

Top Issues in Universal Acceptance

Jim DeLaHunt / IUC42 / 11 September 2018



Universal Acceptance

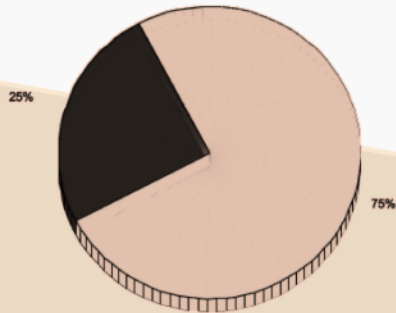
Agenda

Slides: <http://go.jdlh.com/iuc42s6t2>

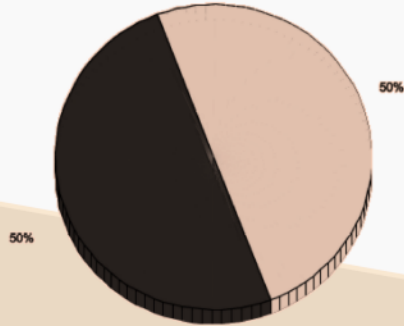
- * Who we are: UASG, Jim DeLaHunt
- * Universal Acceptance successes, so far in 2018
- * Scenario: justify IDN & EAI
- * Scenario: emails going to Junk
- * ~~Scenario: scary news article~~
- * ~~Scenario: support the culture~~
- * ~~Scenario: foreign email addresses~~
- * Q&A

Warm-up Exercise

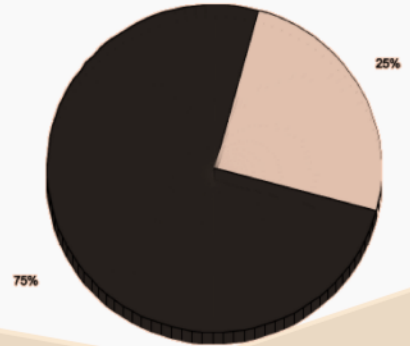
According to *w3techs*, which of the following pie charts most closely represents the fraction of websites on the Internet that are **primarily English language based content**?



25% English
75% All other



50% English
50% All other

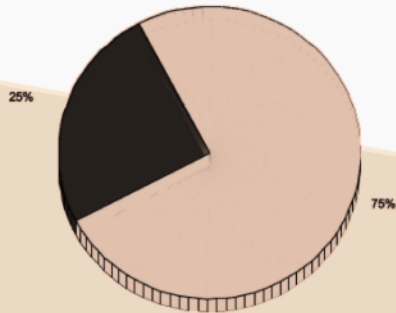


75% English
25% All other

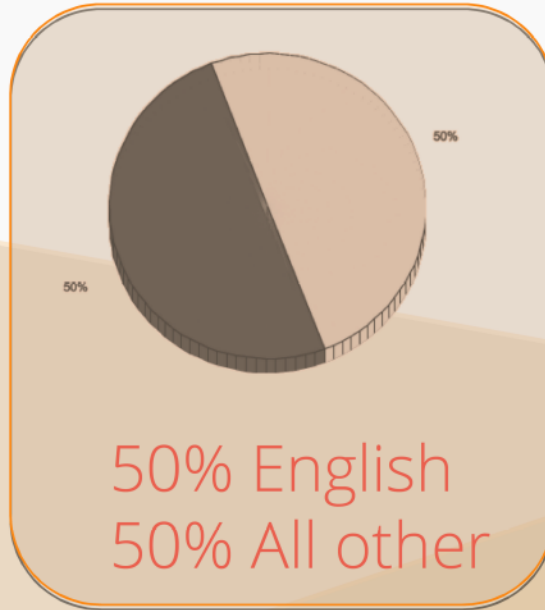
Actual data can be found at https://w3techs.com/technologies/history_overview/content_language/ms/y. Data shown is approximate.

Warm-up Exercise: Solution

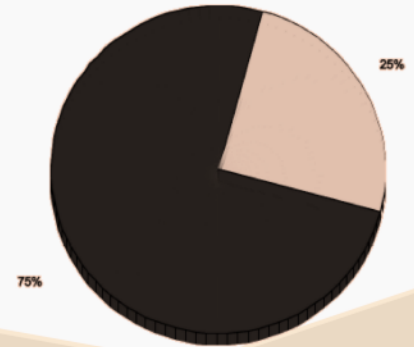
According to w3techs, which of the following pie charts most closely represents the fraction of websites on the Internet that are primarily English language based content?



25% English
75% All other



50% English
50% All other



75% English
25% All other

Actual data can be found at https://w3techs.com/technologies/history_overview/content_language/ms/y. Data shown is approximate.

Warm-up Exercise

Each of the 3 groups below contain lists of Top Level Domains (TLDs) that are valid (approved and delegated by ICANN), except that each list contains one made-up or invalid TLD. **Which TLD in each group is invalid?**

Group A

嘉里
ANALYTICS
BLOCKBUSTER
DIAMONDS
HOTELES
广东
MOVISTAR
இந்தியா
REALLY
政务

Group B

ABC
مصر
ATHLETA
இலங்கை
CANCERRESEARCH
CITIC
新加坡
ESURANCE
FAKE
ไทย

Group C

GMAIL
قطر
JOY
LIKE
لدى
ONYOURSIDE
OOO
فلسطين
SILLY
SUCKS

Warm-up Exercise

The 3 invalid TLDs are highlighted below in **red**. There are 1541 valid TLDs (as of January 2nd, 2018), and the list is continuing to grow.

Group A

嘉里
ANALYTICS
BLOCKBUSTER
DIAMONDS
HOTELES
广东
MOVISTAR
இந்தியா
REALLY
政务

Group B

ABC
مصر
ATHLETA
இலங்கை
CANCERRESEARCH
CITIC
新加坡
ESURANCE
FAKE
ไทย

Group C

GMAIL
قطر
JOY
LIKE
كوسا
ONYOURSIDE
OOO
فلسطين
SILLY
SUCKS

Who we are



Who we are

Universal Acceptance Steering Group (UASG)

- * <http://www.uasg.tech>
- * Community-led initiative, world-wide
- * Raise awareness, identify problems, solve them
- * Project of ICANN, the domain name system organisation

Jim DeLaHunt

- * <http://jdlh.com>, ☎ +1-604-376-8953
- * Vancouver, Canada
- * Consultant in multilingual websites; software engineer
- * UASG volunteer participant

Who you are

This talk is a story of technology evangelism, of wins and obstacles. We assume you generally understand Internationalised Domain Name (IDN) and Email Address Internationalisation (EAI), and don't need to be persuaded of its value.

Primary audience

- * Application developers handling domain name and email addresses
- * Language practitioners developing software tools and architecture
- * Dev, QA, marketers, system administrators, and management

Context

Five Criteria of UA



Accept



Validate



Store



Process



Display

Universal Acceptance is the state where

- * all valid domain names and email addresses are **accepted, validated, stored, processed** and **displayed** correctly and consistently
- * by all Internet-enabled applications, devices and systems.

Why Universal Acceptance matters

1

The next 1,000,000,000 internet users

Compared to the current internet population (about 1bn), the next billion will demand more strongly to have service in their own script and language.

2

Work with the Internet as it is, not as it was

Internationalised Domain Names and Email Address Internationalisation are in use now. Work with it, and be open to the new opportunities there.

3

Better user experience

Engaging customers in their own script and language retains those customers better, and attracts new customers as well.

The next 1,000,000,000 Internet users

Next 1 billion

China, India, Third World.
Large share use non-Latin script.
Little marginal North American,
European increase.
Mostly mobile and small-screen,
lower share on desktop, laptop.
Extending to mid-, lower-educated,
less comfortable with Latin script.

VS

First 1 billion

First world, N. America, Europe.
Large share use Latin script.
Includes large share of North
American, European potential.
Mostly desktop & laptop
computers, mobile only later.
Cream of highly-educated in each
market, the best at Latin script

Universal Acceptance Successes in 2018

EAI in use

More and more major email system providers can send and receive to EAI addresses. A few services can host EAI addresses. Major institutions are committing to EAI.

Hosting EAI addresses

- * Hosting EAI addresses
 - * Coremail, XGenPlus, Raseel, ...
- * Send and receive to EAI addr
 - * Microsoft, Google, US Congress

Mandates for EAI

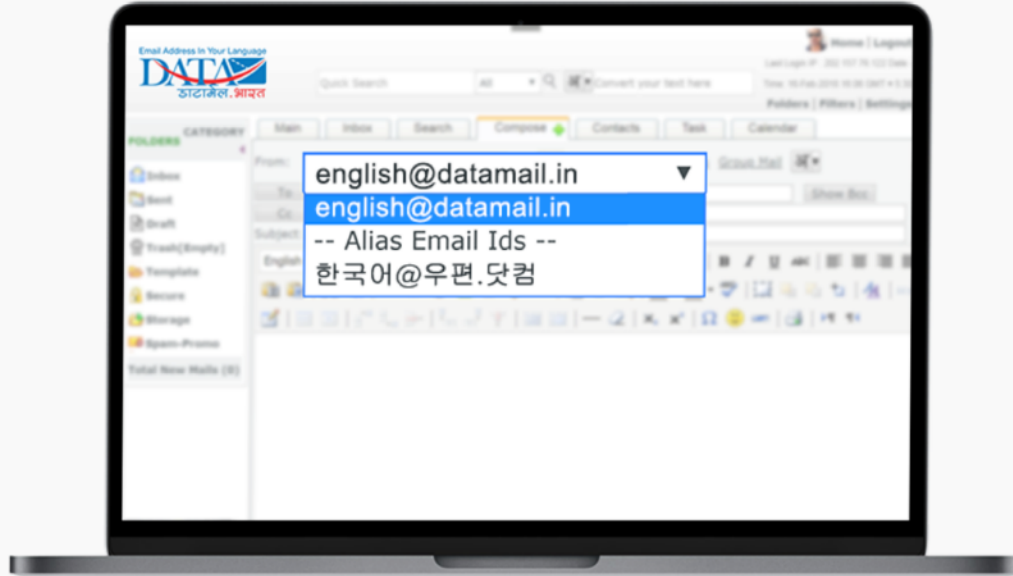
- * Rajasthan State (India)
 - * Hindi address for every resident
- * Government of India
 - * Hindi address for 5m employees

Microsoft EAI announcement, Feb 2018



“Support... email addresses in multiple Indian languages across most of our email apps and services including Office 365, Outlook 2016 client, Outlook.com, Outlook apps for Android and iOS, as well as Exchange Online, and Exchange Online Protection.” [and Chinese]

XgenPlus Korean email, May 2018



DataMail in Korean...
Arabic, Russian, Thai,
Chinese...
Hindi , Marathi, Gujarati,
Bengali, Tamil, Telugu,
Urdu, Punjabi!

Indians pushing for Universal Acceptance

In recent months, Indian thought leaders in UASG have presented to meetings, held conferences, and garnered media coverage for UA. The enthusiasm and accomplishment is remarkable.

Notable events

- * Internet and Mobile Association of India (IAMAI) campaign for UA
- * “more than 60% of Indian citizens have never used the Internet.”
- * ISOC India Mumbai meeting
 - * “Open Mic Event” for storytellers and poets

Promoting Universal Acceptance in India



Internet and Mobile
Association of India
(IMAI) kicks off UA
campaign, 2018-08-28

Useful UASG tools for everyone

EAI Readiness Tester

- * If an address is ready to get EAI
 - * <https://uasg.tech/eai-check/>
 - * Does not send email, just asks

UA problem report form

- * Tell us about UA problems
- * We'll follow up on them
<https://uasg.tech/global-support-centre/>

Articles and papers for key audiences

The UASG has a communications campaign to reach key audiences by means of events earning media, placing articles, and white papers. UA Ambassadors reach out to their own communities.

Key audiences

- * People who **make** UA happen: developers. architects
- * People who **direct** UA to happen: CIOs
- * People who **influence** UA to happen: C* suite, governments, media

“\$10 Billion Opportunity” (white paper)

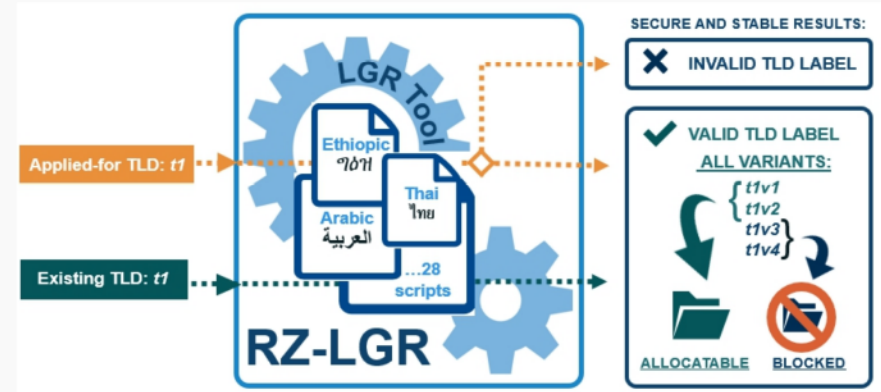


White paper, *Unleashing The Power of All Domains: The Social, Cultural and Economic Benefits of Universal Acceptance*
<https://uasg.tech/whitepaper/>

DNS IDN Variant Label Generation Rules

For DNS root, some 2LDs

- * Prevents confusingly similar labels being registered
- * 28 scripts already covered
 - * +Neo-Brahmi script Generation Panel (NBGP)
- * Required for more IDN gTLDs



<https://www.icann.org/public-comments/managing-idn-variant-tlds-2018-07-25-en>

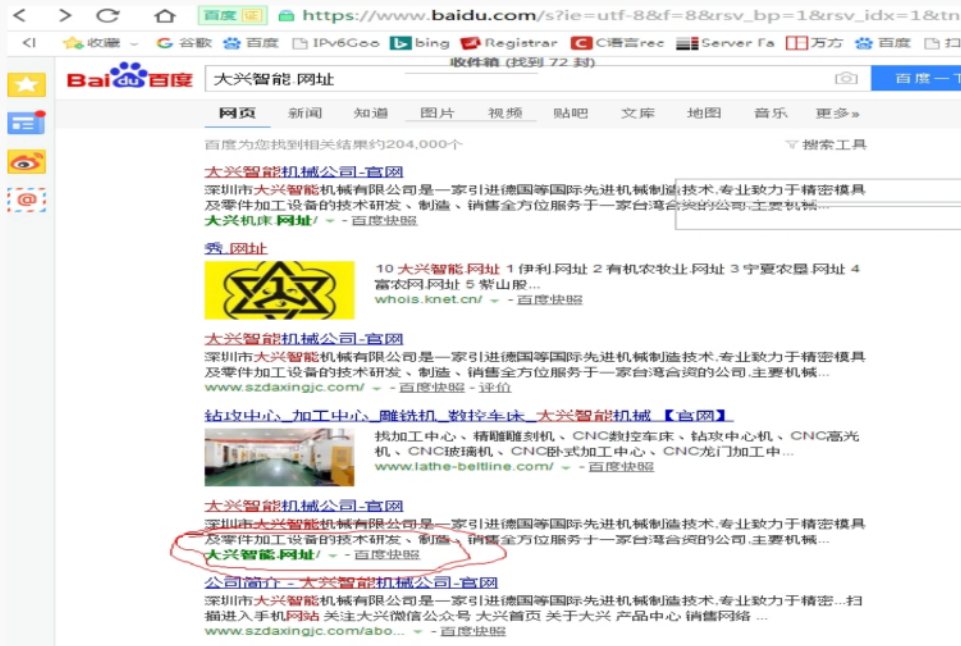
International Domain Names adoption

IDN registration is possible. IDN top-level domains like .pф, . 在线 operate. Country code domains like .ru, .cn, register IDNs. But adoption is stagnant — because UA is not yet widespread.

IDN adoption metrics

- * 11% of TLDs are IDNs (<http://data.iana.org/TLD/tlds-alpha-by-domain.txt>)
- * 7% of domains under ccTLDs are IDNs (APTLD: Status of IDNs in ccTLDs)
 - * 22% in .th. Also, above average in .tw, .hk, .jp
 - * 2.5% average in Europe
- * Microsoft alone holds 645 IDNs.

Baidu returns IDN domains in results!



Search for " 大兴智能 . 网址 " in Baidu (China's major search engine). Results include IDNs.

Scenario: justify IDN & EAI

Why bother, when IDN, EAI are so hard?

A CEO of one of the largest Company of India is asking me to give him three compelling reason on "Why Should he Adopt IDNs" while it create difficulties in marketing, communicating, handling and promoting different domains to his clients. —Ashish Modi

<https://mm.icann.org/pipermail/ua-discuss/2018-August/002690.html>

Opportunity [of adopting UA] & risk [of not]

'If you have global clients, partners, or affiliates they will require you to support EAI or move to a competitor that will....

Users may have EAI email addresses when they [contact] you.

...Opportunity and risk, often two sides of the same coin....

Supporting native language in IDN demonstrates commitment to local markets, quality, good customer support. Lack of support puts the brand at risk of being perceived as poor quality, hard to use, insensitive to local needs.' —Tex Texin

<https://mm.icann.org/pipermail/ua-discuss/2018-August/002707.html>

UA is speed-dial

'Most humans have challenges remembering phone numbers, and IP addresses are even more complex. Humans use speed dial on their phones... so that they don't have to dial all the numbers....

Domain names are like speed dial for accessing and connecting on the Internet. [They] make it possible for humans to reach each other and resources by using something familiar and easy to remember.

Humans have a variety of languages and alphabets that are familiar to them.... IDN unlocks an increased familiarity and affinity for humans.' —Jothan Frakes

<https://mm.icann.org/pipermail/ua-discuss/2018-August/002698.html>

Look at the street signage

'[In India] when I look at advertisement in the streets a large part of it - and in some parts of India "most of it" - is in the local script. My conclusion is that there must be a good reason for it. So why should the internet be the exception?' —Roberto Gaetano

<https://mm.icann.org/pipermail/ua-discuss/2018-August/002693.html>

How unusual they actually are

‘At Microsoft everyone I know speaks and reads English, no matter where on Earth they are from, and many of their families back “in the old country” are ASCII-capable too. They have trouble conceiving how unusual they actually are; your CEO friend might be having a similar dissonance. And “ASCII-capable” doesn’t mean ASCII-loving; I can read Katakana but if you are marketing to me you’d best use ASCII.’ —Mark Svancarek

<https://mm.icann.org/pipermail/ua-discuss/2018-August/002694.html>

Scenario: emails going to Junk



“Email from EAI addrs goes to Junk folder.”

An email admin gets complaint from users: email from EAI addresses goes to Junk folder, even when they label email from that address as “Not Junk”. —(inspired by query from Aman Masjide)

<https://mm.icann.org/pipermail/ua-discuss/2018-August/002675.html>

Percentage legit is “microscopic”

‘Regarding the newer TLDs, the % legitimate traffic registering them is microscopic. I’ve been told by a spammer that periodically domains in these TLDs are available for sale in bulk for cents and therefore very much loved by spammers (.bid, .club as well). Haven’t seen international TLDs abused yet in spam.

...We don’t have a deterministic rule to penalize these TLDs but it could be content filter ML is downgrading them (statistically for the filter it makes sense). [We and others] do have rules that increase the spam scoring of these TLDs or do more scrutiny.’ —webmail developer comments via Mark Svancarek

<https://mm.icann.org/pipermail/ua-discuss/2018-August/002675.html>

Percentage legit is “microscopic”

‘I don't know anyone who's blocking new TLDs just for being new, other than a few hobbyists. But if you think about it for a few minutes, the only people who want bulk TLDs are crooks, for spam or for phish and malware landing pages. It is entirely reasonable to block an otherwise little used TLD if you see a spike of abuse from it.’

—John Levine

<https://mm.icann.org/pipermail/ua-discuss/2018-August/002676.html>

Anti-spam is in scope for UASG

'We see the anti-abuse community as part of our target audience.

- 1) They should be aware of all the TLDs, the dynamic nature of the root zone population, and shouldn't be blocking an entire TLD because their systems are not aware of them.
- 2) If they block an entire TLD because it is a 'shady' TLD, that's NOT a UA Issue.
- 3) The bulk mail operators in the Anti-Abuse community should be aware of EAI Addresses....Our role is not to advocate for TLDs to not be blocked at the top level, but to ensure that those blocking entire TLDs are doing so consciously.' —Don Hollander

<https://mm.icann.org/pipermail/ua-discuss/2018-April/002451.html>

Acceptance without anti-abuse rings hollow

'Personally I have always believed these two areas to be deeply connected - and many new TLDs like Club and others have pretty strong anti-abuse standards.

Asking system administrators to bake in acceptance for new TLDs but then not being able to address the abuse issue - rings hollow for a lot of admins.

I would certainly be in favour of expanding the mandate to include a view on abuse etc.' —Dirk Bhagat

<https://mm.icann.org/pipermail/ua-discuss/2018-April/002453.html>

Getting domains to be worthy of support

'My take is that you cannot separate a charter to get everyone to support all domains equally from getting all domains to be equally worthy of support.

At some point, you'll just replace "structural" barriers to acceptance by black and gray lists.

... I think this group ignores the potential costs/risks of universal acceptance at its peril. "Poor anti-abuse stance" should not be rewarded with acceptance.' —Asmus Freytag

<https://mm.icann.org/pipermail/ua-discuss/2018-April/002454.html>

Block only subdomains

'[F]olks should not block entire TLDs, only subdomains. Which is what happens in all legacy TLDs. I doubt any admin is blocking ALL of .com, .biz, .uk etc.

My suggestion: make sure you accept any TLD by default, but feel free to blacklist any domain name that you consider harmful.'

—Paul Stahura [lightly edited]

<https://mm.icann.org/pipermail/ua-discuss/2018-April/002469.html>

Block only subdomains

'> make sure you accept any TLD by default, but feel free to blacklist any domain name that you consider harmful.

But there are TLDs with badness quotients of well over 50%. I'm all in favor of blacklisting the whole TLD as an "incentive" to come clean.

<https://www.spamhaus.org/statistics/tlds/>

—Asmus Freytag

<https://mm.icann.org/pipermail/ua-discuss/2018-April/002472.html>

In .gq, .cf, .tk, .ml, .ga, over 89% of their domains 'bad'. —spamhaus

Q&A

Thank you!

Q&A

Slides: <http://go.jdlh.com/iuc42s6t2>

Evaluation: <http://unicodeconference.org/eval-sp/>



Top Issues in Universal Acceptance

Jim DeLaHunt / IUC42 / 11 September 2018

Extra slides

Scenario: scary news article

Scenario: support the culture

Scenario: foreign email addresses