



Universal Acceptance

Стандарты и другие документы, касающиеся универсального принятия

09.05.2019



Стандарты и другие документы, касающиеся универсального принятия

UA006

Последние изменения: 09.05.2019

IDNA	
Интернационализованные доменные имена в приложениях (IDNA): Определения и структура документа	https://tools.ietf.org/html/rfc5890
Интернационализованные доменные имена в приложениях (IDNA): Протокол	https://tools.ietf.org/html/rfc5891
Кодовые точки Unicode и интернационализованные доменные имена в приложениях (IDNA)	https://tools.ietf.org/html/rfc5892
Алфавиты с направлением письма справа налево в интернационализованных доменных именах в приложениях (IDNA)	https://tools.ietf.org/html/rfc5893
Интернационализованные доменные имена в приложениях (IDNA): история вопроса, пояснение и обоснование	https://tools.ietf.org/html/rfc5894
Сопоставление символов интернационализованных доменных имен в приложениях (IDNA), 2008	https://tools.ietf.org/html/rfc5895
EAI	
Общие сведения и концепция интернационализации электронной почты	https://tools.ietf.org/html/rfc6530
Расширения протокола SMTP для интернационализации электронной почты	https://tools.ietf.org/html/rfc6531
Заголовки сообщений при интернационализации электронной почты	https://tools.ietf.org/html/rfc6532



Интернационализация уведомлений о доставке и расположении сообщений	https://tools.ietf.org/html/rfc6533
Поддержка IMAP для UTF-8	https://tools.ietf.org/html/rfc6855
Поддержка протокола POP3 для UTF-8	https://tools.ietf.org/html/rfc6856
Понижение версии сообщения после доставки при интернационализации электронной почты	https://tools.ietf.org/html/rfc6857
Упрощенное понижение версии протоколов POP и IMAP при интернационализации электронной почты	https://tools.ietf.org/html/rfc6858
IRI	
Интернационализированные идентификаторы ресурсов (IRI)	https://tools.ietf.org/html/rfc3987
Другое	
Punycode: Bootstring-кодирование Unicode	https://tools.ietf.org/html/rfc3492
Unicode и другие документы	
UAX15 Формы нормализации Unicode	https://www.unicode.org/reports/tr15/
UTS36 Соображения о безопасности Unicode	https://www.unicode.org/reports/tr36/
UTS39 Механизмы безопасности Unicode	https://www.unicode.org/reports/tr39/
UTS46 Сопоставительная обработка IDNA Unicode	https://www.unicode.org/reports/tr46/
Документы Международного консорциума всемирной сети Интернет по оформлению текста	Китайский: https://www.w3.org/TR/2018/WD-clreq-20181220/ Арабский: https://www.w3.org/TR/2018/WD-alreq-20180222/ Индийский: https://www.w3.org/TR/2017/WD-ilreq-20170220/ Хангиль: https://www.w3.org/TR/2015/WD-klreq-20150723/
Пояснение SSAC по использованию эмодзи в доменных именах	https://www.icann.org/en/system/files/files/sac-095-en.pdf



<p>Интернационализованные доменные имена в приложениях (IDNA): Определения и структура документа</p> <p>Выдержка: Данный документ входит в набор документов, которые в совокупности описывают протокол и контекст использования интернационализованных доменных имен в приложениях (IDNA) в очередной редакции, заменяющей собой предыдущую версию. В нем описан набор документов и представлены определения и прочие материалы, общие для всех документов из данного набора.</p>	<p>https://tools.ietf.org/html/rfc5890</p>
<p>Интернационализованные доменные имена в приложениях (IDNA): Протокол</p> <p>Выдержка: Данный документ представляет собой пересмотренные определения протокола интернационализованных доменных имен (IDN). Обоснование внесенных изменений, то, как они относятся к предыдущей версии спецификации, а также важные термины представлены в других документах. В настоящем документе описан механизм данного протокола, который называется «интернационализованные доменные имена в приложениях» (IDNA), предназначенный для записи и подстановки интернационализованных доменных имен без необходимости их изменения. Механизм IDNA предназначен для обработки только доменных имен, а не любого набора символов.</p>	<p>https://tools.ietf.org/html/rfc5891</p>
<p>Элементы Unicode и интернационализованные доменные имена для приложений (IDNA)</p> <p>Выдержка: В данном документе указаны правила принятия решений о том, может ли та или иная кодовая точка с учетом или без учета контекста использоваться в составе</p>	<p>https://tools.ietf.org/html/rfc5892</p>



<p>интернационализованного доменного имени (IDN).</p> <p>Это часть спецификации интернационализованных доменных имен в приложениях от 2008 года (IDNA2008).</p>	
<p>Алфавиты с направлением письма справа налево в интернационализованных доменных именах в приложениях (IDNA)</p> <p>Выдержка: Использование систем письма справа налево в интернационализованных доменных именах (IDN) сопряжено с рядом проблем. В настоящей записке представлено новое правило двунаправленного (Bidi) анализа для меток интернационализованных доменных имен в приложениях (IDNA), составленное с учетом опыта проблем, с которыми столкнулись пользователи некоторых систем письма, а также с учетом недостатков критерия Bidi IDNA от 2003 года.</p>	<p>https://tools.ietf.org/html/rfc5893</p>
<p>Интернационализованные доменные имена в приложениях (IDNA): история вопроса, пояснение и обоснование</p> <p>Выдержка: Прошло уже несколько лет с того времени, как была закончена работа и выполнено развертывание исходного протокола интернационализованных доменных имен.</p> <p>За это время обнаружился ряд проблем, в том числе необходимость обновления этой системы для учета более новых версий Unicode. Некоторые из указанных проблем требуют настройки существующих протоколов и таблиц, от которых они зависят. В этом документе представлен обзор пересмотренной системы и пояснительные материалы к ее компонентам.</p>	<p>https://tools.ietf.org/html/rfc5894</p>



<p>Сопоставление символов интернационализированных доменных имен в приложениях (IDNA), 2008</p> <p>Выдержка: В исходной версии протокола интернационализированных доменных имен в приложениях (IDNA) любая кодовая точка Unicode, полученная с пользовательского ввода, сопоставлялась с набором «понятных» кодовых точек Unicode, а затем кодировалась и передавалась в систему доменных имен (DNS).</p> <p>Протокол IDNA2008 (описанный в документах RFC 5890, 5891, 5892 и 5893) исходит из того, что данные, поступающие на ввод протокола, содержат символы из набора «допустимых» кодовых точек, которые затем кодируются и передаются в DNS. Однако в описании протокола не указано, что нужно делать с результатами пользовательского ввода. В этом документе описываются действия, которые могут выполняться в том или ином примере реализации протокола после получения пользовательского ввода перед передачей допустимых кодовых точек в новый протокол IDNA.</p>	<p>https://tools.ietf.org/html/rfc5895</p>
<p>Общие сведения и концепция интернационализации электронной почты</p> <p>Выдержка: Для полноценного использования электронной почты по всему миру необходимо (помимо прочих ограничений), чтобы люди могли использовать варианты, близкие к написанию своих имен (правильно записанных в используемых ими языках и наборах символов) в качестве названий почтовых ящиков в адресах электронной почты. Этим документом вводится ряд спецификаций, определяющих механизмы и расширения протоколов, которые необходимы для полной</p>	<p>https://tools.ietf.org/html/rfc6530</p>



<p>поддержки интернационализированных адресов электронной почты. Эти изменения включают расширение протокола SMTP и расширение синтаксиса заголовков электронной почты для обеспечения поддержки данных в формате UTF-8. Данный документ также содержит обсуждение ключевых допущений и проблем, связанных с развертыванием полностью интернационализированной электронной почты. Данный документ заменяет собой документ RFC 4952; он отражает дополнительные проблемы, выявленные с момента публикации предыдущего документа.</p>	
<p>Расширения протокола SMTP для интернационализации электронной почты</p> <p>Выдержка: Этот документ определяет расширение протокола SMTP для передачи и доставки сообщений электронной почты с интернационализированными адресами или заголовками.</p>	<p>https://tools.ietf.org/html/rfc6531</p>
<p>Заголовки сообщений при интернационализации электронной почты</p> <p>Выдержка: Изначально почта в Интернете ограничивалась 7-разрядной кодировкой ASCII. С помощью MIME была добавлена поддержка использования 8-битных наборов символов в тексте сообщений, а также определена функция составления кодированных слов, обеспечивающая возможность использования других наборов символов в значениях определенных полей заголовков сообщений. Однако для полной интернационализации электронной почты необходимы дополнительные усовершенствования, которые позволили бы использовать Unicode, в т. ч. символы, не входящие в таблицу ASCII, в адресах электронной почты, а также использовать</p>	<p>https://tools.ietf.org/html/rfc6532</p>



<p>Unicode непосредственно в полях заголовков, таких как «Отправитель», «Кому» и «Тема», без необходимости применять сложные составные конструкции из закодированных слов. В этом документе указаны усовершенствования формата интернет-сообщений и MIME, позволяющие использовать Unicode в адресах электронной почты и значениях большинства полей заголовков.</p> <p>Эта спецификация обновляет раздел 6.4 документа RFC 2045 для снятия ограничения, запрещающего использования закодированных для передачи неидентификационных данных в дочерних типах объекта "message/".</p>	
<p>Интернационализация уведомлений о доставке и расположении сообщений</p> <p>Выдержка: Уведомления о доставке сообщений (DSN) являются жизненно важным фактором надлежащей работы систем электронной почты. Однако существующие проекты стандартов (RFC 3461, RFC 3464, RFC 6522) в настоящее время содержат ограничение, требующее использования только текста в кодировке ASCII в тех частях протокола, которые предназначены для машинного считывания. Этой спецификацией добавляется новый тип интернационализованных адресов электронной почты, позволяющий корректно сохранять исходный адрес получателя, содержащий символы, отличные от ASCII, после понижения версии поддерживаемого стандарта. Кроме того, в ней представлен обновленный список типов данных возвращаемого содержания для уведомлений о доставке и расположении сообщений для поддержки использования нового типа адресов. Этот документ является расширением стандартов RFC 3461, RFC 3464, RFC 3798 и RFC 6522.</p>	<p>https://tools.ietf.org/html/rfc6533</p>



<p>Поддержка IMAP для UTF-8</p> <p>Выдержка: Данная спецификация расширяет протокол IMAP для поддержки для символов различных языков, закодированных в формате UTF-8, в именах пользователей, адресах электронной почты и заголовках сообщений. Данная спецификация заменяет собой документ RFC 5738.</p>	<p>https://tools.ietf.org/html/rfc6855</p>
<p>Поддержка протокола POP3 для UTF-8</p> <p>Выдержка: Данная спецификация расширяет протокол POP3 для поддержки строк, записанных символами различных языков и закодированных в формате UTF-8, в именах пользователей, паролях, адресах электронной почты, заголовках сообщений и текстовых строках уровня протокола.</p>	<p>https://tools.ietf.org/html/rfc6856</p>
<p>Понижение версии сообщения после доставки при интернационализации электронной почты</p> <p>Выдержка: Расширение протокола SMTP для интернационализации адресов электронной почты (SMTPUTF8) позволяет использовать в полях заголовков сообщений электронной почты символы Unicode, закодированные в формате UTF-8 и не совпадающие с символами таблицы ASCII. Обновленные серверы POP и IMAP поддерживают интернационализированные сообщения. Если клиент протокола POP или IMAP не поддерживает интернационализацию адресов электронной почты, сервер POP или IMAP не может доставить клиенту интернационализированные сообщения и удалить их с сервера. Для недопущения подобных ситуаций в данном документе описан механизм преобразования интернационализированных сообщений в традиционный формат сообщений электронной почты. В рамках процесса такого преобразования записываются или удаляются те элементы сообщения, для</p>	<p>https://tools.ietf.org/html/rfc6857</p>



<p>которых необходима поддержка интернационализации электронной почты, что позволит получателю понять, что полученное сообщение содержит такие элементы, даже если клиент не сможет их обработать.</p>	
<p>Упрощенное понижение версии протоколов POP и IMAP при интернационализации электронной почты</p> <p>Выдержка: В этом документе указан метод, с помощью которого серверы IMAP и POP могут обеспечивать доставку интернационализированных сообщений клиентам, не поддерживающим такую возможность. Эта спецификация отличается простотой и легкостью реализации и обеспечивает только самые базовые возможности.</p>	<p>https://tools.ietf.org/html/rfc6858</p>
<p>Интернационализированные идентификаторы ресурсов (IRI)</p> <p>Выдержка: В этом документе определен новый элемент протокола — интернационализированный идентификатор ресурса (Internationalized Resource Identifier, IRI), дополняющий собой унифицированные идентификаторы ресурсов (Uniform Resource Identifier, URI). Идентификатор IRI — это последовательность символов из универсального набора символов стандарта Unicode/ISO 10646. Определены соответствия для сопоставления между собой идентификаторов IRI и URI, что означает, что при необходимости для обозначения ресурсов вместо идентификаторов URI можно использовать идентификаторы IRI.</p> <p>Вместо того, чтобы расширять или менять определение идентификаторов URI, был сделан выбор в пользу подхода, заключавшегося в определении нового элемента протокола. Это было сделано для того, что обеспечить четкое различие между</p>	<p>https://tools.ietf.org/html/rfc3987</p>



<p>этими элементами и не допустить несовместимости с существующим ПО. Представлены рекомендации по использованию идентификаторов IRI в различных протоколах, форматах и программных компонентах, в которых в настоящее время используются идентификаторы URI.</p>	
<p>Punycode: Bootstring-кодирование строк Unicode для интернационализированных доменных имен в приложениях (IDNA)</p> <p>Выдержка: Punycode — это простой и эффективный синтаксис кодирования символов для передачи их при использовании с интернационализированными доменными именами в приложениях (IDNA).</p> <p>Он позволяет преобразовать строку в формате Unicode в уникальную строку в кодировке ASCII и назад. Символы ASCII в строке Unicode представлены в исходном виде, а символы, не входящие в таблицу ASCII, представлены символами ASCII, разрешенными для использования в метках имен узлов (буквы, цифры и дефис). В настоящем документе определен общий алгоритм Bootstring, который позволяет использовать строки базовых кодовых точек для уникального отображения любых строк кодовых точек, относящихся к большому набору кодовых точек. Punycode — это частный случая использования алгоритма Bootstring с определенными значениями параметров, указанными в этом документе, для использования с интернационализированными доменными именами в приложениях (IDNA).</p>	<p>https://tools.ietf.org/html/rfc3492</p>
<p>UAX15 Формы нормализации Unicode</p> <p>В данном приложении описаны формы нормализации текста Unicode. Если при реализации строки остаются в нормализованном виде, совершенно точно, что эквивалентные строки будут иметь двоичную репрезентацию. В данном</p>	<p>https://www.unicode.org/reports/tr15/</p>



<p>приложении также приведены примеры, дополнительные спецификации, касающиеся нормализации текста Unicode, и информация о проверке соответствия форм нормализации Unicode. (Определяет NFC в U-метках)</p>	
<p>UTS36 Соображения о безопасности Unicode</p> <p>Поскольку Unicode содержит огромное количество символов и включает разнообразные системы письма, некорректное использование может подвергнуть программы или системы угрозе атак на безопасность. Это особенно важно в связи с расширением интернационализации продуктов. Данный документ затрагивает ряд вопросов безопасности, которые необходимо принять во внимание программистам, системным аналитикам, разработчикам стандартов и пользователям, и дает конкретные рекомендации, направленные на снижение риска атак.</p>	<p>https://www.unicode.org/reports/tr36/</p>
<p>UTS39 Механизмы безопасности Unicode</p> <p>Поскольку Unicode содержит огромное количество символов и включает разнообразные системы письма, некорректное использование может подвергнуть программы или системы угрозе атак на безопасность. В данном документе определены механизмы, которые можно использовать для выявления возможных проблем безопасности.</p> <p>(Определяет уровни ограничения алфавита, которые мы предлагаем использовать при создании имен почтовых ящиков)</p>	<p>https://www.unicode.org/reports/tr39/</p>
<p>UTS46 Сопоставительная обработка IDNA Unicode</p> <p>Клиентское программное обеспечение, например браузеры и средства рассылки электронной почты, сталкивается с проблемой перехода от версии</p>	<p>https://www.unicode.org/reports/tr46/</p>



<p>интернационализованных доменных имен, утвержденной в 2003 г. (IDNA2003), к редакции, утвержденной в 2010 г. (IDNA2008). В спецификации в настоящем документе приведен механизм, минимизирующий влияние этого перехода на клиентское ПО, что обеспечивает ему доступ к доменам, действующим в любой из этих систем.</p> <p>Спецификация обеспечивает две основных функции. Первая — тщательное сопоставление в поддержку имеющихся ожиданий пользователя в отношении регистра и других вариантов доменных имен (такое сопоставление обеспечивается IDNA2008). Вторая — механизм совместимости, поддерживающий существующие доменные имена, которые предоставлялись в соответствии с IDNA2003. Эта вторая функция необходима для того, чтобы повысить функциональные возможности клиента в переходный период.</p>	
<p>Документы Международного консорциума всемирной сети Интернет по оформлению текста</p> <p>Китайский: https://www.w3.org/TR/2018/WD-clreq-20181220/ Арабский: https://www.w3.org/TR/2018/WD-alreq-20180222/ Индийский: https://www.w3.org/TR/2017/WD-ilreq-20170220/ Хангиль: https://www.w3.org/TR/2015/WD-klreq-20150723/</p>	
<p>Пояснение SSAC по использованию эмодзи в доменных именах – SAC095</p> <p>В связи с высокой популярностью эмодзи возникли вопросы и обсуждения их использования в доменных именах. Консультативный комитет по безопасности и стабильности (SSAC) изучил вопрос и пришел к выводу, что использование эмодзи в каких-либо метках доменных имен</p>	<p>https://www.icann.org/en/system/files/files/sac-095-en.pdf</p>



не рекомендуется. В данном пояснении SSAC излагает свою аргументацию.	
---	--

■ **Целевая аудитория**

Данный документ предназначен для специалистов по программному обеспечению, которым необходимо найти соответствующие исходные стандарты.

■ **История документа**

Версия	Дата	Авторы:	Примечания
1.1	2019-05-06	Дон Холландер (Don Hollander)	Добавлен дополнительный Unicode, ссылки на W3C и SSAC2
1.0	2017-17-11	Дон Холландер (Don Hollander)	Исходный список соответствующих RFC IETF