

Evaluation Studies

UASG Workshop October 27 2017 Abu Dhabi



Universal Acceptance

Measurement & Monitoring

Noting where the software industry is at with respect to
Universal Acceptance & EAI and encouraging progress

Measurement & Monitoring Projects

- * Websites
- * Browsers
- * Linkification
- * EAI



Review of Websites



Websites – Looking under the hood



Website review: Client email address validation

Sara Dickinson / UASG@ICANN60 / 27 October 2017



Universal Acceptance

Companion Study to Main Website Review

*Goal of Study was to understand **WHY***

- *6% websites rejected all email addresses*
- *7% websites accepted all email addresses*

by investigating the underlying implementation.

Also to look for commonality in the sites that had mixed results.



Why so few sites investigated in detail?

- **Can only review implementation of sites that perform client side validation - no visibility of server side validation at all**
- Client side implementations fall into 3 general categories:
 - Straightforward to find validation code
 - Moderately complex to understand how validation is performed
 - Multiple validation points
 - Asynchronous behaviour
 - Implementation is complex and obscure... cannot find validation by code inspection



Observations

1

Regular Expressions (regex)

In all cases the syntactic validation of the email address by the client was done using a regular expression embedded in the client code.

2

Rejects were due to restrictions in Regular Expression

Non-ASCII domain names, Non-ASCII mailbox identifiers,
TLDs longer than 3/4 characters, TLD's on hardcoded list

3

All sites do it differently

Although the regular expressions used show a certain number of common features, they are all unique. Also, there was absolutely no commonality in the way in which each site used JavaScript to validate email.

Recommendations for client side validation? (for discussion)

- Note this is a non-trivial problem:
 - Lack of knowledge or reference material currently available
 - The syntactic validation itself is not trivial (often copied/anecdotal)
 - Server side validation still required (opaque)
 - Many, many websites will need 'fixing'
- Ideally a lightweight library with a standard (reference) implementation of email validation
- Minimal regex that just checks for `<anytext> + @ + <anytext>`
 - Security risks?

Review of Browsers



Review of Linkification



Review of EAI

