

Consideraciones en el uso del alfabeto latino en variantes de dominios de nivel superior internacionalizados

Informe final del Grupo de Estudio VIP de la Corporación para la asignación de números y nombres en Internet (ICANN) para el alfabeto latino

Resumen ejecutivo

El grupo de estudio examinó todos los caracteres de la versión 6.1.0 de la Tabla de códigos de caracteres Unicode asociada con el alfabeto latino y válida en virtud del Protocolo IDNA2008 (Nombres de dominio internacionalizados en aplicaciones). Se identificaron varias formas con 'capacidad de confusión' que podrían requerir de una cuidadosa consideración en la recopilación de un subconjunto de repertorio más amplio para el uso local. No obstante, la resolución de estas cuestiones depende en gran medida de las convenciones ortográficas locales. Frecuentemente, estas cuestiones tratan a los mismos caracteres de diferentes maneras. Las cadenas de caracteres que son confusamente similares en el contexto de un idioma, pueden no tener tales connotaciones en otro.

Tomando en cuenta que el alfabeto latino es utilizado por un mayor número de comunidades lingüísticas separadas que cualquier otra escritura singular, el intentar proporcionar una visión global de las necesidades de todas ellas constituye un esfuerzo no realista. Un intento resumido por hacerlo, resultaría de todos modos culturalmente insensible a las comunidades que aún no se han unido a la discusión sobre nombres de dominio internacionalizados (IDN).

Por tanto el grupo de estudio no encuentra ningún fundamento base para el tratamiento categórico de ningún punto de código asignado a un elemento del alfabeto latino, como equivalente a ningún otro punto de código de este tipo. Tampoco considera que dicho fundamento exista más allá de lo que ya ha sido incorporado en el protocolo IDNA2008. El proceso de solicitud de dominios de nivel superior (TLD) de la ICANN no debe permitir las solicitudes de múltiples cadenas de caracteres latinas bajo la premisa de que unas son variantes de otras. La evaluación de las etiquetas de TLD propuestas para resolver confusiones requiere de un examen cuidadoso, pero eso no las convierte en variantes en cuanto al sentido en que se enfoca el estudio VIP.

El presente informe cuenta con dos tablas de referencia anexas que enumeran e ilustran los puntos de código válidos asociados con el alfabeto latino. La tabla principal enumera aquellos que se pueden tomar como elementos de los sistemas de escritura de cada día y por tanto son de potencial utilidad en los repertorios locales de IDN. La tabla secundaria incluye los caracteres secundarios que carecen de ese atributo o son de otra manera menos adecuados para el uso ilimitado. Aunque la primera puede ser entendida como el universo de puntos de código disponibles para las posibles etiquetas de TLD basadas en el alfabeto latino, el grupo de estudio no recomienda la aceptación incondicional de ningún punto de código sin la demostración específica de garantía para su uso en el contexto lingüístico previsto. De este modo, la tabla de caracteres latinos para la zona raíz puede ser una recopilación descriptiva de los puntos de código que en realidad

aparecen en esa zona, en lugar de un cuadro normativo de los puntos de código que pueden ser incluidos en la misma sin mayor consideración.

1. Introducción

Como parte de sus esfuerzos para ampliar la disponibilidad de IDNs, la ICANN ha puesto en marcha un programa limitado para introducir dominios de nivel superior con código de país en alfabetos no latinos y está a punto de lanzar un programa que permite dominios genéricos de nivel superior en cualquier tipo de escritura/alfabeto. Una de las complejidades de este proceso es que desafía la suposición de que un único dominio se puede representar con una sola etiqueta en el Sistema de nombres de dominio (DNS). Con los IDNs, y el estándar Unicode sobre el cual se construyeron, algunas etiquetas pueden ser representadas de diversas formas. Esto ocurre cuando un carácter abstracto (véase Sección 6.1) en la etiqueta se encuentra en más de un punto de código en la Tabla de códigos de caracteres Unicode o la misma puede ser representada por secuencias alternativas de puntos de código.

Para facilitar el desarrollo de políticas para la asignación y distribución de tales etiquetas múltiples en los procesos de dominio de nivel superior, la ICANN ha creado el "Proyecto de Variantes IDN TLD (Dominios de nivel superior de nombres de dominio internacionalizados)". Su mandato consiste en identificar y analizar las cuestiones que requieren consideración y formular recomendaciones sobre las cuales se podrían basar políticas. Como primer paso, el estudio busca analizar exhaustivamente y llegar a un entendimiento común acerca de aquello que las partes interesadas consideran son variantes y los problemas o cuestiones relacionadas con ellas. Estos enunciados de los problemas se fundieron en un documento sobre las necesidades individuales, contra las cuales se pueden identificar y evaluar las soluciones correspondientes.

Para obtener amplios aportes multisectoriales, se constituyeron seis equipos de estudio de caso: cada uno para un determinado alfabeto utilizado en los sistemas de escritura de un segmento importante de la población mundial. Estos alfabetos son: árabe, chino, cirílico, devanagari, griego y latino. El caso del alfabeto latino aquí analizado, difiere de los otros en varios aspectos. En primer lugar, siendo el alfabeto sobre el cual se basa el DNS, el mismo tiene una historia de casi 30 años de despliegue para respaldar observaciones y conclusiones. La inclusión del alfabeto latino básico "A a la Z" en el conjunto de caracteres ASCII da lugar a las necesidades de muchas comunidades lingüísticas en un grado tolerable. Las fuerzas que impulsan el uso del alfabeto latino para los IDNs pueden por lo tanto diferir de las escrituras no latinas.

En segundo lugar, como un alfabeto con fuertes lazos históricos con caracteres cirílicos y griegos, las similitudes/concordancias entre esas tres escrituras deben ser consideradas. El despliegue práctico de múltiples alfabetos supone la evaluación previa de las cuestiones de confusión y seguridad a través de sus delimitaciones.

Como una semilla para su trabajo, los equipos de casos de estudio recibieron una lista de preguntas preliminares. Si bien estas cuestiones no son prescriptivas, las mismas fueron diseñadas para abordar algunas de las áreas clave que se espera los equipos consideren a fin de proporcionar una base integral para la declaración del problema.

En esencia, se pidió a los equipos de estudio de caso evaluar el concepto de variantes, desde la perspectiva de la experiencia del usuario. Los usuarios del DNS — no sólo los usuarios de computadoras, sino también aquellos que utilizan dominios en teléfonos móviles, los administradores de sistemas que mantienen los sistemas técnicos, los proveedores de software, registros, registradores, etc.— tienen expectativas acerca de cómo los nombres de dominio deberían funcionar. La introducción de variantes se debe basar en llenar las brechas definidas por

las expectativas que deben cumplirse a fin de contar con operaciones previsibles y fiables. Al esbozar esas brechas de expectativas como una declaración del problema, se pueden evaluar las posibles soluciones en términos de su capacidad para abordar esas deficiencias, evaluando sus implicaciones técnicas y políticas.

2. Definiciones

Para asegurar que los grupos de estudio utilicen el mismo vocabulario, se les aconsejó que se adhieran a la terminología específica, siempre que resultase aplicable. Las definiciones en la solicitud de comentarios (RFC) 6365 debían ser consideradas como una base para la terminología de internacionalización, como mejoras para las definiciones brindadas. En los lugares donde la documentación no articule claramente una definición necesaria para este estudio de caso, en el presente documento se agregan mayores detalles, según corresponda.

3. El alfabeto latino

El repertorio fundacional de caracteres para el DNS se brinda en el estándar ASCII. Si bien cualquier octeto puede utilizarse en un nombre del DNS, una regla diferente para los nombres del servidor limita el repertorio a las letras latinas "a" a la "z" en mayúsculas y minúsculas, los dígitos del "0" al "9" y el guión "-" (el "conjunto LDH"). Aunque hay diferencias en la forma en que el DNS trata a las formas minúscula y mayúscula de las letras, en su mayoría son considerados como transparentes para los usuarios. Estos caracteres han sido utilizados para todas las registraciones de nombres del DNS hasta la llegada de los IDNs en 2003.

El protocolo utilizado para la implementación de los IDNs se conoce formalmente como "Nombres de dominio internacionalizados en aplicaciones" o IDNA. Este sistema utiliza una "Codificación ASCII compatible" de caracteres extraídos del repertorio Unicode mucho más amplio. La codificación y decodificación, así como la visualización de los caracteres Unicode depende del apoyo en el nivel de aplicación. Las registraciones correspondientes en el DNS, en sí, continúan basándose en ASCII. La versión más reciente de este protocolo se denomina IDNA2008¹, (a partir de la fecha de su iniciación; se publicó en 2010). Tanto dicho protocolo como el IDNA2003 inicial únicamente se basan en el estándar Unicode² para derivar los puntos de código permitidos.

El estándar Unicode asigna a cada punto de código varias propiedades del carácter, cada una de las cuales indica un alfabeto. Esta propiedad se utiliza para subdividir la tabla completa de caracteres Unicode en bloques encabezados por el indicador de alfabeto, tal como se publicó en el sitio web de Unicode.³

El alfabeto latino se presenta en nueve de dichos bloques, los cuales se encuentran listados en el Apéndice B1.

¹ "Nombres de dominio internacionalizados en aplicaciones (IDNA): Definiciones y marco del documento" (RFC 5890) <http://tools.ietf.org/html/rfc5890>

² <http://unicode.org/standard/standard.html>

³ <http://www.unicode.org/charts/>

El bloque "latino básico" reafirma el repertorio ASCII y por lo tanto, incluye el conjunto LDH a los cuales históricamente se limitaron los nombres del servidor. Las letras latinas más allá del conjunto LDH, así como las diacríticamente marcadas y decoradas de otras maneras se presentan en bloques latinos extendidos y complementarios, con más caracteres latinos en los bloques bajo el título "Símbolos fonéticos".

El protocolo IDNA2008 permite el uso de un gran número de puntos de código a partir de estos bloques (desaprobando completamente sólo aquellos bajo el título "Letras latinas de ancho completo"). Asimismo, impone algunas limitaciones categóricas, por ejemplo, en puntos de código que representan marcas o símbolos de puntuación. Los puntos de código permitidos se denominan "de protocolo válido" o PVALID.

El subconjunto de los puntos de código de protocolo válido asignado al alfabeto latino, que el grupo consideró relevante para el estudio de las variantes latinas, se encuentra listado en el Apéndice B2. En el Apéndice B3 se brinda una segunda lista de puntos de código de protocolo válido que no se utilizan en los sistemas de escritura diaria, pero que sin embargo pueden aparecer en los IDNs en circunstancias especiales. Juntos, proporcionan una enumeración exhaustiva de lo que el grupo ha considerado como los puntos de código del alfabeto latino.

4. Expectativas del usuario

4.1. Sensibilidad a mayúsculas y minúsculas

La base del alfabeto latino para los nombres de dominio no internacionalizados ha dado lugar a una expectativa bien establecida del usuario respecto a cómo funcionan los nombres de dominio con escritura latina. En particular, la falta de distinción entre mayúsculas y minúsculas en las cadenas de caracteres LDH permite su representación artística de una manera que no se extiende a los IDNs. Sin embargo, debido a que los usuarios a menudo escriben las etiquetas LDH en mayúsculas, minúsculas o en forma mixta, es razonable suponer que se espere poder hacerlo también con otros caracteres latinos.

4.2. Navegadores web

Por lo general, el software de navegación web mantiene las letras minúsculas o mayúsculas de un nombre de dominio LDH tal como ha sido introducido, dejando los efectos de cualquier normalización de minúscula o mayúscula transparente para el usuario. Esta característica no se extiende necesariamente al entorno de los IDN y una serie de otros rasgos del navegador que soportan IDNs, carece de homologación en el entorno de trabajo tradicional. Las dificultades que los usuarios podrían tener al acomodarse a esto se ven agravadas por la falta de un enfoque uniforme para la implementación de IDNA por todos los desarrolladores de navegadores. Algunos navegadores muestran la forma codificada en ASCII de un IDN (la "etiqueta-A") en lugar de la forma localizada esperada (la "etiqueta-U") si el TLD no está en la "lista blanca" de un repositorio central mantenido por el desarrollador. Otros navegadores pueden exponer la etiqueta-A en base a la lógica derivada de las configuraciones idiomáticas del sistema operativo subyacente, así como las del navegador.

4.3. Teclados

Los caracteres A-Z y 0-9 están presentes en la mayoría de los teclados informáticos.

Los caracteres latinos extendidos, que incluyen marcas diacríticas o están representados en otras formas especiales, están representados en los teclados de las regiones donde tales caracteres son utilizados. Si bien existen pequeñas diferencias en el diseño de los teclados de alfabeto latino, tal como el teclado QWERTY o AZERTY, el conjunto LDH está ampliamente disponible en todos ellos.

5. Prácticas pertinentes con dominios no internacionalizados

Típicamente, los registros de dominios que permiten el registro de IDNs recopilan pequeños repertorios de puntos de código permitidos a partir del conjunto de protocolos válidos. (De hecho, IDNA2008 espera la compaginación de repertorios locales más pequeños.) Estos suelen ser documentados en forma de tabla, agrupados por alfabeto o idioma, según corresponda a la comunidad objetivo. Muchos son publicados en un repositorio central que mantiene la ICANN en el sitio web de la IANA (Autoridad de números asignados en Internet).⁴

En la práctica anterior no existía ninguna homologación para esta selectividad. El conjunto LDH ofrece veintiséis letras que pueden ser complementadas con diez dígitos y un guión. Las letras pueden ser utilizadas sin distinción entre mayúsculas y minúsculas. El protocolo DNS trata a ambas formas como equivalentes en la medida en que "oldnews" y "OldNews" sean la misma etiqueta. Como se señaló anteriormente, esto ha permitido que etiquetas visualmente similares, tal como "OldNews" versus "01dNews". El uso frecuente de fuentes que ocultan su diferenciación visual puede causar una confusión significativa (es decir, "OldNews" versus "OldNews"). El grupo no tiene conocimiento de intentos para limitar esto, restringiendo los caracteres permitidos para un subconjunto LDH, a excepción de la zona raíz donde los caracteres numéricos están excluidos de los nombres de dominio para evitar confusiones con las direcciones IP.

Tabla: Ambigüedades visuales LDH en el sistema existente⁵

G	6	
I	1	1
O	0	
S	5	
U	V	
Z	2	

5.2 Consideraciones ortográficas

Las diferencias entre la ortografía estadounidense y británica, tal como "organization" y "organisation", son comúnmente notorias aunque el grupo desconoce de su inclusión en las políticas de registro. La registración de una marca en una forma no confiere derechos a la otra, y

⁴ <http://www.iana.org/domains/idn-tables>

⁵ Tabla original de "Punycode: Una sintaxis de codificación de Unicode para los Nombres de dominio internacionalizados en aplicaciones" (RFC 3492), Sección 5. <http://tools.ietf.org/html/rfc3492>

un titular de un nombre que desee prevenir los riesgos de confusión de los usuarios, ha tradicionalmente tenido la libertad de registrar las dos. Una de las razones por las cuales tales cosas no fueron percibidas como una fuente de preocupación es que las etiquetas eran consideradas como una conveniencia nemotécnica y cualquier propiedad relacionada con la palabra que mostraban, era considerada de interés secundario.

De acuerdo con RFC 3454:

"Los usuarios probablemente esperan que todos los equivalentes de ortografía sean equiparados, o ninguno de ellos lo sea. Algunos ejemplos de equivalentes de ortografía son 'theater' versus 'theatre' y 'hemoglobin' versus 'hæmoglobin' en inglés americano vs. inglés británico."⁶

"Equivalencias específicos del idioma tal como 'Aepfel' vs. 'Äpfel', los cuales a veces en alemán se consideran equivalentes, pueden ser considerados como equivalentes en otros idiomas."⁷

En el documento estatutario inicial para el Proyecto de variantes de IDN, se proporcionó un ejemplo en alfabeto latino de un hipotético dominio de nivel superior para la ciudad de Cologne (Köln), Alemania. Las reglas de ortografía del alemán permiten que esto se escriba, ya sea como ".koeln" o ".köln".

A los efectos del estudio VIP, el grado en que un usuario alemán de Internet podría ver a estas dos etiquetas como similares, o para utilizar otro ejemplo: ".strasse" y ".straße", ilustra una situación que merece consideración pero que no es susceptible de solución algorítmica. En todo caso, pone de relieve el grado en que tal equivalencia matemática no puede ser impuesta en el alfabeto latino.

6. Taxonomía de las variantes

Varios términos han sido utilizados para designar el concepto de equivalencia, pero ninguno ha sido brindado con una definición que sea adecuada en todos los contextos en los cuales se necesita. Es probable que esta dificultad terminológica esa causada, al menos en parte, por las distintas percepciones y necesidades que existen de una comunidad lingüística a otra. Esto es cierto incluso entre las comunidades lingüísticas que comparten la misma escritura. Y es particularmente cierto para el alfabeto latino, el cual se utiliza para escribir un mayor número de idiomas que todas las otras escrituras combinadas.

La relación de variante más evidente en el alfabeto latino se encuentra entre las formas mayúscula y minúscula de una determinada letra. Tal como se ha señalado, el DNS evita la necesidad de preocuparse por esto, siempre y cuando el repertorio se limite al conjunto LDH. La forma de igualdad que se transmite a esos caracteres se toma comúnmente como de alguna manera extensible a otras relaciones de variante, reconociendo que en consecuencia se necesitaría de un marco más amplio, aunque no sugiere ninguno.

⁶ "Preparación de cadenas de caracteres internacionalizadas" (RFC 3454).

⁷ <http://tools.ietf.org/html/rfc3454>

Ibídem.

Además, se asume comúnmente que cualquier nivel de complejidad mayor que podría añadirse a la implementación de tal marco especulativo, no presentaría una tarea excesiva para la comunidad técnica. De hecho, la falta de sensibilidad a las formas mayúscula y minúscula de ASCII es respaldada por una operación matemática trivial. Idear los medios para la aplicación algorítmica de cualquier otra forma de relación de variantes —desde una sencilla convertibilidad de las letras del alfabeto latino más allá del conjunto LDH de mayúscula a minúscula, hasta la equivalencia entre caracteres considerados similares en algún sentido—, no constituye una empresa sencilla en ningún aspecto. Por lo menos, los caracteres utilizados en los sistemas de escritura de múltiples idiomas cuentan frecuentemente con relaciones variantes que difieren de un idioma a otro, lo cual de algún modo también necesita una cuantificación universal.

El protocolo IDNA aborda las preocupaciones centrales en las relaciones variantes del alfabeto latino, y se ocupa, por ejemplo, directamente de la sensibilidad de las formas mayúscula y minúscula (aunque en lo que podría ser una forma inesperada). Otras cuestiones de variantes dependen tan profundamente de consideraciones locales que resulta poco probable encontrar una solución general a nivel de protocolo.

6.1. Formas de visualización

Un carácter abstracto, por ejemplo, la LETRA M MINÚSCULA LATINA (utilizando la representación formal Unicode de nombres de caracteres), está representado por uno o más glifos, cada uno con una forma específica de visualización. Es en este nivel que los atributos, tal como remates/terminales, cursiva y grosor del trazo (negrita), son cuantificados:

m **m** *m* *m* **m** *m*

Una fuente es una colección de glifos (por ejemplo, Helvética o Times Roman), que respaldará uno o más alfabetos y que puede utilizar el mismo glifo en escrituras diferentes. Un ejemplo clásico de esto es la LETRA A MINÚSCULA LATINA y la LETRA A MINÚSCULA CIRÍLICA:

a **a**

Algunas fuentes están deliberadamente diseñadas para hacer tales distinciones tan visibles como sea posible, pero el usuario típico dispone de fuentes locales específicas que normalmente no respaldan a subconjuntos extensos del repertorio Unicode. La representación de los diferentes caracteres abstractos con glifos indistinguibles, utilizando fuentes comúnmente disponibles en resoluciones de pantalla normal, es por tanto de gran peso en otras consideraciones específicas de la comunidad al momento de evaluar la relación entre la posible relación de variante entre dos puntos de código.

6.2. Glifos idénticos

Las tablas del alfabeto latino de los Apéndices B2 y B3 (que se explica en la sección 6.4) contienen sólo una instancia de dos caracteres abstractos que suelen compartir el mismo glifo. Estos son la LETRA E MINÚSCULA LATINA ROTADA (U+01DD), y la LETRA SCHWA MINÚSCULA LATINA (U+0259):

ə ə

En cualquier sistema de escritura en el cual aparece alguno de ellos, el otro no está incluido como un sustituto. Esto permite un debate sobre la utilidad de indicar una relación de variante entre ellos en cualquier repertorio de IDN en el cual aparecen. La falta de atributos locales específicos en la zona raíz, indica sin embargo la necesidad de regulación contextual explícita para el uso de estos caracteres en las etiquetas de TLD.

6.3. Distinción/normalización en mayúsculas y minúsculas

La diferencia más inmediata entre la presentación de una cadena de caracteres en letras latinas limitadas al conjunto LDH, y una en un conjunto más amplio válido para IDNA puede ilustrarse mediante el regreso a etiquetas mixtas en mayúsculas y minúsculas. IDNA2008 no funciona en las etiquetas LDH y sólo acepta letras minúsculas en las etiquetas que procesa. Esto significa que las etiquetas "clubamerica" y "ClubAmerica" coinciden en el DNS, pero "ClubAmérica" no es válido y sólo se puede utilizar como "clubamérica"⁸. En este caso, la presencia o ausencia del acento es poco probable que cambie algún significado que la etiqueta tenga intención de transmitir, y el titular del nombre puede atribuir un mayor valor a la flexibilidad en la representación de mayúsculas y minúsculas que a la marca diacrítica.

Existe menos libertad en otras situaciones que no implican más que una sola letra no-ASCII. Alguien con el nombre sueco "Östen" que desee utilizarlo como etiqueta en un nombre de dominio es casi seguro que deba aceptar una restricción a minúsculas en lugar de renunciar a la diéresis y utilizar la etiqueta "osten", que es la palabra sueca para queso. En el idioma alemán, la expansión de una "ö" a "oe" constituye una alternativa establecida, aunque sólo es de aplicación limitada a los nombres propios. Sin embargo, no ocurre lo mismo en el idioma sueco, donde la "ö" no se descompone en dos letras, ni se considera una "o" diacríticamente marcada. Es una letra por derecho propio (la última en un alfabeto de 29 letras) y las dos letras no son variantes la una de la otra. En contextos de entrada de texto, si es una "ö" no está disponible, la alternativa es el uso de una "o".

6.4. Variantes decorativas y contrastivas

En los casos en los cuales las marcas diacríticas clarifican la pronunciación pero no cambian el significado de una palabra: "naivete" versus "naïveté", su uso es "decorativo". En los casos en que el uso de las marcas diacríticas cambian o desambiguan el significado: "resume" versus "résumé", su uso es "contrastivo". Independientemente de la distinción entre el uso decorativo y contrastivo, las letras con marcas diacríticas son comúnmente referidas como "decoradas". El mismo término también se aplica a las letras que aparecen en formas modificadas, ya sea que tengan marcas o no:

n ñ ñ

En las tablas de escritura de los apéndices B2 y B3 se brinda una ilustración extensiva de las letras latinas decoradas disponibles para su uso en los IDNs, junto con las letras de base que no se utilizan en inglés y por lo tanto no incluidas en el conjunto LDH. Estas proporcionan una lista

⁸ IDNA2003 permite entradas en letras mayúsculas, pero las mismas se normalizan a minúsculas. IDNA2008 brinda a los implementadores de software la libertad de incluir la normalización en mayúsculas o minúsculas al procesar los IDNs.

exhaustiva de los puntos de código que son etiquetados como letras en los bloques del alfabeto latino de la versión Unicode 6.1.0 (prevista para ser lanzada en febrero de 2012 y no diferente en ningún aspecto relevante a la versión 6.0.0 actualmente vigente) y son válidos en el protocolo IDNA2008. Los bloques fonéticos también incluyen elementos latinos que se utilizan en los sistemas de escritura que aparecen en el espacio de IDN, o que razonablemente se pueden esperar en el futuro. Las tablas actuales incluyen los puntos de código de los bloques de fonética específicamente necesarios para dar lugar al "Alfabeto africano de referencia" o ARA. El ARA se usa aquí repetidamente en las ilustraciones, tanto porque es muy adecuado para ese propósito como para hacer hincapié en que la preocupación por las variantes de caracteres en el alfabeto latino no está principalmente anclada en los idiomas europeos. Muchos de los puntos de código incluidos en su base también aparecen en otros alfabetos basados en la escritura latina.

Los bloques del alfabeto latino fueron considerados hasta el centro del Alfabeto extendido C, deteniéndose en U+2C74. Los demás puntos de código se utilizan en el discurso analítico sobre el idioma pero no aparecen en ningún sistema de escritura cotidiana conforme al conocimiento del grupo, a excepción de U+A78C que se detalla a continuación. No obstante, estos puntos de código pueden ser incluidos en los repertorios locales, por parte de los registros que han documentado la necesidad de ellos y entienden la manera de integrarlos en las cadenas de caracteres que desambiguan su función (utilizando medios basados en la etiqueta para resolver las preocupaciones a nivel de carácter) y que son capaces de articular las limitaciones contextuales necesarias para evitar su uso inadecuado en otros lugares. Parece seguro afirmar que actualmente la inclusión de estos puntos de código en la zona raíz no tiene ninguna utilidad que lo justifique, y el grupo recomienda que sean evitados por completo en ese contexto.

6.5. Caracteres precompuestos

Los bloques latinos de Unicode incluyen un prodigioso número de puntos de código únicos que representan un carácter base en combinación con una o más marcas diacríticas, o que aparecen en otras formas decoradas. Sin embargo, existen otras marcas que pueden ser aplicadas a estos caracteres precompuestos. En la tabla no se incluye ninguna otra marca combinada pero muchas de las marcas que aparecen en los caracteres precompuestos también son separadamente válidas en el protocolo IDNA, en sus formas de combinación. Por tanto cualquiera que recopile una tabla de caracteres local, es libre de incluir cualquiera de las siguientes formas del mismo carácter:

las precompuestas,

LETRA C MINÚSCULA LATINA CON CEDILLA Y ACENTO AGUDO:

ç

la cual también se puede representar mediante la combinación de caracteres de dos puntos de código por separado:

LETRA C MINÚSCULA LATINA CON CEDILLA + COMBINACIÓN DE ACENTO AGUDO:

ç

LETRA C MINÚSCULA LATINA CON ACENTO AGUDO + COMBINACIÓN DE CEDILLA:

ç

o mediante tres puntos de código por separado:

LETRA MINÚSCULA LATINA C + COMBINACIÓN DE ACENTO AGUDO +
COMBINACIÓN DE CEDILLA: ç

LETRA C MINÚSCULA LATINA + COMBINACIÓN DE CEDILLA + COMBINACIÓN DE
ACENTO AGUDO: ç

Estas representaciones de uno, dos y tres puntos de código son equivalentes en todos los aspectos y por lo tanto constituyen una instancia sin ambigüedades de una recopilación de caracteres variantes. Los caracteres precompuestos en los ejemplos segundo y tercero se pueden representar de manera similar en las dos formas pre y post combinadas, proporcionando dos recopilaciones más. El número de recopilaciones de caracteres variantes que puede ser seleccionado de esta manera en forma recursiva a partir de la tabla, sería realmente abrumador si tales relaciones fuesen indicadas por separado. En términos prácticos, esta preocupación es en gran parte, si no enteramente, obviada por la normalización realizada en IDNA2008.

6.6. Combinación de marcas

Siempre y cuando la combinación de marcas independientes sea excluida de la tabla, no existe la necesidad de indicar la plétora de variantes que pueden generar. Hay, sin embargo, sistemas de escritura que utilizan caracteres no incluidos en la tabla Unicode en forma precompuesta. El ARA incluye numerosas consonantes destacadas, algunas de las cuales están disponibles en forma precompuesta, como la LETRA T MINÚSCULA LATINA CON SUBRAYADO (U+1E6F):

t

mientras que otras requieren de la combinación de marcas separadas, tal como la LETRA C MINÚSCULA LATINA + COMBINACIÓN DE LÍNEA BAJA (U+0063 U+0332):

c

El glosario Unicode se refiere a un carácter compuesto que no existe en forma precompuesta como un "grupo grafema".⁹

También existen situaciones en las cuales marcas similares, tal como la COMBINACIÓN DE LÍNEA BAJA y la COMBINACIÓN DE MACRÓN BAJO, se pueden utilizar indistintamente con marcas distintivas más gráficamente, tal como la COMBINACIÓN DE PUNTO BAJO, constituyendo elementos que podrían ser tratados como una recopilación de variantes funcionales. No obstante, raramente existe una relación variante intrínseca entre las marcas en sí mismas. Esa propiedad está modulada por el carácter de base con el cual se combinan y, además, modulada por el sistema de escritura en el cual los caracteres combinados están siendo utilizados. El tratamiento de las marcas como variantes de cada una en un sistema de escritura, no impide de ninguna manera que puedan ser utilizadas de manera independiente la una de la otra.

Los caracteres combinados pueden fácilmente colocarse en una tabla sin necesidad de una lista separada de marcas combinadas, indicando las relaciones variantes cuando resulte apropiado. Sin embargo, debe existir alguna forma de restricción contextual colocada en el carácter(es) sobre el cual una determinada marca puede ser aplicada. Este es un buen ejemplo de una situación en la

⁹ "Una unidad de texto segmentable horizontalmente, que consiste en algún grafema de base... junto con cualquier número de marcas sin espacio que se le aplique."

cual un problema masivo en la gestión de variantes puede ser evitado en el nivel de protocolo. No obstante, el inicio de tal acción no constituye una preocupación actual. El protocolo IDNA2008 reconoce la necesidad de reducir las formas variantes de este tipo a una sola representación canónica y normaliza algorítmicamente todas las alternativas de múltiples puntos de código a la forma precompuesta, si existe una. Cuando no, aún es necesaria una restricción de política.

Al ser incluidas en una etiqueta-U propuesta, las cinco formas marcadas de la 'c' en el ejemplo anterior aparecen indistintamente como:

U+0061 U+1E09 U+0065	; açe
U+0061 U+00E7 U+0301 U+0065	; açe
U+0061 U+0107 U+0327 U+0065	; açe
U+0061 U+0063 U+0301 U+0327 U+0065	; açe
U+0061 U+0063 U+0327 U+0301 U+0065	; açe

Aunque sólo la primera de ellas es una etiqueta-U válida, las otras formas se normalizan antes de generar lo que luego es la misma etiqueta-A para todas ellas: "xn--ae-ess". Es importante señalar que esta normalización no se lleva a cabo por medio del algoritmo de codificación en sí mismo. Cada una de las cinco etiquetas-U propuestas, válidas e inválidas, tiene su propia forma Punycode:

U+0061 U+1E09 U+0065	; açe	# ae-ess
U+0061 U+00E7 U+0301 U+0065	; açe	# ae-4ia66s
U+0061 U+0107 U+0327 U+0065	; açe	# ae-vla09s
U+0061 U+0063 U+0301 U+0327 U+0065	; açe	# ace-ldc1qd
U+0061 U+0063 U+0327 U+0301 U+0065	; açe	# ace-ldc3q

[Nota: Existen instalaciones de "conversión de IDN" en línea, que simplemente realizan la codificación Punycode y establecen el resultado con el prefijo "xn--". Por tanto, es de esperar que algunos posibles titulares de nombres presenten las solicitudes con la etiqueta-A presentada en forma inválida. Como advertencia adicional, las letras precompuestas son a menudo más tipográficamente estables que sus equivalentes de múltiples puntos de código. Los errores típicos de visualización de estos últimos son la mala colocación de las marcas de combinación, por ejemplo en el caso de açe, con cualquiera de las marcas de combinación cambiando de la 'c' a la 'e', sin cambiar los puntos de código reales. Por tanto es de esperar que las solicitudes de las etiquetas-U también puedan presentarse en forma incorrecta. Es concebible que la aparición simultánea de ambas la etiqueta-A y la etiqueta-U incorrectas hagan que sea imposible identificar la etiqueta real que está siendo solicitada.]

6.7. Puntuación

La única marca de puntuación permitida en el repertorio LDH es el GUIÓN-MENOS (U+002D). El repertorio IDNA2008 incluye una serie de caracteres que no son marcas de puntuación pero que utilizan glifos que no se pueden distinguir fácilmente de ellos (y tienen nombres Unicode que son híbridos letra/puntuación). Obviamente, el uso de tales caracteres requiere de un cuidado extraordinario. Un ejemplo es la LETRA MODIFICADORA COMA ROTADA (U+02BB), la cual se utiliza en numerosos contextos de los cuales uno es como la “ ‘okina”, que indica la oclusión glotal en, por ejemplo, “Hawai‘i”:

Otro ejemplo de un carácter utilizado para indicar una oclusión/pausa glotal, pero el cual es una letra latina y no una marca de puntuación, es la LETRA SALTILLO MINÚSCULA LATINA (U+A78C):“Hawai‘i”:

‘

Otro ejemplo de un carácter utilizado para indicar una oclusión/pausa glotal, pero el cual es una letra latina y no una marca de puntuación, es la LETRA SALTILLO MINÚSCULA LATINA (U+A78C):

’

Su anotación de Unicode establece que: "es ampliamente utilizada en muchos idiomas en México y otras regiones, incluyendo Izere en Nigeria", y por lo tanto, puede aparecer en las tablas localizadas del alfabeto latino. Expectativas similares se aplican a la ‘okina en un contexto de idioma polinesio. La única base para el armado de estos caracteres en una colección de caracteres variantes es a través de su similitud grafémica. Ellos son utilizados de otras maneras en los sistemas de escritura separados y no están mutuamente reconocidos a través de esas delimitaciones. En cualquier caso, lo que sería la variante preferida obvia, el APÓSTROFE (U+0027), no puede aparecer en una etiqueta-U. El problema es que constituye una representación alternativa compartida por las otras formas, las cuales raramente están indicadas en forma directa en los teclados clásicos y para las cuales los usuarios comúnmente sustituyen un apóstrofe. Todo esto garantiza un alto índice de fracaso cuando una etiqueta que contiene un carácter parecido a un apóstrofe se transcribe de la tinta al papel. La latitud para la sustitución engañosa en situaciones de origen digital es muy elevada.

El problema con los caracteres que se asemejan a los apóstrofes se ve agravado por la representación alternativa del carón sobre algunas consonantes, ya sea como la marca nominalmente esperada vista en el primero de los siguientes ejemplos o como la marca parecida a un apóstrofe, como en el segundo ejemplo. Estas dos versiones muy diferentes de la LETRA T MINÚSCULA CON CARÓN (U+0165) proporcionan una posible base para otra forma de colección de variantes:

t̃

t’

Una similitud de letra/puntuación adicional se encuentra en el ARA. La LETRA LATINA DE OCLUSIÓN GLOTA (U+0294) se asemeja al SIGNO DE INTERROGACIÓN (U+003F), el cual no puede aparecer en una etiqueta-U:

ʔ

?

Tomando uno de los cuatro caracteres que se utilizan para representar el chasquido consonántico en varios idiomas que se hablan en el sur de África (y que no están incluidos en el ARA), la LETRA LATINA DE CHASQUIDO RETROFLEJO (U+01C3) se asemeja del mismo modo, y en muchas fuentes comparte el mismo glifo con el SIGNO DE EXCLAMACIÓN (U+0021), el cual también está prohibido en las etiquetas-U:

!

!

Por razones obvias, estos caracteres, junto con el resto de los demás caracteres válidos que se asemejan a las marcas de puntuación inválidas, no deben ser permitidos para su uso como sus sustitutos.

6.8. Puntos de código de los sistemas de escritura no relacionados

Una preocupación similar es causada por las letras que se encuentran en distintos puntos de código, que son utilizadas en los sistemas de escritura no relacionados y que generalmente se representan con glifos similares, si no idénticos. Repitiendo el ejemplo de LETRA E MINÚSCULA LATINA ROTADA (U+01DD), y la LETRA SCHWA MINÚSCULA LATINA (U+0259) que aparece arriba:

ə ə

El uso del mismo glifo para ambos en muchas fuentes ampliamente utilizadas, no implica que los caracteres sean formas variantes el uno del otro, ni es probable que algún sistema de escritura incluya a más de uno de ellos. No obstante, algunos problemas imprevistos podrían resultar si un repertorio que incluye uno de estos puntos de código fuese recopilado sin el conocimiento del otro. Por lo tanto, además de cualquier columna de una tabla de escritura generalizada que indica las relaciones entre puntos de código variantes, se podría incluir una columna separada para indicar los puntos de código que no deberían aparecer en el mismo repertorio localizado sin una cuidadosa consideración. Si de hecho ambos fuesen requeridos, se deberán articular y hacer cumplir claramente las reglas contextuales para su uso.

Lo contrario también se aplica a los caracteres dentro de un sistema de escritura único. El ARA incluye tanto la LETRA A MINÚSCULA LATINA (U+0061) y la LETRA ALFA MINÚSCULA LATINA (U+0251), las cuales no pueden ser tomadas como variantes la una de la otra:

a a

Sin embargo, el segundo carácter claramente necesita ser restringido a su uso en las etiquetas escritas con un alfabeto que incluye a ambas. Una vez más, la inclusión de este último carácter en una etiqueta de TLD requerirá de una regulación contextual explícita.

Existe un desafío de enormes proporciones en el equilibrio necesario para minimizar la posibilidad del abuso generado por la disponibilidad simultánea de dos caracteres como estos, respecto a las expectativas ortográficas legítimas —por no decir requisito— en contextos lingüísticos en los cuales se reconocen y entienden las diferencias lingüísticas entre las dos letras. Esto se aplica a muchos de los otros ejemplos dados aquí.

La experiencia necesaria para enfrentar estas situaciones, sin ninguna apariencia de discriminación cultural, se extiende más allá del conocimiento detallado de primera mano del sistema de escritura de la comunidad lingüística respectiva y de mantener un seguimiento de cambios en sus reglas ortográficas. También se requiere un profundo conocimiento de los entornos de trabajo localizados, particularmente en lo que respecta a la capacidad de los motores de renderizado utilizados para mostrar correctamente los caracteres decorados en una amplia variedad de fuentes, y el conocimiento de cualquier discrepancia que existiese entre el punto de código de un carácter determinado tal como es generado por un teclado clásico y el punto de código Unicode de ese carácter. Además hay más necesidad de cercanía suficiente a la comunidad para poder seguir la evolución en el entorno del usuario y garantizar la rápida adaptación de los sistemas de IDN para reflejar los cambios, a medida que se hacen.

Si existe alguna variación dentro de una comunidad lingüística en cosas tales como el punto de código(s), producida para un carácter dado en un teclado de ingreso de texto, la preparación de un repertorio único de caracteres IDN para esa comunidad también tendría que tomar en cuenta tales diferencias. (Este es un tema importante en la gestión de variantes, pero no de especial preocupación con el alfabeto latino). El entendimiento similar de las condiciones locales es necesario para evaluar la posible necesidad de un tratamiento especial de los caracteres en distintos puntos de código que, al menos en algunos contextos, se muestran regularmente con glifos que pueden no ser distinguibles incluso para los miembros más eruditos de las comunidades lingüísticas que los utilizan.

Más allá de la familiaridad con las condiciones locales, la evaluación y gestión de estos riesgos requiere comprender los efectos que pueden tener sobre los usuarios externos a una comunidad que, por un lado, son cómodamente capaces de leer la escritura básica que utiliza para las etiquetas IDN, pero por el otro, no reconocerá las distinciones cruciales entre puntos de código proyectadas en grafemas familiares.

7. Similitud visual con otros alfabetos

El grupo señala que algunos puntos de código asignados a los alfabetos cirílico y griego, comúnmente se representan con glifos que también son utilizados para la visualización de los puntos de código asignados a la escritura latina. Otros alfabetos no incluidos en el estudio VIP, como el cherokee, tienen la misma forma de solapamiento con el alfabeto latino.

Un ejemplo de esta propiedad es la cadena de caracteres "paypal", escrita toda en alfabeto latino, versus "paypal", utilizando la letra "а" del alfabeto cirílico. Estas dos cadenas de caracteres son visualmente idénticas, no sólo similares, aunque se codifican de manera diferente.¹⁰

Los Equipos de estudio de caso para el alfabeto latino y cirílico celebraron una reunión conjunta para analizar la posibilidad de confusión y perjuicio con el uso mixto de los puntos de código asignados a los dos alfabetos diferentes. Los equipos no estaban al tanto de ningún idioma escrito con el alfabeto latino o cirílico, en el cual se requiera de la mezcla de esos alfabetos para expresar una idea o término, o que tuviese algún otro valor nemotécnico esencial. Sin necesidad de ello, y dado el riesgo de confusión que engendra, los grupos acordaron que tal uso mixto deberá ser evitado. La cuestión más espinosa de "confusiones de alfabeto completo" —la yuxtaposición de etiquetas de un único alfabeto que no puede ser visualmente diferenciada— no fue discutida, ya que estaba fuera del alcance de los estudios específicos de los alfabetos y por lo tanto, espera ser considerada en detalle en la revisión cruzada del estudio.

8. Evaluación de solicitudes de TLD con variantes

La ICANN proporciona una herramienta para probar las cadenas de caracteres mediante el uso del "algoritmo SWORD". Esto ha estado en línea para las pruebas de preproducción durante el período previo al lanzamiento del programa de nuevos gTLD. Su documentación afirma el apoyo en una cantidad de idiomas: *“Este algoritmo en versión de preproducción admite los caracteres más comunes de los alfabetos árabe, chino, cirílico, devanagari, griego, japonés, coreano y*

¹⁰ “Norma Técnica 39 Unicode Preliminar: Mecanismos de seguridad Unicode”, Sección 4. <http://unicode.org/reports/tr39/tr39-1.html>

latino. ... los alfabetos latino y griego pertenecen a la familia de escritura europea y podrían ser comparados. ... Esta versión del algoritmo no incluye la validación de las cadenas de caracteres propuestas respecto al cumplimiento con los protocolos de IDNA." [énfasis añadido]

Hay un interés en que las aplicaciones de TLD sean examinadas a través de este u otros mecanismos automatizados de evaluación de las cadenas de caracteres. Este mecanismo debería señalar e identificar, por ejemplo, a dos candidatos diferentes que compiten por .koeln y .köln, respectivamente. Estas cadenas de caracteres podrían ser consideradas un escenario de disputa, pero esta sugerencia parecía estar más allá del estatuto o el mandato del grupo de redacción inicial.¹¹

9. Impacto de las variantes en las operaciones de registro/registrador

La gestión de las variantes de TLD —es decir, la forma en que un dominio es delegado bajo un TLD variante—, puede plantear problemas en los niveles de registro y registrador. Desde la perspectiva de un registro, los siguientes elementos tendrían que ser adaptados al introducir una variante: políticas y procedimientos, incluyendo aquellos aplicados a su lanzamiento y posible sincronización de los dominios en virtud de los mismos; las tarifas de los registradores; los procedimientos de transferencia, incluyendo las opciones de bloqueo de dominio.

A nivel de registro, cualquier desarrollo en el entorno de variantes deberá prever consultas extendidas con los registradores acreditados del registro que tiene previsto introducir una variante. Existen múltiples razones para esto, entre las cuales la de mayor urgencia se debe a que en la mayoría de los TLDs los dominios son registrados a través de los registradores y no directamente con el registro. El lanzamiento de una variante sin algún registrador involucrado puede dar lugar a oportunidades de beneficio para los usuarios de una minoría privilegiada. Durante la introducción de los IDNs en el segundo nivel, muchos registradores no respaldaron a los alfabetos en forma inmediata, a medida que eran introducidos por un registro.

Por lo tanto, el impacto de las variantes en el ámbito de registro/registrador debe ser cuidadosamente explorado, no únicamente a nivel de operaciones —que en cualquier caso debe ser investigado—, sino particularmente al momento de su introducción, con enfoque principal en los intereses de los usuarios finales.

El impacto de los TLDs variantes en los registros y registradores puede ser altamente dependiente de los diferentes métodos de implementación. Cualquier implementación propuesta requerirá de una amplia participación de las partes interesadas para garantizar que los registros y registradores ofrecen operaciones estables, seguras, consistentes y sin ambigüedades en el DNS. Esto incluye la mayor claridad posible en la comunicación y comprensión de los TLDs variantes, limitación de la confusión del usuario final de IDN, del registratario, del registrador y del registro. Todas las áreas de comportamiento de las aplicaciones, resolución y servicios de registración, servicio de WHOIS y la lógica de negocio, necesitan ser evaluadas con el fin de determinar si estos objetivos son alcanzables.

10. Efectos de las variantes en los procesos de resolución de disputas

¹¹ Cabe señalar que el sitio SWORD genera un error cuando estas dos cadenas de caracteres son comparadas o cuando köln se somete a prueba —“köln: La etiqueta debe estar puramente en letras minúsculas ASCII; se encontró 'ö' (U+00F6) en la posición 1”.

10.1. Reglas vigentes en la resolución de disputas

Muchos registros de TLD siguen la Política uniforme de resolución de disputas de nombres de dominio (la "UDRP"). En virtud de ello, la mayoría de los tipos de disputas de nombres de dominio basadas en la marca registrada debe ser resuelta de común acuerdo, mediante acciones judiciales o de arbitraje antes de que un registrador cancele, suspenda o transfiera un nombre de dominio. Las disputas que supuestamente se derivan del registro abusivo de nombres de dominio (por ejemplo, por ciberocupación) pueden ser tratadas mediante procedimientos administrativos expeditos que el titular de los derechos de marca inicia mediante la presentación de un reclamo a un proveedor aprobado de servicios de resolución de disputas.

Un panel de arbitraje verifica si el dominio en disputa (ya sea toda la cadena de caracteres o un componente de la misma) es idéntica o confusamente similar a una marca de producto o servicio sobre la cual el demandante tiene derechos, y si la misma está creando un riesgo de confusión con la marca del demandante en relación a la fuente, patrocinio, afiliación o promoción del sitio web del actual registratario o la ubicación de un producto o servicio en el sitio web o ubicación del registratario.

Los dominios pueden ser confusamente similares para la marca registrada o una marca de servicio, teniendo en consideración:

- a) la parte del dominio que es la etiqueta que fue registrada (típicamente en el parte de segundo nivel, a la izquierda del dominio de nivel superior); o
- b) en casos excepcionales, tanto en los componentes a la izquierda como a la derecha del punto (en ambos componentes de nivel superior y componentes de segundo nivel del dominio).

10.2. Impacto de las variantes latinas

Es probable que la cantidad de disputas aumente en situaciones en las cuales las variantes de un único TLD son confusamente similares y el registro permite que más de una parte registre dominios que conforman el escenario de disputa. Por ejemplo, si a un registro se le conceden dos TLDs variantes, por ejemplo: ".koeln" y ".köln", y luego delega un dominio bajo .koeln a un solicitante y la misma etiqueta a otro registratario bajo .köln —por ejemplo, "hotel-airport.koeln" registrado por una parte y "hotel-airport.köln" registrado por otra parte—, existe una alta probabilidad de que los casos de resolución de disputa surjan a partir de cualquier política que permita esto.

11. Conclusiones

Las conclusiones del Equipo de estudio de caso del alfabeto latino son las siguientes:

1. En el repertorio de caracteres latinos que pueden ser utilizados en IDNs se encuentran varias formas de similitud visual. Éstas requieren de una evaluación cuidadosa al recopilar un subconjunto de un repertorio completo, para su uso local. No obstante, la resolución de estas cuestiones depende en gran medida de las convenciones ortográficas locales. Frecuentemente, estas cuestiones tratan a los mismos caracteres de diferentes maneras. Las cadenas de caracteres que son confusamente similares en el contexto de un idioma, pueden no tener tales connotaciones en otro.
2. Tomando en cuenta que el alfabeto latino es utilizado por un mayor número de comunidades lingüísticas separadas que cualquier otra escritura singular, el intentar proporcionar una visión global de las necesidades de todas ellas constituye un esfuerzo no realista. Un intento resumido por hacerlo, resultaría de todos modos culturalmente insensible a las comunidades que aún no se han unido a la discusión sobre los IDN.
3. Al considerar los dos puntos precedentes, el grupo de estudio no encuentra ningún fundamento base para el tratamiento categórico de ningún punto de código asignado a un elemento de la escritura latina, como equivalente a ningún otro punto de código de este tipo. Tampoco considera que dicho fundamento exista más allá de lo que ya ha sido incorporado en el protocolo IDNA2008.
4. El proceso de solicitud de TLD de la ICANN no debe permitir las solicitudes de múltiples cadenas de caracteres latinas bajo la premisa de que unas son variantes de otras. La evaluación de las etiquetas de TLD propuestas para resolver confusiones requiere de un examen cuidadoso, pero eso no las convierte en variantes en cuanto al sentido en que se enfoca el estudio VIP.
5. Aunque el protocolo IDNA permite que un gran número de puntos de código aparezcan en las etiquetas de TLD basadas en el alfabeto latino, el grupo de estudio no recomienda la aceptación incondicional de ningún punto de código para su inclusión en la zona raíz del DNS, sin la demostración específica de garantía para su uso en el contexto lingüístico previsto. En contraste con las zonas de TLD que pueden contar con claros atributos específicos locales que permitan evaluar estas cuestiones, la raíz constituye un recurso compartido a nivel mundial y debe adaptarse a todas las comunidades lingüísticas en la manera más equitativa posible.
6. De este modo, la tabla de caracteres latinos para la zona raíz puede ser una recopilación descriptiva de los puntos de código que en realidad aparecen en esa zona, en lugar de un cuadro normativo de los puntos de código que pueden ser incluidos en la misma sin mayor consideración.

APÉNDICES

A. Equipo de estudio de caso del alfabeto latino

El Equipo de estudio de caso del alfabeto latino se conformó y reunió por primera vez en la 41^a reunión de la ICANN, celebrada en Singapur en el mes de junio de 2011. Con posterioridad a esa reunión, se reunió aproximadamente en forma quincenal por teléfono para considerar el conjunto de preguntas preliminares, discutir y analizar su relevancia e implicaciones en relación con el alfabeto latino, y para elaborar el presente informe.

A1. Miembros del equipo

Los miembros del Equipo de estudio de caso son:

Nombre	Función
Jothan Frakes	Coordinador del estudio de caso
Harald Alvestrand	Miembro del equipo
Andrzej Bartosiewicz	Miembro del equipo
Eric Brown	Miembro del equipo
Cary Karp	Miembro del equipo (en representación de la organización patrocinadora)
Nadya Morozova	Miembro del equipo
Francisco Obispo	Miembro del equipo
Giovanni Seppia	Miembro del equipo
Will Shorter	Miembro del equipo
Wil Tan	Miembro del equipo
Avri Doria	Observador
Leo Vegoda	Coordinador del estudio de caso
Francisco Arias	Experto en la materia (Operaciones de registro)
Kim Davies	Experto en la materia (Seguridad)
Nicholas Ostler	Experto en la materia (Lingüística)
Steve Sheng	Experto en la materia (Política)
Andrew Sullivan	Experto en la materia (Protocolos)

A2. Declaraciones de interés

Con el fin de garantizar la transparencia mediante el intercambio de información relevante sobre los intereses que los miembros del equipo tienen en relación con las áreas de estudio, se solicitó a los miembros del equipo la presentación de declaraciones escritas, en establecimiento de sus intereses.

Estas declaraciones están publicadas en línea, en la página web del equipo de estudio de caso.¹²

A3. Agradecimientos

12

El equipo agradece a la Fundación de infraestructura de Internet, la organización patrocinadora para el Estudio de caso del alfabeto latino.

El equipo también reconoce que los elementos clave de este informe se derivan de un Resumen de respaldo introductorio aportado por Cary Karp, en nombre de la organización patrocinadora, con el apoyo del Museo sueco de historia natural.¹³

B. Repertorio de caracteres del alfabeto latino

Con el fin de centrar la labor del Equipo de estudio de caso, se necesitó establecer una definición de cuáles caracteres eran considerados en el alcance.

Conforme las especificaciones del protocolo, se consideran IDNs válidos a un subconjunto limitado de puntos de código Unicode. Estos códigos "de protocolo válido" o "PVALID", limitan los puntos de código que se pueden representar dentro del ámbito del protocolo IDNA.

Este conjunto puede ser más limitado a aquellos puntos de código que reflejan los que pertenecen a los bloques de diferentes puntos de código Unicode que se atribuyen al alfabeto latino.

B1. Asignación de puntos de código del bloque Unicode al alfabeto latino

Los bloques latinos de la versión 6.1.0 de la Tabla de códigos de caracteres Unicode son los siguientes:

Nombre de la Tabla Unicode	Rango del punto de código
Alfabeto latino básico	U+0020 U+007F
Suplemento latino-1	U+0080 U+00FF
Alfabeto latino extendido-A	U+0100 U+017F
Alfabeto latino extendido-B	U+0180 U+024F
Alfabeto latino extendido-C	U+2C60 U+2C7F
Alfabeto latino extendido-D	U+A720 U+A7FF
Alfabeto latino extendido adicional	U+1E00 U+1EFF
Ligaduras del alfabeto latino	U+FB00 U+FB0F
Letras latinas de ancho completo	U+FF00 U+FF5E

B2. Puntos de código de protocolo válido asignados al alfabeto latino

Las tablas de esta sección y la siguiente se compilaron de acuerdo con los principios descritos en la Sección 6.4 del cuerpo principal de este informe. Estas tablas están destinadas a apoyar e ilustrar la acción del estudio y no pueden ser tratadas como autoritativas en ningún sentido. La declaración normativa de los puntos de código de protocolo válido se brinda en la RFC 5892¹⁴ con una tabla de referencia correspondiente en el Registro de Protocolo de la IANA.¹⁵

¹³ <http://mm.icann.org/pipermail/latin-vip/attachments/20111007/ed0ab5a4/latin-support-brief-final-0001.pdf>

¹⁴ <http://tools.ietf.org/html/rfc5892>

Los puntos de código asignados al alfabeto latino que se espera aparezcan en repositorios de caracteres IDN recabados a nivel local, se presentan a continuación en orden alfabético:

U+0061	; a # LETRA A MINÚSCULA LATINA
U+00E0	; à # LETRA A MINÚSCULA LATINA CON ACENTO GRAVE
U+00E1	; á # LETRA A MINÚSCULA LATINA CON ACENTO AGUDO
U+00E2	; â # LETRA A MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO
U+00E3	; ã # LETRA A MINÚSCULA LATINA CON TILDE
U+00E4	; ä # LETRA A MINÚSCULA LATINA CON DIÉRESIS
U+00E5	; â # LETRA A MINÚSCULA LATINA CON ANILLO ARRIBA
U+0101	; ā # LETRA A MINÚSCULA LATINA CON MACRÓN
U+0103	; ǎ # LETRA A MINÚSCULA LATINA CON BREVE
U+0105	; ą # LETRA A MINÚSCULA LATINA CON OGONEK
U+01CE	; ǣ # LETRA A MINÚSCULA LATINA CON CARÓN
U+01DF	; ā̄ # LETRA A MINÚSCULA LATINA CON DIÉRESIS Y MACRÓN
U+01E1	; ā̇ # LETRA A MINÚSCULA LATINA CON PUNTO ARRIBA Y MACRÓN
U+01FB	; ǻ # LETRA A MINÚSCULA LATINA CON ANILLO ARRIBA Y ACENTO AGUDO
U+0201	; â # LETRA A MINÚSCULA LATINA CON ACENTO GRAVE DOBLE
U+0203	; â # LETRA A MINÚSCULA LATINA CON BREVE INVERTIDA
U+1E01	; ą̇ # LETRA A MINÚSCULA LATINA CON ANILLO ABAJO
U+1EA1	; ą̈ # LETRA A MINÚSCULA LATINA CON PUNTO ABAJO
U+1EA3	; ą̇ # LETRA A MINÚSCULA CON GANCHO ARRIBA
U+1EA5	; ǻ̇ # LETRA A MINÚSCULA LATINA CON ACENTOS CIRCUNFLEJO Y AGUDO
U+1EA7	; ǻ̈ # LETRA A MINÚSCULA LATINA CON ACENTOS CIRCUNFLEJO Y GRAVE
U+1EA9	; ǻ̇ # LETRA A MINÚSCULA CON ACENTO CIRCUNFLEJO Y GANCHO
U+1EAB	; ǻ̈ # LETRA A MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO Y TILDE
U+1EAD	; ǻ̇ # LETRA A MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO Y PUNTO ABAJO
U+1EAF	; ǻ̈ # LETRA A MINÚSCULA LATINA CON BREVE Y ACENTO AGUDO
U+1EB1	; ǻ̇ # LETRA A MINÚSCULA LATINA CON BREVE Y ACENTO GRAVE
U+1EB3	; ǻ̈ # LETRA A MINÚSCULA LATINA CON BREVE Y GANCHO ARRIBA
U+1EB5	; ǻ̇ # LETRA A MINÚSCULA LATINA CON BREVE Y TILDE
U+1EB7	; ǻ̈ # LETRA A MINÚSCULA LATINA CON BREVE Y PUNTO ABAJO
U+0227	; à # LETRA A MINÚSCULA LATINA CON PUNTO ARRIBA
U+2C65	; ǻ̇ # LETRA A MINÚSCULA LATINA CON BASTÓN
U+0251	; a # LETRA ALFA MINÚSCULA LATINA
U+00E6	; æ # LETRA AE MINÚSCULA LATINA
U+0062	; b # LETRA B MINÚSCULA LATINA
U+0180	; b̄ # LETRA B MINÚSCULA LATINA CON BARRA
U+0183	; b̆ # LETRA B MINÚSCULA LATINA CON BARRA ARRIBA
U+1E03	; ḃ # LETRA B MINÚSCULA LATINA CON PUNTO ARRIBA
U+1E05	; b̈ # LETRA B MINÚSCULA LATINA CON PUNTO ABAJO
U+1E07	; b̆ # LETRA B MINÚSCULA LATINA CON LÍNEA ABAJO
U+0253	; ḃ # LETRA B MINÚSCULA LATINA CON GANCHO
U+0063	; c # LETRA C MINÚSCULA LATINA
U+00E7	; ç # LETRA C MINÚSCULA LATINA CON CEDILLA
U+0107	; ć # LETRA C MINÚSCULA LATINA CON ACENTO AGUDO
U+0109	; ȇ # LETRA C MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO
U+010B	; ȇ # LETRA C MINÚSCULA LATINA CON PUNTO ARRIBA

U+010D ; ċ # LETRA C MINÚSCULA LATINA CON CARÓN
 U+0188 ; ċ # LETRA C MINÚSCULA LATINA CON GANCHO
 U+023C ; ċ # LETRA C MINÚSCULA LATINA CON BASTÓN
 U+0255 ; ċ # LETRA C MINÚSCULA LATINA CON BUCLE
 U+A793 ; □ # LETRA C MINÚSCULA LATINA CON BARRA
 U+1E09 ; ċ # LETRA C MINÚSCULA LATINA CON CEDILLA Y ACENTO AGUDO
 U+2184 ; ɔ # LETRA C MINÚSCULA LATINA INVERTIDA
 U+0064 ; d # LETRA D MINÚSCULA LATINA
 U+010F ; ḋ # LETRA D MINÚSCULA LATINA CON CARÓN
 U+0111 ; d̄ # LETRA D MINÚSCULA LATINA CON BARRA
 U+018C ; d̄ # LETRA D MINÚSCULA LATINA CON BARRA ARRIBA
 U+0221 ; ḍ # LETRA D MINÚSCULA LATINA CON BUCLE
 U+1E0D ; ḍ # LETRA D MINÚSCULA LATINA CON PUNTO ABAJO
 U+1E0F ; ḍ # LETRA D MINÚSCULA LATINA CON LÍNEA ABAJO
 U+1E11 ; ḍ # LETRA D MINÚSCULA LATINA CON CEDILLA
 U+1E13 ; ḍ # LETRA D MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO ABAJO
 U+1E0B ; ḍ # LETRA D MINÚSCULA LATINA CON PUNTO ARRIBA
 U+0256 ; ḍ # LETRA D MINÚSCULA LATINA CON COLA
 U+0257 ; ḍ # LETRA D MINÚSCULA LATINA CON GANCHO
 U+1E9F ; □ # LETRA DELTA MINÚSCULA LATINA
 U+00F0 ; ð # LETRA ETH MINÚSCULA LATINA
 U+0065 ; e # LETRA E MINÚSCULA LATINA
 U+00E8 ; è # LETRA E MINÚSCULA LATINA CON ACENTO GRAVE
 U+00E9 ; é # LETRA E MINÚSCULA LATINA CON ACENTO AGUDO
 U+00EA ; ê # LETRA E MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO
 U+00EB ; ë # LETRA E MINÚSCULA LATINA CON DIÉRESIS
 U+0113 ; ē # LETRA E MINÚSCULA LATINA CON MACRÓN
 U+0115 ; ě # LETRA E MINÚSCULA LATINA CON BREVE
 U+0117 ; è # LETRA E MINÚSCULA LATINA CON PUNTO ARRIBA
 U+0119 ; ę # LETRA E MINÚSCULA LATINA CON OGONEK
 U+011B ; ě # LETRA E MINÚSCULA LATINA CON CARÓN
 U+0205 ; è # LETRA E MINÚSCULA LATINA CON ACENTO GRAVE DOBLE
 U+0207 ; ê # LETRA E MINÚSCULA LATINA CON BREVE INVERTIDA
 U+1E15 ; è # LETRA E MINÚSCULA LATINA CON MACRÓN Y ACENTO GRAVE
 U+1E17 ; é # LETRA E MINÚSCULA LATINA CON MACRÓN Y ACENTO AGUDO
 U+1E19 ; ę # LETRA E MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO ABAJO
 U+1E1B ; ẽ # LETRA E MINÚSCULA LATINA CON TILDE ABAJO
 U+1E1D ; ě # LETRA E MINÚSCULA LATINA CON CEDILLA Y BREVE
 U+1EB9 ; ę # LETRA E MINÚSCULA LATINA CON PUNTO ABAJO
 U+1EBB ; ê # LETRA E MINÚSCULA LATINA CON GANCHO ARRIBA
 U+1EBD ; ẽ # LETRA E MINÚSCULA LATINA CON TILDE
 U+1EBF ; é # LETRA E MINÚSCULA LATINA CON ACENTOS CIRCUNFLEJO Y AGUDO
 U+1EC1 ; è # LETRA E MINÚSCULA LATINA CON ACENTOS CIRCUNFLEJO Y GRAVE
 U+1EC3 ; ê # LETRA E MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO Y GANCHO
 ARRIBA
 U+1EC5 ; ẽ # LETRA E MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO Y TILDE
 U+1EC7 ; ê # LETRA E MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO Y PUNTO ABAJO
 U+0229 ; ę # LETRA E MINÚSCULA LATINA CON CEDILLA
 U+0247 ; ċ # LETRA E MINÚSCULA LATINA CON BASTÓN
 U+025B ; ε # LETRA E MINÚSCULA LATINA ABIERTA
 U+01DD ; ə # LETRA E MINÚSCULA LATINA ROTADA
 U+0259 ; ə # LETRA SCHWA MINÚSCULA LATINA
 U+0066 ; f # LETRA F MINÚSCULA LATINA

U+0192 ; f # LETRA F MINÚSCULA LATINA CON GANCHO
 U+1E1F ; ḟ # LETRA F MINÚSCULA LATINA CON PUNTO ARRIBA
 U+0067 ; g # LETRA G MINÚSCULA LATINA
 U+011D ; ĝ # LETRA G MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO
 U+011F ; ğ # LETRA G MINÚSCULA LATINA CON BREVE
 U+0121 ; ḡ # LETRA G MINÚSCULA LATINA CON PUNTO ARRIBA
 U+0123 ; ḡ̇ # LETRA G MINÚSCULA LATINA CON CEDILLA
 U+01E5 ; ḡ̄ # LETRA G MINÚSCULA LATINA CON BARRA
 U+01E7 ; ḡ̆ # LETRA G MINÚSCULA LATINA CON CARÓN
 U+01F5 ; ḡ̈ # LETRA G MINÚSCULA LATINA CON ACENTO AGUDO
 U+0260 ; ğ̣ # LETRA G MINÚSCULA LATINA CON GANCHO
 U+1E21 ; ḡ̅ # LETRA G MINÚSCULA LATINA CON MACRÓN
 U+1D77 ; ḡ̸ # LETRA G MINÚSCULA LATINA G ROTADA
 U+0263 ; γ # LETRA GAMMA MINÚSCULA LATINA
 U+0068 ; h # LETRA H MINÚSCULA LATINA
 U+0125 ; ĥ # LETRA H MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO
 U+0127 ; ĥ̄ # LETRA H MINÚSCULA LATINA CON BARRA
 U+021F ; ĥ̆ # LETRA H MINÚSCULA LATINA CON CARÓN
 U+1E23 ; ĥ̇ # LETRA H MINÚSCULA LATINA CON PUNTO ARRIBA
 U+1E25 ; ĥ̈ # LETRA H MINÚSCULA LATINA CON PUNTO ABAJO
 U+1E27 ; ĥ̉ # LETRA H MINÚSCULA LATINA CON DIÉRESES
 U+1E29 ; ĥ̊ # LETRA H MINÚSCULA LATINA CON CEDILLA
 U+1E2B ; ĥ̋ # LETRA H MINÚSCULA LATINA CON BREVE ABAJO
 U+2C68 ; ĥ̌ # LETRA H MINÚSCULA LATINA CON COMA
 U+1E96 ; ĥ̍ # LETRA H MINÚSCULA LATINA CON LÍNEA ABAJO
 U+0266 ; ĥ̎ # LETRA H MINÚSCULA LATINA CON GANCHO
 U+02AE ; ĥ̏ # LETRA H MINÚSCULA LATINA ROTADA CON ANZUELO
 U+02AF ; ĥ̐ # LETRA H MINÚSCULA LATINA ROTADA CON ANZUELO Y COLA
 U+0069 ; i # LETRA I MINÚSCULA LATINA
 U+00EC ; ì # LETRA I MINÚSCULA LATINA CON ACENTO GRAVE
 U+00ED ; í # LETRA I MINÚSCULA LATINA CON ACENTO AGUDO
 U+00EE ; î # LETRA I MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO
 U+00EF ; ï # LETRA I MINÚSCULA LATINA CON DIÉRESIS
 U+0129 ; ï̇ # LETRA I MINÚSCULA LATINA CON TILDE
 U+012B ; ï̈ # LETRA I MINÚSCULA LATINA CON MACRÓN
 U+012D ; ï̉ # LETRA I MINÚSCULA LATINA CON BREVE
 U+012F ; ï̊ # LETRA I MINÚSCULA LATINA CON OGONEK
 U+0131 ; i̋ # LETRA I MINÚSCULA LATINA SIN PUNTO
 U+01D0 ; ï̌ # LETRA I MINÚSCULA LATINA CON CARÓN
 U+0209 ; ï̍ # LETRA I MINÚSCULA LATINA CON ACENTO GRAVE DOBLE
 U+020B ; ï̎ # LETRA I MINÚSCULA LATINA CON BREVE INVERTIDA
 U+1E2D ; ï̏ # LETRA I MINÚSCULA LATINA CON TILDE ABAJO
 U+1E2F ; ï̐ # LETRA I MINÚSCULA LATINA CON DIÉRESIS Y ACENTO AGUDO
 U+1EC9 ; i̇ # LETRA I MINÚSCULA LATINA CON GANCHO ARRIBA
 U+1ECB ; ï # LETRA I MINÚSCULA LATINA CON PUNTO ABAJO
 U+0269 ; ι # LETRA IOTA MINÚSCULA LATINA
 U+006A ; j # LETRA J MINÚSCULA LATINA
 U+0135 ; ĵ # LETRA J MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO
 U+01F0 ; ĵ̇ # LETRA J MINÚSCULA LATINA CON CARÓN
 U+0237 ; j̈ # LETRA J MINÚSCULA LATINA SIN PUNTO
 U+0249 ; j̉ # LETRA J MINÚSCULA LATINA CON BARRA
 U+025F ; j̊ # LETRA J MINÚSCULA LATINA SIN PUNTO CON BARRA
 U+006B ; k # LETRA K MINÚSCULA LATINA

U+0137 ; ꞥ # LETRA K MINÚSCULA LATINA CON CEDILLA
 U+0199 ; Ꞧ # LETRA K MINÚSCULA LATINA CON GANCHO
 U+01E9 ; ꞧ # LETRA K MINÚSCULA LATINA CON CARÓN
 U+1E31 ; Ꞩ # LETRA K MINÚSCULA LATINA CON ACENTO AGUDO
 U+1E33 ; ꞩ # LETRA K MINÚSCULA LATINA CON PUNTO ABAJO
 U+1E35 ; Ɦ # LETRA K MINÚSCULA LATINA CON LÍNEA ABAJO
 U+2C6A ; Ɜ # LETRA K MINÚSCULA LATINA CON COMA
 U+0138 ; Ɡ # LETRA KRA MINÚSCULA LATINA
 U+006C ; Ɬ # LETRA L MINÚSCULA LATINA
 U+013A ; Ɪ # LETRA L MINÚSCULA LATINA CON ACENTO AGUDO
 U+013C ; ꞯ # LETRA L MINÚSCULA LATINA CON CEDILLA
 U+013E ; Ʞ # LETRA L MINÚSCULA LATINA CON CARÓN
 U+0142 ; Ʇ # LETRA L MINÚSCULA LATINA CON BASTÓN
 U+019A ; Ʝ # LETRA L MINÚSCULA LATINA CON BARRA
 U+0234 ; Ꭓ # LETRA L MINÚSCULA LATINA CON BUCLE
 U+1E37 ; Ꞵ # LETRA L MINÚSCULA LATINA CON PUNTO ABAJO
 U+1E39 ; ꞵ # LETRA L MINÚSCULA LATINA CON PUNTO ABAJO Y MACRÓN
 U+1E3B ; Ꞷ # LETRA L MINÚSCULA LATINA CON LÍNEA ABAJO
 U+1E3D ; ꞷ # LETRA L MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO ABAJO
 U+2C61 ; Ꞹ # LETRA L MINÚSCULA LATINA CON BARRA DOBLE
 U+006D ; ꞹ # LETRA M MINÚSCULA LATINA
 U+1E3F ; Ꞻ # LETRA M MINÚSCULA LATINA CON ACENTO AGUDO
 U+1E41 ; ꞻ # LETRA M MINÚSCULA LATINA CON PUNTO ARRIBA
 U+1E43 ; Ꞽ # LETRA M MINÚSCULA LATINA CON PUNTO ABAJO
 U+006E ; ꞽ # LETRA N MINÚSCULA LATINA
 U+00F1 ; Ꞿ # LETRA N MINÚSCULA LATINA CON TILDE
 U+0144 ; ꞿ # LETRA N MINÚSCULA LATINA CON ACENTO AGUDO
 U+0146 ; Ꟁ # LETRA N MINÚSCULA LATINA CON CEDILLA
 U+0148 ; ꟁ # LETRA N MINÚSCULA LATINA CON CARÓN
 U+019E ; Ꟃ # LETRA N MINÚSCULA LATINA CON PIERNA DERECHA LARGA
 U+01F9 ; ꟃ # LETRA N MINÚSCULA LATINA CON ACENTO GRAVE
 U+0235 ; Ꞔ # LETRA N MINÚSCULA LATINA CON BUCLE
 U+1E45 ; Ʂ # LETRA N MINÚSCULA LATINA CON PUNTO ARRIBA
 U+1E47 ; Ᶎ # LETRA N MINÚSCULA LATINA CON PUNTO ABAJO
 U+1E49 ; Ꟈ # LETRA N MINÚSCULA LATINA CON LÍNEA ABAJO
 U+1E4B ; ꟈ # LETRA N MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO ABAJO
 U+0272 ; Ꟊ # LETRA N MINÚSCULA LATINA CON GANCHO IZQUIERDO
 U+014B ; ꟊ # LETRA ENG MINÚSCULA LATINA
 U+006F ; Ɤ # LETRA O MINÚSCULA LATINA
 U+00F2 ; Ꟍ # LETRA O MINÚSCULA LATINA CON ACENTO GRAVE
 U+00F3 ; ꟍ # LETRA O MINÚSCULA LATINA CON ACENTO AGUDO
 U+00F4 ; ꟎ # LETRA O MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO
 U+00F5 ; ꟏ # LETRA O MINÚSCULA LATINA CON TILDE
 U+00F6 ; Ꟑ # LETRA O MINÚSCULA LATINA CON DIÉRESIS
 U+00F8 ; ꟑ # LETRA O MINÚSCULA LATINA CON BASTÓN
 U+014D ; ꟒ # LETRA O MINÚSCULA LATINA CON MACRÓN
 U+014F ; ꟓ # LETRA O MINÚSCULA LATINA CON BREVE
 U+0151 ; ꟔ # LETRA O MINÚSCULA LATINA CON ACENTO AGUDO DOBLE
 U+01A1 ; ꟕ # LETRA O MINÚSCULA LATINA CON ASTA
 U+01D2 ; Ꟗ # LETRA O MINÚSCULA LATINA CON CARÓN
 U+01EB ; ꟗ # LETRA O MINÚSCULA LATINA CON OGONEK
 U+01ED ; Ꟙ # LETRA O MINÚSCULA LATINA CON OGONEK Y MACRÓN
 U+01FF ; ꟙ # LETRA O MINÚSCULA LATINA CON BASTÓN Y ACENTO AGUDO

U+020D ; ò # LETRA O MINÚSCULA LATINA CON ACENTO GRAVE DOBLE
 U+020F ; ô # LETRA O MINÚSCULA CON BREVE INVERTIDA
 U+022B ; õ # LETRA O MINÚSCULA LATINA CON DIÉRESIS Y MACRÓN
 U+022D ; õ̃ # LETRA O MINÚSCULA LATINA CON TILDE Y MACRÓN
 U+022F ; ö # LETRA O MINÚSCULA LATINA CON PUNTO ARRIBA
 U+0231 ; ȭ # LETRA O MINÚSCULA LATINA CON PUNTO ARRIBA Y MACRÓN
 U+1E4D ; ó̃ # LETRA O MINÚSCULA LATINA CON TILDE Y ACENTO AGUDO
 U+1E4F ; õ̆ # LETRA O MINÚSCULA LATINA CON TILDE Y DIÉRESIS
 U+1E51 ; ò̄ # LETRA O MINÚSCULA LATINA CON MACRÓN Y ACENTO GRAVE
 U+1E53 ; ó̄ # LETRA O MINÚSCULA LATINA CON TILDE Y ACENTO AGUDO
 U+1ECD ; ɔ # LETRA O MINÚSCULA LATINA CON PUNTO ABAJO
 U+1ECF ; ô̇ # LETRA O MINÚSCULA LATINA CON GANCHO ARRIBA
 U+1ED1 ; ó̇ # LETRA O MINÚSCULA LATINA CON ACENTOS CIRCUNFLEJO Y AGUDO
 U+1ED3 ; ò̈ # LETRA O MINÚSCULA LATINA CON ACENTOS CIRCUNFLEJO Y GRAVE
 U+1ED5 ; ô̈ # LETRA O MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO Y GANCHO
 ARRIBA
 U+1ED7 ; õ̉ # LETRA O MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO Y TILDE
 U+1ED9 ; ổ # LETRA O MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO Y PUNTO
 ABAJO
 U+1EDB ; ó̊ # LETRA O MINÚSCULA LATINA CON ASTA Y ACENTO AGUDO
 U+1EDD ; ò̊ # LETRA O MINÚSCULA LATINA CON ASTA Y ACENTO GRAVE
 U+1EDF ; ô̋ # LETRA O MINÚSCULA LATINA CON ASTA Y GANCHO ARRIBA
 U+1EE1 ; õ̌ # LETRA O MINÚSCULA LATINA CON ASTA Y TILDE
 U+1EE3 ; ɔ̋ # LETRA O MINÚSCULA LATINA CON ASTA Y PUNTO ABAJO
 U+0254 ; ɔ̄ # LETRA O MINÚSCULA LATINA ABIERTA
 U+0275 ; ø # LETRA O MINÚSCULA LATINA CON BARRA
 U+0153 ; œ # LETRA OE MINÚSCULA LATINA CON LIGADURA
 U+0070 ; p # LETRA P MINÚSCULA LATINA
 U+01A5 ; þ # LETRA P MINÚSCULA LATINA CON GANCHO
 U+1E55 ; þ̇ # LETRA P MINÚSCULA LATINA CON ACENTO AGUDO
 U+1E57 ; þ̈ # LETRA P MINÚSCULA LATINA CON PUNTO ARRIBA
 U+0071 ; q # LETRA Q MINÚSCULA LATINA
 U+024B ; q̄ # LETRA Q MINÚSCULA LATINA CON COLA
 U+0072 ; r # LETRA R MINÚSCULA LATINA
 U+0155 ; ɾ # LETRA R MINÚSCULA LATINA CON ACENTO AGUDO
 U+0157 ; ɽ # LETRA R MINÚSCULA LATINA CON CEDILLA
 U+0159 ; ɿ # LETRA R MINÚSCULA LATINA CON CARÓN
 U+0211 ; ɿ̄ # LETRA R MINÚSCULA LATINA CON ACENTO GRAVE DOBLE
 U+0213 ; ɿ̇ # LETRA R MINÚSCULA LATINA CON BREVE INVERTIDA
 U+1E59 ; ɿ̈ # LETRA R MINÚSCULA LATINA CON PUNTO ARRIBA
 U+1E5B ; ɿ̊ # LETRA R MINÚSCULA LATINA CON PUNTO ABAJO
 U+1E5D ; ɿ̋ # LETRA R MINÚSCULA LATINA CON PUNTO ABAJO Y MACRÓN
 U+1E5F ; ɿ̌ # LETRA R MINÚSCULA LATINA CON LÍNEA ABAJO
 U+024D ; ɿ̍ # LETRA R MINÚSCULA LATINA CON BARRA
 U+027D ; ɿ̎ # LETRA R MINÚSCULA LATINA CON COLA
 U+027F ; ɿ̏ # LETRA R MINÚSCULA LATINA INVERTIDA CON ANZUELO
 U+0073 ; s # LETRA S MINÚSCULA LATINA
 U+015B ; ś # LETRA S MINÚSCULA LATINA CON ACENTO AGUDO
 U+015D ; ŝ # LETRA S MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO
 U+015F ; ş # LETRA S MINÚSCULA LATINA CON CEDILLA
 U+0161 ; ŝ̄ # LETRA S MINÚSCULA LATINA CON CARÓN
 U+0219 ; ş̇ # LETRA S MINÚSCULA LATINA CON COMA ABAJO
 U+1E61 ; ş̈ # LETRA S MINÚSCULA LATINA CON PUNTO ARRIBA

U+1E63 ; § # LETRA S MINÚSCULA LATINA CON PUNTO ABAJO
 U+1E65 ; ŝ # LETRA S MINÚSCULA LATINA CON ACENTO AGUDO Y PUNTO ARRIBA
 U+1E67 ; š # LETRA S MINÚSCULA LATINA CON CARÓN Y PUNTO ARRIBA
 U+1E69 ; ſ # LETRA S MINÚSCULA LATINA CON PUNTOS ABAJO Y ARRIBA
 U+023F ; ꝛ # LETRA S MINÚSCULA LATINA CON COLA ORNAMENTAL
 U+00DF ; ß # LETRA S MINÚSCULA LATINA AGUDA
 U+1E9C ; Ꝑ # LETRA S MINÚSCULA LATINA LARGA CON BARRA DIAGONAL
 U+1E9D ; ꝑ # LETRA S MINÚSCULA LATINA LARGA CON BARRA ALTA
 U+0283 ; ƒ # LETRA ESH MINÚSCULA LATINA
 U+0074 ; t # LETRA T MINÚSCULA LATINA
 U+0163 ; ƚ # LETRA T MINÚSCULA LATINA CON CEDILLA
 U+0165 ; ƥ # LETRA T MINÚSCULA LATINA CON CARÓN
 U+0167 ; Ƨ # LETRA T MINÚSCULA LATINA CON BARRA
 U+01AB ; ƨ # LETRA T MINÚSCULA LATINA CON GANCHO PALATAL
 U+01AD ; Ʃ # LETRA T MINÚSCULA LATINA CON GANCHO
 U+1E6B ; ƪ # LETRA T MINÚSCULA LATINA CON PUNTO ARRIBA
 U+1E6D ; ƫ # LETRA T MINÚSCULA LATINA CON PUNTO ABAJO
 U+1E6F ; Ƭ # LETRA T MINÚSCULA LATINA CON LÍNEA ABAJO
 U+1E71 ; Ʒ # LETRA T MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO ABAJO
 U+0236 ; Ƨ # LETRA T MINÚSCULA LATINA CON BUCLE
 U+021B ; ƨ # LETRA T MINÚSCULA LATINA CON COMA ABAJO
 U+1E97 ; Ʃ # LETRA T MINÚSCULA LATINA CON DIÉRESIS
 U+2C66 ; ƪ # LETRA T MINÚSCULA LATINA CON BASTÓN DIAGONAL
 U+0288 ; ƫ # LETRA T MINÚSCULA LATINA CON GANCHO RETROFLEJO
 U+00FE ; þ # LETRA THORN MINÚSCULA LATINA
 U+0075 ; u # LETRA U MINÚSCULA LATINA
 U+00F9 ; ù # LETRA U MINÚSCULA LATINA CON ACENTO GRAVE
 U+00FA ; ú # LETRA U MINÚSCULA LATINA CON ACENTO AGUDO
 U+00FB ; û # LETRA U MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO
 U+00FC ; ü # LETRA U MINÚSCULA LATINA CON DIÉRESIS
 U+0169 ; ŭ # LETRA U MINÚSCULA LATINA CON TILDE
 U+016B ; ū # LETRA U MINÚSCULA LATINA CON MACRÓN
 U+016D ; ŭ̄ # LETRA U MINÚSCULA LATINA CON BREVE
 U+016F ; ŭ̇ # LETRA U MINÚSCULA LATINA CON ANILLO ARRIBA
 U+0171 ; ŭ̂ # LETRA U MINÚSCULA LATINA CON ACENTO AGUDO DOBLE
 U+0173 ; ƚ # LETRA U MINÚSCULA LATINA CON OGONEK
 U+01B0 ; ƚ # LETRA U MINÚSCULA LATINA CON ASTA
 U+01D4 ; ŭ̃ # LETRA U MINÚSCULA LATINA CON CARÓN
 U+01D6 ; ŭ̄̇ # LETRA U MINÚSCULA LATINA CON DIÉRESIS Y MACRÓN
 U+01D8 ; ŭ̂̇ # LETRA U MINÚSCULA LATINA CON DIÉRESIS Y ACENTO AGUDO
 U+01DA ; ŭ̄̃ # LETRA U MINÚSCULA LATINA CON DIÉRESIS Y CARÓN
 U+01DC ; ŭ̄̂̇ # LETRA U MINÚSCULA LATINA CON DIÉRESIS Y ACENTO GRAVE
 U+0215 ; ŭ̄̂̇ # LETRA U MINÚSCULA LATINA CON ACENTO GRAVE DOBLE
 U+0217 ; ŭ̄̂̇ # LETRA U MINÚSCULA LATINA CON BREVE INVERTIDA
 U+1E73 ; ƚ # LETRA U MINÚSCULA LATINA CON DIÉRESIS ABAJO
 U+1E75 ; ƚ # LETRA U MINÚSCULA LATINA CON TILDE ABAJO
 U+1E77 ; ƚ # LETRA U MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO ABAJO
 U+1E79 ; ŭ̇ # LETRA U MINÚSCULA LATINA CON TILDE Y ACENTO AGUDO
 U+1E7B ; ŭ̄̇ # LETRA U MINÚSCULA LATINA CON MACRÓN Y DIÉRESIS
 U+1EE5 ; ƚ # LETRA U MINÚSCULA LATINA CON PUNTO ABAJO
 U+1EE7 ; ŭ̇ # LETRA U MINÚSCULA LATINA CON GANCHO ARRIBA
 U+1EE9 ; ŭ̇ # LETRA U MINÚSCULA LATINA CON ASTA Y ACENTO AGUDO
 U+1EEB ; ŭ̇ # LETRA U MINÚSCULA LATINA CON ASTA Y ACENTO GRAVE

U+1EED ; ũ# LETRA U MINÚSCULA LATINA CON ASTA Y GANCHO ARRIBA
 U+1EEF ; ũ̃# LETRA U MINÚSCULA LATINA CON ASTA Y TILDE
 U+1EF1 ; ũ̇# LETRA U MINÚSCULA LATINA CON ASTA Y PUNTO ABAJO
 U+0076 ; v# LETRA V MINÚSCULA LATINA
 U+1E7D ; ñ# LETRA V MINÚSCULA LATINA CON TILDE
 U+1E7F ; v̇# LETRA V MINÚSCULA LATINA CON PUNTO ABAJO
 U+1EFD ; □# LETRA V MINÚSCULA LATINA DE GALÉS MEDIO
 U+2C71 ; v# LETRA V MINÚSCULA LATINA CON GANCHO DERECHO
 U+2C74 ; v̄# LETRA V MINÚSCULA LATINA CON BUCLE
 U+028B ; v̂# LETRA V MINÚSCULA LATINA CON GANCHO
 U+028C ; v̂# LETRA V MINÚSCULA LATINA ROTADA
 U+0077 ; w# LETRA W MINÚSCULA LATINA
 U+0175 ; ŵ# LETRA W MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO
 U+1E81 ; w̃# LETRA W MINÚSCULA LATINA CON ACENTO GRAVE
 U+1E83 ; ẇ# LETRA W MINÚSCULA LATINA CON ACENTO AGUDO
 U+1E85 ; ẅ# LETRA W MINÚSCULA LATINA CON DIÉRESIS
 U+1E87 ; w̄# LETRA W MINÚSCULA LATINA CON PUNTO ARRIBA
 U+1E89 ; ẇ# LETRA W MINÚSCULA LATINA CON PUNTO ABAJO
 U+1E98 ; ŵ# LETRA W MINÚSCULA LATINA CON ANILLO ARRIBA
 U+2C73 ; w# LETRA W MINÚSCULA LATINA CON GANCHO
 U+01BF ; p# LETRA WYNN LATINA
 U+0078 ; x# LETRA X MINÚSCULA LATINA
 U+1E8B ; ẋ# LETRA X MINÚSCULA LATINA CON PUNTO ARRIBA
 U+1E8D ; ẍ# LETRA X MINÚSCULA LATINA CON DIÉRESIS
 U+0079 ; y# LETRA Y MINÚSCULA LATINA
 U+00FD ; ý# LETRA Y MINÚSCULA LATINA CON ACENTO AGUDO
 U+00FF ; ý̈# LETRA Y MINÚSCULA LATINA CON DIÉRESIS
 U+0177 ; ý̂# LETRA Y MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO
 U+01B4 ; ý# LETRA Y MINÚSCULA LATINA CON GANCHO
 U+0233 ; ÿ# LETRA Y MINÚSCULA LATINA CON MACRÓN
 U+024F ; y# LETRA Y MINÚSCULA LATINA CON BARRA
 U+1E99 ; ý̂# LETRA Y MINÚSCULA LATINA CON ANILLO ARRIBA
 U+1EF3 ; ý̃# LETRA Y MINÚSCULA LATINA CON ACENTO GRAVE
 U+1EF5 ; ý̇# LETRA Y MINÚSCULA LATINA CON PUNTO ABAJO
 U+1EF7 ; ý̄# LETRA Y MINÚSCULA LATINA CON GANCHO ARRIBA
 U+1EF9 ; ý̄# LETRA Y MINÚSCULA LATINA CON TILDE
 U+1EFF ; □# LETRA Y MINÚSCULA LATINA CON BUCLE
 U+1E8F ; ý̇# LETRA Y MINÚSCULA LATINA CON PUNTO ARRIBA
 U+021D ; ʒ# LETRA YOGH MINÚSCULA LATINA
 U+007A ; z# LAETRA Z MINÚSCULA LATINA
 U+017A ; ž# LETRA Z MINÚSCULA LATINA CON ACENTO AGUDO
 U+017C ; ž̇# LETRA Z MINÚSCULA LATINA CON PUNTO ARRIBA
 U+017E ; ž̂# LETRA Z MINÚSCULA LATINA CON CARÓN
 U+01B6 ; z# LETRA Z MINÚSCULA LATINA CON BARRA
 U+0225 ; z# LETRA Z MINÚSCULA LATINA CON GANCHO
 U+0240 ; ʒ# LETRA Z MINÚSCULA LATINA CON COLA ORNAMENTAL
 U+1E91 ; ž̂# LETRA Z MINÚSCULA LATINA CON ACENTO CIRCUNFLEJO
 U+1E93 ; ż# LETRA Z MINÚSCULA LATINA CON PUNTO ABAJO
 U+1E95 ; z̄# LETRA Z MINÚSCULA LATINA CON LÍNEA ABAJO
 U+2C6C ; z# LETRA Z MINÚSCULA LATINA CON COMA
 U+0292 ; ʒ# LETRA EZH MINÚSCULA LATINA
 U+01B9 ; ʒ̂# LETRA EZH MINÚSCULA LATINA INVERTIDA
 U+01BA ; ʒ̂# LETRA EZH MINÚSCULA LATINA CON COLA

U+01EF	; ǰ # LETRA EZH MINÚSCULA LATINA CON CARÓN
U+01C0	; # LETRA LATINA DE CHASQUIDO DENTAL
U+01C1	; # LETRA LATINA DE CHASQUIDO LATERAL
U+01C2	; † # LETRA LATINA DE CHASQUIDO ALVEOLAR
U+01C3	; ! # LETRA LATINA DE CHASQUIDO RETROFLEJO
U+0294	; ʔ # LETRA LATINA DE OCLUSIÓN GLOTAL

B3. Puntos de código de reserva

Los siguientes puntos de código también están disponibles para su inclusión en los IDNs de alfabeto latino, aunque el grupo de estudio sólo espera que aparezcan en circunstancias especiales —que no incluyen el uso de la zona raíz del DNS—, por lo cual los ha colocado en una tabla separada. Estos puntos de código están listados en orden numérico en virtud de sus títulos o subtítulos de bloque Unicode:

EXTENSIONES IPA (alfabeto fonético internacional)

U+0250	; v # LETRA A MINÚSCULA LATINA ROTADA
U+0252	; v̥ # LETRA ALFA MINÚSCULA LATINA ROTADA
U+0258	; ɘ # LETRA E MINÚSCULA LATINA INVERTIDA
U+025A	; ə # LETRA SCHWA MINÚSCULA LATINA CON GANCHO
U+025C	; ɜ # LETRA E MINÚSCULA LATINA ABIERTA INVERTIDA
U+025D	; ɜ̥ # LETRA E MINÚSCULA LATINA ABIERTA INVERTIDA CON GANCHO
U+025E	; ɞ # LETRA E MINÚSCULA LATINA ABIERTA INVERTIDA CERRADA
U+0264	; ɣ # LETRA ASTA DE CARNERO MINÚSCULA LATINA
U+0265	; ɥ # LETRA H MINÚSCULA LATINA ROTADA
U+0267	; ɦ # LETRA HENG MINÚSCULA LATINA CON GANCHO
U+0268	; ī # LETRA I MINÚSCULA LATINA CON BARRA
U+026A	; ɪ # LETRA I MAYÚSCULA LATINA PEQUEÑA
U+026B	; ɩ̄ # LETRA L MINÚSCULA LATINA CON TILDE MEDIO
U+026C	; ɬ # LETRA L MINÚSCULA LATINA CON CINTO
U+026D	; ɭ # LETRA L MINÚSCULA LATINA CON GANCHO RETROFLEJO
U+026E	; ɮ # LETRA LEZH MINÚSCULA LATINA
U+026F	; ɯ # LETRA M MINÚSCULA LATINA ROTADA
U+0270	; ɰ # LETRA M MINÚSCULA LATINA ROTADA CON PIERNA LARGA
U+0271	; ɱ # LETRA M MINÚSCULA LATINA CON GANCHO
U+0273	; ɲ # LETRA N MINÚSCULA LATINA CON GANCHO RETROFLEJO
U+0274	; ɳ # LETRA N MAYÚSCULA LATINA PEQUEÑA
U+0276	; œ # LETRA OE MAYÚSCULA LATINA PEQUEÑA
U+0277	; ɶ # LETRA OMEGA MINÚSCULA LATINA CERRADA
U+0278	; ϕ # LETRA PHI MINÚSCULA LATINA
U+0279	; ɽ # LETRA R MINÚSCULA LATINA ROTADA
U+027A	; ɽ̥ # LETRA R MINÚSCULA LATINA ROTADA CON PIERNA LARGA
U+027B	; ɽ̥ # LETRA R MINÚSCULA LATINA ROTADA CON GANCHO
U+027C	; ɽ̥ # LETRA R MINÚSCULA LATINA CON PIERNA LARGA
U+027E	; ɾ # LETRA R MINÚSCULA LATINA CON ANZUELO
U+0280	; ɿ # LETRA R MAYÚSCULA LATINA PEQUEÑA

U+0281	; ʀ # LETRA R MAYÚSCULA LATINA PEQUEÑA INVERTIDA
U+0282	; ſ # LETRA S MINÚSCULA LATINA CON GANCHO
U+0284	; ꝥ # LETRA J MINÚSCULA LATINA SIN PUNTO CON BARRA Y GANCHO
U+0285	; ʀ # LETRA ESH MINÚSCULA LATINA INVERTIDA ACHAPARRADA
U+0286	; Ꝧ # LETRA ESH MINÚSCULA LATINA CON BUCLE
U+0287	; ʀ # LETRA T MINÚSCULA LATINA ROTADA
U+0289	; ʀ # LETRA U MINÚSCULA LATINA CON BARRA
U+028A	; ʀ # LETRA ÍPSILON MINÚSCULA LATINA
U+028D	; ʀ # LETRA W MINÚSCULA LATINA ROTADA
U+028E	; ʀ # LETRA Y MINÚSCULA LATINA ROTADA
U+028F	; ʀ # LETRA Y MAYÚSCULA LATINA PEQUEÑA
U+0290	; ʀ # LETRA Z MINÚSCULA LATINA CON GANCHO RETROFLEJO
U+0291	; ʀ # LETRA Z MINÚSCULA LATINA CON BUCLE
U+0293	; ʀ # LETRA EZH MINÚSCULA LATINA CON BUCLE
U+0295	; ʀ # LETRA LATINA CON PRONUNCIACIÓN FARÍNGEA FRICATIVA
U+0296	; ʀ # LETRA LATINA CON OBSTRUCCIÓN GLOTA INVERTIDA
U+0297	; ʀ # LETRA C LATINA ESTIRADA
U+0298	; ʀ # LETRA LATINA CON CHASQUIDO BILABIAL
U+0299	; ʀ # LETRA B MAYÚSCULA LATINA PEQUEÑA
U+029A	; ʀ # LETRA E MINÚSCULA LATINA ABIERTA CERRADA
U+029B	; ʀ # LETRA G MAYÚSCULA LATINA PEQUEÑA CON GANCHO
U+029C	; ʀ # LETRA H MAYÚSCULA LATINA PEQUEÑA
U+029D	; ʀ # LETRA J MAYÚSCULA LATINA PEQUEÑA CON COLA CRUZADA
U+029E	; ʀ # LETRA K MINÚSCULA LATINA ROTADA
U+029F	; ʀ # LETRA L MAYÚSCULA LATINA PEQUEÑA
U+02A0	; ʀ # LETRA Q MINÚSCULA LATINA CON GANCHO
U+02A1	; ʀ # LETRA LATINA DE OCLUSIÓN GLOTA CON BARRA
U+02A2	; ʀ # LETRA LATINA DE OCLUSIÓN GLOTA INVERTIDA CON BARRA
U+02A3	; ʀ # LETRA DZ MINÚSCULA LATINA DÍGRAFO
U+02A4	; ʀ # LETRA DEZH MINÚSCULA LATINA DÍGRAFO
U+02A5	; ʀ # LETRA DZ MINÚSCULA LATINA DÍGRAFO CON BUCLE
U+02A6	; ʀ # LETRA TS MINÚSCULA LATINA DÍGRAFO
U+02A7	; ʀ # LETRA TESH MINÚSCULA LATINA DÍGRAFO
U+02A8	; ʀ # LETRA TC MINÚSCULA LATINA DÍGRAFO CON BUCLE
U+02A9	; ʀ # LETRA FENG MINÚSCULA LATINA DÍGRAFO
U+02AA	; ʀ # LETRA LS MINÚSCULA LATINA DÍGRAFO
U+02AB	; ʀ # LETRA LZ MINÚSCULA LATINA DÍGRAFO
U+02AC	; ʀ # LETRA LATINA BILABIAL PERCUSIVA
U+02AD	; ʀ # LETRA LATINA BIDENTAL PERCUSIVA

EXTENSIONES FONÉTICAS

U+1D6B	; ʀ # LETRA UE MINÚSCULA LATINA
U+1D6C	; ʀ # LETRA B MINÚSCULA LATINA CON TILDE MEDIO
U+1D6D	; ʀ # LETRA D MINÚSCULA LATINA CON TILDE MEDIO
U+1D6E	; ʀ # LETRA F MINÚSCULA LATINA CON TILDE MEDIO
U+1D6F	; ʀ # LETRA M MINÚSCULA LATINA CON TILDE MEDIO
U+1D70	; ʀ # LETRA N MINÚSCULA LATINA CON TILDE MEDIO
U+1D71	; ʀ # LETRA P MINÚSCULA LATINA CON TILDE MEDIO

U+1D72	; ɣ # LETRA R MINÚSCULA LATINA CON TILDE MEDIO
U+1D73	; ɥ # LETRA R MINÚSCULA LATINA CON ANZUELO Y TILDE MEDIO
U+1D74	; ȝ # LETRA S MINÚSCULA LATINA CON TILDE MEDIO
U+1D75	; ȥ # LETRA T MINÚSCULA LATINA CON TILDE MEDIO
U+1D76	; ȥ # LETRA Z MINÚSCULA LATINA CON TILDE MEDIO
U+1D79	; ȥ # LETRA G MINÚSCULA LATINA INSULAR
U+1D7A	; ȥ # LETRA TH MINÚSCULA LATINA CON BASTÓN ATRAVESADO
U+1D7B	; ı # LETRA I MAYÚSCULA LATINA PEQUEÑA CON BARRA
U+1D7C	; ı # LETRA IOTA MINÚSCULA LATINA CON BARRA
U+1D7D	; ƀ # LETRA P MINÚSCULA LATINA CON BARRA
U+1D7E	; ƀ # LETRA U MAYÚSCULA LATINA PEQUEÑA CON BARRA
U+1D7F	; ƀ # LETRA IPSILON MINÚSCULA LATINA CON BARRA
U+1D80	; Ɓ # LETRA B MINÚSCULA LATINA CON GANCHO PALATAL
U+1D81	; Ɓ # LETRA D MINÚSCULA LATINA CON GANCHO PALATAL
U+1D82	; Ɓ # LETRA F MINÚSCULA LATINA CON GANCHO PALATAL
U+1D83	; Ɓ # LETRA G MINÚSCULA LATINA CON GANCHO PALATAL
U+1D84	; Ɓ # LETRA K MINÚSCULA LATINA CON GANCHO PALATAL
U+1D85	; Ɓ # LETRA L MINÚSCULA LATINA CON GANCHO PALATAL
U+1D86	; Ɓ # LETRA M MINÚSCULA LATINA CON GANCHO PALATAL
U+1D87	; Ɓ # LETRA N MINÚSCULA LATINA CON GANCHO PALATAL
U+1D88	; Ɓ # LETRA P MINÚSCULA LATINA CON GANCHO PALATAL
U+1D89	; Ɓ # LETRA R MINÚSCULA LATINA CON GANCHO PALATAL
U+1D8A	; Ɓ # LETRA S MINÚSCULA LATINA CON GANCHO PALATAL
U+1D8B	; Ɓ # LETRA ESH MINÚSCULA LATINA CON GANCHO PALATAL
U+1D8C	; Ɓ # LETRA V MINÚSCULA LATINA CON GANCHO PALATAL
U+1D8D	; Ɓ # LETRA X MINÚSCULA LATINA CON GANCHO PALATAL
U+1D8E	; Ɓ # LETRA Z MINÚSCULA LATINA CON GANCHO PALATAL
U+1D8F	; ǂ # LETRA A MINÚSCULA LATINA CON GANCHO RETROFLEJO
U+1D90	; ǂ # LETRA ALFA MINÚSCULA LATINA CON GANCHO RETROFLEJO
U+1D91	; ǂ # LETRA D MINÚSCULA LATINA CON GANCHO Y COLA
U+1D92	; ǂ # LETRA E MINÚSCULA LATINA CON GANCHO RETROFLEJO
U+1D93	; ǂ # LETRA E MINÚSCULA LATINA ABIERTA CON GANCHO RETROFLEJO
U+1D94	; ǂ # LETRA E MINÚSCULA LATINA ABIERTA INVERTIDA CON GANCHO RETROFLEJO
U+1D95	; ǂ # LETRA SCHWA MINÚSCULA LATINA CON GANCHO RETROFLEJO
U+1D96	; ǂ # LETRA I MINÚSCULA LATINA CON GANCHO RETROFLEJO
U+1D97	; ǂ # LETRA O MINÚSCULA LATINA ABIERTA CON GANCHO RETROFLEJO
U+1D98	; ǂ # LETRA ESH MINÚSCULA LATINA CON GANCHO RETROFLEJO
U+1D99	; ǂ # LETRA U MINÚSCULA LATINA CON GANCHO RETROFLEJO
U+1D9A	; ǂ # LETRA EZH MINÚSCULA LATINA CON GANCHO RETROFLEJO

EXTENSIÓN C DEL ALFABETO LATINO

Letras claudias

U+2C76 ; ɀ # LETRA MEDIA H MINÚSCULA LATINA

Adiciones para UPA (alfabeto fonético urálico)

U+2C77 ; ɀ # LETRA PHI MINÚSCULA LATINA SIN COLA

U+2C78 ; ɀ # LETRA E MINÚSCULA LATINA CON MUESCA

U+2C79 ; ☐ # LETRA R MINÚSCULA LATINA CON COLA
U+2C7A ; ☐ # LETRA O MINÚSCULA LATINA CON ANILLO BAJO ADENTRO
U+2C7B ; ☐ # LETRA E MAYÚSCULA LATINA PEQUEÑA ROTADA

EXTENSIÓN D DEL ALFABETO LATINO

Ediciones egipcias

U+A723 ; ☐ # LETRA ALEF EGIPTOLÓGICA MINÚSCULA LATINA
U+A725 ; ☐ # LETRA AIN EGIPTOLÓGICA MINÚSCULA LATINA

Adiciones mayas

U+A727 ; ☐ # LETRA HENG MINÚSCULA LATINA
U+A729 ; ☐ # LETRA TZ MINÚSCULA LATINA
U+A72B ; ☐ # LETRA TRESILLO MINÚSCULA LATINA
U+A72D ; ☐ # LETRA CUATRILLO MINÚSCULA LATINA
U+A72F ; ☐ # LETRA CUATRILLO MINÚSCULA LATINA CON COMA

Adiciones medievales

U+A730 ; ☐ # LETRA F MAYÚSCULA LATINA PEQUEÑA
U+A731 ; ☐ # LETRA S MAYÚSCULA LATINA PEQUEÑA
U+A733 ; ☐ # LETRA AA MINÚSCULA LATINA
U+A735 ; ☐ # LETRA AO MINÚSCULA LATINA
U+A737 ; ☐ # LETRA AU MINÚSCULA LATINA
U+A739 ; ☐ # LETRA AV MINÚSCULA LATINA
U+A73B ; ☐ # LETRA AV MINÚSCULA LATINA CON BARRA HORIZONTAL
U+A73D ; ☐ # LETRA AY MINÚSCULA LATINA
U+A73F ; ☐ # LETRA C MINÚSCULA LATINA INVERTIDA CON PUNTO
U+A741 ; ☐ # LETRA K MINÚSCULA LATINA CON BARRA
U+A743 ; ☐ # LETRA K MINÚSCULA LATINA CON BASTÓN DIAGONAL
U+A745 ; ☐ # LETRA K MINÚSCULA LATINA CON BARRA Y BASTÓN DIAGONAL
U+A747 ; ☐ # LETRA L MINÚSCULA LATINA PARTIDA
U+A749 ; ☐ # LETRA L MINÚSCULA LATINA CON BARRA ALTA
U+A74B ; ☐ # LETRA O MINÚSCULA LATINA CON BARRA LARGA SOBREPUESTA
U+A74D ; ☐ # LETRA O MINÚSCULA PEQUEÑA CON BUCLE
U+A74F ; ☐ # LETRA OO MINÚSCULA PEQUEÑA
U+A751 ; ☐ # LETRA P MINÚSCULA PEQUEÑA CON BARRA ATRAVESADA
U+A753 ; ☐ # LETRA P MINÚSCULA LATINA CON FLORITURA
U+A755 ; ☐ # LETRA P MINÚSCULA LATINA CON COLA DE ARDILLA
U+A757 ; ☐ # LETRA Q MINÚSCULA LATINA CON BARRA ATRAVESADA
U+A759 ; ☐ # LETRA Q MINÚSCULA LATINA CON BASTÓN DIAGONAL
U+A75B ; ☐ # LETRA R MINÚSCULA LATINA CON ROTONDA
U+A75D ; ☐ # LETRA RUM MINÚSCULA LATINA CON ROTONDA
U+A75F ; ☐ # LETRA V MINÚSCULA LATINA CON BASTÓN DIAGONAL
U+A761 ; ☐ # LETRA VY MINÚSCULA LATINA
U+A763 ; ☐ # LETRA Z VISIGÓTICA MINÚSCULA LATINA
U+A765 ; ☐ # LETRA THORN MINÚSCULA LATINA CON BARRA
U+A767 ; ☐ # LETRA THORN MINÚSCULA LATINA CON BARRA ATRAVESADA
U+A769 ; ☐ # LETRA VEND MINÚSCULA LATINA
U+A76B ; ☐ # LETRA ET MINÚSCULA LATINA

U+A76D ; ☐ # LETRA IS MINNÚSCULA LATINA
U+A76F ; ☐ # LETRA CON MINÚSCULA LATINA
U+A771 ; ☐ # LETRA DUM MINÚSCULA LATINA
U+A772 ; ☐ # LETRA LUM MINÚSCULA LATINA
U+A773 ; ☐ # LETRA MUM MINÚSCULA LATINA
U+A774 ; ☐ # LETRA NUM MINÚSCULA LATINA
U+A775 ; ☐ # LETRA RUM MINÚSCULA LATINA
U+A776 ; ☐ # LETRA RUM MAYÚSCULA LATINA PEQUEÑA
U+A777 ; ☐ # LETRA TUM MINÚSCULA LATINA
U+A778 ; ☐ # LETRA UM MINÚSCULA LATINA

Letras insulares y celtas

U+A77A ; ☐ # LETRA D MINÚSCULA LATINA INSULAR
U+A77C ; ☐ # LETRA F MINÚSCULA LATINA INSULAR
U+A77F ; ☐ # LETRA G MINÚSCULA LATINA INSULAR ROTADA
U+A781 ; ☐ # LETRA L MINÚSCULA LATINA ROTADA
U+A783 ; ☐ # LETRA R MINÚSCULA LATINA INSULAR
U+A785 ; ☐ # LETRA S MINÚSCULA LATINA INSULAR
U+A787 ; ☐ # LETRA T MINÚSCULA LATINA INSULAR

Letras ortográficas para sonidos glotales

U+A78C ; ' # LETRA SALTILLO MINÚSCULA LATINA

Antiguas letras epigráficas romanas

U+A7FB ; ☐ # LETRA F EPIGRÁFICA LATINA INVERTIDA
U+A7FC ; ☐ # LETRA P EPIGRÁFICA LATINA INVERTIDA
U+A7FD ; ☐ # LETRA M EPIGRÁFICA LATINA INVERTIDA
U+A7FE ; ☐ # LETRA I LONGA EPIGRÁFICA LATINA
U+A7FF ; ☐ # LETRA M ARCAICA EPIGRÁFICA LATINA

Símbolos fonéticos

U+A78E ; ☐ # LETRA L MINÚSCULA LATINA CON GANCHO CIRCUNFLEJO Y CINTO

Letras de janalif

U+A791 ; ☐ # LETRA N MINÚSCULA LATINA CON COMA

Letras del idioma oficial de Letonia para ortografía previa a 1921

U+A7A1 ; ☐ # LETRA G MINÚSCULA LATINA CON BARRA OBLÍCUA
U+A7A3 ; ☐ # LETRA K MINÚSCULA LATINA CON BARRA OBLÍCUA
U+A7A5 ; ☐ # LETRA N MINÚSCULA LATINA CON BARRA OBLÍCUA
U+A7A7 ; ☐ # LETRA R MINÚSCULA LATINA CON BARRA OBLÍCUA
U+A7A9 ; ☐ # LETRA S MINÚSCULA LATINA CON BARRA OBLÍCUA

Agregado para UPA (alfabeto fonético urálico)

U+A7FA ; ☐ # LETRA M MAYÚSCULA PEQUEÑA LATINA ROTADA