



# Enabling Access to a Multilingual Internet

## Deploying Domain Names and Email Addresses in Local Languages

# Introduction

- **Champika Wijayatunga**  
Technical Engagement Director – Asia Pacific

# Need for Internationalization of Domain Names and Email Addresses

# The Domain Name System Has Evolved



## Domain Name System

Domain names used to navigate the internet



## Internationalized Domain Names

Community need for domain names in all languages (IDNs) so IETF developed standards



## Internationalized Email Address

Standards also updated to support email address in local languages (EAI) by IETF

---

Originally domain names only using letters a-z, digits 0-9 and hyphen.

---

Since 2003, ICANN and the community enabled IDNs, with top-level domains available since 2009

---

Email addresses have been available in local languages

# Rules for Forming Domain Names

ASCII encodes a-z, digits, other symbols and keyboard functions

	0	1	2	3	4	5	6	7
0	NUL	DLE	space	0	@	P	,	p
1	SOH	DC1 XON	!	1	A	Q	a	q
2	STX	DC2	"	2	B	R	b	r
3	ETX	DC3 XOFF	#	3	C	S	c	s
4	EOT	DC4	\$	4	D	T	d	t
5	ENQ	NAK	%	5	E	U	e	u
6	ACK	SYN	&	6	F	V	f	v
7	BEL	ETB	'	7	G	W	g	w
8	BS	CAN	(	8	H	X	h	x
9	HT	EM	)	9	I	Y	i	y
A	LF	SUB	*	:	J	Z	j	z
B	VT	ESC	+	;	K	[	k	{
C	FF	FS	,	<	L	\	l	
D	CR	GS	-	=	M	]	m	}
E	SO	RS	.	>	N	^	n	~
F	SI	US	/	?	O	_	o	del

Unicode standard encodes multiple scripts

	060	061	062	063	064	065	066	067	068	069	06A	06B	06C	06D	06E	06F
0	ا	ب	ت	ث	ج	ح	خ	د	ذ	ر	ز	س	ش	ص	ض	ط
1	ق	ك	گ	گ	گ	گ	گ	گ	گ	گ	گ	گ	گ	گ	گ	گ
2	پ	پ	پ	پ	پ	پ	پ	پ	پ	پ	پ	پ	پ	پ	پ	پ
3	ف	ف	ف	ف	ف	ف	ف	ف	ف	ف	ف	ف	ف	ف	ف	ف
4	ب	ب	ب	ب	ب	ب	ب	ب	ب	ب	ب	ب	ب	ب	ب	ب
5	پ	پ	پ	پ	پ	پ	پ	پ	پ	پ	پ	پ	پ	پ	پ	پ
6	ت	ت	ت	ت	ت	ت	ت	ت	ت	ت	ت	ت	ت	ت	ت	ت

HEX	C	J	K	V	HEX	C	J	K	V
4E90	乐	禾	禾	禾	4E94	芝			芝
4E91	禾	兵	兵	兵	4E95	芝			芝
4E92	兵	兵	兵	兵	4E96	书			书
4E93	兵	兵	兵	兵	4E97	艺			艺
4E94	乔	乔	乔	乔	4E98	乱			乱
4E95	席		席	席	4E99	乱			乱
4E96	乖	乖	乖	乖	4E9A	乱			乱
4E97	乖	乖	乖	乖	4E9B	乱			乱
4E98	乘	乘	乘	乘	4E9C	乱			乱
4E99	乘	乘	乘	乘	4E9D	乱			乱
4E9A	乘	乘	乘	乘	4E9E	乱			乱
4E9B	乘	乘	乘	乘	4E9F	乱			乱
4E9C	乘	乘	乘	乘	4E9D	乱			乱
4E9D	乘	乘	乘	乘	4E9E	乱			乱
4E9E	乘	乘	乘	乘	4E9F	乱			乱
4E9F	乘	乘	乘	乘	4E9D	乱			乱

	120	121	122	123	124	125	126	127	128	129	12A	12B
0	ॠ	ॡ	ॢ	ॣ	।	॥	०	१	२	३	४	५
1	०	१	२	३	४	५	६	७	८	९	५	५
2	५	६	७	८	९	५	५	५	५	५	५	५
3	५	६	७	८	९	५	५	५	५	५	५	५
4	५	६	७	८	९	५	५	५	५	५	५	५
5	५	६	७	८	९	५	५	५	५	५	५	५

	090	091	092	093	094	095	096	097
0	ॠ	ॡ	ॢ	ॣ	।	॥	०	१
1	ॠ	ॡ	ॢ	ॣ	।	॥	०	१
2	ॠ	ॡ	ॢ	ॣ	।	॥	०	१
3	ॠ	ॡ	ॢ	ॣ	।	॥	०	१
4	ॠ	ॡ	ॢ	ॣ	।	॥	०	१
5	ॠ	ॡ	ॢ	ॣ	।	॥	०	१
6	ॠ	ॡ	ॢ	ॣ	।	॥	०	१
7	ॠ	ॡ	ॢ	ॣ	।	॥	०	१

# Security and Usability with Variant Labels

.epic

0065 0070 0069 0063

.epic

0435 0440 0456 0441

.澳門

6FB3 95E8

.澳門

6FB3 9580

.السعودية

.السعودية

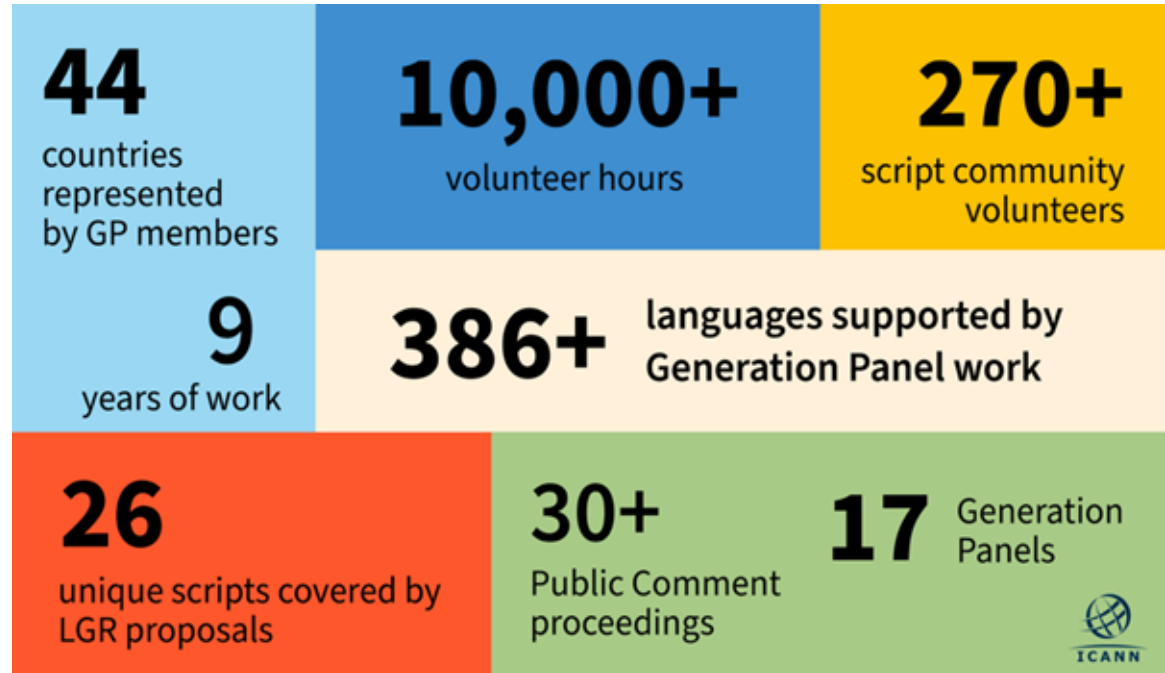
0627 0644 0633 0639 0648 062F 064A 0629

0627 0644 0633 0639 0648 062F 06CC 06C3

# Root Zone Label Generation Rules

RZ-LGR Version 6 was published in 2025 and integrates the following scripts:

- Arabic, Armenian, Bangla, Chinese (Han), Cyrillic, Devanagari, Ethiopic, Georgian, Greek, Gujarati, Gurmukhi, Hebrew, Japanese (Hiragana, Katakana, and Kanji [Han]), Kannada, Khmer, Korean (Hangul and Hanja [Han]), Lao, Latin, Malayalam, Myanmar, Oriya, Sinhala, Tamil, Telugu, Thaana and Thai.



# Internationalized Domain Name (IDN) and Email Address Internationalization (EAI)

# Internationalized Domain Name (IDN) Labels

example-123.com



**Second Level Domain (SLD)**, example-123, is formed with Letters, Digits and Hyphen (LDH). Additional rules, e.g. may not start with hyphen

Label length of 63 octets

**Top Level Domain (TLD)**, com, must be alphabetic (RFC 1123)

ยูเอตดสอบ.ไทย



## Syntax of IDN Labels

Unicode code points as constrained by the “LDH” scheme within IDNA 2008

## Syntax of IDN Labels

further constrained by the “letter” principle for TLDs

# Punycode

- Encoding system used in the DNS to represent non-ASCII characters (e.g. Arabic, Chinese, accented Latin letters, etc.) using only ASCII characters (a–z, 0–9, and hyphens)
- Users see and type a domain name in their local script
- DNS internally converts it into a Punycode form
- That Punycode version is what DNS actually resolves
- Punycode-encoded labels always start with **xn--**

# Two Equivalent Forms of IDNs

- U-Label – for the front-end interfaces, used by humans: (e.g. ไทย)
  - Unicode string in UTF-8 format.
  - Normalized to Unicode Normalization Form C (NFC).
    - è = U+00E8
    - e + ` = è = U+0065 + U+0300
  - Follow IDNA2008 protocol (not older IDNA2003 version)
- A-Label – for the back-end processing, used by machines (e.g. xn--o3cw4h)
  - Formed with equivalent *punycode* (RFC3492) prefixed by “xn--”.
  - Maximum length (RFC1035): label 63, domain name 255 octets.
  - Convert from U-label in NFC form.
  - Example:
    - امارات → xn--mgbam7a8h
    - امارات.مرحبا → xn--mgbop3h.xn--mgbam7a8h

# Examples of Internationalized Domain Names

**Armenian**

համընդհանուր-ընկալում-թեստ.հայ

ଯୁନିଭରସାଲ-ଏକସେପ୍ଟ-ଟେଷ୍ଟ.ଭାରତ

**Oriya**

**Georgian**

უნივერსალური-თავსობადობის-ტესტი.გე

다국어도메인이용환경테스트.한국

**Korean**

**Malayalam**

സാർവത്രിക-സ്വീകാര്യതാ-പരിശോധന.ഭാരതം

تجربة-القبول-الشامل.موريتانيا

**Arabic**

# Email Address Internationalization (EAI)

- What is EAI
  - Having UTF8 support for:
    - mailbox name (before the @ sign)
    - domain name (after the @ sign)
- What is not EAI
  - Having UTF8 support in:
    - Subject line
    - Address comments
    - Message body
  - MIME provides all these in conventional mail
  - Use of any character set other than UTF-8

# Email Address Internationalization (EAI)

**name@example.com**

Mailbox name is the label before the @ sign

Domain name appears after the @ sign

In an email address both mailbox and domain name are in ASCII

Mailbox name syntax governed by the rules by the the email administrator

**อีเมลทดสอบ@ยูเอ  
ทดสอบ.ไทย**

Mailbox is in Unicode UTF-8 format (no equivalent ASCII format)

Domain name may be an IDN. If IDN, it can be in U-label or A-label format

- อีเมลทดสอบ@ยูเอทดสอบ.ไทย
- (อีเมลทดสอบ@xn--l3cfk3a5bpd5gxc.xn--o3cw4h)

Mailbox naming rules for each script are set by the email administrator

# Examples of Internationalized Email Addresses

**Armenian**

Էլփոստ-թեստ@համընդհանուր-ընկալում-թեստ.հայ

电子邮件测试@普遍适用测试.我爱你

**Chinese**

**Devanagari**

ईमेल-परीक्षण@सार्वभौमिक-स्वीकृति-परीक्षण.संगठन

تجربة-بريد-الالكتروني@تجربة-القبول-الشامل-موريتانيا

**Arabic**

**Greek**

ηλεκτρονικό-μήνυμα-δοκιμή@καθολική-αποδοχή-δοκιμή.ευ

மின்னஞ்சல்-சோதனை@பொது-ஏற்பு-சோதனை.சிங்கப்பூர்

**Tamil**

# Adapting Domain Name and Email Deployment Process

# Domain Name and Email Deployment Process

## Select a Domain Name

Input and Display  
Domain name considerations  
Suggestions

## Register a Domain Name

Shopping Cart  
Registration Management  
Registration Data

## Deploy a Domain Name

Domain records  
Certificate  
Website

## Deploy an Email Address

Email server  
Email client  
Email address considerations



# Deploy a Domain Name

## Records

Use A-label to update DNS records with hosting application

---

A, MX, PTR, TXT  
(SPF/DKIM/DMARC)  
records

## Certificate

Generate certificates using A-labels

---

Use A-label in the Common Name and Subject Alternate Name fields

## Website

Web servers support IDN configuration for headers and redirects

---

Requests, logs and database to support UTF-8

## CMS

Support IDN in URL, permalinks, and other configurations

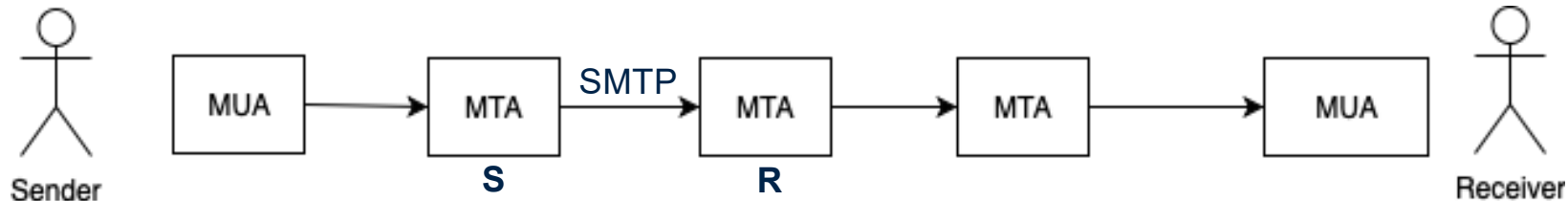
---

Use IDNA 2008-compliant libraries to handle user input and display

# Process of using IDNs

- User enters in a browser: `https://පිරික්සුම.ලංකා`
- Browser does normalization on the user entry.
- Browser converts `පිරික්සුම.ලංකා` in an ASCII compatible representation called Punycode [RFC3492], and adds 'xn--' in front of it.
  - `පිරික්සුම.ලංකා` becomes: `xn--3zc5cnk5b2b1cbv.xn--fzc2c9e2c`
- Browser calls the DNS to get the IP address of `xn--3zc5cnk5b2b1cbv.xn--fzc2c9e2c`
- Browser connects to the HTTP server at the received IP address.

# Update for Supporting EAI



S: <connect>

R: 220 receive.net ESMTP

S: EHLO sender.org

R: 250-8BITMIME

R: 250-**SMTPUTF8**

R: 250 PIPELINING

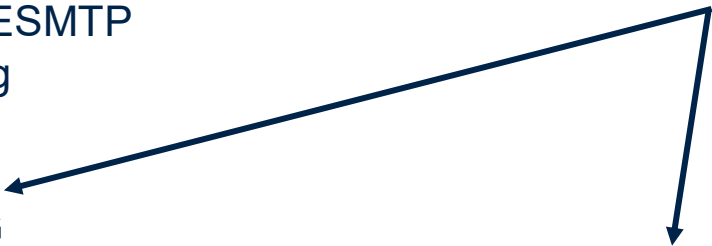
S: MAIL FROM:<猫王@普遍接受-测试.世界> **SMTPUTF8**

R: 250 Sender accepted

S:RCPT TO:<ray@receive.net>

R:250 Recipient accepted

Specific SMTPUTF8 Signaling (EAI support)



# Deploy an Email Address

## Server supporting EAI

Deploy server  
supporting  
SMTPUTF8 flag

---

Allow Unicode UTF-8  
mailbox name and IDN  
in email address

## Client supporting EAI

Update webmail client  
and dashboard to  
input and display EAI

---

Ensure that display  
of domain name is in  
U-label format

## Consideration for EAI

Set characters and  
rules for forming  
mailbox names

---

Manage rules to  
manage end-user  
confusion and  
security

## Support customers

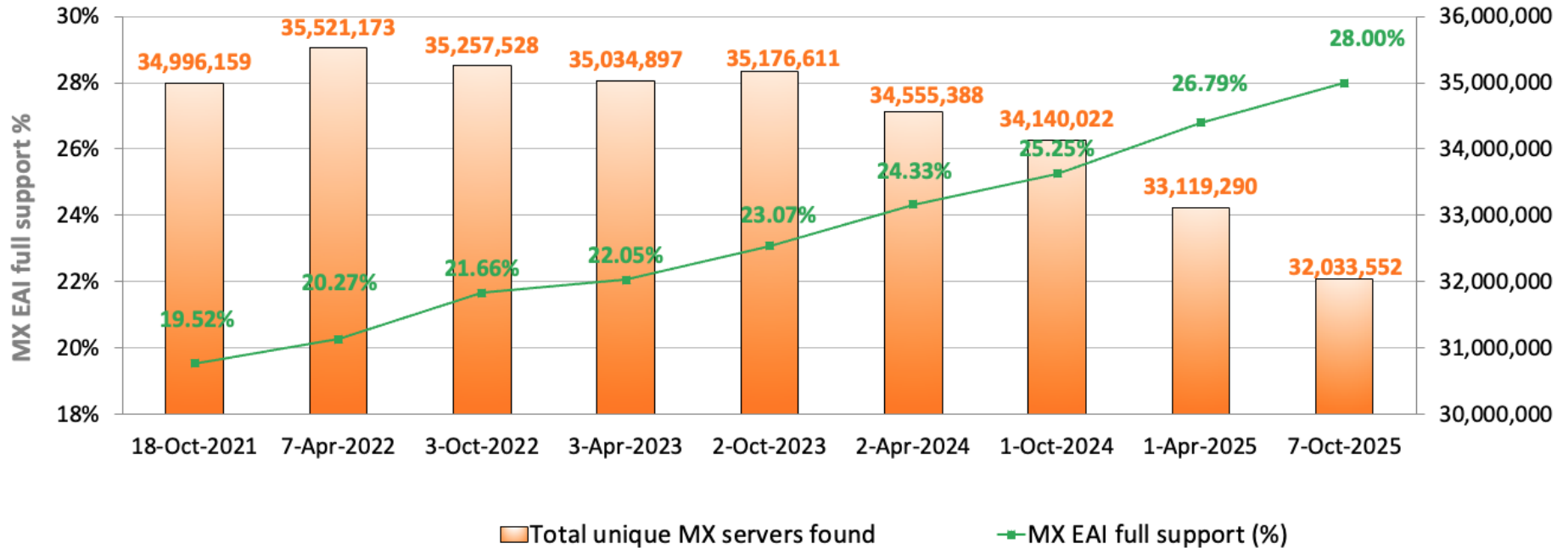
Provide guidelines and  
customer support to  
configure for EAI

---

Provide support for  
issues related to EAI  
acceptance by other  
email recipients

# EAI Support

## EAI Support Percentage vs Total Unique MX Servers Tested



Source: icann.org

# ICANN Webpage and Social Media Links



[icann.org](https://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.com/company/icann](https://linkedin.com/company/icann)



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