Universal Acceptance Steering Group
Quick Guide to Linkification – UASG010

10 May 2016

Background

Universal Acceptance (UA) 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.

Due to the rapidly changing domain name landscape, many systems do not recognize or appropriately process new domain names, primarily because they may be more than three characters in length or in a non-ASCII format. The same is true for email addresses that incorporate these new extensions.

The Universal Acceptance Steering Group (UASG), supported by Internet Corporation for Assigned Names and Numbers (ICANN), is a community-led, Internet industry-wide initiative working on creating awareness and identifying and resolving problems associated with the universal acceptance of domain names. The purpose of these efforts is to help ensure a consistent and positive experience for Internet users globally.

For more information on the UASG and recent development, visit www.uasg.tech.

Linkification and Its Challenges

Modern software sometimes allows a user to automatically create a hyperlink simply by typing in a string that looks like a web address, email name or network path. For example, typing "www.icann.org" into an email message may result in a clickable link to `http://www.icann.org` being automatically created if the application treats "www." as a special prefix or ".org" as a special suffix.

Linkification should work consistently for all well-formed web addresses, email names or network paths.

Linkification is the action where an application accepts a string and dynamically determines whether it should create a hyperlink to an Internet Location (URL) or an email address (mailto:)

Linkification uses algorithms and rules created by software developers to determine whether a string should be deemed a link – or not. Related to this is how people can identify a string as a domain name. While browsers, email clients and word processors are obvious places, there are many more applications that make these decisions.

Good Practice Recommendations

1. Attempt to linkify based on explicit protocol prefixes (e.g. "http://", ftp://", "mailto:") but only complete the action if the rest of the string is well formed

Example String	Expected Behavior/ Result
example.com	No linkification because protocol is absent and not inferred.
http://example.com	Create hyperlink because protocol is explicit
http:example.com	No linkification because of bad syntax (missing //)
http://example.a	No linkification because ICANN Policies require TLD to be at least two characters. NB: This syntax could be supported within an internal network.
http://example..ab	No linkification because of bad syntax (consecutive dots)
http:// 普遍接受-测试.世界	Create hyperlink because protocol is explicit.

2. Attempt to linkify based on implicit protocol prefixes (e.g. "www" infers "http://www")

Example String	Expected Behavior/ Result
www.example.com	Create hyperlink because protocol is implied[1]
label@example.com	Create mailto: label@example.com because protocol is implied.

3. Map the Ideographic Full Stop "。" (U+3002) to Full Stop "." (U+002E) (e.g. http://田中。com à http://田中.com) if string is otherwise well formed.

4. If TLDs are used as a 'special suffix' to determine linkability, then all TLDs must be included. A list of valid TLDs should be updated dynamically on a frequent basis.

---

[1] Note: it might be the case that the actual website requires that end users type https:// instead of http://. If this is the case, then the hyperlink may not resolve or may return an error page.