

ICANN73 | Semana de preparación – Actualización sobre el Programa de Nombres de Dominio

Internacionalizados (IDN)

Martes, 22 de febrero de 2022 – 09:00 a 10:00 AST

PITINAN KOOARMORNPATANA: Gracias. Bienvenidos a todos. Bienvenidos a la primera sesión de todas las sesiones de la semana de preparación de ICANN73. Buenos días, buenas tardes, buenas noches. Yo soy Pitinan Kooarmornpatana, del equipo del programa de IDN-UA. Quiero darles la bienvenida a esta sesión sobre actualización del programa de IDN.

Será una sesión de 60 minutos que cubrirá los siguientes temas. En primer lugar tendremos una actualización sobre las pautas para la implementación de IDN. Luego pasaremos al tema de códigos de escritura adicionales para IDN. Luego tendremos una actualización sobre el estado de RZ-LGR, información general, actualización del panel de generación de Myanmar y también tendremos a un miembro del panel, Marc Blanchet, que nos dará una actualización sobre RZ-LGR versión 5. Sin más, vamos a pasar a la primera diapositiva.

En cuanto a las pautas para la implementación de IDN, esta es la información general. Son las políticas y prácticas para el registro de IDN de segundo nivel. El objetivo es minimizar el riesgo de ciberseguridad, de ciberocupación y confusión entre los consumidores. Para los gTLD, este es un acuerdo. Hay obligaciones vinculantes. Tenemos el acuerdo base de registros de gTLD, especificación 6, sección 1.4. Pautas para la implementación de IDN.

Nota: El contenido de este documento es producto resultante de la transcripción de un archivo de audio a un archivo de texto. Si bien la transcripción es fiel al audio en su mayor proporción, en algunos casos puede hallarse incompleta o inexacta por falta de fidelidad del audio, como también puede haber sido corregida gramaticalmente para mejorar la calidad y comprensión del texto. Esta transcripción es proporcionada como material adicional al archivo, pero no debe ser considerada como registro autoritativo.

También tenemos el acuerdo de acreditación de registradores, especificación de operación de registradores adicionales. Los registradores deberán cumplir con las pautas de IDN. También deben cumplir los administradores, tal como figura indicado en el proceso.

Si bien se están actualizando las pautas en este momento, básicamente la versión anterior era la versión 3 publicada en 2011. Se han realizado análisis adicionales. Tenemos datos adicionales, experiencia operativa con los IDN. Por ejemplo, la terminología para IDN, la representación de la tabla que tendremos a disposición pronto en RFC 7940.

También tenemos el diseño de las tablas de IDN, que también se publicó en estos RFC 6912 y también 8228. Además, la organización de la ICANN desarrolló los LGR de referencia para el segundo nivel. También tenemos el repertorio máximo inicial o MSR. También tenemos las reglas de generación de etiquetas para la zona raíz. También tenemos una propuesta del proyecto desarrollado por la comunidad que indica cómo utilizar el código de escritura.

Las comunidades crearon un grupo de trabajo para actualizar las pautas y llegamos a la versión 4. La versión 4 tiene siete temas y 19 pautas. Hay un grupo de transición. Tenemos los temas aquí. El formato de tablas de IDN, uniformidad de tablas y prácticas de IDN, etiquetas variantes de IDN, similitud y confusión de etiquetas así como publicación de política y registros y finalmente la terminología.

Estas pautas han sido desarrolladas por el grupo de trabajo. Luego atravesaron un periodo de comentarios públicos y la versión C4.0 se publicó el 10 de mayo de 2018. En 2019 la GNSO le solicitó a la junta

directiva de la ICANN que le permitiera estudiar las pautas antes de la implementación. La junta directiva consideró esta solicitud y en mayo de 2021 la GNSO inició un proceso acelerado de desarrollo de políticas o EPDP sobre los IDN para abordar las cuestiones adicionales relacionadas con los IDN y que no están cubiertas en el informe SubPro. Esto incluye la forma de administrar las variantes y de actualizar las pautas para la implementación de IDN en el futuro.

Posteriormente, en agosto, la GNSO también le solicitó a la junta que la adopción de las pautas se postergara hasta que hubiera finalizado el EPDP de IDN debido a que había superposiciones. La junta respondió en octubre solicitando a la GNSO que analizara e identificara las pautas que se superponen con los temas y en enero la GNSO respondió apoyando la sugerencia de la junta e identificó las siguientes pautas que presentan superposiciones con los temas incluidos en la carta orgánica.

Actualmente la junta está considerando esta respuesta para ver cómo avanzamos. Este es el estado actual hasta el momento con respecto a las pautas para la implementación. Voy a hacer una pausa. Vamos a ver si hay algún comentario, aporte o pregunta antes de que avancemos.

SARMAD HUSSAIN: No hay preguntas en el chat. Gracias, Pitinan.

PITINAN KOOARMORNPATANA: Vamos a dejar espacio para preguntas y respuestas al final de forma tal que podemos volver a este tema si es necesarios. Pasemos

ahora al siguiente tema. Los códigos de escritura adicionales para IDN. Como información general, según la misión de la ICANN estamos trabajando en los nombres de dominio, específicamente en los IDN. Los nombres de dominio internacionalizados y su estándar, que es IDNA, está basado en el estándar Unicode. Con el IDNA 2008 tenemos un estándar que utiliza un mecanismo de algoritmos para los puntos de código y también pone limitaciones adicionales según el análisis.

Sin embargo, no lleva a cabo una revisión basada en el código de escritura para su inclusión. La ICANN ha estado trabajando con las comunidades de los códigos de escritura para desarrollar soluciones seguras y estables. Se ha utilizado el procedimiento RZ-LGR. Recién hemos finalizado la primera ronda. Tenemos 28 códigos de escritura que es el repertorio máximo inicial. Son los que figuran aquí. Árabe, armenio, bengalí, etc. Estos son los 28 códigos de escritura identificados para la zona raíz. Son los mismos 28 códigos que se utilizan para desarrollar el LGR de referencia para el segundo nivel también así como para los gTLD.

En el Unicode a partir de ahora la versión 14 tiene 159 códigos que incluyen también algunos de los códigos históricos que no se utilizan hoy en día. No todos los códigos de escritura quizá sean adecuados para los identificadores. Es necesario identificar cuáles son estos. La pregunta que tenemos que responder es cuáles son. De estos 159 códigos, cuáles serían los códigos que debe apoyar la organización de la ICANN. Lo vamos a considerar en el alto nivel y también en el segundo nivel.

Vamos a ver ahora en mayor detalle el Unicode. De estos 159 códigos de escritura, el estándar Unicode tiene el anexo 31, que es el identificador. Identifica estos códigos en tres categorías: los códigos excluidos, los de uso limitado y los recomendados. Con respecto a los excluidos tenemos 94 códigos. Básicamente son códigos que no se utilizan hoy en día, que quizá son históricos o tienen un uso muy limitado. Se utilizan en comunidades muy pequeñas. Además, alguno de estos códigos también tiene algunos problemas de arquitectura no resueltos. Por ejemplo, problemas de representación, etc. Algunos ejemplos son Brahmi, Coptic, jeroglíficos egipcios, etc. Estos son entonces códigos de escritura de uso excluidos.

Los códigos de escritura limitados son códigos de escritura modernos que tienen un uso limitado. Para evitar problemas de seguridad quizá algunas implementaciones no los permitan pero su uso sigue siendo posible. Por ejemplo, balinés, silábicas aborígenes canadienses, cherokee, etc.

Luego, la última categoría son los códigos de escritura recomendados. Tenemos 29 y también tenemos códigos de escritura comunes y heredados. En este caso tenemos códigos de uso amplio y generalizado en la actualidad. Aquí tenemos algunos ejemplos como árabe, armenio, etc. Estos son códigos de escritura con los que ya estamos trabajando.

Para responder la pregunta anterior, la ICANN contrató a especialistas para realizar una revisión del UAX 31 en el contexto de IDN. También podemos vincularlo con el informe de esta presentación pero a modo de resumen tenemos estas categorías. Los códigos de escritura

recomendados, que son adecuados para la zona raíz y también para el segundo nivel. Si tenemos un código de escritura de uso limitado, quizá no sea adecuado para la zona raíz pero sí puede ser utilizado en el segundo nivel. Debemos analizarlo caso por caso. En cuanto a los códigos excluidos, no son adecuados para la zona raíz ni tampoco para el segundo nivel.

Continuamos. Aquí tenemos la evaluación de los códigos de escritura. ¿Por qué debemos considerar los códigos de escritura de uso limitado caso por caso? Esto se debe a que el código de escritura podría presentar potencialmente algunos problemas de seguridad o estabilidad. Por lo tanto, la comunidad de los códigos de escritura debe ayudar a desarrollar una solución segura. Debido a que es uso limitado, a veces quizá no podamos encontrar una comunidad con la cual consultar este tema para desarrollar reglas de generación de etiquetas. Quizá no contemos con información suficiente para desarrollar los datos de referencia. Sin información suficiente no podríamos desarrollar una referencia que nos sirva para poder utilizar este código de escritura de forma eficiente.

Estos son algunos ejemplos de riesgos que podrían surgir con códigos de escritura de uso limitado. Aquí tenemos los códigos de escritura limitados como cherokee, kayah, lisu, etc. Aquí tenemos un ejemplo. Tomamos algunas etiquetas. Algunas quizá parezca como mayúsculas en caracteres latinos pero tienen puntos de código diferentes. Si ven aquí los puntos de código verán que son distintos de la letra mayúscula latina A. Esto por lo tanto genera confusiones. Además, algunos quizá se parezcan a las letras minúsculas latinas como CEO, o

como NEW. Básicamente se parecen pero por detrás tenemos puntos de código diferentes.

Otro ejemplo similar es el que tenemos aquí con respecto a la confusión. Tenemos un código de uso limitado que podría parecerse al código de escritura recomendado. Por ejemplo, aquí New Tai Lue puede parecerse al de Myanmar o Malayalam. Podemos crear etiquetas para New Tai Lue y otra etiqueta para Myanmar que se parecen bastante pero son códigos puntos de código diferentes. Lo mismo ocurre aquí con Bangla. Este es el código de escritura recomendado ya utilizado en la zona raíz. Tenemos Syloti Nagri que podría darnos un resultado similar.

Otro ejemplo de riesgos son los problemas de representación debido al uso limitado. Por su naturaleza algunos de estos sistemas quizá no tengan el mecanismo de representación adecuado como para gestionar este código de escritura. Por ejemplo, aquí, en el código de Newa tenemos un código de uso limitado. Si copiamos este carácter de aquí, de Unicode, tenemos el glifo 11420. En este caso en pantalla vemos que aparece algo diferente. Por lo tanto, no se puede reproducir y no genera un signo de interrogación. Este es un código de escritura de uso limitado que no tiene un soporte adecuado para su representación o reproducción y podría crear problemas. Esta lista incluye algunos códigos de escritura de uso limitado con los que hicimos pruebas y en los que detectamos e identificamos problemas.

Así llegamos a la conclusión. Tenemos un resumen de alto nivel. Los códigos excluidos quizá no sean adecuados para alto nivel o el segundo nivel. Los códigos de uso limitado quizá sean adecuados para

el segundo nivel y hay que analizarlo caso por caso. Todo esto estará incluido en el informe. Publicamos ya el informe para comentarios públicos. Este periodo se abrió el 5 de enero de este año. Finalizó la semana pasada, el 15 de febrero. El informe se publicará a principios de marzo. Estamos recibiendo los aportes, analizando los aportes que recibimos y luego continuaremos con los siguientes pasos. Voy a hacer una pausa ahora para ver si hay algún comentario. Creo que vi algo en el chat.

SARMAD HUSSAIN:

Hay algunas preguntas generales. Las voy a leer. Hay una pregunta de Nabeel Yasin. “¿Ustedes creen que los IDN llevarán al aislamiento o a una fragmentación de Internet?” Luego hay otra pregunta de Prince Andrew Zutah: “¿Cómo vamos a resolver la barrera del idioma cuando recibimos una pregunta de un dominio de IDN en un idioma diferente?” Hay algunas respuestas a esto en el chat ya. También hay una pregunta de Raymond Mamattah. “¿Cuáles son las diferencias principales entre la aceptación universal y los IDN?” Estas preguntas están siendo respondidas en el chat también pero, Pitinan, si quiere responder, le pedimos que tome la palabra. Vamos a tratar de responderlas en vivo.

PITINAN KOOARMORNPATANA:

Gracias. Voy a elegir una. Si no, podemos continuar. Vamos a hablar de la diferencia entre IDN y aceptación universal. Los IDN tienen más que ver con que nosotros preparamos la tecnología y quizá no tengamos desarrolladas las soluciones que se necesitan en una comunidad con un código de escritura específico. ICANN ahora facilita

el trabajo comunitario con un esfuerzo comunitario y esta solución ahora está disponible.

El próximo paso es que si se puede utilizar este código de escritura técnicamente podría tratarse de un proveedor que esté en el ecosistema de Internet y que no lo sepa. Por ejemplo, Si tenemos emails en los nombres de dominio internacionalizados que llamamos EAI, es decir, dirección de email internacionalizada, si queremos suscribirnos a algo en un sitio web, ese sitio web puede chequear y verificar que el email no es correcto. El próximo paso más allá de los IDN. Yo diría que cuando lo técnico esté listo podemos hacer que sea aceptado en las distintas plataformas y en los distintos sistemas.

SARMAD HUSSAIN:

Gracias, Pitinan. Se ha compartido también un link a una sesión sobre aceptación universal que va a ocurrir mañana. Hay también una pregunta de Dennis. “¿Los problemas de rendering no estarían más bien vinculados a la aceptación universal?”

PITINAN KOOARMORNPATANA:

Sí. Es una buena pregunta. Creo que tendríamos que encontrar la respuesta a este rendering en conjunto. Para lograr que el sistema acepte los nombres de dominio ASCII no tradicionales, eso sería algo que sería parte de la aceptación universal. Muy bien. Vamos a continuar. Si hay tiempo, al final podemos volver a alguna de estas preguntas. Luego entonces veremos una actualización sobre las generaciones de etiquetas en la zona raíz. Le voy a dar entonces la palabra al próximo orador, que es el copresidente del panel de

generación de etiquetas, quien nos contará sobre este tema. Luego le daremos la palabra a Marc para la versión 5 de las reglas de generación de etiquetas.

Vamos a ver una breve historia de este proyecto. Este proyecto comenzó a partir de los requisitos de la comunidad, cuando hubo una necesidad de variantes en los nombres de dominio internacionalizados. Esto ocurrió en el año 2009-2010. En ese momento, la comunidad se reunió y trabajó en un informe de temas integrados que identificó qué es lo que se debe hacer para permitir los TLD con variante.

Lo primero que se identificó es que no había una definición de qué es una variante y básicamente esto depende de los distintos códigos de escritura. En segundo lugar, no hay un mecanismo sobre cómo gestionar estas variantes para luego entonces establecer una definición de variantes, la comunidad desarrolló unos procesos de LGR donde se establece cómo se desarrollan las reglas de generación de etiquetas.

Estas son las etiquetas para la zona raíz. Dentro de estas reglas se identifica la definición de variante. Este procedimiento ha sido adoptado por la junta directiva en 2013 y desde ese momento la comunidad para este código de escritura ha estado trabajando en su propio código y desarrolló las LGR para la zona raíz. La organización de la ICANN también comenzó a desarrollar las recomendaciones sobre cómo gestionar los TLD con variante.

Luego, en el año 2019, la junta directiva de la ICANN resolvió que la GNSO y la ccNSO deben tomar en cuenta las recomendaciones para la

gestión de los TLD con variante de IDN. En estas recomendaciones también se integraron las reglas de generación de etiquetas para la zona raíz y ahora esto está siendo tenido en cuenta en el desarrollo de políticas para la ccNSO y la GNSO.

También en el año 2020, con la disponibilidad de las distintas propuestas de la comunidad sobre las reglas de generación de etiquetas para las distintas cadenas de caracteres, se resolvió que la GNSO y la ccNSO tomen en cuenta las recomendaciones de la utilización técnica de la generación de etiquetas para la zona raíz. En el 2021, la GNSO también publicó el informe sobre la ronda de procedimientos posteriores para los nuevos gTLD que también incorpora el uso de las reglas de generación de etiquetas para la zona raíz.

Este es un panorama del trabajo que hemos hecho, que ha hecho la comunidad en los 28 códigos de escritura que identificamos para esta primera ronda del trabajo. Ahora tenemos entonces 26 códigos de escritura y podemos ver parte del trabajo en los últimos años hubo una gran diversidad. A principios de este año estamos llegando al final del panel de generación de etiquetas. Todavía tenemos que trabajar con el código de escritura tibetano para incluirlo en las reglas de generación de etiquetas a posteriori. Con esto entonces le voy a dar la palabra a Yin May Oo, que es copresidente del panel de generación de etiquetas, para que nos dé una actualización sobre las etiquetas.

YIN MAY OO:

Hola a todos. Espero que estén bien. Ahora voy a apagar mi vídeo y voy a explicar solamente las diapositivas. Me complace poder reunirme

con todos ustedes. Voy a comenzar con mi presentación. Hola a todos. Estos son los idiomas en los que nos focalizamos en base a las distintas variantes y los distintos usos de alrededor de 500 millones de personas que utilizan estos idiomas. Estamos hablando aquí de las distintas comunidades que usan los distintos códigos de escritura. Estos códigos de escritura se utilizaron en distintos países y hemos tomado mucho vocabulario de los distintos idiomas. Aquí tenemos un conjunto de toda la costa este del sudeste asiático.

Con los años hemos adoptado distintas culturas y distintos idiomas, y los hemos mezclado. Tenemos distintos usuarios de los distintos idiomas y hay otros idiomas históricos que quizá no hayan sido descubiertos hasta ahora. De ahí es de donde provinimos. Como todos ustedes saben, este código de escritura es un sistema que va desde la India hasta el sudeste asiático. También hemos adoptado otros códigos de escritura que hemos incluido como [inaudible]. Ahí hay muchas consonantes y muchos diacríticos. Estos diacríticos van desde seis o siete a una sola sílaba. Siguiente diapositiva, por favor.

Para una sílaba podría ser una consonante. Podemos llegar a otras consonantes con mediales y podemos agregar otra consonante luego para utilizar un sonido diferente y para hacerlo en un tono más alto o pronunciarlo de forma diferente. Hay muchas posibilidades y tenemos que asegurarnos de que el Unicode sea el adecuado y que no estemos con una representación equivocada. Tenemos que asegurarnos de que la etiqueta sea segura. Esta etiqueta, como pueden ver, el significado es capacidad humana. Tenemos una consonante. Luego cambiamos. Ponemos otro signo y el significado puede ampliarse.

Como pueden ver, en la región tenemos más de 10 códigos de escritura en uso. Sin embargo, nos centramos en seis idiomas y excluimos algunos. Los signos de exclamación, los caracteres obsoletos, para poder cubrir seis idiomas, seis lenguajes. Tuvimos que dejar de lado algunos lenguajes que tienen pocos usuarios y para los que no tenemos valores de referencia. Pasemos a la siguiente diapositiva.

Analizamos los puntos de código para las etiquetas y luego estudiamos otros idiomas para ver cuáles eran las variantes posibles y nos encontramos con latino, cirílico, armenio. Tenemos el código de Myanmar. En este caso, se decidió utilizar este carácter circular como glifo. A partir de esta letra descubrimos que muchos idiomas en el mundo también utilizan este círculo. Aquí tenemos la lista que creamos y como tenemos esta forma de círculo también tenemos aquí otros círculos que tienen la parte superior abierta, la parte inferior abierta o la del costado. Estas son las otras formas que identificamos y que podrían ser variantes de otros códigos de escritura. Muchas gracias. La siguiente diapositiva, por favor.

Tenemos variantes basados en distintos caracteres. Como las variantes están cerca de los idiomas occidentales y como nuestro idioma tiene muchos símbolos diacríticos, hay distintos tipos de símbolos diacríticos, acentos vocales, mediales, tuvimos que especificarlo y dividirlo por grupos y cada grupo tiene un orden determinado con la grafía adecuada. Tuvimos que clasificarlos entonces y crear las reglas de acuerdo con la letra de cada código. La siguiente, por favor.

Nosotros entonces tuvimos que crear las reglas iniciales para cada diacrítico. El diacrítico no puede venir después del otro diacrítico porque esto generaría confusión y problemas de representación o reproducción. Teníamos que asegurarnos entonces de utilizarlo correctamente. Algunos usuarios, la mayoría de los usuarios, podrían ser personas que tienen recursos limitados en nuestros países, en algunas regiones. Tenemos gente que no conoce todos los idiomas pero quizá quieran utilizarlos. Nuestras etiquetas no deberían generar confusión y no deberían impedir la lectura. Aquí tenemos las etiquetas para cada regla. Tenemos que ver toda la lista. Tenemos que asegurarnos de que no se forme nada que genere confusión. La siguiente diapositiva, por favor.

A lo largo de los años, empezamos en 2018, formamos el grupo, finalizamos nuestra propuesta. Este es el periodo de comentarios públicos que se abrió en enero de 2022 y todavía esperamos recibir más comentarios con respecto a las distintas situaciones. Esperamos poder llegar a terminar con el código de Myanmar. Esperamos que los usuarios de Birmania entiendan que deben ayudarnos ya que Internet no es solo para los angloparlantes. Si tienen aportes para hacer, esperamos recibirlos para poder ver las diferentes situaciones. Esperamos que la comunidad sepa que Internet está lista para derribar las barreras idiomáticas. Gracias.

PITINAN KOOARMORNPATANA: Muchas gracias, Yin May Oo. ¿Hay alguna pregunta o algún comentario para Yin May Oo?

Internacionalizados (IDN)

SARMAD HUSSAIN: Gracias, Pitinan. Hay una pregunta en el chat que no está directamente relacionada con Yin May Oo sino con una parte anterior de la presentación. Es una pregunta que hizo Andrei. “¿Por qué se utilizan alfabetos históricos no utilizados?” Si es que entendí correctamente. Por qué se hace lo necesario para incluirlos.

PITINAN KOOARMORNPATANA: Gracias, Andrei, por la pregunta. Nosotros nos planteamos una pregunta similar y la planteamos en el periodo de comentarios públicos que acaba de finalizar. Gracias por su pregunta. Si tiene una respuesta, sería bueno que nos la transmita. Quizá podríamos pegar aquí el vínculo para comentarios públicos, se lo agradecería, para recibir información con respecto a cómo manejar las diferentes categorías de códigos de escritura incluidas en el informe. ¿Hay alguna otra pregunta?

SARMAD HUSSAIN: Hay un par de preguntas más que se están respondiendo en el chat también. Hay una pregunta que hizo Mouloud Khelif. “¿Cuál es la relación que ustedes tienen con aceptación universal y específicamente el grupo directivo de aceptación universal?” Hay otra pregunta de la misma persona. “Hasta el momento, ¿cuántos TLD en IDN se han delegado en la zona raíz? TLD con variantes de IDN”.

PITINAN KOOARMORNPATANA: El grupo directivo de aceptación universal es el grupo liderado por la comunidad. Trabajamos con el programa de IDN UA de la ICANN y UASG. El grupo de UASG también tiene múltiples grupos de trabajo

por debajo de nosotros y nosotros tenemos una sesión específica sobre aceptación universal mañana y también la próxima semana podrían asistir a esa sesión si les interesa el tema pero básicamente el programa de IDN y de UA trabajan también con UASG que es el grupo liderado por la comunidad. La otra pregunta era cuántos TLD con variantes de IDN se han delegado. Aproximadamente 150. No sé el número exacto para ccTLD y gTLD. Después les voy a decir el número exacto y lo voy a publicar en el chat. Muchas gracias. 154. Muchas gracias. Continuemos. Le voy a dar la palabra a Marc Blanchet, que es miembro del panel de integración y va a compartir con nosotros una actualización sobre la RZ-LGR versión 5.

MARC BLANCHET:

Muchas gracias, Pitinan. Esta presentación se refiere a nuestro trabajo en el panel de integración y todo lo que hicimos desde la última reunión de la ICANN. El objetivo es hacer aquello de lo que se trata el panel, que es la integración. Por lo tanto, tomamos todos los LGR del panel de integración y los integramos para formar el LGR para la zona raíz que es utilizado por la comunidad. La próxima diapositiva, por favor.

Actualmente estamos trabajando en la quinta versión del LGR para la zona raíz. La cuarta versión finalizó en otoño de 2020. RZ-LGR cuatro contenía 18 códigos de escritura como árabe, bangla, han, devanagari, hebreo, gurmukhi, gujarati, georgiano, khmer, lao, malayalam, oriya, sinhala, tamil, telugu y thai. El objetivo es integrar seis códigos de escritura nuevos que son griego, y hiragana y katakana de japonés, que ya estaban en la versión cuatro, coreano, que incluye algunos

códigos específicos, hangul y también han, que ya está en la versión cuatro, latín y Myanmar.

Lo interesante es que agregamos dos códigos de escritura que ya tenemos desde hace muchos años: armenio y cirílico, pero decidimos postergarlos, aplazarlos. Es decir, en ese momento no los integramos a la zona raíz porque estaban muy relacionados con latín y griego. Estábamos esperando hasta obtener los cuatro códigos para tener la versión final de forma tal que los pudiéramos integrar porque compartimos variantes y a través del proceso de integración vamos imponiendo las distintas variantes y por eso debíamos esperar hasta tenerlos a todos.

También tuvimos que esperar porque queríamos asegurarnos de que hubiera coherencia entre los diferentes códigos de escritura vinculados. Por lo tanto, LGR 5 ahora tiene dos conjuntos de código relacionado, que están totalmente integrados. Japonés, chino, coreano y latín, griego, cirílico y armenio. De cierta forma, LGR 5 será la versión más compleja que hayamos hecho. Será la integración más compleja y por lo tanto estamos dedicándole mucho tiempo a la verificación y a asegurarnos de que estemos haciendo la integración correctamente. La siguiente diapositiva, por favor. Creo que se saltaron una. Gracias. Muy bien. Muy bien. Por favor, ¿podemos retroceder una?

LGR 5, la versión 5, que es la versión en la que estamos trabajando actualmente, incluirá básicamente 25 LGR elementales que significa que son específicos para un código de escritura. ¿Por qué tenemos 25? Porque tenemos 26 códigos de escritura pero Japón tiene dos códigos

de escritura en un LGR y por eso son 26 menos 1. También vamos a incluir un LGR común, que es un LGR generado mecánicamente.

Apuntamos al 24 de marzo para lanzar el LGR 5 y presentarlo para comentarios públicos. Por lo tanto, esperamos tener la versión de Myanmar. Esperamos recibirla pronto para poder integrarla. Quisiéramos que la gente dedique el tiempo a hacer una revisión de LGR 5 porque, una vez más, será la más compleja y también la que con más probabilidad se utilizará en el futuro. Esto finalizará básicamente en la integración de todo el trabajo del panel de generación con las LGR que recibimos. La siguiente diapositiva.

Con respecto a las variantes con código de escritura cruzado, esto puede ser muy significativo en lo que se refiere a la publicación. Estas variantes van a estar completas en la próxima LGR. Allí vamos a tomar todos los códigos de escritura cruzados entre las distintas LGR y los vamos a combinar con transitividad y así vamos a crear el LGR común. Este LGR común es el que se va a utilizar para los testeos de colisión y para las etiquetas válidas.

Lo que vemos más abajo es lo que hemos decidido como la forma... No lo que decidimos sino lo que es realmente en lo que se refiere al procesamiento de variantes. Hay una variante que está definida en el LGR de un elemento y que va a ser integrado. Todo código de escritura en el repertorio siempre va a estar indicado en la LGR afectada. Las variantes ASCII siempre están indicadas también y los LGR de un elemento en los códigos de escritura armenio y latino van a incluir también todas las variantes pero solamente las LGR latinas van a incluir variantes por fuera de ese set. Las variantes armenia y cirílica

que han sido impuestas también se van a agregar y va a haber otros elementos de LGR como el malayalam, que han sido seleccionados como variantes que se imponen en los nuevos códigos de escritura en las reglas.

Aquí hay un llamado a que los paneles de generación que están involucrados en todo esto, por ejemplo el cirílico y el armenio, el malayo y otros, a ellos les va a resultar muy útil si se puede verificar que el trabajo de integración que hicimos, que de hecho agrega nuevas variantes en el LGR con un elemento, ese elemento tenemos que asegurarnos de que no se haya agregado ningún otro problema en el proceso de integración. Este es un llamado a que toda la comunidad verifique nuestro trabajo y que todo esté en forma para los próximos pasos para estas reglas de generación de etiquetas de la zona raíz.

Esto fue presentado la última vez. No voy a entrar mucho en detalle pero quiero solamente decir que típicamente, cuando queremos procesar o verificar una etiqueta dentro del LGR de zona raíz, primero hay que verificar que la etiqueta sea válida utilizando el LGR de un elemento. Luego utilizaremos el LGR común para verificar las colisiones y, si está en buen estado, vamos a utilizar el elemento LGR de un código de escritura para computar la variante y luego podremos tener ya esas variantes para las distintas etiquetas que son las variantes de la etiqueta original. Para esto utilizaremos el LGR de zona raíz. Tenemos una herramienta que ha sido puesta a disposición por ICANN, que permite cuidar todo este proceso a través de esa herramienta.

Vamos a ver ahora los próximos pasos en el panel de integración. Nuevamente tenemos la intención de publicar el RZ-LGR versión 5 el 24 de marzo. Esto es dentro de aproximadamente un mes. Lo que quisiéramos realmente es poder ver que la comunidad, y especialmente los paneles de generación, puedan revisar y verificar el trabajo que vamos a estar haciendo. Se trata de un comentario muy útil para nosotros, si es que ustedes quizá no encuentren problemas pero sí puedan enviar un mensaje al portal de comentarios donde digan: “Hemos verificado nuestro código de escritura y vemos que todo está bien”. Esto va a resultar muy valioso para todos nosotros. Por eso, si se toman el tiempo para verificar, cuéntenos.

Tentativamente, nuestro objetivo es tener finalizado el RZ-LGR versión 5 para el segundo trimestre del año calendario 2022. Habrá una posible próxima versión de la LGR para la zona raíz una vez que se forme el panel y se presente el LGR a nosotros para que haya una integración final. Como ha dicho Pitinan, estos paneles todavía no se han formado. Por lo tanto, no tenemos por el momento ninguna fecha para esta nueva versión que muy probablemente va a ser la LGR 6. No tenemos entonces ninguna fecha para ese trabajo. Creo, ahora sí, que esta es mi última diapositiva. Es el fin de mi presentación. Gracias.

PITINAN KOOARMORNPATANA: Gracias, Marc. Vamos a ver ahora si hay alguna pregunta.

SARMAD HUSSAIN: Sí. Tenemos algunas preguntas. Hay una de Bill Jouris. “¿Por qué no listar las variantes cruzadas entre el cirílico y el armenio, incluso si el

GP latino no las toma como variantes?” Quizá podemos responder esta pregunta y luego ocuparnos de otras más.

MARC BLANCHET: Creo que esa pregunta estaba en el chat. Déjenme la verifique, a ver si la puedo entender. Muy bien. La política que hemos realizado indica que todas las variantes cruzadas en la LGR van a estar allí presentes, no solo para que tengamos todas las LGR con respecto a las variantes que no están vinculadas al código de escritura en sí sino que hemos decidido restringir la posición en algunos LGR de elemento para la legibilidad de los códigos de escritura pero siempre vamos a usar el LGR común para tener el listado total de las variantes con el código de escritura cruzada.

SARMAD HUSSAIN: Gracias, Marc. Tenemos otra pregunta más que es de Vadim Mikhaylov. “Si yo entiendo correctamente, el latino en LGR 5 no es solo latín básico sino que se ha extendido al latín también. En ese caso, ¿qué extensiones, A, B, C de los estándares de Unicode han sido incluidas?”

MARC BLANCHET: Esa pregunta está más vinculada a las reglas de generación de etiquetas latinas. Creo que se podría responder allí. Cada panel de generación lo que hace es básicamente identificar todos los idiomas con cierto nivel de uso, con un estándar específico conocido. Son ellos quienes definen el alcance de todo ese trabajo en términos de todos los idiomas que van a tener soporte en esa LGR y en ese código de escritura. Así es como se establecen los puntos de código que son los

que se revisan, lo que se verifican. Al mismo tiempo tienen que mirar los MSR con el repertorio de inicio, que suele ser el set más grande, donde todas las cuestiones históricas que ya no se utilizan son removidas. De ese modo se crea un set inicial de los puntos de código que son los que se verifican. Todos los caracteres latinos son los que tienen soporte en más de 200 idiomas.

SARMAD HUSSAIN: Gracias. Hay otra pregunta de Satish Babu. “¿La versión 5 va a ser compatible con la versión 4?”

MARC BLANCHET: Sí. La versión final todavía no está lista pero la intención es que sea completamente compatible y no hay ninguna razón por el momento para pensar que no lo va a ser. Si eso sucede, va a estar escrito en los comentarios públicos para que la gente lo pueda ver.

SARMAD HUSSAIN: Gracias. Hay otra pregunta que también ha sido respondida en el chat. La voy a leer: “¿Todos los sistemas de DNS como el BIND pueden traducir los dominios de IDN a las IP?” Hubo un comentario en el chat que dice que en ese sistema de DNS todas las etiquetas están en el formato A y por lo tanto pueden vincularse con los caracteres ASCII para procesamiento de resolución interna. Marc, si quiere agregar algo.

MARC BLANCHET: La pregunta tiene más que ver con el IDN-A en general. Hay dos formas de nombres de dominio en los nombres de dominio internacionalizados. Una es la que está dentro del código de escritura, el verdadero valor que el usuario va a ver. Esa cadena de caracteres es la que se le presenta al usuario típicamente en un navegador pero esa cadena de caracteres no es la que se codifica en el sistema de DNS. El sistema de DNS sigue siendo ASCII y lo diseñamos precisamente con el objetivo de no tener que cambiar todos los sistemas de DNS en el mundo.

Esta cadena de caracteres que el usuario escribe en su propio código de escritura se va a traducir, como Michael lo estaba escribiendo, en lo que nosotros denominamos una etiqueta A con un algoritmo que se llama Punycode, que comienza con xn-- y el resto es una lista de caracteres ASCII y, por lo tanto, el usuario no lo va a ver pero eso es lo que es el DNS o, mejor dicho, es lo que DNS utiliza para la resolución de direcciones IP, por ejemplo.

SARMAD HUSSAIN: Gracias. Hay una pregunta de Raymond Mamattah. “¿Hay algún trabajo que se esté haciendo con respecto a los caracteres Ewe? Este es una de los mayores idiomas que se hablan en África oriental. Especialmente en Togo y en Ghana”.

MARC BLANCHET: No tengo idea sobre este. No sé si Pitinan o Sarmad lo saben pero yo no lo sé. Aquí Michael dice que ha sido considerado en nuestro trabajo para los caracteres latinos.

PITINAN KOOARMORNPATANA: Sí. Es parte del código de escritura latino. Tenemos todavía un minuto más para ir entonces ya cerrando. Vamos a darles un breve resumen con respecto a los lineamientos todavía tenemos que ver qué es lo que van a decir los expertos. El comentario público ya ha cerrado para los IDN de soporte. Si quieren presentar otros comentarios pueden hacerlo pero en este momento estamos generando los comentarios finales. Para los comentarios públicos del código de escritura de Myanmar va a estar abierto hasta el 3 de marzo y si bien parece una fecha un poco remota tenemos que incluir algunos de los comentarios. Quizá les pueda resultar interesante verlos.

Finalmente, para las reglas de generación de etiquetas de versión 5, nosotros esperamos poder publicarlo alrededor del final de marzo. Por eso quisiéramos recibir los aportes para poder revisarlo. Creo que esto es todo lo que tenemos que decirles hoy. Siéntanse en libertad de contactarse con nosotros en idnprogram@icann.org, si quieren enviar preguntas o comentarios. Les agradecemos a todos por participar de esta sesión. Esperamos que tengan una buena semana y una exitosa reunión de la ICANN. Ya podemos detener la grabación. Gracias.

[FIN DE LA TRANSCRIPCIÓN]