Standards and Other Documents Related to Universal Acceptance

9 May 2019

Standards and Other Documents Related to Universal Acceptance

UA006
Last Revised 2019-05-09

IDNA	
Internationalized Domain Names for Applications (IDNA): Definitions and Document Framework	https://tools.ietf.org/html/rfc5890
Internationalized Domain Names in Applications (IDNA): Protocol	https://tools.ietf.org/html/rfc5891
The Unicode Code Points and Internationalized Domain Names for Applications (IDNA)	https://tools.ietf.org/html/rfc5892
Right-to-Left Scripts for Internationalized Domain Names for Applications (IDNA)	https://tools.ietf.org/html/rfc5893
Internationalized Domain Names for Applications (IDNA): Background, Explanation, and Rationale	https://tools.ietf.org/html/rfc5894
Mapping Characters for Internationalized Domain Names in Applications (IDNA) 2008	https://tools.ietf.org/html/rfc5895
EAI	
Overview and Framework for Internationalized Email	https://tools.ietf.org/html/rfc6530
SMTP Extension for Internationalized Email	https://tools.ietf.org/html/rfc6531
Internationalized Email Headers	https://tools.ietf.org/html/rfc6532
Internationalized Delivery Status and Disposition Notifications	https://tools.ietf.org/html/rfc6533
IMAP Support for UTF-8	https://tools.ietf.org/html/rfc6855

Post Office Protocol Version 3 (POP3) Support for UTF-8	https://tools.ietf.org/html/rfc6856
Post-Delivery Message Downgrading for Internationalized Email Messages	https://tools.ietf.org/html/rfc6857
Simplified POP and IMAP Downgrading for Internationalized Email	https://tools.ietf.org/html/rfc6858
IRIs	
Internationalized Resource Identifiers (IRIs)	https://tools.ietf.org/html/rfc3987
Other	
Punycode: A Bootstring Encoding of Unicode	https://tools.ietf.org/html/rfc3492
Unicode & Other Documents	
UAX15 Unicode Normalization Forms	https://www.unicode.org/reports/tr15/
UTS36 Unicode Security Considerations	https://www.unicode.org/reports/tr36/
UTS39 Unicode Security Mechanisms	https://www.unicode.org/reports/tr39/
UTS46 Unicode IDNA Compatibility Processing	https://www.unicode.org/reports/tr46/
W3C Documents on Text Layout	Chinese: https://www.w3.org/TR/2018/WD-clreq-20181220/ Arabic: https://www.w3.org/TR/2018/WD-alreq-20180222/ Indic: https://www.w3.org/TR/2017/WD-ilreq-20170220/ Hangul: https://www.w3.org/TR/2015/WD-klreq-20150723/
SSAC Advisory on the Use of Emoji in Domain Name	https://www.icann.org/en/system/files/files/sac-095-en.pdf

Internationalized Domain Names for Applications (IDNA): Definitions and Document Framework Abstract: This document is part of a collection that, together, describes the protocol and usage	https://tools.ietf.org/html/rfc5890

context for a revision of Internationalized Domain Names for Applications (IDNA), superseding the earlier version. It describes the document collection and provides definitions and other materials that are common to the set.	
Internationalized Domain Names in Applications (IDNA): Protocol Abstract: This document is the revised protocol definition for Internationalized Domain Names (IDNs). The rationale for changes, relationship to the older specification, and important terminology are provided in other documents. This document specifies the protocol mechanism, called Internationalized Domain Names in Applications (IDNA), for registering and looking up IDNs in a way that does not require changes to the DNS itself. IDNA is only meant for processing domain names, not free text.	https://tools.ietf.org/html/rfc5891
The Unicode Code Points and Internationalized Domain Names for Applications (IDNA) Abstract: This document specifies rules for deciding whether a code point, considered in isolation or in context, is a candidate for inclusion in an Internationalized Domain Name (IDN). It is part of the specification of Internationalizing Domain Names in Applications 2008 (IDNA2008).	https://tools.ietf.org/html/rfc5892
Right-to-Left Scripts for Internationalized Domain Names for Applications (IDNA) Abstract: The use of right-to-left scripts in Internationalized Domain Names (IDNs) has presented several challenges. This memo provides a new Bidi rule for Internationalized Domain Names for Applications (IDNA) labels, based on the problems encountered with some scripts as well as shortcomings in the 2003 IDNA Bidi criterion.	https://tools.ietf.org/html/rfc5893

Internationalized Domain Names for Applications (IDNA): Background, Explanation, and Rationale Abstract: Several years have passed since the original protocol for Internationalized Domain Names (IDNs) was completed and deployed. During that time, a number of issues have surfaced, including the need to update the system to deal with newer versions of Unicode. Some of these issues require tuning of the existing protocols and tables on which they depend. This document provides an overview of a revised system and provides explanatory material for its components.	https://tools.ietf.org/html/rfc5894
Mapping Characters for Internationalized Domain Names in Applications (IDNA) 2008 Abstract: In the original version of the Internationalized Domain Names in Applications (IDNA) protocol, any Unicode code points taken from user input were mapped into a set of Unicode code points that "made sense" and then encoded and passed to the Domain Name System (DNS). The IDNA2008 protocol (described in RFCs 5890, 5891, 5892, and 5893) presumes that the input to the protocol comes from a set of "permitted" code points, which it then encodes and passes to the DNS, but does not specify what to do with the result of user input. This document describes the actions that can be taken by an implementation between receiving user input and passing permitted code points to the new IDNA protocol.	https://tools.ietf.org/html/rfc5895
Overview and Framework for Internationalized Email Abstract: Full use of electronic mail throughout the world requires that (subject to other constraints) people be able to use close variations on their own names (written correctly in their own languages and scripts) as	https://tools.ietf.org/html/rfc6530

mailbox names in email addresses. This document introduces a series of specifications that define mechanisms and protocol extensions needed to fully support internationalized email addresses. These changes include an SMTP extension and extension of email header syntax to accommodate UTF-8 data. The document set also includes discussion of key assumptions and issues in deploying fully internationalized email. This document is a replacement for RFC 4952; it reflects additional issues identified since that document was published.	
SMTP Extension for Internationalized Email Abstract: This document specifies an SMTP extension for transport and delivery of email messages with internationalized email addresses or header information.	https://tools.ietf.org/html/rfc6531
Internationalized Email Headers Abstract: Internet mail was originally limited to 7-bit ASCII. MIME added support for the use of 8-bit character sets in body parts, and also defined an encoded-word construct so other character sets could be used in certain header field values. However, full internationalization of electronic mail requires additional enhancements to allow the use of Unicode, including characters outside the ASCII repertoire, in mail addresses as well as direct use of Unicode in header fields like "From:", "To:", and "Subject:", without requiring the use of complex encoded-word constructs. This document specifies an enhancement to the Internet Message Format and to MIME that allows use of Unicode in mail addresses and most header field content. This specification updates Section 6.4 of RFC 2045 to eliminate the restriction prohibiting the use of non-identity content-transfer- encodings on subtypes of "message/".	https://tools.ietf.org/html/rfc6532
Internationalized Delivery Status and Disposition Notifications	https://tools.ietf.org/html/rfc6533

Abstract: Delivery status notifications (DSNs) are critical to the correct operation of an email system. However, the existing Draft Standards (RFC 3461, RFC 3464, RFC 6522) are currently limited to ASCII text in the machine-readable portions of the protocol. This specification adds a new address type for international email addresses so an original recipient address with non-ASCII characters can be correctly preserved even after downgrading. This also provides updated content return media types for delivery status notifications and message disposition notifications to support use of the new address type. This document extends RFC 3461, RFC 3464, RFC 3798, and RFC 6522.	
IMAP Support for UTF-8 Abstract: This specification extends the Internet Message Access Protocol (IMAP) to support UTF-8 encoded international characters in user names, mail addresses, and message headers. This specification replaces RFC 5738.	https://tools.ietf.org/html/rfc6855
Post Office Protocol Version 3 (POP3) Support for UTF-8 Abstract: This specification extends the Post Office Protocol version 3 (POP3) to support international strings encoded in UTF-8 in usernames, passwords, mail addresses, message headers, and protocol-level text strings.	https://tools.ietf.org/html/rfc6856
Post-Delivery Message Downgrading for Internationalized Email Messages Abstract: The Email Address Internationalization (SMTPUTF8) extension to SMTP allows Unicode characters encoded in UTF-8 and outside the ASCII repertoire in mail header fields. Upgraded POP and IMAP servers support internationalized messages. If a POP or IMAP client does not support Email Address Internationalization, a POP or IMAP server	https://tools.ietf.org/html/rfc6857

cannot deliver internationalized messages to the client and cannot remove the message. To avoid that situation, this document describes a mechanism for converting internationalized messages into the traditional message format. As part of the conversion process, message elements that require internationalized treatment are recoded or removed, and receivers are able to recognize that they received messages containing such elements, even if they cannot process the internationalized elements.	
Simplified POP and IMAP Downgrading for Internationalized Email Abstract: This document specifies a method for IMAP and POP servers to serve internationalized messages to conventional clients. The specification is simple, easy to implement, and provides only rudimentary results.	https://tools.ietf.org/html/rfc6858
Internationalized Resource Identifiers (IRIs) Abstract: This document defines a new protocol element, the Internationalized Resource Identifier (IRI), as a complement to the Uniform Resource Identifier (URI). An IRI is a sequence of characters from the Universal Character Set (Unicode/ISO 10646). A mapping from IRIs to URIs is defined, which means that IRIs can be used instead of URIs, where appropriate, to identify resources. The approach of defining a new protocol element was chosen instead of extending or changing the definition of URIs. This was done in order to allow a clear distinction and to avoid incompatibilities with existing software. Guidelines are provided for the use and deployment of IRIs in various protocols, formats, and software components that currently deal with URIs.	https://tools.ietf.org/html/rfc3987
Punycode: A Bootstring Encoding of Unicode for Internationalized Domain Names in Applications (IDNA)	https://tools.ietf.org/html/rfc3492

Abstract: Punycode is a simple and efficient transfer encoding syntax designed for use with Internationalized Domain Names in Applications (IDNA). It uniquely and reversibly transforms a Unicode string into an ASCII string. ASCII characters in the Unicode string are represented literally, and non-ASCII characters are represented by ASCII characters that are allowed in host name labels (letters, digits, and hyphens). This document defines a general algorithm called Bootstring that allows a string of basic code points to uniquely represent any string of code points drawn from a larger set. Punycode is an instance of Bootstring that uses particular parameter values specified by this document, appropriate for IDNA.	
UAX15 Unicode Normalization Forms This annex describes normalization forms for Unicode text. When implementations keep strings in a normalized form, they can be assured that equivalent strings have a unique binary representation. This annex also provides examples, additional specifications regarding normalization of Unicode text, and information about conformance testing for Unicode normalization forms. (Defines NFC used in U-labels)	https://www.unicode.org/reports/tr15/
UTS36 Unicode Security Considerations Because Unicode contains such a large number of characters and incorporates the varied writing systems of the world, incorrect usage can expose programs or systems to possible security attacks. This is especially important as more products are internationalized. This document describes some of the security considerations that programmers, system analysts, standards developers, and users should take into account, and provides specific recommendations to reduce the risk of attacks.	https://www.unicode.org/reports/tr36/
UTS39 Unicode Security Mechanisms	https://www.unicode.org/reports/tr39/

Because Unicode contains such a large number of characters and incorporates the varied writing systems of the world, incorrect usage can expose programs or systems to possible security attacks. This document specifies mechanisms that can be used to detect possible security problems. (Defines the script restriction levels we suggest people use when creating mailbox names)	
UTS46 Unicode IDNA Compatibility Processing Client software, such as browsers and emailers, face a difficult transition from the version of Internationalized Domain Names approved in 2003 (IDNA2003), to the revision approved in 2010 (IDNA2008). The specification in this document provides a mechanism that minimizes the impact of this transition for client software, allowing client software to access domains that are valid under either system. The specification provides two main features. One is a comprehensive mapping to support current user expectations for casing and other variants of domain names (such mapping is allowed by IDNA2008). The second is a compatibility mechanism that supports the existing domain names that were allowed under IDNA2003. This second feature is intended to improve client behavior during the transitional period.	https://www.unicode.org/reports/tr46/
W3C Documents on Text Layout Chinese: https://www.w3.org/TR/2018/WD-clreq-20181220/ Arabic: https://www.w3.org/TR/2018/WD-alreq-20180222/ Indic: https://www.w3.org/TR/2017/WD-ilreq-20170220/ Hangul: https://www.w3.org/TR/2015/WD-klreq-20150723/	
SSAC Advisory on the Use of Emoji in Domain Names – SAC095	https://www.icann.org/en/system/files/files/sac-095-en.pdf

Due to their popularity, there have been questions and discussions on the use of emoji in domain names. The Security and Stability Advisory Committee (SSAC) has studied this question and concluded that the use of emoji in any label of a domain name should be discouraged. In this advisory, the SSAC explains its reasoning.	

◼ Audience

This document is aimed at software engineers who need to find relevant original standards.

◼ Document History

Version	Date	Authors	Notes
1.1	2019-05-06	Don Hollander	Added additional Unicode, W3C and SSAC references
1.0	2017-17-11	Don Hollander	Initial list of relevant IETF RFCs