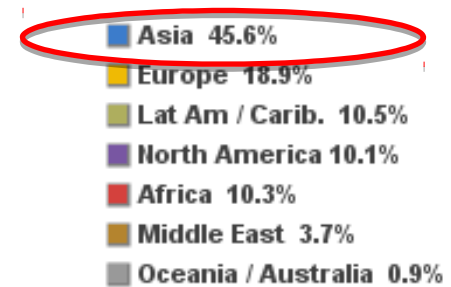
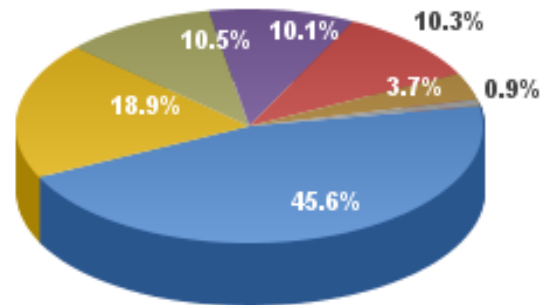


Internationalized Domain Names

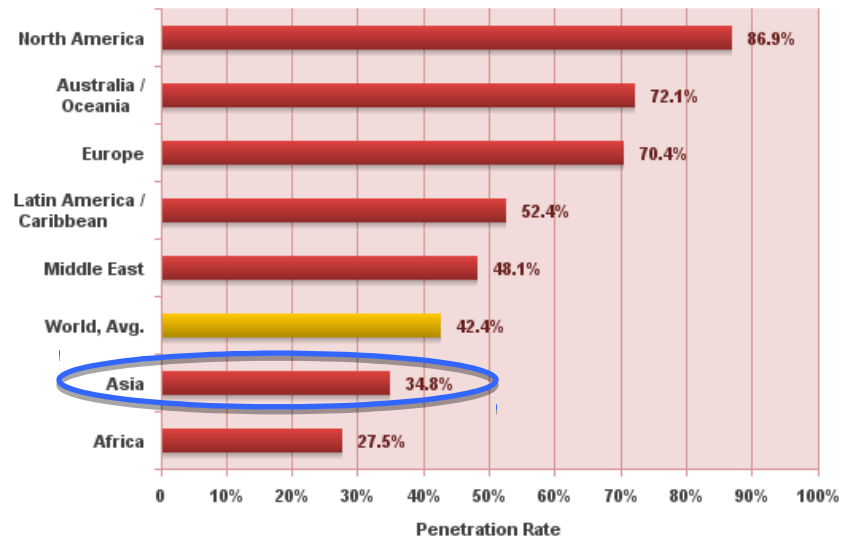
Satish Babu | APIGA2017 | 9 Aug 2017

Asia and the Next Billion Internet Users

- 3.8 billion Internet users in the world



- Asia Internet penetration still low



- Asia - majority of next billion people coming online

Source: www.internetworldstats.com; Euromonitor

The Context for IDN

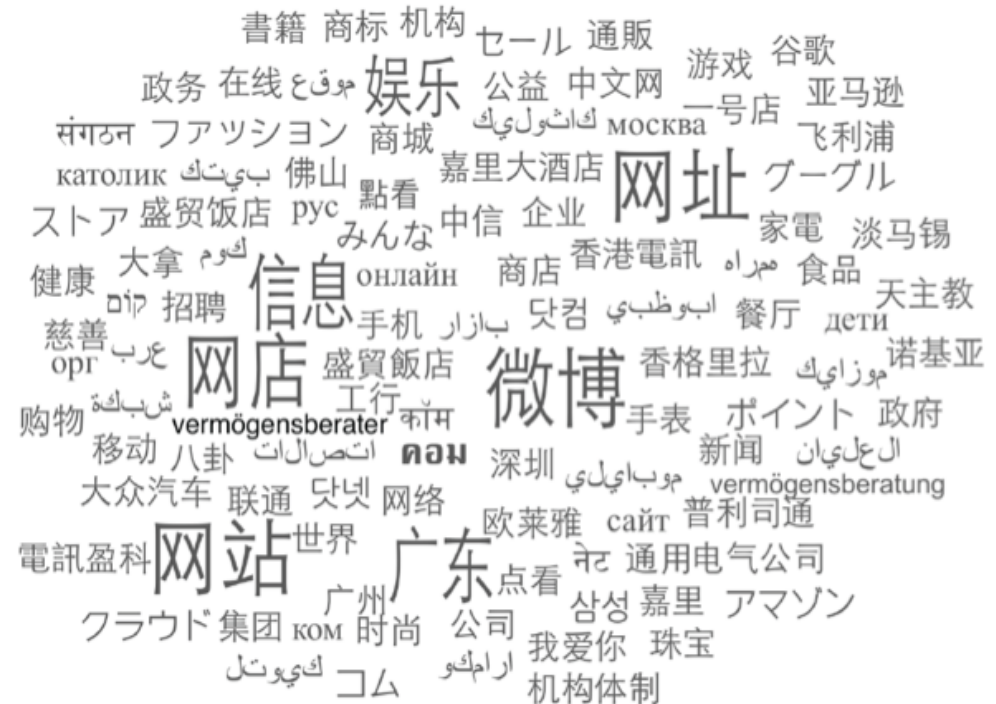
- A multilingual world with diverse writing systems/scripts
- More than 7000 living languages in the world, of which about 50% spoken in the Asia Pacific region
- Numerous writing systems/scripts (fewer than languages)
 - Examples: Han, Hindi and Arabic
- IDNs are meant to serve the online needs of a world that is diverse in languages and script

Milestones

- IDNA Protocol developed by IETF in 2003
- IDN at the second level introduced in 2004
- IDN ccTLD introduced in 2010
- Dot Asia's sunrise period for CJK in 2011
- IDN gTLD in 2013

Internationalized Domain Names (IDNs)

- ◎ IDN – Domain Names that can be represented in scripts, in addition to Latin, e.g. Chinese, Thai, Devanagari and Arabic
- ◎ IDNs enable opportunities for local communities to come online with their own content



ASCII and IDNs

www.cafe.com



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



Syntax of ASCII TLD
Use only **Letters [a-z]**

xn--11b1bsj2byb5g.xn--h2brj9c

Punycode Representation

नमस्कार.भारत



IDN Second Level Domain

IDN Level Domain (TLD)

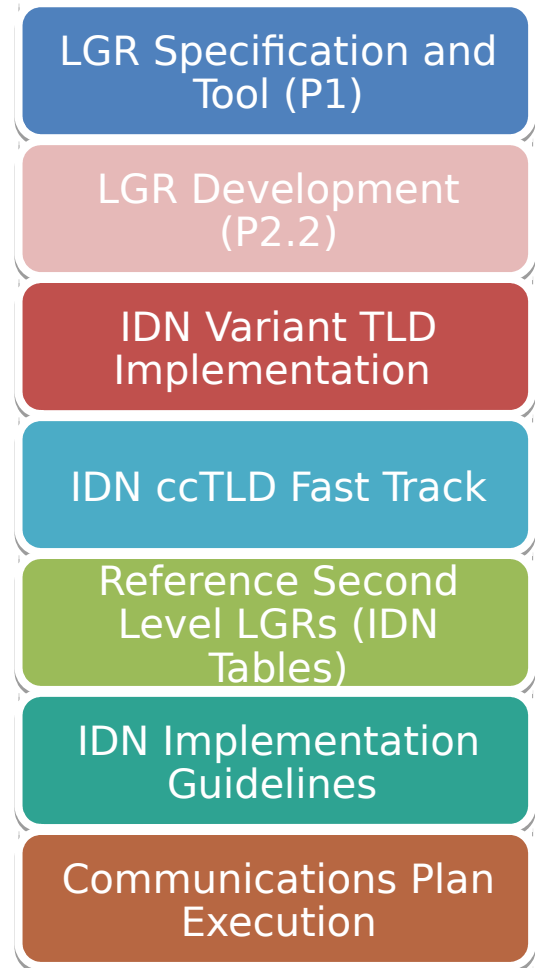


Syntax of IDN TLD
U-Label, constrained by the **“letter” principle**

What is a “letter” in a script?
Are all “letters” unique in a script?

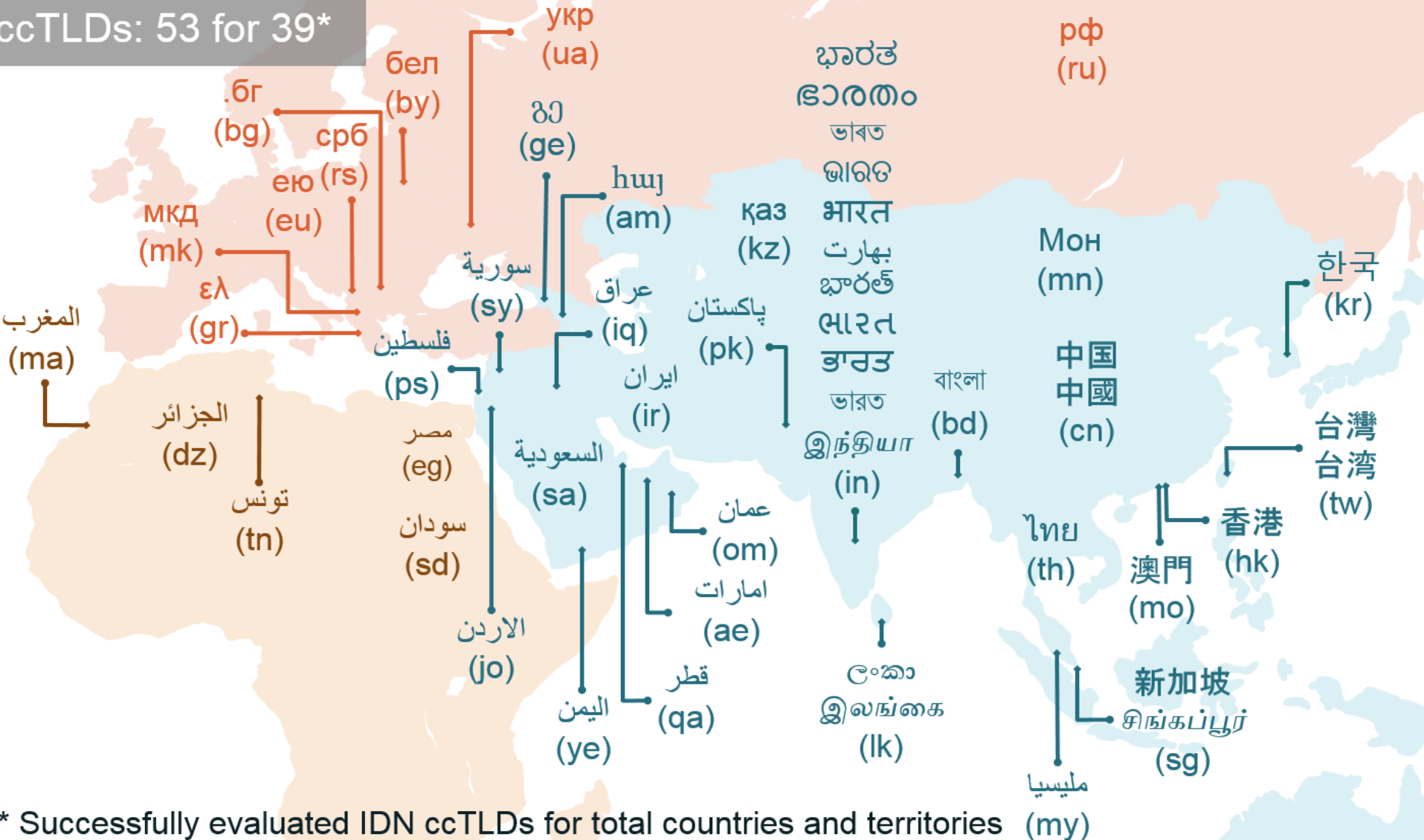
Overview of IDN Program at ICANN

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



IDN Country Code Top-Level Domains

ccTLDs: 53 for 39*



* Successfully evaluated IDN ccTLDs for total countries and territories

IDN TLD Program: Label Generation Rules (LGR) for the Root Zone

- Need for community to determine valid and unique IDN Top Level Domains (TLDs)
- For each script or writing system, community to define:
 - Which code points are valid for use?
 - Are any of these code points variants of each other?
 - Are there any additional constraints on the labels?

Label Generation Rules (LGR)

06B2	06B3	06B4	06B5	06B6	06B7	06B8	06B9	06BA	06BB
ك	ڪ	۳	۴	۵	۶	۷	۸	۹	۰
0643	0653	0663	0673	0683	0693	06A3	06B3	06C3	06D3
ل	ل	۴	۵	۶	۷	۸	۹	۰	۱
0644	0654	0664	0674	0684	0694	06A4	06B4	06C4	06D4
م	م	۵	۶	۷	۸	۹	۰	۱	۲
0645	0655	0665	0675	0685	0695	06A5	06B5	06C5	06D5
ن	ن	۶	۷	۸	۹	۰	۱	۲	۳
0646	0656	0666	0676	0686	0696	06A6	06B6	06C6	06D6
ه	ه	۷	۸	۹	۰	۱	۲	۳	۴
0647	0657	0667	0677	0687	0697	06A7	06B7	06C7	06D7
و	و	۸	۹	۰	۱	۲	۳	۴	۵
0648	0658	0668	0678	0688	0698	06A8	06B8	06C8	06D8
ی	ی	۹	۰	۱	۲	۳	۴	۵	۶
0649	0659	0669	0679	0689	0699	06A9	06B9	06C9	06D9
ي	ي	۰	۱	۲	۳	۴	۵	۶	۷
064A	065A	066A	067A	068A	069A	06AA	06BA	06CA	06DA
ک	ک	۱	۲	۳	۴	۵	۶	۷	۸
064B	065B	066B	067B	068B	069B	06AB	06BB	06CB	06DB

- Valid code points
- Variants code points

پاکستان

پاکستان

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

ڪلڪتہ ✓

كلڪتہ ✓

ڪلڪتہ x

كلڪتہ x

Root Zone LGR Development Status

June 2016

Label Generation Rules (LGR)

Generation Panel Status

Finalizing

Seated

Active

Forming

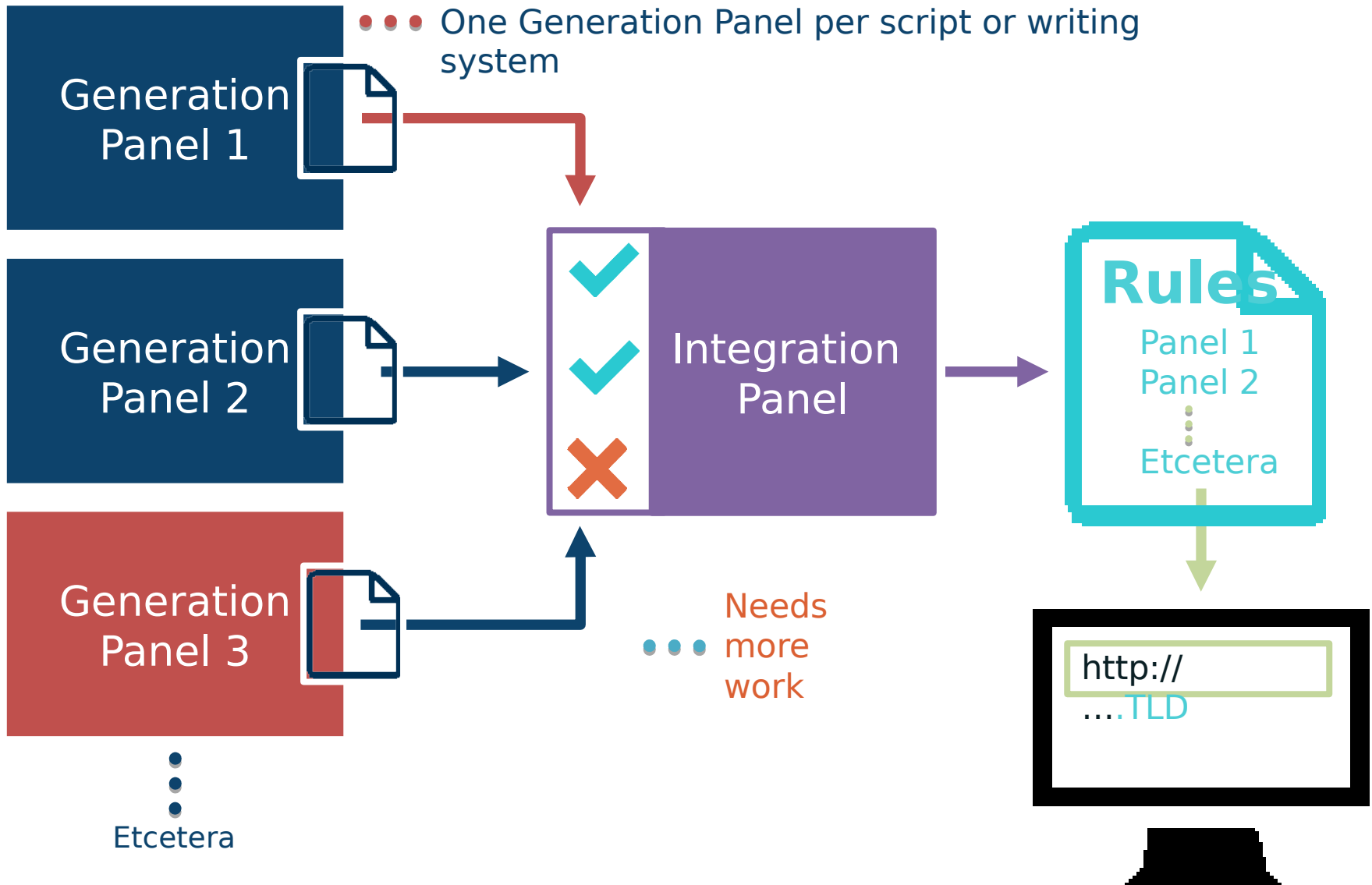
- 28+ Scripts
- 19+ Generation Panels

- Other:
- Hebrew
 - Myanmar
 - Sinhala
 - Thaana

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

Maximal Starting Repertoire (MSR)

Root Zone LGR Procedure



Coordination Challenges

- Some writing systems share same script, so need to coordinate for variant sets
 - Chinese, Japanese and Korean all use Han script
- Some writing systems have cross-script variant letters, so need to coordinate to identify them
 - Some letters among Khmer, Lao and Thai look similar
 - Some letters among Armenian, Cyrillic, Greek and Latin look the same as letters in other scripts in this set

- Security issues
 - Ability to mix scripts to simulate letters
 - Homographs and user confusion
paypa1.com
 - How will end-users be protected ?

Other End-User Issues (2/3)

- Users would prefer emails in their own languages
- IDN-based emails: Quite hard at this time
 - സതീഷ് @ നമസ്കാര.भारत
- There is limited support for such email IDs in most common email systems
- Special (and paid) services are required

Other End-User Issues (3/3)

- Emoji-based domain names: Good or bad?

🍷.WS

🍺🍺🍺.WS

🍷.WS

😊🎬.WS

ryan♥jessica.ws

How to get involved?

Volunteer

Volunteer for your script Generation Panel (GP)

To contribute expertise, contribute to the GP for your script. You can get involved by simply emailing your CV and a brief statement of interest to identlds@icann.org

Review

Review work through public comments

Sign up for the IDN mailing list vip@icann.org (to sign up, visit <https://mm.icann.org/listinfo/vip>) and participated in the review of IDN work being done at ICANN through the public comments

Listen

Keep yourself updated

Attend regular IDN Program Update sessions at ICANN meetings and sign up on the IDN mailing list vip@icann.org to get updates on the IDN Program at ICANN

Engage with ICANN



Thank You and Questions

Reach us at: IDNProgram@icann.org

Website: icann.org/idn



twitter.com/icann



[gplus.to/icann](https://plus.google.com/icann)



facebook.com/icannorg



weibo.com/ICANNorg



linkedin.com/company/icann



flickr.com/photos/icann



youtube.com/user/icannnews



slideshare.net/icannpresentations