

UA Tech WG Meeting Notes

13 May 2024

Attendees

Satish Babu

Abdulkarim Oloyede

Anna Bagdasaryan

Gopal Tadepalli

Harsha Wijayawardhana

Herve Hounzandji

Jim DeLaHunt

Letsatsi Lekhooa

Maud Adjeley Ashong Elliot

Waqar Ahmad

Mohamed Elnour Abdelhafez

Arnt Gulbrandsen

Seda Akbulut

(Apologies by Jabhera Matgoro due to ICANN80 preparation from AFRALO.)

Meeting Agenda:

1. Welcome and roll call
2. Final inputs before the call for the proposals: [SOW](#) for T1 action item: identify the technology stacks for UA testing: “Javascript libraries React, Angular and Node.” Identify UA testing with these stacks.
3. Signing off the [5-year action plan](#) for Tech WG to achieve the [UASG’s 5-year strategic plan](#) (linkification goal?) Discuss on what to take up next from the action plan
 1. Write a paper that syntax check is useless and recommend a validation method by sending an email to confirm - Jim ([see topic 2](#))
 2. Develop a framework for what it means to be UA-ready for websites-Harsha
4. Tentative timeline for finalizing the [Tech WG Survey](#). What’s next?
5. Next scheduled meeting: Monday 27 May 2024, 15:00 UTC
6. Tentative ICANN80 Prep Week UA Community Update Session: 29 May at 15:00 UTC

Meeting Recording: [Link](#); Passcode: .&wCr2Nx17

Meeting Notes:

Seda presented the agenda and suggested WG take a look at the SOW and focus on the received comments. Satish led the discussions.

Agenda#2: [SOW](#) of Java Script Libraries React, Angular and Nodes

Jim has posted comments with some details on the SOW. Seda has incorporated the comments, and some required follow-up discussions.

At section ‘Description of Work’:

UASG 004 is the document with two versions, one for text file (UASG004A) and one for pdf; and the email addresses and domain names could be picked from it. Jim explained that this document provides information of valid EAI email addresses of various scripts. Copying addresses from PDF format is bad for some scripts, and to bypass the problem, a text file of all the addresses as UASG 004A is provided.

Satish asked if there were any restrictions or guidelines for the mailbox part on an EAI email address. Arnt said anything possible from the list but there are guidelines for each label not to mix scripts and/or writing directions. So he suggested not including those mixed one on the local part and the domain part; so domain name and mailbox should be in the same script.

Gopal said there is automatic transliteration between Tamil, Telugu and Kannada, and email addresses of both languages could possibly be interfering with each other. Satish said this could happen when the scripts are within the Neo-Brahmi family.

Jim noted in the chat: “If there is an Arabic practice of mixing directions, e.g. <shop124@arabic.tld>, and UASG004 does not contain that case, then that sounds to me like a deficiency in UASG004 which we should fix by extending UASG004.”

Arnt said if we limit ourselves to use the addresses only from the UASG 004 document, the writing directions of the addresses in the examples do not mix. Arabic script could be right-to-left, and when it has numbers, it is bidirectional. However, the tests would not need those. Satish said testing within the given examples could be sufficient. If there is a real use case, it should be incorporated into the document. Jim said to fix the UASG 004 is a separate issue, and suggested focusing on the SOW.

UASG 004 Test Cases for UA Readiness Evaluation EN

<https://uasg.tech/download/uasg-004-use-cases-for-ua-readiness-evaluation-en/>

UASG 004A Test Cases for UA Readiness Evaluation - Data EN

<https://uasg.tech/download/uasg-004a-use-cases-for-ua-readiness-evaluation-data-en/>

Jim said the SOW mentioned UASG 004 to be reviewed, and asked if there is a specific purpose and timeline for this. Seda presented the list for WG to review. Satish explained that for email addresses, the review should be on the both mailbox name side and domain name side. Seda pointed out that some email addresses may be difficult to read. Arnt suggested testing the addresses of item 32 to item 40 and one email address with Chinese characters, item 56. Arnt briefly explained the issues. Seda added this into the SOW document.

Gopal noted in the chat the following and then shared he would send an [additional information](#) to the group about it: "I am sorry for the weak audio. The RHS of the @ is in our court. We can think of a "testing space". To my understanding the LHS of the @ is synthetic and we presume that they operate much like what we do with the RHS. The synthetic part of the e-mail Id needs a different kind of space model. This is the parlance we used some time back."

Satish responded to Gopal's comment on LTR vs RTL text directions that the test is about the encoding but not rendering, which will not affect the direction of text. Jim said the username part of the email address has no rules and it could appear directionless when it comes to emoji or other characters which do not have clear directions.

At section ‘Deliverables’:

Seda shared and answered Herve Hounzandji’s comment, and then resolved it. The SOW will be shared on the ICANN Wiki page of the working group, and proposals will be received and reviewed. The selected proposal will be contracted, and the relevant provider will be called ‘contractor’.

At section ‘Conflict of Interest’:

Jim commented to add explanation to the abbreviation ‘COI’. Seda added the information and resolved the comment. COI will be periodically checked and renewed accordingly, and shared with the contractor.

Comment on Markdown:

There is a suggestion by Jim on the ‘markdown’ information.

Arnt shared two steps of linkifying automatically:

- 1) Test if an IDN is treated like a traditional domain name (eg. [.org])
- 2) Test if an IDN is detected like traditional domain name is detected

Jim explained the markdown is the text around an IDN or an EAI which is different from linkification, which would be a kind of text information to indicate something is a domain name/URL or an email address (eg. “()” or “[]”). Satish agreed with the explanation and summarized the points that are related.

Satish asked if we had any documents on Linkification. Seda shared UASG010-Quick-Guide-to-Linkification-EN-2017, and upon that Satish mentioned that it could be updated and be more detailed.

Agenda#3: Signing off the 5-year action plan

Seda mentioned that there was no linkification goal stated in the action plan, and asked if WG would like to add. Satish agreed. The goal is added to the Consolidated 5-year Action Plan as item 15: “Test for identifying the linkification problems, and consider remediation of the issues.”

This would also be in the annual plan of the Tech WG for the upcoming year. Jim said sharing the annual plan document on the Tech WG's wiki page would be helpful. Satish said once the list is finalized, it would be shared. Seda would be adding the plan documents for all WG.

Satish also mentioned that there was a suggestion through mailing-list about linkification by having a static list of domain names. Arnt said that was not the case, the list updating would be some months later than ICANN.

Satish said that the general concern of the plan is the cutoff time. The process may need to be reconsidered as it progresses.

On the document hub, the UASG010 'Quick Guide to Linkification' document can be found as a reference. WG would look into this document for guidance.

Agenda#4: [Discuss on the next the action plan](#)

HTML5 validation of email addresses needs to be updated. This was mentioned through the email thread for both Tech WG and EAI WG. W3C was approached to cover this Regex-based email syntax verification.

Jim answered that the reliable validation would be the software which received the input email address sending a verification email to it, and receiving a response, possibly with a verification link within the email body. Syntax validation is less useful. Email validation should be regarded as send-receive activity of the backend.

Arnt added that some really need to verify an email address as sending emails could be potentially costly. If a small percentage of emails were being sent to non-email-addresses, or spam-traps, their servers would be in a lot of pain. To encourage accepting all kinds of email addresses would not pass the real world. They have to use heuristic ways to verify if an email address is the kind that people use without having to send out an email. Arnt said there are significant participants in the email community who would agree with him.

Satish said even in all ASCII, there is a format of well-formed email addresses. Satish agreed with checking the well-formedness of the addresses, however, the term validation is used loosely. Arnt said he would suggest avoiding Regex, however, using a different rule set would be alright. HTML5 is using ASCII-specific Regex, and not achieving agreement on extended-Regex.

Satish said more discussion on this topic is required, and starting to write a paper would be a good idea. Harsha said he did some work on verifying the well-formedness of Sinhala, Tamil and English email addresses. Harsha said he has a powerpoint presentation on this topic for the next meeting, with guidelines and explanations. Another thing is to support UTF-8 email addresses, and also in the contact forms. There are still some issues with them. Harsha agreed to present his work in the next meeting.

Meeting adjourned.

Next Meeting: 27 May 2024, Monday 15:00 UTC

Action Items:

No	Action Item	Owner
1	Seda to add consolidated plan on the wiki page to all WGs	Seda
2	Review UASG010 for Linkification and identify what needs to be added	Tech WG
3	Sharing with Presentation on Email Verification and accepting UTF-8 email addresses in the next meeting	Harsha