

Universal Acceptance of Domain Names

What is Universal Acceptance of Domain Names?

Internet users expect that any valid domain name and email address will be accepted, validated, stored, processed and displayed rationally and consistently by all Internet-enabled applications, devices and systems. However, due to the rapidly changing domain name landscape, many Internet systems do not recognize or appropriately process many of the new domain names that are being introduced into the root zone, primarily because they may be more than three characters in length or in a non-ASCII format. The same holds true for email addresses that incorporate these new extensions. This could result in an unsatisfactory user-experience and have a long-lasting negative effect on the domain name system and those who depend on it. The Universal Acceptance Steering Group (UASG), a community-led, industry-wide initiative, is working on identifying and resolving problems associated with Universal Acceptance of Domain Names to help ensure a consistent and positive experience for Internet users globally.

What is a Domain Name?

A domain name is a text string used as a human-friendly identifier for computers and networks on the Internet (e.g. www.domain.tld). The top-level domain (TLD), often called “suffix” or “extension” is at the right most end of the domain name (“.tld” in the above example). Domain names are also used in email addresses (e.g. user@domain.tld), which are often used as login names. An email address contains two parts, the local part that is directly before the “@” sign, followed by the domain name after the “@” sign.

Why is Universal Acceptance an issue now?

In the 1980s and 90s, the technical coordination of the Domain Name System (DNS) root zone – where all TLDs are recorded – followed a relatively simple pattern. General use top-level domain names were limited to a small number of three character names such as “.com” and “.edu,” as well as two-letter country codes top-level domains or (ccTLDs) to represent countries or territories. For example the ccTLD for Japan is “.jp.”

Between 2001 and 2010, the Internet community directed ICANN (Internet Corporation for Assigned Names and Numbers) to facilitate the introduction of new TLDs, many longer than three characters. In addition, TLDs in different scripts, called Internationalized Domain Names or IDNs were added to the DNS root starting in 2010. An example of an IDN is “.рф” for “Russia” in the Cyrillic language. In 2011, the New gTLD Program was launched, attracting hundreds of applications for new gTLDs (including IDN gTLDs) to be added to the DNS root. As of November 2015, nearly 800 new gTLDs have been added to the root zone.

While Universal Acceptance issues are not new, the addition of new TLDs brought over-simplified assumptions many software developers used in their applications to the forefront. The Universal Acceptance issue is analogous to the Y2K scenario that faced the technology industry in the late 1990s in that problems existed but were basically ignored until the potential consequences became too important to do so. Today, advancements in the DNS and the lack of Universal Acceptance are constraining Internet growth, consumer choice and consumer trust of the DNS. The UASG and Internet community must take steps now to ensure all valid domain names to function properly in all scenarios to support the continued expansion of the Internet.

What DNS advancements do developers need to be aware of?

There are several changes that have occurred and are driving the need for developers to reframe assumptions related to the DNS. These include:

- *The set of TLDs changes over time:* Many developers assume that the list of TLDs is relatively static. In

reality, this list is changing rapidly and there are now over 1,000 valid TLDs in the root zone.

Design consideration: The list of valid TLDs is now dynamic. Do your systems hard-code a list of valid TLDs? Or do they update from an authoritative source, such as ICANN?

- **The length of TLDs varies:** General use TLDs may now be longer than 2 or 3 characters in length. In fact, TLDs can be anywhere from 2 to 63 characters long.

Design consideration: Do your systems assume that all valid domain names and email addresses must have a TLD length at most 3 characters?

- **Multilingual domain names:** Prior to the introduction of IDNs, the character repertoire for valid domain names was limited to a subset of ASCII (e.g. a-z, 0-9 and hyphen '-').

Design consideration: Fields that accept domain names as input, such as URLs, email addresses, nameservers, etc., need to accept domain names in different scripts, languages and writing systems (left to right and right to left). Does your software accept, validate, store, process, and display domain names with internationalized characters?

How are TLDs and domain names verified?

If the application has access to the Internet, a straightforward way is simply to perform a DNS query. This ensures the most accurate and up-to-date data is returned from the most authoritative source – the DNS itself.

If it is not possible to use the DNS directly for validation, ICANN provides guidance on alternate methods. For example, the program could use a regularly updated list of valid top-level domains to perform the check. ICANN offers sample programming code that software developers can use for this purpose under an open source license at: <https://github.com/icann>.

What are some Do's & Don'ts developers should be aware of?

- **DO** use an IDN library to properly identify and handle Internationalized Domain Names. IDN support requires the application of specific algorithms to convert between an ASCII-encoded version of the name (A-Label) and a Unicode version (U-Label). There are many software libraries that can be used to implement this functionality. Make sure the library supports the most current standards (e.g. IDNA2008), as older standards introduce compatibility issues.
- **DO** make sure TLD lists are regularly updated using appropriate methodology. List should be updated weekly or more frequently. ICANN provides an authoritative source.
- **DO** avoid over-simplification of assumptions. The Universal Acceptance Steering Group can provide advice to software developers and implementers on what is needed. Contact us at: tld-acceptance@icann.org and/or join the Universal Acceptance discussion at <https://mm.icann.org/mailman/listinfo/ua-discuss>.
- **DO NOT** use a hard-coded list of TLDs in your application. If you need to check if a domain exists, the best way to do it is to query the DNS directly. A live DNS query happens very quickly and will provide your application with the best data available.
- **DO NOT** simply rely on the length of a domain to determine validity. Domain names can be up to 63 characters long.

More information

To learn more about the effort, visit us at: <http://www.icann.org/universalacceptance>.

To share your ideas and suggestions on the topic email us at: tld-acceptance@icann.org.