



Ускоренная программа ввода доменов IDN ccTLD Предлагаемые детали внедрения, касающиеся: разработки и использования таблиц имен IDN и вариантов символов для строк доменов высшего и второго уровней

Дата публикации:

4 февраля 2008 г.

История – ускоренный процесс ввода доменов IDN ccTLD

Одной из важных инноваций в Интернете с момента его создания станет ввод многоязычных доменных имен (IDN) на высшем уровне. Имена IDN открывают много новых возможностей и преимуществ для пользователей Интернета, говорящих на разных языках по всему миру, позволяя создавать и использовать домены на родном языке и алфавите.

Тема имен IDN обсуждается сообществом ICANN уже несколько лет. Изначально разработка акцентировалась на предоставлении возможности ввода имен IDN путем их регистрации в существующих доменах TLD, но за прошедший год фокус сместился на расширение набора символов, доступных для использования в строках высшего уровня. Ускоренный процесс ввода доменов IDN ccTLD является одним из процессов, над которыми работает корпорация ICANN для обеспечения этого ввода. Процесс внедрения новых доменов gTLD будет также поддерживать многоязычные домены высшего уровня, входящие в [программу ввода новых доменов gTLD](#).

Первые шаги ввода доменов IDN ccTLD были [инициированы Советом директоров корпорации ICANN](#) на собрании в Сан-Паулу (декабрь 2006 г.) Во время консультаций и обсуждений совместной рабочей группы по именам IDN комитета GAC и организации ccNSO стало понятно, что у ряда стран и территорий существует крайняя необходимость в доменах IDN ccTLD. Ускоренный процесс ввода доменов IDN ccTLD в особенности направлен на удовлетворение краткосрочных потребностей и получение опыта работы с механизмами выбора и авторизации подобных доменов TLD, который позволит предоставлять информацию для проводимого долгосрочного процесса разработки политики

Внедрение ускоренного процесса ввода доменов IDN ccTLD идет полным ходом и основывается на [итоговом отчете рабочей группы IDNC](#), в котором представлены рекомендации по механизмам ввода ограниченного числа неконфликтных доменов IDN ccTLD, связанных с двухбуквенными кодами стандарта ISO 3166-1. В [первом предварительном плане внедрения](#) для ускоренного процесса ввода имен IDN ccTLD перечислены открытые вопросы, для которых требуются сведения от сообщества и которые должны быть разрешены для завершения внедрения.

Данный документ входит в серию документов, описывающих предлагаемые решения этих открытых вопросов. Предлагаемые решения основываются на полученных общественных комментариях и отзывах, полученных во время встреч, например проведенных во время собрания корпорации ICANN в Каире, Египет, 3-7 ноября 2008 г. Эти документы публикуются вместе с обновленным предварительным планом внедрения для дальнейшего взаимодействия сообщества в частности до и во время собрания корпорации ICANN в Мехико, Мексика, 1-6 марта 2009 г. Начинается период общественного обсуждения этих документов с целью получения и документирования этих обсуждений сообщества. Полученные комментарии будут в дальнейшем использоваться для редактирования документов при подготовке итогового плана внедрения.

Обратите внимание, это только предварительная версия, предназначенная для обсуждения. Потенциальным отправителям запросов на домены IDN ccTLD не следует полагаться на точность описания подробностей программы, поскольку она будет подвергаться обсуждению и пересмотру.

Полный обзор деятельности, связанной с ускоренным процессом ввода доменов IDN ccTLD и их внедрению, можно найти по адресу <http://www.icann.org/en/topics/idn/fast-track/>

Сводка основных положений данного документа

- Таблицы имен IDN и соответствующие варианты символов разработаны с целью уменьшения потенциально возросшей путаницы, которая в противном случае может возникнуть у пользователей при вводе новых имен IDN.
- Представлено объяснение и предложены рекомендации по процессу разработки таблиц имен IDN и в частности по определению и использованию символов с вариантами в строках доменов второго и высшего уровней.
- Настоятельно рекомендуется совместная разработка таблиц имен IDN и соответствующих политик сообществами с общим алфавитом или в случае существования определенной путаницы символов в используемых языках. Это обеспечит предоставление всем языковым сообществам равных возможностей по обеспечению доступности языков для регистрации доменных имен
- Предлагаемые рекомендации не изменяют предыдущий способ разработки таблиц имен IDN. Ответственность за разработку таблиц несут реестры доменов TLD.

I. Сводная информация

Тема таблиц имен IDN и соответствующих символов с вариантами обсуждалась на нескольких сессиях во время собрания корпорации ICANN в Каире, Египет, в ноябре 2008 г. В результате разъяснительная информация была включена в предварительный план внедрения по адресу <http://www.icann.org/en/announcements/announcement-26nov08-en.htm>

В этом документе представлена дополнительная информация по таблицам имен IDN и их преимуществам для реестров доменов TLD, планирующих ввод имен IDN (на втором или высшем уровне). В данном документе в общих чертах описана разработка таблицы имен IDN и методология использования корпорацией ICANN таблиц имен IDN, представленных реестрами, для выделения и управления доменами TLD.

В общем:

Draft—for discussion only—please refer to the disclaimer on the title page of this document.

1. Таблица имен IDN представляет собой список всех символов, доступных для регистрации доменного имени, в табличной форме.
2. У реестра домена TLD может быть несколько таких таблиц, например, по одной для каждого языка. Таблица может быть основана на: языке, наборе языков или алфавите (в соответствии с указаниями по именам IDN).
3. Символы с вариантами представляют собой два или более символов, означающих «одно и то же» при использовании для регистрации доменных имен.
4. Таблицы имен IDN, определяющие символы с вариантами, весьма полезны, поскольку они позволяют уменьшить потенциальную путаницу, которая может возникнуть в связи с типографским сходством имен.
5. В данном документе предложены процедуры разработки таблиц имен IDN. В этих процедурах кандидатам настоятельно рекомендуется объединить свои усилия, если может возникнуть потенциальная путаница с языками других стран и территорий.
 - a. Одни и те же языки могут иногда использоваться граничащими территориями. Иногда это может привести к путанице среди пользователей соответствующего языкового сообщества или сообщества алфавита.
 - b. В некоторых случаях может также существовать визуальная путаница между различными алфавитами (например, греческий, кириллица и латинский). Таблица имен IDN с взаимозависимостями между определенными символами с вариантами может уменьшить вероятность возникновения этой путаницы при использовании нескольких алфавитов в одном домене TLD.
6. Ограниченная роль корпорации ICANN в связи с разработкой таблиц имен IDN заключается в предоставлении поддержки кандидатам в случае отправки соответствующего запроса.
7. В этом документе предлагается, что корпорация ICANN будет использовать все отправленные таблицы имен IDN при рассмотрении запроса на строки доменов высшего уровня. Эти таблицы будут использоваться в качестве руководства для определения вероятности возникновения путаницы запрашиваемой и существующих строк. Когда путаница может возникнуть вследствие использования символа с вариантами, запрашиваемая строка не будет передана в корневую зону.

Публикация данного документа корпорацией ICANN означает необходимость в комментариях по этой теме. Отзывы будут играть ключевую роль при подготовке окончательных планов внедрения, предназначенных для представления на собрании корпорации ICANN в Сиднее (июнь 2009 г.)

II. Определение таблицы имен IDN

Таблица имен IDN представляет список всех символов, поддерживаемых реестром определенного домена TLD помимо двадцати шести букв базового латинского алфавита (a-z), десяти цифр (0-9) и дефиса (-). Если символы в таблице считаются вариантами друг друга (означая «то же самое»), это указывается рядом со всеми

Draft—for discussion only—please refer to the disclaimer on the title page of this document.

символами из группы вариантов. Термин «вариант» обозначает орфографический эквивалент на уровне символов, например, между символами «æ» и «ae» в словах «encyclopædia» и «encyclopaedia», но не в более широком смысле, подразумевающем варианты правописания слов, например «encyclopædia» и «encyclopaedia» или «color» и «colour».

Обычно в таблице имен IDN содержатся символы, представляющие определенный язык или взятые из определенного алфавита без определенной ссылки на языки, в которых он используется. Используемый здесь термин «таблица имен IDN» соответствует ранее использованным терминам: «таблица вариантов», «таблица вариантов языка», «таблица языка» и «таблица алфавита».

Необходимо провести лингвистическую и орфографическую экспертизу для определения того, может ли символ считаться вариантом другого символа. Одни и те же символы алфавита могут по-разному трактоваться в разных языках. (Обратимся опять к примеру с символами «æ» и «ae» в таблице английского языка, первый будет вероятно считаться вариантом последнего. В таблице датского языка символ «æ» является отдельной буквой алфавита.) Представленные здесь рекомендации не изменяют этого подхода.

III. Преимущества обладания таблицами имен IDN

Когда количество символов, которые можно включить в доменные имена, существенно возрастает с 37 до порядка 100 000 символов из различных алфавитов, существенно увеличивается и риск возникновения путаницы в связи с типографским сходством имен. Несмотря на то что компьютер, например, может с легкостью распознать различие между символами «a» (латинский), «α» (греческий) и «а» (кириллица), человеческий глаз не может это сделать. Ситуация осложняется различием между шрифтами, их размером при отображении и времени, необходимым для обработки и запоминания используемого символа.

Чтобы снизить возросший риск возникновения путаницы, (в соответствии с рекомендациями итогового отчета рабочей группы IDNC) политика регистрации реестра домена TLD, посвященная именам IDN, должна описывать создание таблиц имен IDN. Таким образом политики регистрации имен IDN реестра домена TLD основываются на четко определенном наборе символов. Используя таблицы имен IDN, структурированные одинаковым образом, реестры доменов TLD поддерживают адекватную базу для определения символов, доступных для регистрации, и специальные термины, используемые для символов, трактуемых как варианты друг друга.

Несмотря на то что весь опыт в данной области относится к регистрации имен IDN на втором уровне в существующих доменах TLD и регистрациях на более низких уровнях, базовая концепция применима и становится все более значимой для строк доменов TLD. Это позволит избежать приводящих к путанице схожих строк, добавленных в корень, в частности схожих строк, находящихся в управлении разных организаций.

Исторически таблицы имен IDN разрабатывались реестрами доменов TLD. Хотя агентство IANA и размещает эти таблицы в репозитории Интернета с целью предоставления одного источника информации, корпорация ICANN и агентство IANA не проверяет их содержание. При этом, чтобы таблицы считались таблицами имен IDN, они должны удовлетворять требованиям в указаниях по именам IDN и правилам форматирования, входящим в требования процедуры репозитория имен

IDN агентства IANA. В свою очередь, будут внесены поправки в [указания по именам IDN](#) и [процедуры репозитория имен IDN агентства IANA](#) после обсуждения настоящего предложения и его реализации.

IV. Разработка таблиц имен IDN

Уровень сложности разработки таблицы имен IDN и определения вариантов содержащихся в ней символов зависит от количества символов в ней и языка или алфавита, который она представляет. Например, если таблица содержит символы из одного алфавита, который поддерживает один язык, определение трактовки схожести языковым сообществом может не составить труда. Однако если символы алфавита используются для нескольких языков и/или реестр домена TLD намеревается поддерживать несколько языков, это значительно усложняет адекватную оценку соответствующих лингвистических элементов.

Фундаментальные различия между системами письма приводят к возникновению ситуаций, когда определенный элемент алфавита используются по-разному в разных языках, повышая риск появления путаницы у тех, у кого нет глубокого понимания вариации орфографических практик. Эту ситуацию необходимо учесть в именах IDN так же скрупулезно, как и в других контекстах, где появляется письменный язык. Несмотря на это, сообщество пользователей получит значительное преимущество от минимизации потенциальной путаницы. Первый вклад в разработку политик, основанных на алфавитах и предназначенных для многочисленных языковых сообществ, привел к созданию указаний совместной инженерной группой (JET) по регистрации и администрированию многоязычных доменных имен (IDN) для китайского, японского и корейского языков, которые можно найти по адресу <http://www.ietf.org/rfc/rfc3743.txt>.

Аналогичные инициативы сейчас реализуются и другими языковыми сообществами, использующими другие алфавиты, например, [рабочей группой по именам IDN арабского алфавита \(ASIWG\)](#).

Арабский алфавит широко используются несколькими языками, возникшими на Ближнем Востоке, в Африке и Азии. У каждого из этих языковых сообществ свой взгляд на структурирование своей таблицы имен IDN. Единственный способ убедиться, что интересы всех сообществ отражены в соответствии с их общим алфавитом, опубликованным в пространстве имен IDN – это участвовать в совместной разработке таблиц, вне зависимости от того, разрабатывается ли одна таблица имен IDN для алфавита или несколько таблиц для одного или нескольких языков. Единственная альтернатива подразумевает подвержение риску ненамеренного появления несоответствий в трактовке одного элемента в таблицах разных языков, что негативно сказывается на языковых сообществах, использующих этот алфавит.

Другим примером подобной инициативы и сложной ситуации могут послужить индийские языки, которых насчитывают порядка 20, при этом они используют около 13 алфавитов, и некоторые языки используют в письме несколько алфавитов. Несмотря на то что размеры соответствующих языковых сообществ различны, ни один из языков не имеет более высокий формальный статус, чем другие. При подготовке общей таблицы имен IDN или нескольких таблиц вместе должны быть рассмотрены соответствующие лингвистические элементы из всех языков, использующих этот алфавит. Это же следует сделать и при появлении вероятности визуальной путаницы. Этот подход позволит обеспечить справедливую поддержку всех индийских языков.

Вне зависимости от языка или алфавита, доменные имена не всегда представляют слова из словаря, и ничто, присущее метке, не свидетельствует о языке или алфавите, который она должна представлять. В дальнейшем внимание необходимо обратить на способ использования алфавита при письме на других языках, которые могут быть схоже отражены в именах IDN (например, как показано здесь на примерах). Без подобных действий относящиеся к языку детали, принятые одним реестром, могут оказаться несоответствующими политикам другого реестра, поддерживающего другой язык, но использующий этот же алфавит, а это, в свою очередь, приведет к возможной путанице среди более широкого сообщества пользователей Интернета.

Использование таблиц имен IDN и символов с вариантами при регистрации доменных имен

Существует несколько способов обработки символов с вариантами при регистрации доменных имен. Ниже представлено краткое описание самых распространенных из них. Исторически реестр домена TLD самостоятельно осуществляет выбор способа. Рекомендации данного документа не изменяют этого подхода. Следующие сведения представлены только в информационных целях.

1. Пакетные регистрации – для владельца регистрации символы с вариантами приведут к проведению двух или более регистраций (доменные имена с вариантами) автоматически и за ту же цену.
2. Заблокированные регистрации – символы с вариантами приведут к блокировке доменных имен с вариантами. Блокировка доменного имени означает, что никогда не сможет быть зарегистрировано.
3. Зарезервированные регистрации – символы с вариантами приведут к резервированию доменных имен с вариантами. Резервирование обычно означает, что только один владелец регистрации сможет зарегистрировать зарезервированное доменное имя.

Предлагаемая процедура таблиц имен IDN для использования при регистрации доменов SLD

В соответствии с рекомендациями итогового отчета рабочей группы IDNC должна быть доступна одна или несколько таблиц для каждой заявки на домен IDN ccTLD в ускоренном процессе. Такая же информация представлена и в указаниях по именам IDN для реестров, желающих обеспечивать поддержку имен IDN при регистрации доменных имен.

Представленная ниже процедура предназначена для дополнительного разъяснения разработки таблиц имен IDN. Данное предложение предназначено для всех реестров доменов TLD, желающих поддерживать имена IDN на втором уровне.

Основной целью следующего предложения является предоставление всем языковым сообществам равных возможностей по обеспечению доступности языков для регистрации доменных имен.

1. Отправитель запроса на участие в ускоренном процессе ввода доменов IDN ccTLD самостоятельно определяет символы, которые могут быть включены в метки доменов SLD, со своей осмотрительностью, учитывая советы и комментарии правительственных учреждений и целевого

Draft—for discussion only—please refer to the disclaimer on the title page of this document.

- сообщества.
2. Отправитель запроса на участие в ускоренном процессе ввода доменов IDN ccTLD оценивает увеличение количества символов в результирующем списке за счет их появления в запросах на домены IDN ccTLD, отправленных другими странами.
 - а. При отсутствии такой вероятности отправитель запроса решит, должны ли некоторые символы обозначены в качестве вариантов при разработке соответствующих таблиц имен IDN. (В любом случае рекомендуется посоветоваться с экспертами-лингвистами, хорошо знакомыми с языком или алфавитом.)
 - б. При существовании вероятности появления символов в других запросах отправителю запроса следует скоординировать действия по разработке таблиц имен IDN и, соответственно, списка символов с вариантами с другими странами или территориями. Участники взаимодействия должны решить, следует ли создать одну общую таблицу или несколько отдельных таблиц. Только совместные усилия позволяют предотвратить ненамеренное возникновение путаницы и предоставить вербальное объяснение для общего сообщества пользователей, в случае если подобной неопределенности не удастся избежать.
 3. Поскольку отправитель запроса в пункте 2) должен уметь определить другие страны или территории, с которыми необходимо взаимодействовать в ходе ускоренного процесса, корпорация ICANN облегчит взаимодействие отправителей запросов с органами, обладающими соответствующей лингвистической экспертизой, при необходимости такой помощи.

Предлагаемое использование таблиц имен IDN для регистрации доменов TLD

В соответствии с рекомендациями итогового отчета рабочей группы IDNC должна быть доступна одна или несколько таблиц для каждой заявки на домен IDN ccTLD в ускоренном процессе. Такая же информация представлена и в указаниях по именам IDN для реестров, желающих обеспечивать поддержку имен IDN при регистрации доменных имен.

Символы и варианты, представленные в таблице имен IDN, поддерживающей регистрацию доменов SLD, будут также использоваться для высшего уровня. Корпорация ICANN будет использовать эти таблицы имен IDN при проверке запросов, поэтому отправителям запросов следует это учитывать при подготовке своих таблиц имен IDN и меток доменов TLD.

Могут возникнуть ситуации, когда у отправителя запроса на домен IDN ccTLD могут быть обоснованные причины для получения нескольких меток для запрашиваемого домена, отличающихся кодировкой, которая плохо распознается при отображении или вследствие других более очевидных орфографических проблем (называемых «строками с вариантами»). Тем не менее, на данный момент отсутствует стандарт или механизм, с помощью которого можно реализовать такое использование вариантов на корневом уровне, и ускоренный режим не предоставляет возможность передачи нескольких меток на одном языке и алфавите для одного домена IDN ccTLD.

Корпорация ICANN предлагает выделение или блокировку строк с вариантами для регистрации, следуя нижеперечисленным логичным аргументам и требованиям:

- a. Для выделения строки с вариантами должны удовлетворять тем же требованиям ускоренного процесса, что и запрашиваемые строки.
- b. Несмотря на то что итоговый отчет корпорации IDNC о ускоренном процессе содержит рекомендацию об «одной строке на одну территорию и на один официальный язык», от нее пришлось отказаться в связи с концепцией строк с вариантами.
- c. Необходимо расширить концепцию количества имен, чтобы разрешить странам и территориям обладать выделенными строками с вариантами. В противном случае не удастся достичь целей ускоренного процесса в удовлетворении потребностей сообщества, и появится ненужная неразбериха между определенными обществами, если строки с вариантами не будут разрешены.
- d. Строки с вариантами будут выделены, только если есть согласие в их трактовке в качестве вариантов запрашиваемой строки
- e. Строки с вариантами будут добавлены в качестве отдельных элементов в корневую зону системы DNS.
- f. Поскольку на момент подготовки этого документа не известен стандарт или механизм, с помощью которого можно реализовать такое использование вариантов на корневом уровне, отправитель запросов должен включить в свои реализации доменов IDN TLD механизм, который обеспечит использование определенных строк с вариантами для запрашиваемой строки.

Строки с вариантами, удовлетворяющие этим требованиям, также должны быть запрошены кандидатами на домены IDN ccTLD, с отдельным вниманием к следующему:

- Отправители запросов на домены IDN TLD в своих запросах на домены IDN ccTLD должны представить заявление о поддержке органа власти страны или территории с признанной экспертизой орфографии языка, на котором представлена метка домена TLD. Этот эксперт должен быть знаком с системами письма других языков, использующих этот же алфавит, чтобы быть в состоянии подтвердить, что в метке домена TLD используется алфавит таким образом, при котором не возникает конфликта с его использованием в другой метке домена TLD, представляющей другой язык, или указать на потенциальное увеличение неопределенности.

Другие механизмы предотвращения путаницы на корневом уровне описаны в модуле 7 пересмотренного предварительного плана внедрения, выпущенного вместе с этим документом.

Строки с вариантами, не удовлетворяющие указанным выше требованиям, будут заблокированы для дальнейшего выделения в системе DNS. Это будет соответствовать практикам, которые используют менеджеры доменов TLD для регистрации имен IDN второго уровня в настоящий момент. Заблокированные строки будут считаться «существующими строками» при проверке входящих заявок на конфликты с существующими доменами TLD. Поэтому будут отклонены все последующие заявки на ту же строку.

Draft—for discussion only—please refer to the disclaimer on the title page of this document.

Как указано в начале этого документа, корпорация ICANN ожидает получить комментарии по этой важной теме. Отзывы будут играть ключевую роль при подготовке окончательных планов внедрения, предназначенных для представления на собрании корпорации ICANN в Сиднее (июнь 2009 г.)