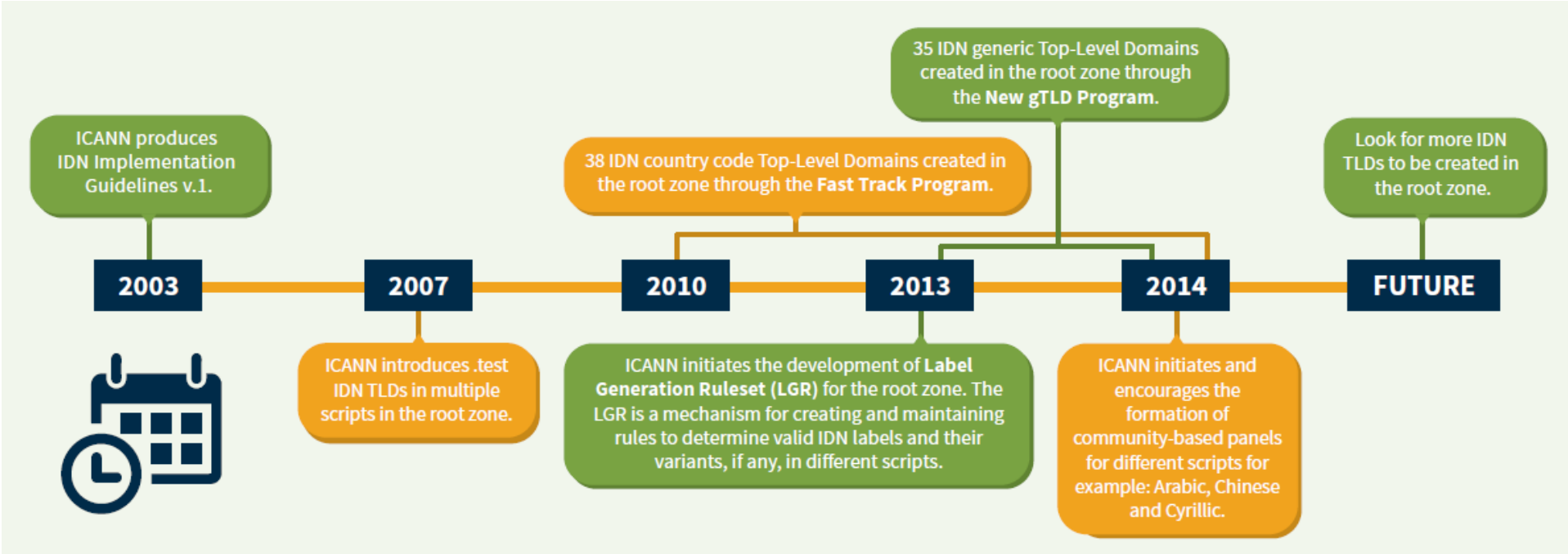
Valid A-Label - "xn--" followed by punycode of U-Label of length 59

Syntax of IDN Labels

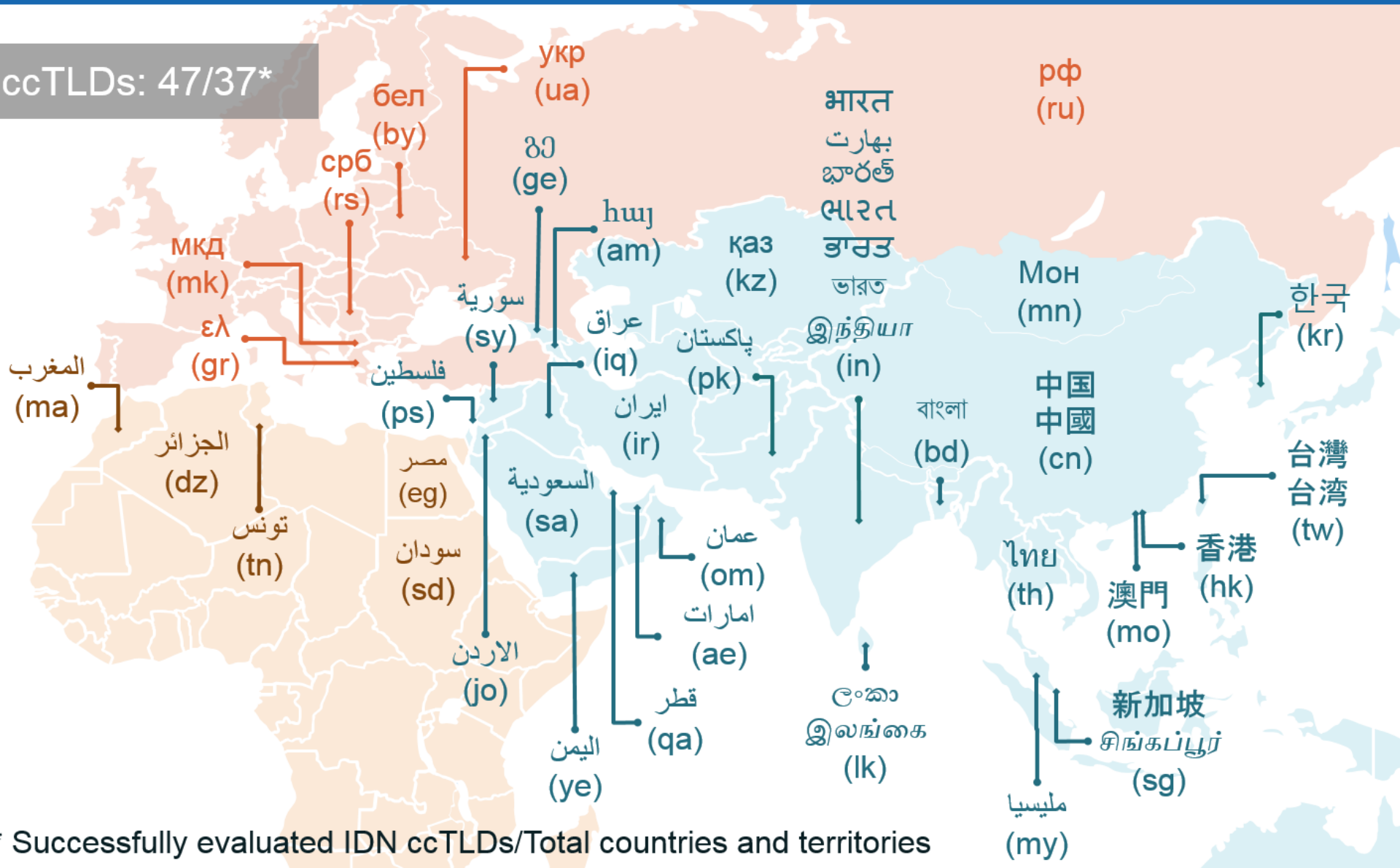
Valid U-Label, further constrained by the "letter" principle for TLDs

Valid A-Label

Timeline of ICANN's Progress on IDN TLDs



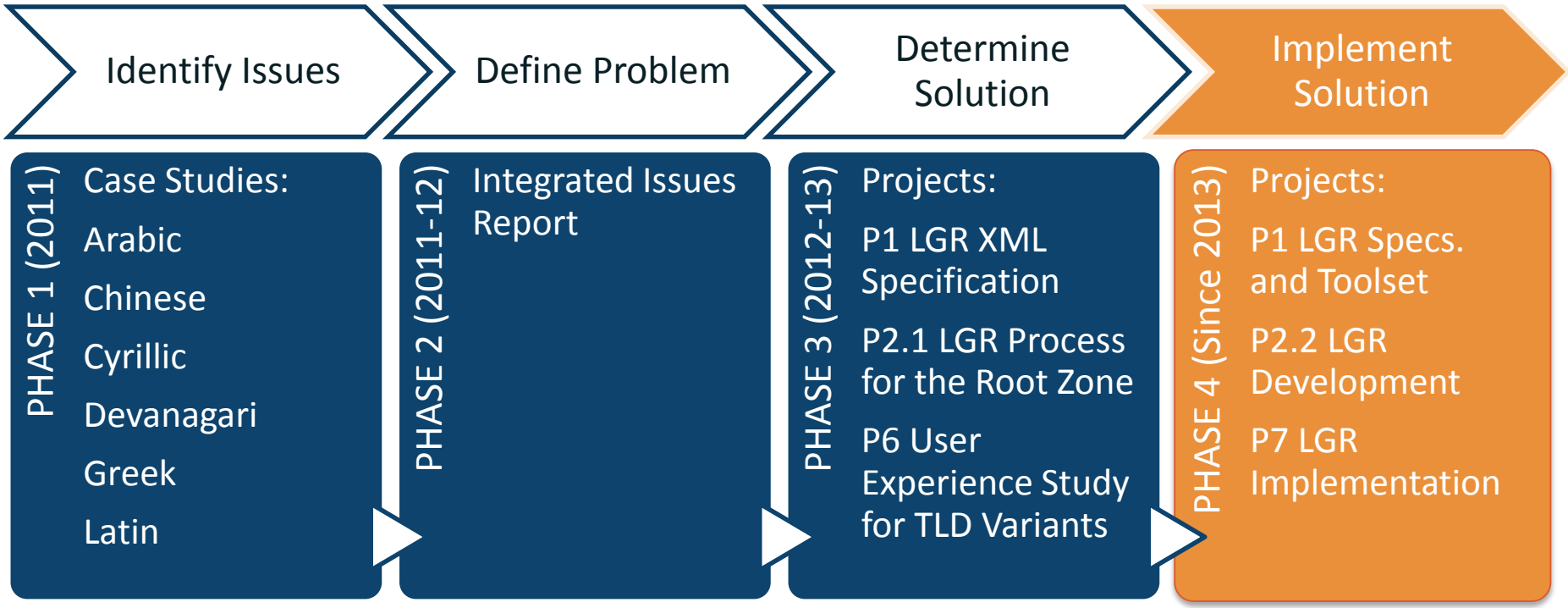
IDN Country Code Top-Level Domains



* Successfully evaluated IDN ccTLDs/Total countries and territories


IDN TLD Program

IDN TLD Program



Community agreed to define a Label Generation Rules (LGR)

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

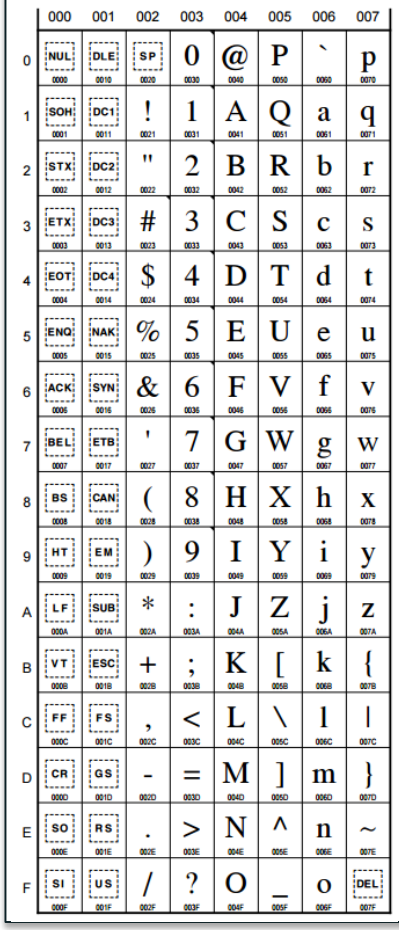
Label Generation Rules for the Root Zone

- 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



	000	001	002	003	004	005	006	007
0	NUL 0000	DLE 0010	SP 0020	0 0030	@ 0040	P 0050	` 0060	p 0070
1	SOH 0001	DC1 0011	! 0021	1 0031	A 0041	Q 0051	a 0061	q 0071
2	STX 0002	DC2 0012	" 0022	2 0032	B 0042	R 0052	b 0062	r 0072
3	ETX 0003	DC3 0013	# 0023	3 0033	C 0043	S 0053	c 0063	s 0073
4	EOT 0004	DC4 0014	$ 0024	4 0034	D 0044	T 0054	d 0064	t 0074
5	ENQ 0005	NAK 0015	% 0025	5 0035	E 0045	U 0055	e 0065	u 0075
6	ACK 0006	SYN 0016	& 0026	6 0036	F 0046	V 0056	f 0066	v 0076
7	BEL 0007	ETB 0017	' 0027	7 0037	G 0047	W 0057	g 0067	w 0077
8	BS 0008	CAN 0018	(0028	8 0038	H 0048	X 0058	h 0068	x 0078
9	HT 0009	EM 0019	) 0029	9 0039	I 0049	Y 0059	i 0069	y 0079
A	LF 000A	SUB 001A	* 002A	: 003A	J 004A	Z 005A	j 006A	z 007A
B	VT 000B	ESC 001B	+ 002B	; 003B	K 004B	[005B	k 006B	{ 007B
C	FF 000C	FS 001C	, 002C	< 003C	L 004C	\ 005C	l 006C	| 007C
D	CR 000D	GS 001D	- 002D	= 003D	M 004D	] 005D	m 006D	} 007D
E	SO 000E	RS 001E	. 002E	> 003E	N 004E	^ 005E	n 006E	~ 007E
F	SI 000F	US 001F	/ 002F	? 003F	O 004F	_ 005F	o 006F	DEL 007F

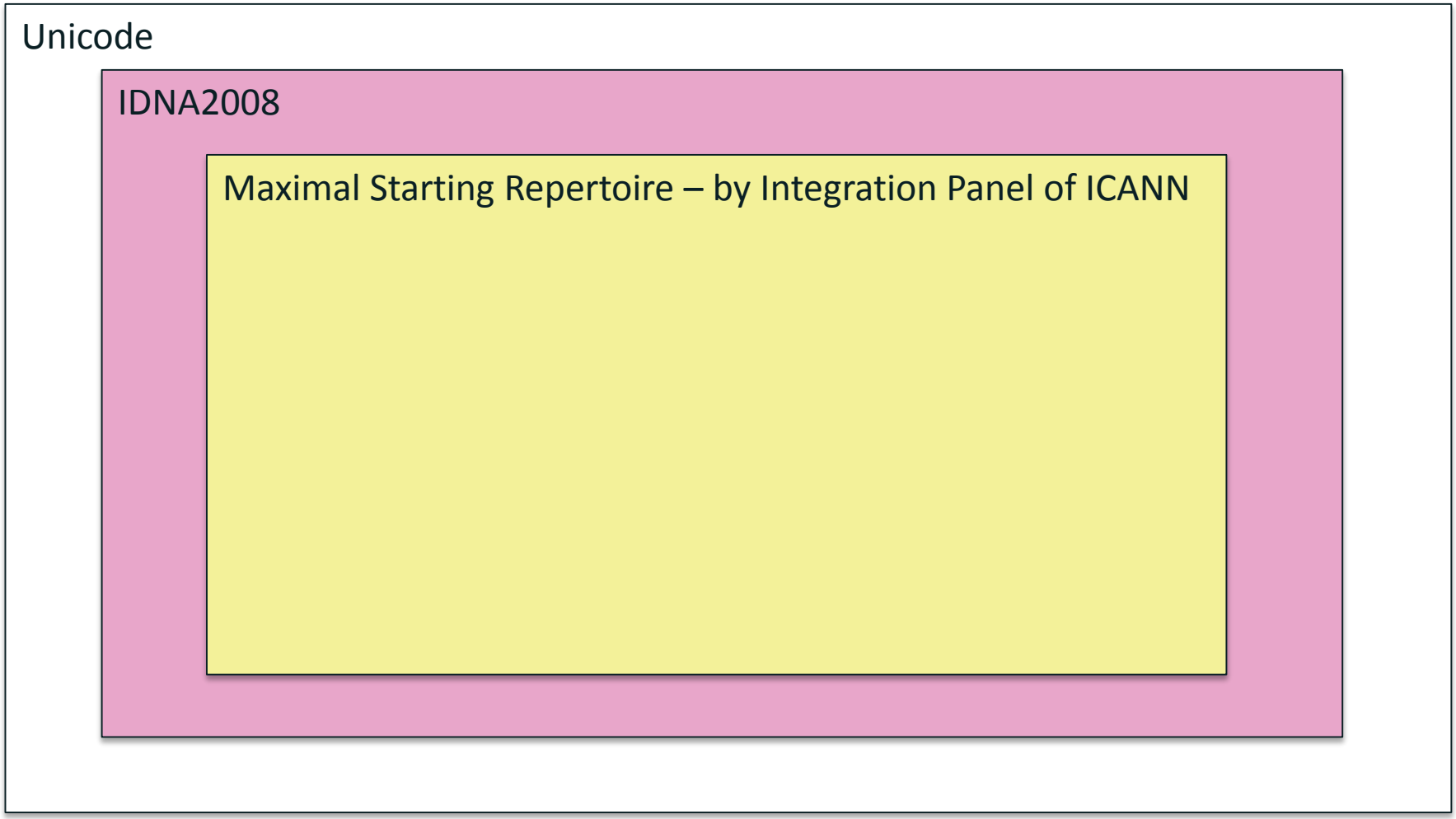
...

	0E8	0E9	0EA	0EB	0EC	0ED	0EE	0EF
0				ะ 0EB0	เ 0EC0	๐ 0ED0		
1	ก 0E81		ม 0EA1	ั 0EB1	แ 0EC1	๑ 0ED1		
2	ข 0E82		ย 0EA2	า 0EB2	โ 0EC2	๒ 0ED2		
3			ร 0EA3	ำ 0EB3	ใ 0EC3	๓ 0ED3		
4	ค 0E84	ด 0E94		ิ 0EB4	ไ 0EC4	๔ 0ED4		
5		ต 0E95	ล 0EA5	ี 0EB5		๕ 0ED5		
6		ถ 0E96		ึ 0EB6	ๆ 0EC6	๖ 0ED6		
7	ง 0E87	ท 0E97	ว 0EA7	ื 0EB7		๗ 0ED7		
8	จ 0E88			ุ 0EB8	◌ 0EC8	๘ 0ED8		
9		น 0E99		ู 0EB9	◌ 0EC9	๙ 0ED9		
A	ช 0E8A	บ 0E9A	ส 0EAA		◌ 0ECA			
B		ป 0E9B	ห 0EAB	◌ 0EBB	◌ 0ECB			
C		ผ 0E9C		◌ 0EBC	◌ 0ECC	ห 0EDC		
D	ย 0E8D	ฝ 0E9D	อ 0EAD	ฯ 0EBD	◌ 0ECD	ห 0EDD		
E		พ 0E9E	ฮ 0EAE			๚ 0EDE		
F		ฟ 0E9F	ฯ 0EAF			๛ 0EDF		

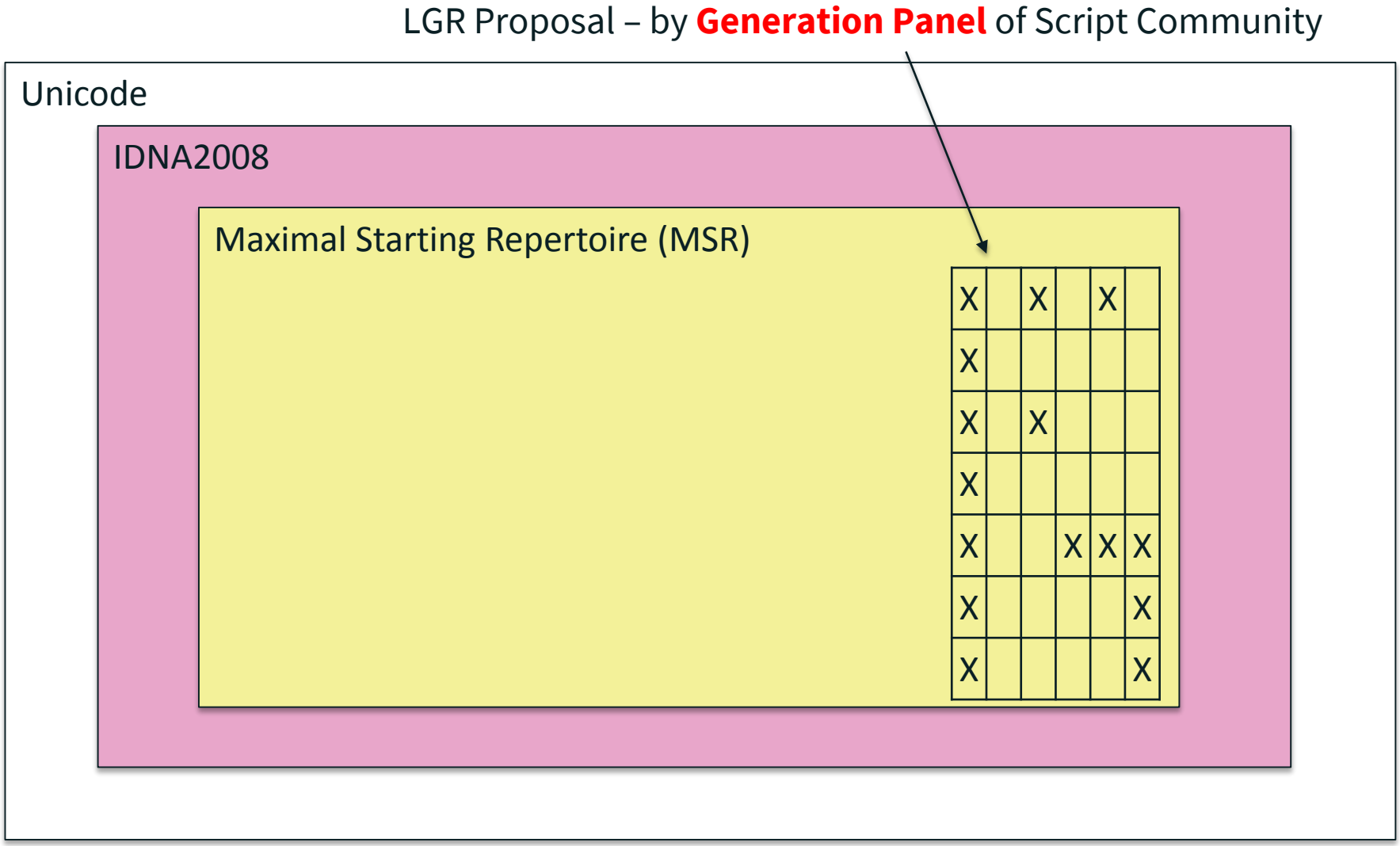
...

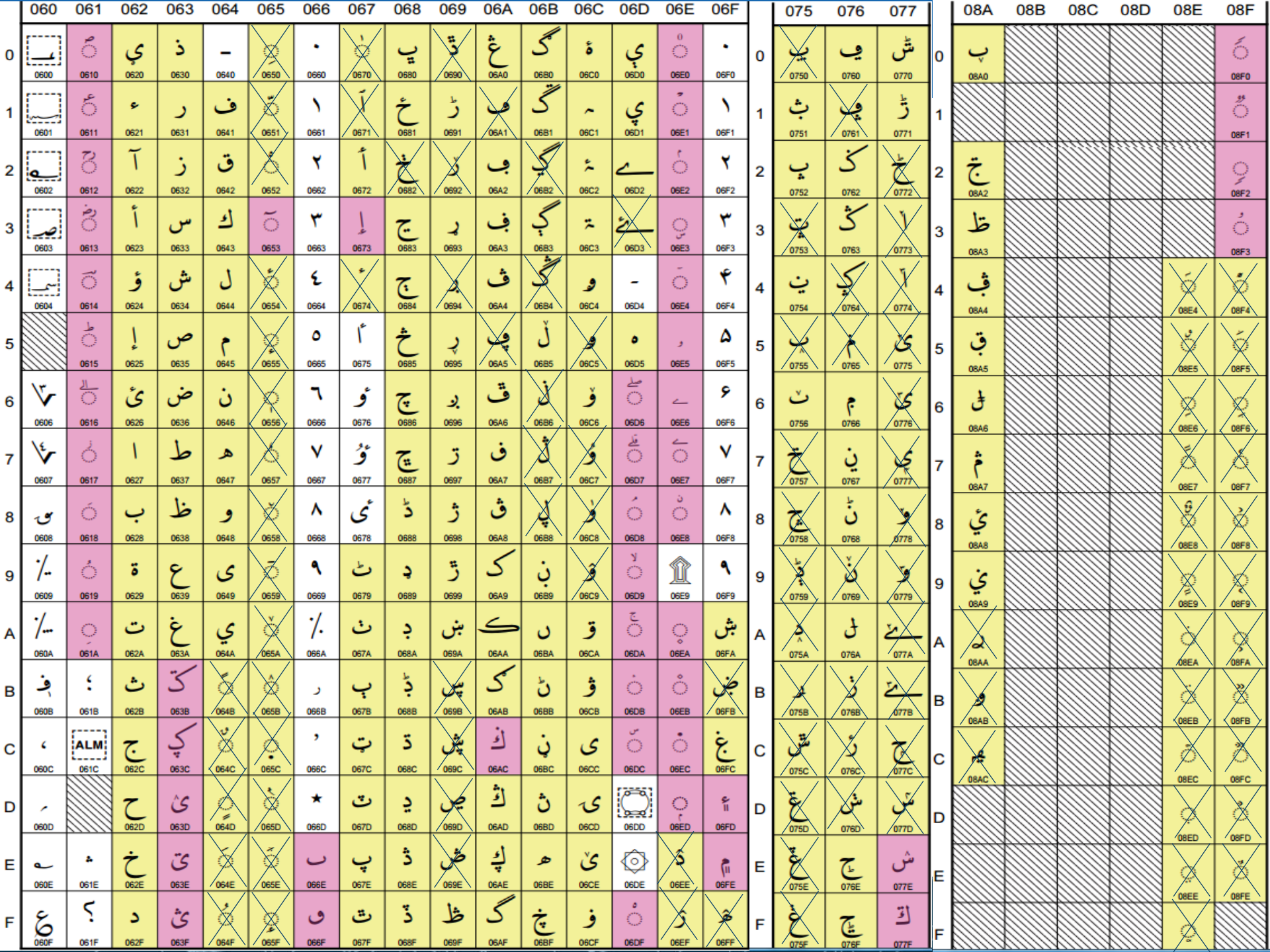
LGR for the Root Zone

LGR for the Root Zone



LGR for the Root Zone





Label Generation Rules (LGR)

- Valid code points
- Variants code points

پاکستان

پاكستان

- Label constraints
  - Cannot mix ک and ك in a label

✓ کلکته

✓ كلكته

کلكته X

كلکته X

LGR Specification

- LGR machine-readable specifications at https://datatracker.ietf.org/doc/draft-davies-idntables

- Example: excerpt from MSR-2 XML file

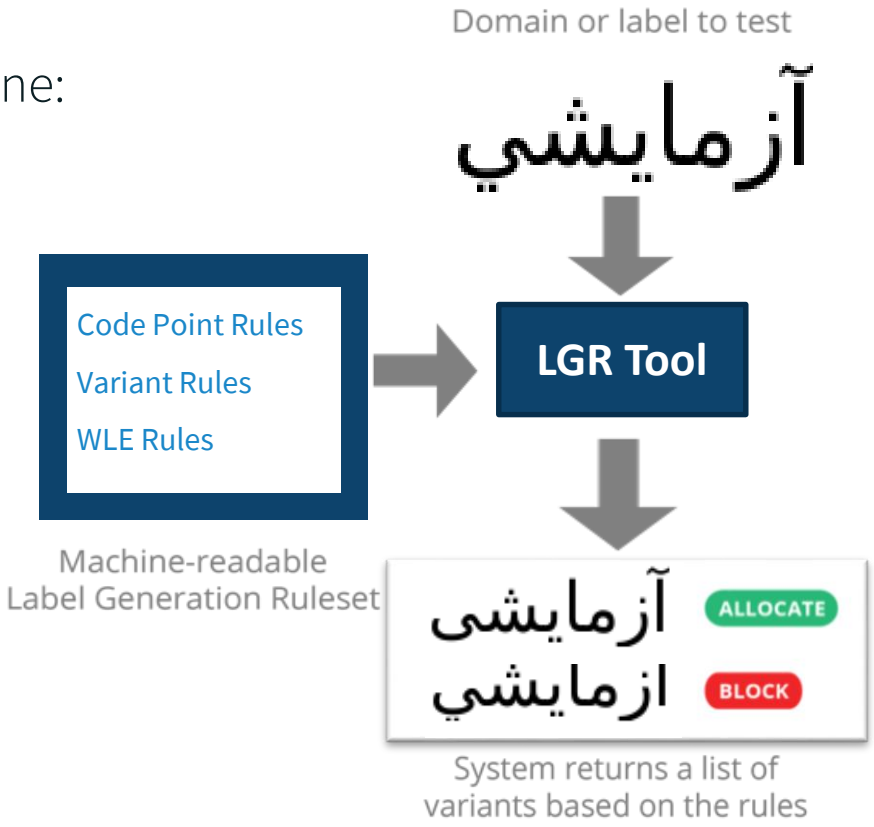
```
…
<range first-cp="0780" last-cp="07B0" tag="sc:Thaa" ref="3"/>
<char cp="07B1" tag="sc:Thaa" ref="5"/>
<char cp="08A0" tag="sc:Arab" ref="12"/>
<range first-cp="08A2" last-cp="08AC" tag="sc:Arab" ref="12"/>
<range first-cp="08E4" last-cp="08EF" tag="sc:Arab" ref="12"/>
<range first-cp="08F4" last-cp="08FE" tag="sc:Arab" ref="12"/>
<range first-cp="0901" last-cp="0903" tag="sc:Deva" ref="0"/>
<char cp="0904" tag="sc:Deva" ref="6"/>
<range first-cp="0905" last-cp="0939" tag="sc:Deva" ref="0"/>
<range first-cp="093A" last-cp="093B" tag="sc:Deva" ref="11"/>
<char cp="093C" tag="sc:Deva" ref="0"/>
<range first-cp="093E" last-cp="094D" tag="sc:Deva" ref="0"/>
<char cp="094F" tag="sc:Deva" ref="11"/>
<range first-cp="0956" last-cp="0957" tag="sc:Deva" ref="11"/>
<char cp="0972" tag="sc:Deva" ref="9"/>
<range first-cp="0973" last-cp="0977" tag="sc:Deva" ref="11"/>
<range first-cp="0979" last-cp="097A" tag="sc:Deva" ref="10"/>
<range first-cp="097B" last-cp="097C" tag="sc:Deva" ref="8"/>
<range first-cp="097E" last-cp="097F" tag="sc:Deva" ref="8"/>
<range first-cp="0981" last-cp="0983" tag="sc:Beng" ref="0"/>
…
```

LGR Specification and Toolset

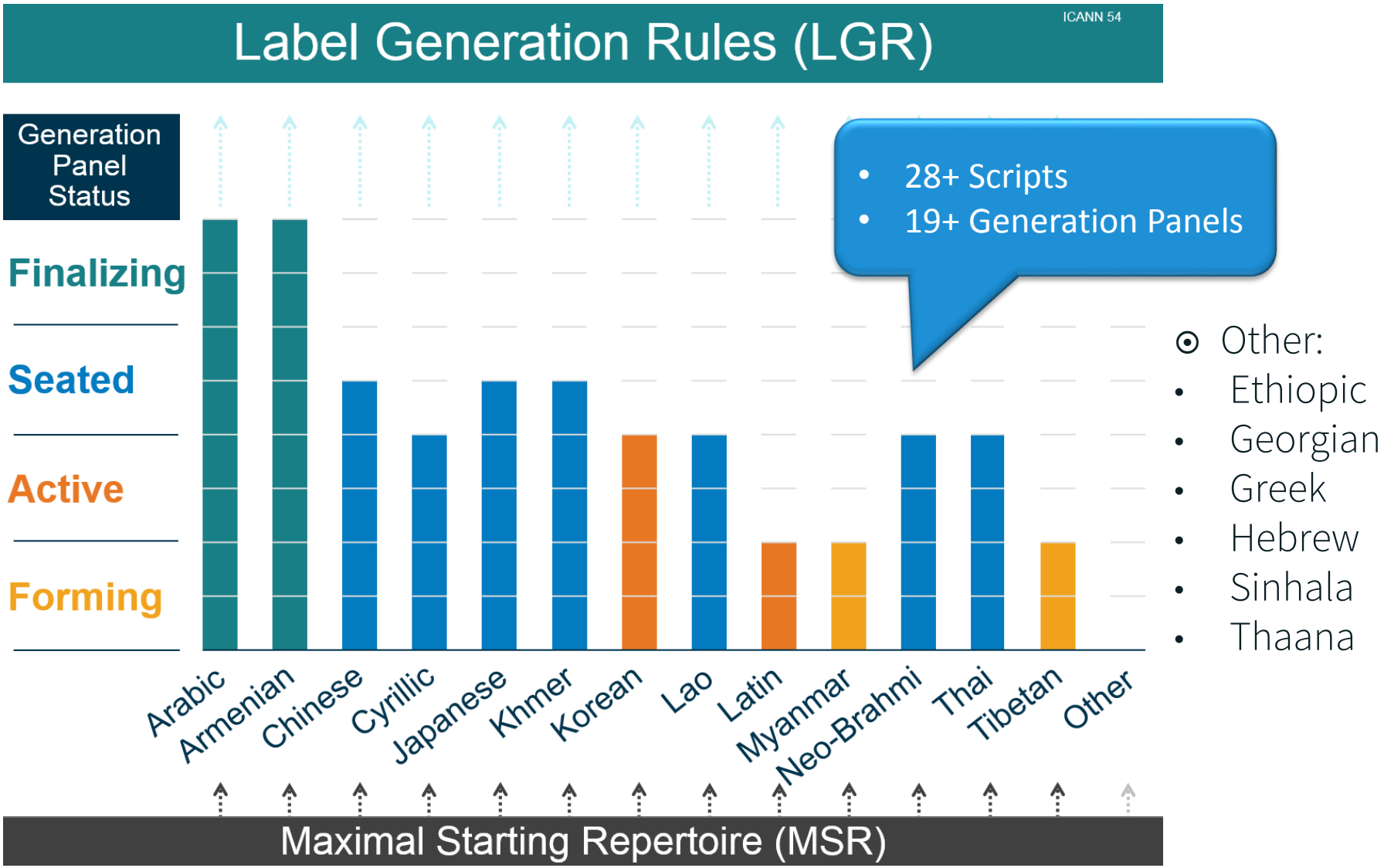
- LGR machine-readable specifications at:
  https://datatracker.ietf.org/doc/draft-davies-idntables
  - LAGER WG at IETF

- Open source LGR Toolset tentative timeline:
  - Create LGR - available
  - Use LGR – 12/15
  - Manage LGRs – 3/16

```
...
<char cp="06CC" >
        <var cp="0649" type="blocked" />
        <var cp="064A" type="allocatable" />
</char>
...
```



Status of LGR Development



XML representation of Arabic script LGR

- https://www.icann.org/en/system/files/files/proposed-arabic-lgr-22aug15-en.xml

- Example of Arabic Script:

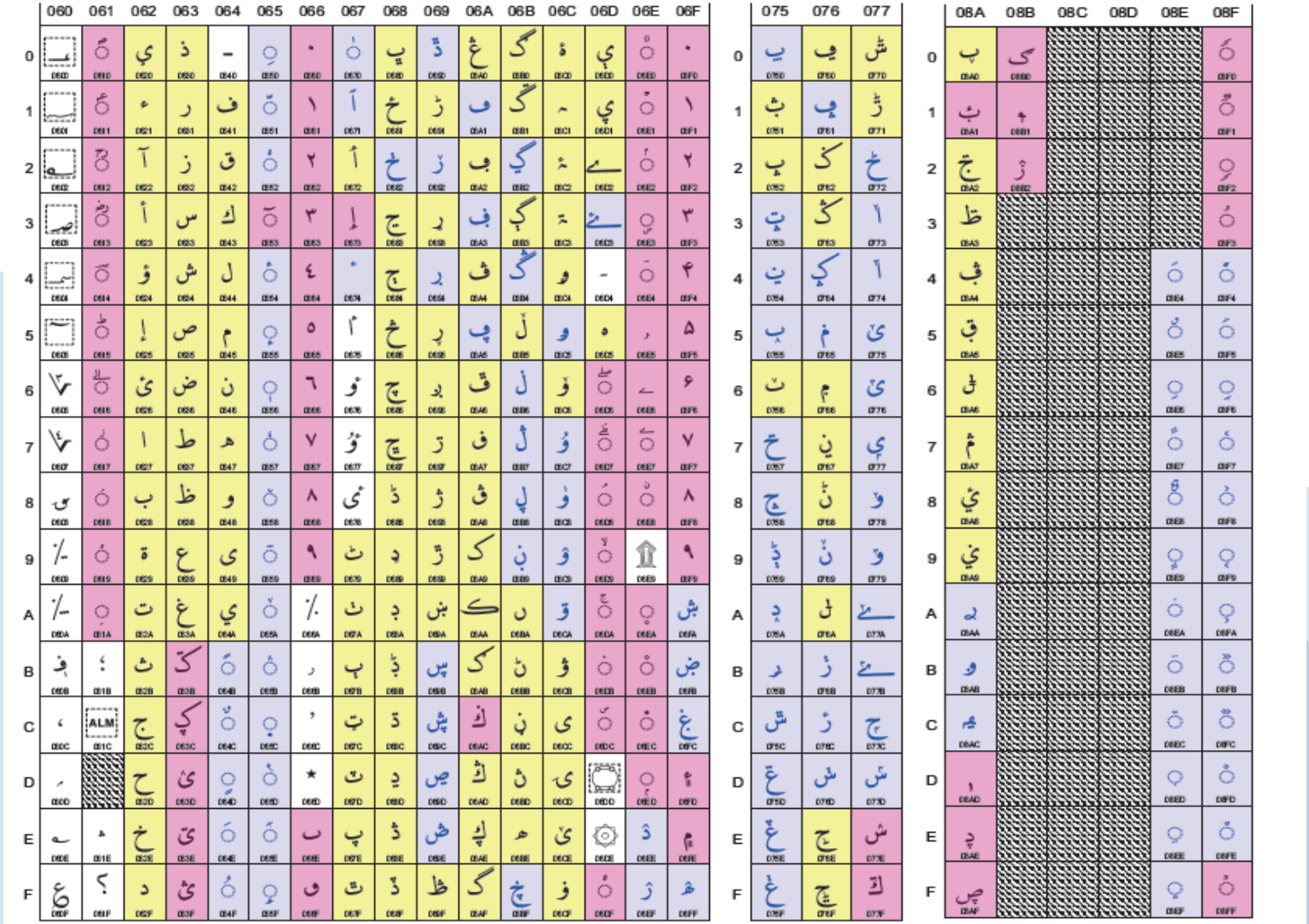
Code point definition:

```
<char cp="0647" tag="sc:Arab" ref="0 100">
    <var cp="0629" type="blocked" />
    <var cp="06BE" type="blocked" />
    <var cp="06C0" type="blocked" />
    <var cp="06C1" type="allocatable" />
    <var cp="06C2" type="blocked" />
    <var cp="06C3" type="blocked" />
    <var cp="06D5" type="blocked" />
  </char>
```

WLE rule:

```
<rule name="no-mix-alef-maksura-farsi-yeh">
    <choice>
     <rule>
      <char cp="0649" />
      <any count="0+" />
      <char cp="06CC" />
     </rule>
     <rule>
      <char cp="06CC" />
      <any count="0+" />
      <char cp="0649" />
     </rule>
    </choice>
  </rule>
```

Code Points



Example of Variant Code Points

Conventional Arabic Orthography	Western (African) Orthography
ف	ڤ
ق	ڧ
ن	ں
ك	ک
ى	ے

Example of Variant Code Points

Unicode Code Point	Isolated Form	Initial Form	Final Form	Medial Form	Applicable Principle No.
0649	ى	بـ	ببى	ببب	1
06CC	ى	يـ	بيى	بيب	1
064A	ي	يـ	بيي	بيب	1
06D0	ې	پـ	بيې	بيپ	4
067B	ټ	پـ	بټ	بيپ	4
06CD	ۍ		بيۍ		3
06D2	ے		بے		2 (06CC), 5 (064A)
0626	ئ	ئـ	بئ	بئب	3

Example of Whole Label Evaluation Rules

S.No.	Code points cannot co-occur within a label	Notes
1	0643 and 06A9	ك and ک
2	0643 and 06AA	ك and ک
3	06CC and 0649	ی and ى
4	06C1 and 0647	ہ and ه
5	0647 and 06D5	ه and ە
6	0647 and 06BE	ه and ھ
7	0629 and 06C3	ة and ۃ
8	06D1 and 06BD	ۑ and ڽ
9	067E and 06BD	پ and ڽ
10	0641 and 06A2	ف and ڢ
11	0642 and 06A7	ق and ڧ

Summary of the Arabic LGR Proposal

Number of code points: 128.

Variants:
Total number of variants: 192 (this is more than the code points as the variants are directional)
Number of variants for type 'allocatable': 26.
Number of variants for type 'blocked': 166.

WLE Rules:
Number of rules defined: 17.

LGR at work !

For the IDN ccTLD of Iran: ایران

Code point sequence: 0627 06CC 0631 0627 0646

Variants Generated : 400

allocatable=4 (including 1 original)
blocked=396

Questions?!

Sarmad Hussain
Senior Manager, IDN Program
E: sarmad.hussain@icann.org
W: http://icann.org/

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