



КООРДИНАЦИОННЫЙ ЦЕНТР
НАЦИОНАЛЬНОГО ДОМЕНА
СЕТИ ИНТЕРНЕТ

IDN Universal Acceptance

Состояние вопроса

Андрей Воробьев
a.vorobyev@cctld.ru



ПРИНЯТИЕ



ПРОВЕРКА



ХРАНЕНИЕ



ОБРАБОТКА



ОТОБРАЖЕНИЕ

Универсальное принятие (UA) — состояние, когда все интернет-приложения, устройства и системы правильно и единообразно принимают, проверяют, хранят, обрабатывают и отображают все корректные доменные имена и адреса электронной почты.

ПРИНЯТИЕ



Принятие — это получение адреса электронной почты или доменного имени в виде строки символов из пользовательского интерфейса, файла или API (интерфейса программирования приложений)

Рекомендации UASG

Любой пользовательский интерфейс обязательно должен поддерживать Unicode и строки длиной до 256 символов.

Пользователям следует разрешать вводить текст в ASCII-совместимой кодировке вместо Unicode. Для отображения по умолчанию следует использовать Unicode, а текст в кодировке Punycode показывать пользователю только тогда, когда это приносит выгоду.

Текущее состояние

Крупнейшие платформы, обслуживающие 95% мировых пользователей интернета*, корректно отображают Unicode и позволяют вводить ASCII-совместимые значения

*по данным UASG

ПРОВЕРКА



Проверка синтаксической правильности полученного или отправленного адреса электронной почты или доменного имени. Часто основана на предположении, что в домене ограниченное количество символов, и все они ASCII. Сейчас это не применимо, так как появились IDN и New gTLD.

Рекомендации UASG

- Проверка должна выполняться только если это необходимо для работы сервиса
- Проверьте соответствие правой части доменного имени официальной таблице на ресурсах ICANN, IANAСм.
- Отправьте в систему доменных имен (DNS) запрос для проверки доменного имени.
- Для предотвращения опечаток требуйте повторного ввода адреса электронной почты.

Текущее состояние

В настоящее время даже некоторые крупные интернет-компании не принимают адреса, отличные от традиционных



I K R A

ХРАНЕНИЕ



Независимо от периода хранения данных, они должны (предпочтительно) храниться в форматах, определяемых стандартами RFC, или в других форматах, которые могут быть преобразованы в форматы RFC.

Рекомендации UASG

- Приложения и сервисы должны поддерживать стандарты Unicode.
- Информацию следует хранить по мере возможности в UTF-8. Других форматов следует избегать.
- Возможно, целесообразно хранить в файле или базе данных только U-метки, поскольку это упрощает поиск и сортировку. Однако преобразование может оказать влияние при взаимодействии с более старыми приложениями и сервисами, не поддерживающими Unicode. Рассмотрите возможность хранения в обоих форматах

Текущее состояние

Часто интернет-сервисы хранят адреса только в Punicode-виде, что является источником ошибок при сравнении адресов с базой данных

ОБРАБОТКА



Обработка происходит всякий раз, когда адрес электронной почты или доменное имя используется приложением или сервисом для выполнения действия (например, поиска или сортировки списка) или преобразуется в альтернативный формат (например, при хранении ASCII как Unicode).

Рекомендации UASG

- Стандарт Unicode расширяется, следует своевременно обновлять настройки. По мере возможности выполняйте обработку в формате UTF-8.
- Убедитесь, что продукт или функция обрабатывает числа так, как это предусмотрено. Например, числовые символы ASCII и азиатские иероглифы, представляющие числа, должны обрабатываться как числа.
- Обновляйте приложения и серверы/сервисы одновременно. Иначе придется преобразовывать данные из одной кодовой страницы в другую при выполнении каждой операции передачи данных между сервером и клиентской частью.

Текущее состояние

Самые частые ошибки разработки – несинхронное обновление, обработка в других форматах, несвоевременное преобразование

ОТОБРАЖЕНИЕ



Отображение – обработка адреса или домена в пользовательском интерфейсе. Отображение не вызывает затруднений, когда используемые алфавиты поддерживаются базовой ОС, а строки хранятся в Unicode; однако в других случаях может возникнуть необходимость преобразований, зависящих от типа приложения.

Рекомендации UASG

- Необходимо обеспечить полную поддержку Unicode для коллекции шрифтов, доступных в операционной системе.
- При разработке обеспечивайте охват всех языков ОС и приложениями.
- По умолчанию отображайте Unicode. Punycode используется только тогда, когда это обосновано необходимо.

Текущее состояние

Одна из главных проблем IDN-доменов и нелатинских адресов электронной почты – ошибка в их отображении, что ухудшает пользовательский опыт

Спасибо!