Configuring for Internationalized Email Addresses (EAI)

APAC Universal Acceptance Training Program
An APTLD-ICANN Collaboration

Sarmad Hussain and Champika Wijayatunga

20 January 2021
<table>
<thead>
<tr>
<th>Date</th>
<th>Session</th>
<th>Audience</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 January 2021</td>
<td>Configuring for Email Address Internationalization (EAI)</td>
<td>Technical (system engineers)</td>
<td>A detailed training on how to configure email systems to support EAI</td>
</tr>
<tr>
<td>7 April 2021</td>
<td>Programming for supporting Universal Acceptance</td>
<td>Technical (programmers)</td>
<td>A detailed training on how to design and develop applications and systems to support UA</td>
</tr>
<tr>
<td>5 May 2021</td>
<td>Universal Acceptance: Its Impact and Next Steps</td>
<td>ccTLD managers, local regulators, tech, and business</td>
<td>A dialogue on how UA issues impact the APAC community, how best to address these issues, and highlight business opportunities by being UA-ready</td>
</tr>
</tbody>
</table>

- Additional follow-up webinars or face-to-face meetings (when possible) may be scheduled based on the discussions held during these training sessions

- Further details published by [APTLD](#) and [ICANN](#)
Overview of Training (5 mins)

Introduction to EAI (20 mins)
  - Quiz (5 mins)

Configuring for EAI (20 mins)

EAI support by email tools and services (5 mins)
  - Quiz (5 mins)

Considerations for mailbox names using EAI (15 mins)
  - Quiz (5 mins)

Q&A (10 mins)
Introduction to Email Address Internationalization
Universal Acceptance of Domain Names and Emails

Goal
All domain names and email addresses work in all software applications

Accept  Validate  Process  Store  Display

Impact
Promote consumer choice, improve competition, and provide broader access to end users
Categories of Domain Names and Email Addresses

- Now possible to have domain names and email addresses in local languages
  - Internationalized domain names (IDNs)
  - Email Address Internationalization (EAI)
  - UTF8 format by Unicode used for IDNs and EAI

- Domain Names
  - Newer top-level domain names: example.sky
  - Longer top-level domain names: example.abudhabi
  - Internationalized domain names: 普遍接受-测试.世界

- Internationalized email addresses (EAI)
  - ASCII@IDN: marc@société.org
  - UTF8@ASCII: ईमेल@example.com
  - UTF8@IDN: 测试@普遍接受-测试.世界
  - UTF@IDN; right to left scripts: ای-میل@مثال.موقع
Acceptance of Email Addresses in Websites Globally

For details, see UASG027

<table>
<thead>
<tr>
<th>Email Address</th>
<th>2017</th>
<th>2019</th>
<th>2020</th>
</tr>
</thead>
<tbody>
<tr>
<td>arabic.arabic@arabic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><a href="mailto:chinese@chinese.chinese">chinese@chinese.chinese</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><a href="mailto:chinese@ascii.ascii">chinese@ascii.ascii</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><a href="mailto:ascii@chinese.ascii">ascii@chinese.ascii</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><a href="mailto:ascii@chinese.ascii">ascii@chinese.ascii</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><a href="mailto:ascii@ascii.newlong">ascii@ascii.newlong</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><a href="mailto:ascii@ascii.newshort">ascii@ascii.newshort</a></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Estimated Support of EAI in Email Systems Under All TLDs

Only 9.7% of the domains sampled were EAI ready in 2019 based on mail servers found through MX records in zones of All TLDs.

For details on methodology, see UASG021D: EAI Readiness in TLDs.
Email Messages

- **Envelope** – Information that accompanies a message in transit, including the address(es) it is being sent to, and the return address to which error or failure reports can be sent.

- **Message Header** – A series of structured fields with a header name such as From: To: or Subject: followed by the contents of the header:
  - free-format, such as the Subject:
  - fixed-format, such as the Date: and Message-ID:
  - A combination of fixed- and free-format, such as the To:, From: and Cc: with fixed-format addresses with free-format comment text.

- **Message Body** – The contents of a message, which may be unformatted text, or it may be one or more formatted or encoded MIME parts.

See [EAI: A Technical Overview](#) for details.
Email Systems

- **MUA** – Mail User Agent - A client program that a person uses to send, receive, and manage mail

- **MSA** – Mail Submission Agent - A server program that receives mail from a MUA and prepares it for transmission and delivery

- **MTA** – Mail Transmission Agent - A server program that sends and receives mail to and from other Internet hosts. An MTA may receive mail from an MSA and/or deliver mail to an MDA

- **MDA** – Mail Delivery Agent - A server program that handles incoming mail and typically stores it in a mailbox or folder

These agents create and process the email envelope, message header and message body and need to be enhanced to handle Unicode text in UTF8 format to support EAI. See [EAI: A Technical Overview](#) for details
Email Address Internationalization

- **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
Levels of EAI Implementation

- **No EAI support** - only ASCII email addresses supported by the tools and services

- **Level 1** - can exchange email with EAI addresses
  - Receive email from an EAI address
  - Send email to an EAI address
  - Cannot create mailbox and domain name in UTF8

- **Level 2** - Level 1 + can create EAI addresses
  - Receive email from an EAI address
  - Send email to an EAI address
  - Create mailbox and domain name in UTF8
Quiz
Applications present Internationalized Domain Names (IDNs) and Email Address Internationalization (EAI) to users in the following encoding format(s):

a. ASCII
b. UTF-8
c. UTF-16
d. UTF-32
e. Some of the above
Is .ieee a valid Top Level Domain?
Configuring for EAI
Pre-Requisites for Setting up EAI

- Normalize Unicode (UTF-8) string before processing, storing, etc. For IDNs use **NFC form**: e + ` (è: U+0065 U+0300) → è (U+00E8)

- Support both **representations of IDN labels**: U-label and A-label. U-Label is used for displaying and comparing; A-label for processing
  - exâmple => exmple-xta => xn--exmple-xta

- Always use IDNA2008, not the older IDNA2003 version

- Do not use code/libraries that have a static list of top-level domains (TLDs) as these change often. See [IANA list for TLDs](https://www.iana.org/domains/root), with regular updates

- Do not use regex for user input validation of internationalized identifiers. Use IDNA2008 libraries for IDN; EAI local part may be difficult to validate
Email Protocol Changes for EAI

- **SMTP**
  - Is augmented to support EAI
  - Has a signaling flag (SMTPUTF8) to specify support of EAI
  - All SMTP servers in the path must support EAI to successfully deliver the email

- **POP/IMAP**
  - Are augmented to properly support EAI
  - Have a signaling flag to specify support of EAI
SMTPUTF8 Example

Server S forwarding an email to server R

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)
SMTPUTF8 Example

S: DATA
R: 354 Send your message
S: From: 猫王 <猫王@普遍接受-测试.世界>
S: To: ray@receive.net
S: Subject: 我们要吃午饭吗？
S: 
S: How about lunch at 12:30?
S: 
R: 250 Message accepted 389dck343fg34
S: QUIT
R: 221 Sayonara

Email itself
To send and receive an email with EAI:
- All email parties involved in the delivery path have to be updated for EAI support
- If a single SMTP server in the path does not support EAI, then the email is not delivered
Protocol Changes, Delivery Path Considerations

- What happens when one email (SMTP) server in the path does not support EAI?
  - The last server trying to send to the next hop:
    - Sends back to the sender user a report of unable to deliver.
    - Drops the email
  - Similar to reports that a sender receives when an email address does not exist
Case folding:
- In ASCII, email users expect the equivalence of lowercase and uppercase. For example, PETER@example.com and peter@example.com will be delivered to the same mailbox.
- Typically for EAI, such case folding functionality is not automatically implemented in most EAI-ready software.

SPAM:
- EAI emails may be considered as spam by spam filtering software even when proper SPF/DKIM records are enabled.

Software/Services:
- Not every server/client software and services support EAI.
EAI Support by Email Tools and Services
## EAI Support by Email Tools and Services

<table>
<thead>
<tr>
<th>Name</th>
<th>MUA</th>
<th>MSA</th>
<th>MTA</th>
<th>MDA</th>
<th>MSP</th>
<th>Web mail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coremail</td>
<td>Few</td>
<td>All L2</td>
<td>Most L2</td>
<td>Few</td>
<td>All L2</td>
<td>Most L2</td>
</tr>
<tr>
<td>MS Outlook.com</td>
<td>Most L1</td>
<td>Most L1</td>
<td>Most L1</td>
<td>None</td>
<td>None</td>
<td>Most L1</td>
</tr>
<tr>
<td>Yandex Mail</td>
<td>Few</td>
<td>None</td>
<td>None</td>
<td>Few</td>
<td>Part</td>
<td>Few</td>
</tr>
<tr>
<td>Roundcube</td>
<td>Most L2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apple Mail</td>
<td>Few</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apple iOS Mail 14.x</td>
<td>Most L2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mozilla Thunderbird</td>
<td>Few</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS Outlook</td>
<td>Most L1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS Exchange Server</td>
<td>All L1</td>
<td>All L1</td>
<td>Few</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exim</td>
<td>Most L2</td>
<td>All L2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Postfix</td>
<td>All L2</td>
<td>All L2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sendmail</td>
<td>Not tested</td>
<td>Not tested</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fetchmail</td>
<td>Not tested</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Courier</td>
<td>All L2</td>
<td>All L2</td>
<td>All L2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gmail</td>
<td>All L1</td>
<td>All L1</td>
<td>All L1</td>
<td>Few</td>
<td></td>
<td></td>
</tr>
<tr>
<td>XgenPlus</td>
<td>Not tested</td>
<td>Not tested</td>
<td>Not tested</td>
<td>All L2</td>
<td>Not tested</td>
<td></td>
</tr>
</tbody>
</table>

See detailed testing results in **UASG030: Evaluation of EAI Support in Email Software and Services Report**
Quiz
Is the following statement True or False?

For Email Address Internationalization (EAI) to work all MTAs should support SMTPUTF8 signaling flag.
Which of the following statement(s) are True when sending and receiving an email with EAI:

a) All email parties/nodes involved in the delivery path have to be updated for EAI support

b) If a single SMTP server in the path does not support EAI, then the email will not be delivered

c) POP/IMAP servers could provide downgraded emails to non-EAI conforming email clients but this is not recommended

d) None of the above
Considerations for Mailbox Names Using EAI
Considerations for Mailbox Names Using EAI

- **UASG028** - Considerations for Naming Internationalized Email Mailboxes

- **Supported Scripts**
  - Know user expectations for writing systems for mailbox name and domain name portion
  - Understand complexities involved for additional scripts (e.g. security, confusion, etc.)

- **Length of a Mailbox Name String**
  - Know constraints of your system and user expectations
  - Consider same or a similar policy as for ASCII mailbox names

- **Script Mixing**
  - Allow limited script mixing only when clear user need based on local practice
  - Consider security and confusion due to script mixing for mailbox and domain name
Considerations for Mailbox Names Using EAI

- Preventing Invalid and Unstably-Rendered Strings
  - Check if Reference IDN tables meet desired mailbox string and update as needed
  - Use a string validation tool (e.g. LGR Tool) to validate the mailbox strings

- Right-to-Left (RTL) Script Consideration
  - Avoid script mixing with right-to-left scripts to avoid confusability and security issues

- Aliases and Display Names Consideration
  - Consider alias-creation option for the user interface during the mailbox name selection process. ASCII alias can be allowed with an EAI mailbox name
  - Optionally allow the user to add additional aliases at a later time
Considerations for Mailbox Names Using EAI

- **Signs and Symbols**
  - Avoid using signs and symbols, especially that do not exist on keyboard/input devices
  - If required for your market, the dot (.), underscore (_), hyphen (-) and plus sign (+) are commonly used
  - Review any additional signs (if needed) and ensure it does not cause a security issue

- **Unicode Character Normalization**
  - Understand the normalization type of your email system
  - Ensure that your email program does normalized-form-independent name comparisons
  - If it's possible to select normalization, prefer to use NFC form
Equivalence Considerations

- Define a policy for determining the “same” or equivalent mailbox names based on the writing system, user expectations, and technical capabilities of your implementation.
- Examine the IDN tables, case-folding, separators, numerals, and symbols for policy.
- Avoid creating different mailboxes using names which are equivalent to each other.
- Share your policy to let end users understand which characters and combinations will be considered valid and which ones might have equivalence.

Other considerations

- Spell domain names with their internationalized non-ASCII names. Avoid displaying the “xn--” alternative name.
- Some email clients might not automatically link the U-label and A-label forms of email address mailbox names, so ensure that both labels are mapped to each other.
Quiz
Which of the following statement(s) are False:

a) Universal Acceptance (UA) is the state in which all valid domain names and email addresses are accepted, validated, stored, processed, and displayed correctly and consistently

b) To achieve Universal Acceptance, Internet applications and systems must treat all Top Level Domains (TLDs) in a consistent manner, including new generic TLDs and all internationalized TLDs

c) All domain names should be validated against the Internationalized Domain Names in Applications IDNA2003 Standard

d) The A-label represents a domain label in Unicode UTF8 format

e) The A-label format is used to represent mailbox names in EAI
What is the A-label of 测试？

Hint: You can use the tool: https://lgrtool.icann.org/
Get Involved!
Get Involved!

- Join APAC EAI Implementers’ Group mailing list for technical support (by THNIC)

- For more information, email info@uasg.tech or UAProgram@icann.org

- Access all UA documents and presentations at website: https://uasg.tech

- Access details of ongoing work from wiki pages: https://community.icann.org/display/TUA

- Register to participate or listen in the UA discussion list at: https://uasg.tech/subscribe

- Register to participate in UA working groups here
Some Relevant Materials

- See [https://uasg.tech](https://uasg.tech) for a complete list of reports.
  - Universal Acceptance Quick Guide: UASG005
  - Introduction to Universal Acceptance: UASG007
  - Quick Guide to EAI: UASG014
  - EAI – A Technical Overview: UASG012
  - EAI – Evaluation of Major Email Software and Services: UASG021B
  - Universal Acceptance Readiness Framework: UASG026
  - Considerations for Naming Internationalized Email Mailboxes: UASG028
  - Evaluation of EAI Support in Email Software and Services Report: UASG030
Engage with ICANN – Thank You and Questions

Visit us at icann.org

Email: champika.wijayatunga@icann.org
      sarmad.hussain@icann.org

@icann
facebook.com/icannorg
youtube.com/icannnews
flickr.com/icann
linkedin/company/icann
slideshare/icannpresentations
soundcloud/icann