

# UA Measurement WG Meeting

24 August 2023

## Attendees

Nabil Benamar

Dessalegn Yehuala

Abdullah Qamar

Anna Bagdasaryan

Alassane Malick

Harsha Wijayawardhana

Jim DeLaHunt

Krislin Goulbourne-Harry

Kunle Olorundare

Sushanta Sinha

Sarmad Hussain

Seda Akbulut

## Meeting Agenda:

1. Welcome and Roll Call
2. Updates from Dessalegn about the [UA Curricula for Academic Programs](#) prepared based on the [statement of work](#)

**Meeting recording:** Link, **password** rm9J.\*Xrhj

## Meeting Notes

Nabil allowed Dessalegn to reconnect what was presented last week, where were the changes and briefing on topics to continue.

Dessalegn briefly retracted what was presented last week as the first part of the UA Curriculum, Module 1 to 5, and continued from Module 6 onwards. Dessalegn reminded that the scope is limited with time constraints and designed for beginners with a slight touch of internationalization and localization but not too deep in details for those subjects. This module's lightweight scope would cover five use-cases of universal acceptance particularly.

Sarmad added that this micro module provides flexibility, because adding a separate full-course module with academic credits would require longer process time and approvals, which sometimes is not feasible. This module would patch up as additional to computer science and IT courses without expanding credit hours.

The curriculum would integrate UA awareness into existing courses without extending credit hours and would be divided into microlearning modules for flexibility. The team discussed modules 6 and beyond, covering various topics

including normalization and IDNA2008 checking in programming using internationalized domain names. Potential issues with the documentation regarding Module 6 were raised, and it was agreed that the curriculum design process would be an active and iterative one.

**Module 6:** This module would provide the knowledge of programming with IDNs. Dessalegn invited for additional topics if required, and in case of those additional topics requiring detailed focus and more credit time, those could be in the more advanced module of the next step.

Sarmad said Module 6 provides IDNA2008 compatible programming languages and normalization. Adding a little more details as part of the programming would be useful for the programming context. Dessalegn agreed on adding some programming libraries as discussion.

Harsha asked about expanding the list of programming languages to Python, Java, PHP, since the libraries compatible with IDNA2008 are already there. In addition, normalization and comparing the strings could be discussed on this topic. For the pitfalls (6.b), Harsha suggested adding examples of Sinhala and Urdu languages.

Dessalegn noted and also notified that programming details could be discussed; however, it would only be a slight touch, and not very much in the details. Dessalegn explained that the content ordering may have missed some topics and we may come back to these points, however, we may not exhaust the process by iterating the similar topics. Harsha said when there is not much documentation to refer to, but some issues are there to address for (6.b). (6.a) has no issue.

Jim added suggestion text for the Module 6; Sarmad agreed with Jim's comment, which is relevant, and in a more general way, it could expose students to new TLDs with ASCII longs and ASCII shorts, which also have issues sometimes. This would be more generic than IDN-only focus. Nabil agreed with the comment as well, and would like to put this before the module starts, since this would not be the first time for the students to see the IDNs.

Jim Suggested adding the list of currently active IDN top level domains for people to read 'xn-' entries at <https://data.iana.org/TLD/tlds-alpha-by-domain.txt> . Students may find the number of IDN TLDs surprising and motivating. Sarmad agreed to Jim's comment and questioned whether we should focus on just IDN or keep it generic with including ASCII short and long TLDs.

Jim said another aspect of pointing people to the actual list of TLDs would be connecting the students directly to the machinery of the actual domain name system, which may make the subject less theoretical and more hands-on with manipulating the real machinery to see what is going on. For engineering students, this might be motivating, and make use of Module 5's network troubleshooting lessons. These would not take more time but are more motivating. Nabil agreed and suggested putting these before the Module 6.

Abdullah commented that Javascript could be added along with Python and Java, since some Javascript libraries support Unicode as well. Dessalegn acknowledged that the Javascript libraries support unicode to ASCII conversions.

Jim and Abdullah's conversation in chat:

*Abdullah: One more thing we should not focus on one programming language. I think the choice of programming language is instructor/teacher choice depending on student background knowledge and interest. leaning concepts ,problem solving ,critical thinking is more important*

*Jim: This was exactly my university's approach to teaching computer science and software engineering. They proudly refused to teach courses on technologies just because they were popular. They told us students that they teach us fundamentals, and those fundamentals help us learn each new language or technology on our own.*

*With the benefit of my 40 years in software, I think that approach was correct.*

**Module 7:** The concepts will cover Label Generation Rules (LGRs), understanding the LGR formats (RFC7940; RFC 4690, and RFC 3743 as earlier text based formats) and familiarize the student with concepts of LGR. There would be a programming exercise to verify a label and identify its variants.

Harsha also suggested talking about variant labels before the LGR. Dessalegn said it would be in the components of LGR.

Sarmad added that these are additional constraints, and could be expanded the explanation in the Module.

Jim asked for clarification of the word 'additional constraints by IDNA2008', Dessalegn answered that the wordings could be changed. Sarmad explained that

IDNA2008 is the baseline and the LGR proposed character-based constraints and variant handling rules. It should say ‘additional constraints on top of IDNA2008’.

Malick commented in the chat suggesting to swap the order of Module 7 and Module 8.

**Module 8:** This is introduction to Email Address Internationalization (EAI).

Jim suggested the topics of Adoption challenges and Communication Etiquette:

1. Adoption challenges in a partially-compatible world - spam filters blocking EAI addresses, delivery failures due to EAI incompatible email systems, providing non-EAI alternate addresses as a workaround.
2. Communication challenges in a multilingual world - when sending from one language to another, what does a recipient do with a sender address which they cannot read?

Sarmad followed up with Jim saying that the Communication Etiquette seemed more like a social issue rather than a technical issue, and asked if this could be considered a technical-related challenge. Jim said a good engineering curriculum should reflect the requirements of the real world. In a different domain, physically posted mails also have similar problems, and the postal union has conventions about how to address letters and parcels. Jim suggested taking a look at real world problem solving examples for the engineering solutions.

Harsha talked about the similar social issue in the IT field when people have letters with Sinhala or Tamil languages, there are some proposals coming on this. Harsha has noticed that the email-interface translator in Gmail works well when an email is completely written in one language to the receiver’s default language. Harsha said other proposed solutions are on the way for this language barrier. Harsha said this part may not fit into the scope but we can discuss this later.

Sarmad advised adding EAI WG’s work on best practices in this module on how to set up local language email addresses. Such as the LGR topic on the availability and constraints, when it comes to the mailbox name part of female addresses, there are considerations to be discussed. This would be L2 work.

**Module 9:**

Harsha asked which programming languages are preferred and Sarmad answered that the codes were developed for both Java and Python. There are others like Javascript and PHP, therefore, the coverage could be discussed later. Sarmad said the basic ones could be covered and let the students explore onwards.

Also for the EAI features, Harsha asked if it covers the EAI self-certification guide contents, and suggested covering some basic components. Dessalegn answered that not all but the basic ones could be covered.

Jim suggested the additional topics:

9.5. Considerations for setting up email mailbox names, e.g. as in <https://uasg.tech/download/uasg-028-considerations-for-naming-internationalized-email-mailboxes-en/>

9.6. Defining scope of EAI features, e.g. using UASG EAI Self-Certification Guide, e.g. <https://uasg.tech/eai-certification/>

Nabil explained the reasons for choosing Java and Python initially, and potential to change to other languages in the future. Although Java is not a top programming language anymore, Python still is, and Javascript is as well. To cover the two different domains; Python is for general purpose or data analysis or AI, and Javascript is for web development.

Dessalegn responded that the students at his university learn Java, or C++ as basic programming languages. When they start learning Internet-based programming, they will start with Javascript. Sarmad said from an academia perspective, it is useful to know what actually is taught in their curriculum.

Nabil suggested not putting Java and JavaScript in the same category as they cover different domains, and they would complement each other. Abdullah said Javascript is an introductory programming language for O-Level students in Pakistan as the demand from the industry is shifting to web-based.

### **Module 10:**

The development framework for EAI would be chosen based on the programming language. This topic would give an overview of programming languages like Python and Java and discuss their libraries. The code for this should be able to run on different platforms including the mobile devices.

### **Module 11:**

Three major topics are prerequisites for this module: Unicode security, DNS security and IDN security. Jim suggested a subtopic under the Unicode security. Jim said above all kinds of security risk, email/IDN phishing is comparatively small, however, other risks not related to DNS are much larger.

Nabil said he agreed with the existence of many other security risks, EAI email addresses used to be blocked as they were suspects of phishing attacks and such.

Nabil recommended covering these. Dessalegn agreed and asked Nabil to add the comments in the note. Nabil added four more subtopics under Unicode Security: Mitigating EAI addresses blocking, DNS Security, ASCII Homograph, and Phishing, Spamming, etc.

### **Module 12:**

This module will cover Unicode Support in Operating systems.

Jim commented that the well known operating systems have the file systems which are not case sensitive, they would rather preserve the other case for the file names of one of the cases. Jim thinks Module 12 is off topic of universal acceptance. Jim's comment in the document:

*12.2 Working with Unicode: Case-sensitive vs case-insensitive vs case-insensitive but case-preserving filename handling.*

There is no objection to Jim's suggestion.

Dessalegn shared the module mapping table near the end of the document. The UA module is mapped to the IT/Computer Program Course's Curriculum. Sarmad said the modules 7, 8 and 9 are not mapped to any coursework, and asked WG to recommend any ideas.

Jim said although this micro course is supposed to be added to an existing curriculum, the topics sound otherwise. Jim suggested explaining at Module 4 about adapting this UA course into an existing course. Dessalegn accepts the idea.

Regarding the scope, Jim said we cannot create new courses for Internationalization and Localization, however, it would be possible to drop it into the existing topics, and there is believed to have mutual benefit for contents of Internationalization and universal acceptance. Therefore, Jim said it might be worth mentioning something about internationalization in the course document. Dessalegn agreed and Sarmad recommended capturing the ideas.

Sarmad said the next step would be developing the content of each module in this course. Sarmad asked if the WG would like to be involved in this work to add future input. Sarmad suggested Dessalegn manage this with a smaller team and then come back with finalized contents on these modules. There would be broader input on the current design. Meeting participants agreed with the current module design.

Nabil thanked everyone for the participation and ended the meeting.

**Next meeting:** Thursday, 07 September 2023 at UTC 1600

**Action items**

<b>No.</b>	<b>Action Item</b>	<b>Owner</b>
1	Come back with updated UA Curriculum	Dessaegn and smaller team