An Introduction to Universal Acceptance

Don Hollander / Presentation to ALAC September 2017
UA in a Nutshell

Universal Acceptance (UA) ensures that all domain names and email addresses can be used by all Internet-enabled applications, devices and systems.
Anatomy of an email address

username@example.com

测试5@普遍接受-测试.世界

* **Username/ 测试5**
  * Mailbox name
  * Can be in ASCII or Unicode
* **Second level domain name (example/ 普遍接受-测试)**
  * Can be ASCII or Unicode.
  * Unicode can be represented as Unicode or Punycode
* **Top Level Domain name (.com/.世界)**
  * Can be ASCII or Unicode
  * Unicode can be represented as Unicode or Punycode
  * Can be 2 to 63 characters long
  * Can ONLY be from an authoritative list that is dynamic and has more than 1,500 choices
Five Verbs to UA Readiness

Accept
Validate
Store
Process
Display
Our Target Audiences

Doers

*Developers & Systems Architects*

Directors

*CIOs and senior IT Management*

Influencers

*C suite, Thought Leaders, Government Ministers and Officials*
UA Documents: uasg.tech/documents

* UASG003 – One page Fact Sheet
* UASG004 – Use Cases for UA Testing
* UASG005 – Quick Guide to Universal Acceptance – 11 Languages
* UASG007 – Introduction to Universal Acceptance – Technical
* UASG008 – Local Engagement
* UASG010 – Quick Guide to Linkification
* UASG014 – Quick Guide to EAI
* UA At a Glance
* What it Addresses
* Why it Matters
* Aimed at non-Geeks
UASG011 – FAQs

* UA At a Glance
* What it Addresses
* Why it Matters
* Aimed at Media
Model of a Local UA Initiative
Multi-Party
We’re looking for local partners to engage local communities
We can provide content and collateral
We can also provide limited funds – See Help Wanted Ad
Aimed at At-Large structures, ISOC Chapters, Local Computer Societies
* 11 Languages
* Accept, Validate, Store, Process & Display
* Aimed at Technical Managers
UASG004 – Use Cases

* Working Websites
* Working Email Addresses
* Provides Test Suite for Developers
* Aimed at Geeks
* Extensive Technical Document – 70 Pages
* English and Chinese
* Aimed at Developers and Architects
* Lists relevant RFCs
* Aimed at Geeks
* Commissioned Report by Analysis Mason
* Looks at Social, Cultural and Financial Implications
* Estimates nearly US$10 Billion added economic Benefit
* Estimates millions of additional users
* Aimed at C*Os, Boards, Ministers
* Identifies the issues around automatic creation of Links in Applications
* Linkification Evaluation of Popular Social Media currently underway
* Aimed at Application Owners
* Introduces Email Address Internationalisation (EA)
* Identifies the Issues
* Aimed at email Geeks & SysAdmins
UASG015 – CIO Blueprint

* Aimed at CIO and IT Management
* Outlines the Issues to address and how
* Includes Checklist of Popular Internet Industry applications
* Additional Blueprints planned for other Industries
* Provides sample tendering and contractual clauses
* Includes IPv6 and DNSSEC
* Aimed at IT Management & Procurement
* APNIC
* Xgen Technologies
* ICANN
* THNIC
* Others to be developed
* Being published September 2017
* Looked at most popular browsers – Desktop & Mobile
* One passed all our tests
* Most others were pretty good
* Looked at 1,000+ popular websites
* Found just 7% accepted all test cases
* Found just 7% rejected all test cases
* Peeked under the hood of the code
* Targeted at web geeks and Accessibility Advocates
Thank You & Questions
The process by which an email address or domain name is received as a string of characters from a user interface, file or API.

UASG Recommendations

* User interface elements must support:
  * Unicode.
  * Strings up to 256 characters.

* ASCII Compatible Encoded text ("Punycoded") in place of Unicode.
  * Unicode shown by default.
  * Punycoded text shown *only* when it provides a benefit.
The process by which an email address or domain name – received or emitted – is checked for syntax correctness.

UASG Recommendations

* Easiest way to ensure all valid domain names are accepted.

* Should not occur unless required. If yes:
  * Verify TLD against authoritative table.
  * Query domain name against DNS.
  * Require repeated entry of email address.
  * Validate characters - no “disallowed” code points.
  * Limit to few, whole-label rules defined in RFCs.
  * If string contains ‘。’ convert to ‘.’
The long-term and/or transient storage of domain names and email addresses.

UASG Recommendations

* Apps / services should support Unicode
* Information stored in UTF-8 whenever possible
* Consider end-to-end scenarios before converting between A-Labels & U-Labels
  * Consider storing in both formats
* Clearly mark email addresses and domain names during storage
Process

UASG Recommendations

* Check code points not defined when application / service was created – shouldn’t “break” user experience.

* Use supported Unicode-enabled APIs.

* Use latest IDNA Protocol & Tables documents for Internationalized Domain Names.

* Process in UTF-8 wherever possible.
Occurs whenever an email address or domain name is used by an application or service to perform an activity, or is transformed into an alternate format.

UASG Recommendations

* Ensure numbers are handled as expected
* Treat ASCII numerals & Asian ideographic number representations as numbers
* Upgrade apps & servers/services together
* Perform code reviews to avoid buffer overflow attacks
Display occurs whenever an email address or a domain name is rendered within a user interface.

UASG Recommendations

* Display all Unicode code points supported by underlying operating system.

* When developing app/service, or operating a registry, consider languages supported.

* Convert non-Unicode data to Unicode before display.

* End user should see “everyone.みんな” vs. “everyone.xn--q9jyb4c.”
Display occurs whenever an email address or a domain name is rendered within a user interface.

UASG Recommendations

* Display Unicode by default
* Use Punycoded text only when it provides a benefit
* Consider that mixed-script addresses will become more common
* Use Unicode IDNA Compatibility Processing to match user expectations
* Be aware of unassigned & disallowed characters
Tools & Resources for Developers

Authoritative Tables:
* http://www.internic.net/domain/root.zone
* http://www.dns.icann.org/services/authoritative-dns/index.html
* http://data.iana.org/TLD/tlds-alpha-by-domain.txt
* See also SAC070: https://tinyurl.com/sac070

Internationalized Domain Names for Applications:

Unicode:
* Security Considerations: http://unicode.org/reports/tr36/
* IDNA Compatibility Processing: http://unicode.org/reports/tr46/

Universal Acceptance Steering Group info & recent developments: www.uasg.tech
* Read the documents at www.uasg.tech/documents
  * UASG003 – Fact Sheet
  * UASG005 – Quick Guide
  * UASG007 – Introduction to UASG
  * UASG011 – FAQs
* Subscribe to the UASG Discussion list www.uasg.tech/subscribe
* Get your own systems UA Ready
* Spread the word…