

ICANN73 | Semaine de préparation – Actualités du programme pour les noms de domaine internationalisés (IDN)
Mardi 22 février 2022 – 9h00 à 10h00 AST

PITINAN KOOARMORNPATANA : Bonjour, merci de lancer l'enregistrement.

Bienvenue à la première séance de la semaine de préparation de l'ICANN73. Je m'appelle Pitinan de l'équipe du programme des IDN et de l'acceptation universelle. Je veux vous souhaiter la bienvenue à cette séance sur les actualités du programme des IDN, une séance de 60 minutes qui abordera les questions suivantes.

En premier lieu, il y aura une mise à jour des directives de mise en œuvre puis les scripts supplémentaires à prendre en charge dans les IDN. Par la suite, il y aura une mise à jour de l'état des règles de génération d'étiquettes pour la zone racine, le contexte, la mise à jour du panel de génération du script du Myanmar et aussi un membre du panel qui nous donnera une mise à jour sur la zone racine LGR version 5.

Nous passons donc à la première diapositive. Pour ce qui est de la mise en œuvre des IDN, voici l'information générale des directives sur les politiques et les pratiques pour les IDN et sur minimiser les risques de cybersécurité pour les consommateurs.

Pour les gTLD, c'est un accord. Il y a des obligations contraignantes. Il y a la section 1.4 de la spécification 6, les directives pour les IDN. Il y a

Remarque : Le présent document est le résultat de la transcription d'un fichier audio à un fichier de texte. Dans son ensemble, la transcription est fidèle au fichier audio. Toutefois, dans certains cas il est possible qu'elle soit incomplète ou qu'il y ait des inexactitudes dues à la qualité du fichier audio, parfois inaudible ; il faut noter également que des corrections grammaticales y ont été incorporées pour améliorer la qualité du texte ainsi que pour faciliter sa compréhension. Cette transcription doit être considérée comme un supplément du fichier mais pas comme registre faisant autorité.

le contrat d'affectation des bureaux d'enregistrement, la clause 3 de la spécification supplémentaire qui mentionne que les bureaux d'enregistrement doivent aussi se conformer aux directives des IDN. Il y a aussi des IDN des domaines qui sont indiqués dans le processus.

Bien qu'on mette à jour les directives en ce moment, la version précédente est la version 3 qui a été publiée en 2011. On a fait des analyses supplémentaires, nous avons des données supplémentaires d'une expérience opérationnelle avec les IDN, par exemple la terminologie pour les IDN, la représentation des tables d'IDN qui seront disponibles prochainement. Nous avons également la conception des tables IDN publiées dans le RFC 6112 et 8228. L'organisation ICANN a également élaboré le LGR de référence et aussi, il y a le répertoire maximal initial ou MSR et les règles de génération d'étiquettes pour la zone racine. Il y a aussi une proposition d'un projet pour les scripts.

La communauté a créé un groupe de travail et nous arrivons à la version 4, qui a sept sujets et 19 directives avec des notes supplémentaires. Il y a un groupe de transition, le format des tables d'IDN, l'uniformité des tables des pratiques d'IDN, les étiquettes des variantes d'IDN, la similarité ou la confusion des étiquettes ainsi que la publication des politiques et des règles d'enregistrement des IDN.

Ces directives ont été élaborées par le groupe de travail qui a organisé un période de consultation publique et la version 4.0 a été publiée en mai 2018. Depuis lors, le 30 avril 2019, l'organisation de soutien aux extensions génériques a demandé au Conseil d'Administration de l'ICANN de lui permettre d'étudier les directives avant leur mise en

œuvre. Le Conseil d'Administration a considéré cette demande et en mai 2021, la GNSO a lancé un processus accéléré d'élaboration de politique ou EPDP sur les IDN pour traiter des questions supplémentaires liées aux IDN non abordées dans le rapport des SubPro, notamment comment définir et gérer en toute sécurité et de manière fiable les gTLD des variantes et comment mettre à jour les directives sur les IDN à l'avenir.

La GNSO a demandé en août que l'adoption des directives de mise en œuvre des IDN soit retardée jusqu'à ce que le PDP sur les IDN soit conclu en raison d'un certain chevauchement. En octobre, le Conseil a répondu et a demandé à la GNSO d'analyser et d'identifier les directives qui chevauchent les sujets inclus dans le PDP sur les IDN. Actuellement, le Conseil considère la réponse pour voir comment pouvoir avancer.

Voici donc l'état actuel pour ce qui est des directives de mise en œuvre. Je vais maintenant faire une petite pause. Nous allons voir s'il y a des questions, des commentaires avant de poursuivre notre travail.

SARMAD HUSSAIN : Pas de questions dans le chat, merci.

PITINAN KOOARMORNPATANA : On va laisser du temps pour les questions et réponses à la fin, alors on peut y revenir si cela s'avère nécessaire.

Nous passons maintenant au thème suivant : les codes d'écriture supplémentaires pour les IDN.

Comme information générale, selon la mission de l'ICANN, on travaille sur les noms de domaine, spécifiquement les IDN. Les IDN et leurs standards IDN sont basés sur l'Unicode. La norme IDNA 2008 utilise un mécanisme algorithmique pour les points de code, mais ils demandent des limitations supplémentaires, l'adresse pour créer des étiquettes de noms de domaine stables.

L'ICANN a travaillé avec cette unité, avec les communautés concernées pour développer des solutions sécurisées et stables. À l'aide de la procédure RZ LGR, jusqu'à présent, l'ICANN a pris en charge la première série et a travaillé avec 28 scripts, à savoir le répertoire maximal initial. Et les voilà : l'arabe, l'arménien, le bengali, etc. Les mêmes 28 scripts sont utilisés pour développer des LGR de référence pour le second niveau et sont également utilisés pour évaluer les tables IDN pour les gTLD.

À partir d'aujourd'hui, la version 14 d'Unicode dispose de 159 scripts, y compris des scripts contemporains et historiques qui n'ont pas été utilisés à ce jour. Et aussi, on peut dire que pas tous les scripts ne peuvent être adéquats pour les identificateurs. Il est nécessaire de savoir quels sont ces 159 codes, quels seraient les codes que l'ICANN devrait soutenir. On va le considérer au premier niveau, au second niveau et à d'autres.

Nous allons voir maintenant plus en détail ces 159 scripts dans les identificateurs. L'annexe 31, on a l'identificateur qui identifie ces

codes en trois catégories : les scripts exclus, les scripts à usage limité et les scripts recommandés.

Pour ce qui est des scripts exclus, il y en a 94 qui sont des scripts qui ne sont pas utilisés dans l'usage moderne habituel et peuvent donc vouloir exclure des identificateurs, inclure des scripts historiques et obsolètes, etc. Certains scripts présentent également des problèmes d'architecture non résolus, quelques exemples : le brahmi, le copte, les hiéroglyphes égyptiens, etc. Voici donc des scripts exclus.

Ce sont les codes d'écriture à utilisation limitée. Ce sont des scripts modernes dont l'utilisation est plus limitée. Pour éviter des problèmes de sécurité, certaines implémentations peuvent souhaiter interdire les scripts à usage limité dans les identificateurs, par exemple le balinais, les syllabiques autochtones canadiens, le cherokee, etc.

Par la suite, les scripts recommandés, il y en a 29. Dans ce cas, nous avons des codes à usage coutumier et répandu à l'heure actuelle et des scripts régionaux pour des grandes communautés. On a comme exemples l'arabe, l'arménien, etc. Ce sont des scripts sur lesquels on travaille.

Pour répondre à la question précédente, l'ICANN a engagé des experts Unicode et IDN pour examiner l'UAX31 dans le contexte des IDN. On peut mettre ceci en rapport avec le rapport de cette présentation. Mais voici donc les catégories : les scripts recommandés appropriés pour la zone racine et pour le second niveau, les scripts à usage limité qui ne sont pas appropriés pour la zone racine mais qui peuvent être utilisés au second niveau mais il faut analyser au cas par cas, puis les

scripts exclus qui ne sont pas appropriés pour la zone racine ou pour le second niveau.

On continue donc et ici, on a l'évaluation des scripts Unicode à utiliser dans les IDN. Pourquoi utiliser des scripts à usage limités au cas par cas ? Ceci parce que le script pourrait présenter des problèmes de sécurité et de stabilité importants. La communauté des scripts n'est pas disponible pour la consultation afin de développer des règles de génération d'étiquettes pour le script. Il y a peu d'informations disponibles. Parfois, on ne pourra pas trouver une communauté avec laquelle consulter pour les règles de génération d'étiquettes. Alors, peut-être qu'on n'aura pas d'informations suffisantes sur le script et sans l'information suffisante, nous ne pourrions pas élaborer ou développer une référence qui soit utile pour pouvoir utiliser ce script de manière efficace.

Voici des exemples de risques qui pourraient apparaître avec des scripts à utilisation limitée. En voici quelques-uns, comme le Cherokee. Il y a plusieurs exemples. On a ici les étiquettes qui peuvent apparaître comme majuscule en caractères latins mais qui ont des points de code différents. Ici, vous verrez que les points de code sont différents de la majuscule latine A. Ceci provoque donc des confusions. Il y en a qui peuvent ressembler aux minuscules latines. En voici des exemples : ceci se ressemble, mais les codes sont différents.

Un autre exemple similaire, c'est celui que nous avons ici pour ce qui est de la confusion. On a un code à usage limité qui pourrait ressembler au script recommandé. Ceci pourrait ressembler au script du Myanmar ou au malayalam. Alors on a des étiquettes pour le tai lue

et une autre pour le script du Myanmar qui se ressemblent mais qui ont des points de code différents. La même chose se passe ici avec le bangla. Voici le script recommandé de la zone racine. Et aussi, on pourrait avoir un résultat similaire.

Un autre exemple de risques avec des scripts d'utilisation limitée, ce sont les problèmes de rendu instable. Ces systèmes n'ont peut-être pas le mécanisme de représentation approprié pour administrer ou gérer ce script. Ici, vous avez un code à usage limité et si nous copions ce caractère unique, dans ce cas, on a le glyphe ou plus 11420 et vous voyez que là, on voit quelque chose de différent et qui ne s'affiche pas dans la barre d'adresse du navigateur. Ces codes d'écriture à usage limitée n'ont pas d'utilisation appropriée pour leur reproduction et pourraient créer des problèmes. Cette liste inclut des scripts à usage limité avec lesquels on a fait des tests et on a identifié des problèmes.

Nous arrivons donc à la conclusion. On a le résumé de haut niveau, les codes exclus ne sont peut-être pas appropriés pour le premier niveau. Il y en a d'autres qui sont appropriés pour le second niveau. Il faut analyser au cas par cas et ceci sera inclus dans le rapport publié pour consultation publique. La période a été ouverte le 5 janvier 2022 et s'est terminée le 15 février et le rapport sera publié le 8 mars. Nous analysons les contributions reçues et nous continuerons avec les démarches suivantes.

Je ferai une pause maintenant pour voir s'il y a des commentaires.

SARMAD HUSSAIN : Je crois que j'ai vu quelque chose sur le chat. Je vais donc lire. Il y a une question de Nabeel Yasin : « Croyez-vous que les IDN feront qu'il y aura un isolement ou une fragmentation de l'internet ? »

Et il y a une autre question du prince Andrew Zutah : « Comment allons-nous résoudre la barrière des langues lorsque nous recevrons une question d'un domaine IDN dans une langue différente ? » Il y a quelques réponses à cela sur le chat.

Et il y a aussi une question de Raymond Mamattah : « Quelles sont les différences principales entre l'acceptation universelle et les IDN ? »

Ces questions ont des réponses sur le chat aussi, mais si quelqu'un veut y répondre, je vous demande de prendre la parole et nous pouvons le faire en direct.

PITINAN KOOARMORNPATANA : Je vais en choisir une et comme cela, nous pourrons continuer. Nous allons parler de la différence entre les IDN et l'acceptation universelle.

Les IDN se rapportent plutôt au fait que nous préparons la technologie et nous n'avons peut-être pas encore développé les solutions nécessaires pour une communauté avec un script spécifique. L'organisation ICANN facilite le travail communautaire dans un effort communautaire et cette solution est maintenant disponible.

La prochaine étape est la suivante. Si l'on peut utiliser ce script au point de vue technique, il pourrait s'agir d'un fournisseur qui serait sur l'écosystème de l'internet et qui ne le saurait pas. Par exemple, si nous avons des emails dans les noms de domaine internationalisés que

nous appelons EIA, adresses email internationalisées, et si nous avons ce EIA et nous voulons nous abonner à quelque chose sur un site web, ce site web peut vérifier que l'adresse email n'est pas correcte.

La prochaine étape, au-delà de la question des IDN, je dirais que lorsque l'aspect technique sera résolu, nous pourrions accepter tout cela dans les différentes plateformes et dans les différents systèmes.

SARMAD HUSSAIN :

Merci Pitinan. On a partagé aussi un lien sur l'acceptation universelle. Il y aura une réunion demain.

Il y a une question de Dennis Tan : « Ceci ne se rapporterait pas plutôt à la question de l'acceptation universelle ? »

PITINAN KOOARMORNPATANA : Oui, c'est une bonne question. Je crois que nous devrions trouver une réponse pour ce rendu ensemble pour que le système accepte des noms de domaine ASCII non traditionnels. Cela ferait partie de l'acceptation universelle. Nous allons poursuivre et si nous avons le temps de le faire, nous pourrions revenir sur ces questions.

Nous allons donc voir une mise à jour sur le panel de génération d'étiquettes pour la zone racine. Je vais donc céder la parole au prochain orateur, le coprésident du panel de génération d'étiquettes, qui parlera sur ce thème. Et je céderai la parole à Marc pour la version 5 des règles de génération d'étiquettes.

Nous allons voir un bref aperçu, un historique de ce projet qui a commencé à partir du besoin des variantes de la communauté lorsqu'il y a eu un besoin de variantes de domaines de premier niveau internationalisés. Ceci a eu lieu en 2009-2010. À ce moment-là, la communauté s'est réunie et a travaillé sur un rapport intégré qui a identifié ce qu'il faut faire pour définir des variables de TLD comme une condition préalable.

La première chose que l'on a identifiée, c'est qu'on n'avait pas de définition de ce qu'est une variante et ceci dépend des différents scripts. Deuxièmement, il n'y a pas de mécanisme de définition de gestion des variantes. Pour établir une définition des variantes, la communauté a spécifié la procédure LGR pour développer les règles de génération d'étiquettes pour la zone racine. Voilà donc les étiquettes pour la zone racine et à l'intérieur de cette règle, on voit la définition des variantes. Cette procédure a été acceptée et adoptée en 2013 par le Conseil d'Administration et à partir de ce moment-là, la communauté pour ce script a travaillé sur son propre script et a développé les LRG pour la zone racine. L'organisation de l'ICANN a commencé aussi à développer les recommandations sur la gestion des TLD ayant des variantes.

En 2019, le Conseil d'Administration de l'ICANN a décidé que la GNSO et la ccNSO devraient tenir compte des recommandations pour la gestion des TLD de variantes IDN. Dans ces recommandations, on a intégré l'utilisation des règles de génération d'étiquettes pour la zone racine et cela est considéré dans le processus d'élaboration de politiques pour la ccNSO et pour la GNSO.

En 2020, par rapport à la disponibilité des différentes propositions de la communauté des règles de génération d'étiquettes pour les différentes chaînes de caractères, le Conseil d'Administration de l'ICANN a décidé que la GNSO et la ccNSO devraient tenir compte des recommandations pour l'utilisation technique de la génération d'étiquettes de la zone racine. En 2021, la GNSO a publié aussi le rapport sur la série de procédures pour les séries ultérieures de nouveaux gTLD qui a mis en avant ces nouveaux jeux d'étiquettes pour la zone racine.

Voilà donc un aperçu du travail du panel de génération pour les 28 scripts que nous avons identifiés pour cette première série, cette première étape du travail. Il y a maintenant donc 26 codes d'écriture et nous pouvons voir une partie du travail. Au cours des dernières années, il y a eu une grande diversité. Depuis le début de cette année, nous en sommes presque à la fin du panel de génération d'étiquettes. Nous devons encore travailler sur le script tibétain pour l'inclure dans les règles de génération d'étiquettes plus tard.

Je vais donc maintenant céder la parole à Yin May Oo, qui est coprésident du panel de génération d'étiquettes, pour qu'il nous présente sa mise à jour de ce panel de génération.

YIN MAY OO :

Bonjour à tous. J'espère que vous allez bien. Je vais maintenant éteindre ma caméra et je vais expliquer seulement les diapositives. C'est un plaisir de pouvoir être ici avec vous et je vais commencer ma présentation.

Bonjour à tous. Voilà les langues sur lesquelles nous avons concentré notre attention sur la base des différentes variantes et différents usages d'environ 500 millions de personnes qui se servent de ces langues-là. Nous parlons ici des différentes communautés qui se servent de différents scripts.

Ces scripts sont utilisés dans différents pays et nous avons pris un grand nombre de mots, un lexique assez abondant, des différentes langues. Et nous avons ici un ensemble de toute la Côte Est du Sud-Est asiatique.

Avec le temps, nous avons adopté différentes cultures et différentes langues. Nous avons différents utilisateurs des différentes langues. Il y a d'autres langues historiques qui n'ont pas été découvertes peut-être jusqu'ici.

Voilà donc d'où nous venons. Et comme vous le savez, ce script est un système qui va de l'Inde jusqu'à l'Asie du Sud-Est. Nous avons aussi adopté d'autres scripts que nous avons inclus comme par exemple le [inaudible]. Il y a beaucoup de consonnes et beaucoup d'éléments diacritiques et ces éléments diacritiques concernent entre six ou sept ou même une seule syllabe. Prochaine diapositive s'il vous plaît.

Pour une syllabe, on pourrait avoir une consonne et nous pouvons parvenir à d'autres consonnes comme les médiales. Nous pouvons ajouter une autre consonne ensuite pour utiliser un son différent et pour le faire sur un ton plus haut ou pour le prononcer différemment. Il y a de nombreuses possibilités et nous devons nous assurer que l'Unicode sera l'Unicode approprié et que nous ne serons pas sur une

représentation erronée. Nous devons nous assurer que l'étiquette est correcte. Cette étiquette, comme vous pouvez le voir, a pour un sens capacité humaine. Nous avons une consonne ensuite. Cela change et on ajoute un autre signe et le sens peut être élargi.

Comme vous pouvez le voir, dans la région, nous avons plus de 10 scripts utilisés, mais nous nous sommes centrés sur six langues en fait et nous avons exclu quelques langues, les points d'exclamation, les numéros, les caractères obsolètes, pour pouvoir couvrir six langues. Nous avons dû mettre de côté quelques langues ayant peu d'utilisateurs pour lesquelles nous n'avons pas de valeur de référence. Nous allons maintenant à la prochaine diapositive.

Nous avons analysé les points de code pour les étiquettes et ensuite, nous avons étudié d'autres langues pour voir quelles étaient les variantes possibles et nous avons trouvé le latin, le cyrillique, l'arménien. Il y a le code du Myanmar et dans ce cas particulier, on a décidé d'utiliser ce caractère circulaire que vous voyez comme glyphe et à partir de cela, nous avons découvert que de nombreuses langues du monde utilisent ce cercle.

Ici, nous avons la liste que nous avons créée et nous avons ce petit cercle. Et aussi, nous avons ici d'autres cercles avec la partie inférieure ou supérieure ouverte ou la partie du côté gauche ou droit. Voici d'autres manières d'identifier des variantes, d'autres codes d'écriture ou d'autres scripts. La diapositive suivante s'il vous plaît.

Nous avons des variantes basées sur différents caractères. Comme les variantes sont proches des langues occidentales et comme notre

langue a beaucoup de symboles diacritiques de différents types, de voyelles, de médiales, nous avons dû spécifier ceci et le diviser par groupe, chacun de ces groupes ayant un ordre déterminé. Avec la graphie appropriée, il a fallu les classer et créer les règles suivant les lettres de chacun des codes. Diapositive suivante s'il vous plaît.

Nous avons dû créer les règles initiales pour chaque élément diacritique qui ne peut pas venir après un autre élément diacritique parce que ceci provoquerait des confusions et aussi des problèmes de reproduction. Il fallait donc nous assurer de l'utiliser correctement. La plupart des utilisateurs pourraient être des personnes à ressources limitées. Dans certains pays ou dans certaines régions. Il peut y avoir des gens qui ne connaissent pas toutes les langues mais qui souhaitent peut-être les utiliser. Ces étiquettes ne devraient pas créer des confusions ni empêcher la lecture. Voici donc les étiquettes pour chacune des règles. Il faut voir toute la liste et nous assurer d'éviter la confusion. Diapositive suivante s'il vous plaît.

Au fil des ans, on a commencé en 2018, on a formé le groupe, on a fini notre proposition, on l'a terminée. Voici la période de consultation publique ouverte en janvier 2022 et nous nous attendons à recevoir davantage de commentaires pour ce qui est des différentes situations. Nous espérons bien pouvoir parvenir à la finalisation du code du Myanmar nécessaire pour que les gens comprennent qu'ils doivent nous aider parce que ceci n'est pas une question d'intérêt seulement pour les anglophones. S'ils ont des contributions à faire, nous espérons donc les recevoir afin d'évaluer les différentes situations.

internationalisés (IDN)

Nous espérons bien que la communauté sait que l'internet est prêt pour éliminer les barrières idiomatiques.

Merci.

PITINAN KOOARMORNPATANA : Merci beaucoup cher collègue.

Avez-vous des questions ou des commentaires pour Yin May ?

SARMAD HUSSAIN :

Merci. Il y a une question sur le chat qui n'est pas directement liée à ce qu'a dit notre orateur mais plutôt avec une partie précédente de la présentation. C'est une question posée par Andrey : « Pourquoi on utilise des alphabets historiques non utilisés ? Si j'ai bien compris, pourquoi est-il nécessaire de les inclure ? »

PITINAN KOOARMORNPATANA : Merci de votre question.

Nous nous sommes posés une question similaire au cours de la période de consultation publique qui vient de finir. Merci de votre question. Si vous avez une question, ce serait bien de nous la transmettre. Nous pourrions peut-être coller ici le lien pour les commentaires publics, je vous en serais très reconnaissante, pour recevoir des informations pour voir comment gérer différentes catégories de scripts incluses dans le rapport.

internationalisés (IDN)

SARMAD HUSSAIN : Il y a quelques autres questions qui sont répondues dans le chat également. Il y a une question de Mouloud Khelif : « Quelle est votre relation avec l'acceptation universelle et spécifiquement avec le groupe directeur de l'acceptation universelle ? » Il y a une autre question : « Jusqu'à maintenant, combien de TLD d'IDN ont été délégués dans la zone racine ? »

PITINAN KOOARMORNPATANA : Le groupe directeur de l'acceptation universelle est dirigé par la communauté. Nous travaillons avec le programme d'acceptation universelle d'IDN de l'ICANN et l'UASG. Le groupe UASG a de nombreux groupes de travail qui sont sous notre groupe. Et nous avons une séance spécifique pour l'acceptation universelle demain et la semaine prochaine et vous pourriez y assister si le thème vous intéresse. Mais le programme des IDN et de l'UA travaille aussi avec l'UASG, qui est le groupe dirigé par la communauté.

L'autre question concernait combien de TLD d'IDN ont été délégués. Voilà, c'est environ 150, mais je ne connais pas le chiffre exact pour les ccTLD et les gTLD. Je vous dirai ce chiffre exact après et je vais le publier sur le chat. 154, merci.

Nous continuons donc. Je vais passer la parole à Marc Blanchet qui est membre du panel d'intégration qui va partager avec nous une mise à jour sur le RZ LGR version 5.

MARC BLANCHET : Merci Pitinan.

Cette présentation a trait à notre travail dans le panel d'intégration et tout ce que l'on a fait depuis la dernière réunion de l'ICANN. L'objectif est de faire le travail du panel, à savoir l'intégration. En conséquence, nous avons pris tous les LGR du panel d'intégration, on les intégrés pour former le LGR pour la zone racine qui est utilisé par la communauté. Diapositive suivante s'il vous plaît.

Actuellement, on travaille sur la cinquième version du RZ LGR. La quatrième version a été finie automne 2020. Le RZ LGR 4 contenait 18 scripts comme l'arabe, le bangla, le devanagari, l'hébreu, le gujarati, le géorgien, le lao, le malayalam, le tamil, le telugu et le thaï.

L'objectif est d'intégrer six nouveaux scripts, à savoir le grec, l'hiragana et le katakana pour le japonais, le han qui était déjà dans la version 4, le coréen qui inclus des codes spécifiques, le hangeul et aussi le han qui se trouve aussi dans la version 4, le latin et le script du Myanmar.

Nous avons ajouté deux scripts dont on dispose depuis des années, mais on a décidé de remettre cela à plus tard. À ce moment-là, nous ne les avons pas intégrés à la zone racine parce qu'ils étaient en relation avec le latin et le grec. Alors on attendait à avoir les quatre codes pour avoir la version finale de manière à pouvoir les intégrer, parce que nous partageons des variantes et à travers le processus d'intégration, nous imposons au fur et à mesure les différentes variantes. Nous avons aussi dû attendre parce que nous voulions nous assurer de la cohérence entre les différents scripts en relation.

Le LGR 5 a deux ensembles de codes absolument intégrés : le japonais, le chinois, le coréen, le latin, le grec, le cyrillique et l'arménien. D'une certaine manière, le LGR 5 sera l'intégration la plus complexe et on consacre donc beaucoup de temps à la vérification et au fait de nous assurer que l'intégration soit bien faite. Diapositive suivante s'il vous plaît. Je crois que vous en avez sauté une, merci. Très bien.

Nous revenons un tout petit peu en arrière. La version 5 sur laquelle nous travaillons à ce jour inclura notamment 25 LGR élémentaires, à savoir qui sont spécifiques pour un code d'écriture. Pourquoi 25 ? Parce que nous avons 26 scripts, mais le Japon a deux scripts dans un LGR, c'est donc 26 moins un. Nous allons inclure également un LGR créé de manière mécanique.

Nous espérons que le 24 mars nous pourrions lancer le LGR 5 et le présenter pour consultation publique. Nous espérons donc recevoir la version du Myanmar prochainement pour pouvoir l'intégrer.

Nous voudrions que les gens consacrent leur temps à faire une révision du LGR 5 parce qu'encore une fois, ce sera la plus complexe et celle qui sera utilisée à l'avenir. Ceci finira l'intégration de tout le travail du panel de génération d'étiquettes avec les LGR que nous avons reçus. Diapositive suivante s'il vous plaît.

Pour ce qui est des variantes communes à différents scripts, ceci peut être vraiment significatif pour la publication. Ces variantes ne sont complètes que dans le LGR commun. Là, nous prendrons toutes les variantes communes à différents scripts et nous allons les combiner

de manière transitive pour créer le LGR commun. Ce LGR commun est utilisé pour les tests de collision entre des étiquettes valides.

Ce que nous voyons au-dessous, ce sont les conditions requises pour le traitement des variantes. Il y a une variante qui sera intégrée. Toute variante en script en répertoire est toujours répertoriée dans le LGR concerné. Les variantes ASCII sont toujours répertoriées par politique et les LGR des éléments de l'ensemble de scripts arménien et latin répertorient toutes les variantes de l'ensemble, mais seuls les LGR latins répertorient les variantes en dehors de cet ensemble.

Les scripts cyrillique et arménien qui ont été imposés seront ajoutés et il y aura d'autres d'éléments de LGR comme le malayalam qui ont été sélectionnés comme des variantes imposées dans les nouveaux scripts pour les règles. Il y a là un appel pour les panels de génération qui sont concernés, par exemple le cyrillique ou l'arménien, le malayalam et d'autres ; pour eux, ce sera vraiment très utile de vérifier le travail d'intégration que nous avons fait qui ajoute en fait de nouveaux scripts dans le LGR avec un élément, cet élément-là. Il faut que nous nous assurions qu'on n'a pas ajouté de problèmes dans le processus d'intégration, donc nous demandons à toute la communauté de bien vouloir vérifier notre travail et de nous assurer que tout est en forme pour les prochaines étapes pour ces règles de génération d'étiquettes de la zone racine.

Voilà la présentation qui a été faite la dernière fois. Je ne rentrerai pas trop dans le détail, mais je veux dire que lorsque nous voulons traiter ou vérifier une étiquette dans le LGR de la zone racine, il faut d'abord vérifier que l'étiquette est valable en utilisant le LGR d'un élément.

Ensuite, nous utiliserons le LGR commun pour vérifier les collisions. Et si tout est correct, nous allons utiliser le LGR de l'élément d'un script pour voir la variante et nous pouvons ensuite avoir ces variantes-là pour les différentes étiquettes qui sont des variantes de l'étiquette originale. Nous allons nous servir pour cela du LGR de la zone racine. Nous avons un outil qui a été mis à disposition par l'ICANN qui permet de travailler sur tout ce processus, qui travaille justement sur cette question.

Nous allons maintenant voir les prochaines étapes du panel d'intégration. Nous avons l'intention de publier le RZ LGR version 5 le 24 mars, c'est-à-dire d'ici un mois. Ce que nous voudrions, c'est voir que la communauté et en particulier les panels de génération aient pu vérifier et réviser notre travail. Il s'agit d'un commentaire très utile pour nous. Si vous ne trouvez peut-être pas de problème, vous pourriez peut-être nous envoyer un message au portail de consultation publique où vous direz que vous avez vérifié votre script et que tout est correct. C'est une information précieuse pour nous. Donc si vous prenez le temps de vérifier et de commenter cela, nous vous en serons reconnaissants. En principe, notre objectif provisoire pour la finalisation du LGR 5 est pour le deuxième trimestre de l'an 2022. Il y aura probablement une prochaine version du LGR de la zone racine une fois que l'on aura formé les panels et que le LGR nous aura été présenté pour une proposition finale.

Comme Pitinan l'a dit, ces panels n'ont pas encore été formés. Nous n'avons pas donc pour le moment de date pour cette nouvelle version

internationalisés (IDN)

qui sera probablement LGR 6. Nous n'avons pas encore de date pour ce travail-là.

Je crois maintenant que c'est ma dernière diapositive et c'est la fin de ma présentation. Merci beaucoup.

PITINAN KOOARMORNPATANA : Merci Marc. Nous allons voir s'il y a des questions.

SARMAD HUSSAIN : Oui, nous avons des questions. Il y en a une de Bill Jouris : « Pourquoi ne pas établir de liste des variantes entre le cyrillique et l'arménien si le GP latin n'est pas utile ? » Nous pouvons peut-être répondre à cette question et en prendre une autre après.

MARC BLANCHET : Je crois que cette question vient du chat. Permettez-moi de vérifier cela pour voir si je peux comprendre ce que l'on veut demander.

La politique que nous avons adoptée indique que toutes les variantes communes dans les LGR seront présentes seulement pour que nous disposions de toutes les LGR par rapport aux variantes qui ne sont pas liées aux scripts en soi. Nous avons décidé de restreindre aussi la position de certains LGR d'éléments pour la lisibilité des codes de lecture. Mais nous allons toujours nous servir du LGR commun pour avoir la liste complète des variantes communes à différents scripts.

internationalisés (IDN)

SARMAD HUSSAIN :

Merci Marc.

Nous avons une autre question de Vadim Mikhaylov : « Si je comprends bien, le latin en LGR 5 comprend non seulement le latin basique, mais il a été utilisé aussi au latin étendu. Alors quelles sont les extensions, A, B, C, les règles d'Unicode, etc. qui sont incluses ? »

MARC BLANCHET :

Cette question est plutôt liée aux règles de génération d'étiquettes latines. Je crois qu'on pourrait y répondre là.

Chaque panel de génération d'étiquettes identifie toutes les langues ayant un certain niveau d'usage avec une règle spécifique connue. C'est eux qui définissent la portée de notre travail en termes de toutes les langues qui disposeront de support dans ces LGR et dans ce script. C'est comme cela que l'on établit les points de code que l'on vérifie.

Il faut en même temps voir les MSR avec le répertoire initial qui est en général l'ensemble le plus étendu où tous les glyphes historiques qui ne sont plus utilisées sont écartés ou sont éliminés. On crée ainsi un ensemble initial des points de code qui sont les points de code à vérifier. Tous les caractères latins sont ceux qui reçoivent un support dans différentes langues.

SARMAD HUSSAIN :

Merci.

Il y a une autre question de Satish Babu : « La version 5 sera-t-elle compatible avec la version 4 ? »

MARC BLANCHET : Oui. La version finale n'est pas encore finie, mais l'intention est qu'elle soit complètement compatible et il n'y a pas de raison de penser que ce ne sera pas le cas. Si cela arrivait, cela apparaîtra dans les commentaires publics pour que tout le monde puisse le voir.

SARMAD HUSSAIN : Il y a une autre question qui a une réponse sur le chat, mais je vais la lire, de Nabeel Yasin : « Est-ce que tous les systèmes de DNS comme par exemple le BIND peuvent traduire les domaines IDN aux IP ? » Il y a eu un commentaire sur le chat qui dit que dans ce système de DNS, toutes les étiquettes sont sous le format a et elles peuvent donc être liées aux caractères ASCII pour le traitement de résolution interne.

Marc, est-ce que vous voulez ajouter quelque chose ?

MARC BLANCHET : La question se rapporte plutôt à l'IDNA en général. Il y a deux modalités des noms de domaine internationalisés. L'une d'elles, c'est celle qui se trouve dans le script, la vraie valeur que l'utilisateur va voir. Et cette chaîne de caractères est celle que l'on présente à l'utilisateur dans un navigateur. Mais cette chaîne de caractères n'est pas celle qui est codifiée dans le système du DNS. Le système du DNS est toujours ASCII et nous l'avons conçu justement pour ne pas avoir à changer tous les systèmes de DNS du monde. Cette chaîne de caractères que l'utilisateur utilise dans son propre code script sera traduite dans ce que nous appelons une étiquette a avec un

internationalisés (IDN)

algorithme qui s'appelle Punycode qui commence par xn-- et ensuite, il y a une liste de caractères ASCII. Donc l'utilisateur ne voit pas cela. C'est ce que le DNS utilise pour la résolution des adresses IP par exemple.

SARMAD HUSSAIN :

Merci.

Il y a une question de Raymond Mamattah : « Y a-t-il un travail en cours quant aux caractères ewe ? C'est l'une des langues les plus largement parlées en Afrique occidentale, en particulier au Togo et au Ghana. »

MARC BLANCHET :

Je n'ai pas d'idée à ce propos. Je ne sais pas si Pitinan ou Sarmad le savent, mais moi, je ne le sais pas, excusez-moi. Michael dit que cela a été considéré dans le travail concernant les caractères latins.

PITINAN KOOARMORNPATANA : Oui, cela fait partie du script latin.

Nous avons encore une minute de plus. Pour finir, nous allons faire un bref résumé pour ce qui est des lignes directrices. Nous devons encore voir ce que diront les experts. La période de consultation publique est close pour les IDN de support. Si vous voulez présenter d'autres commentaires, vous pouvez le faire. Mais en ce moment, nous sommes en train de peaufiner les commentaires finaux. Pour la consultation publique pour le script du Myanmar, elle sera ouverte jusqu'au 3 mars. Et bien que l'on sente que c'est une date un peu

lointaine, il faut inclure quelques commentaires et c'est peut-être intéressant pour vous de le voir. Et pour les règles de génération d'étiquettes version 5, nous espérons bien pouvoir les publier vers la fin du mois de mars. C'est pourquoi nous voudrions recevoir vos commentaires pour pouvoir réviser tout cela.

Je crois que c'est tout ce que nous avons à vous dire aujourd'hui. Nous sommes disponibles, vous pouvez nous contacter à idnprogram@icann.org. Nous vous remercions tous d'avoir participé à cette séance. Nous vous souhaitons une bonne semaine et une réunion vraiment réussie à l'ICANN.

Nous pouvons arrêter l'enregistrement. Merci beaucoup.

[FIN DE LA TRANSCRIPTION]