



Note explicative sur le programme des nouveaux gTLD

Discussions sur l'exigence d'une chaîne à 3 caractères

Date de publication :

30 mai 2009

Contexte - Programme des nouveaux gTLD

Depuis sa création il y a 10 ans en tant qu'organisation multipartite à but non lucratif dédiée à la coordination du système d'adressage de noms sur Internet, l'ICANN compte, parmi ses principes fondamentaux, la promotion de la concurrence sur le marché des noms de domaine et le maintien de la sécurité et de la stabilité d'Internet - un principe reconnu notamment par les États-Unis et d'autres gouvernements. L'évolution annoncée permettra une plus grande liberté d'innovation, de choix et de modification dans le système d'adressage d'Internet, qui se limite aujourd'hui à seulement 21 noms de domaine générique de premier niveau. Dans une communauté Internet grandissante qui compte déjà plus de 1,5 milliard d'utilisateurs, la diversité, le choix et la concurrence sont essentiels à la pérennité du succès et de l'accessibilité du réseau mondial.

La décision du lancement des prochaines sessions de candidature aux nouveaux gTLD fait suite à une longue période de consultation approfondie, menée auprès de l'ensemble des regroupements de la communauté Internet mondiale. Des représentants d'un grand nombre de parties prenantes (gouvernements, individus, société civile, regroupements commerciaux et sur la propriété intellectuelle, communauté technologique) ont participé aux discussions pendant plus de 18 mois. En octobre 2007, l'Organisation de soutien aux politiques des noms génériques (GNSO), un des groupes chargés, au sein de l'ICANN, de coordonner les politiques mondiales relatives à Internet, a achevé ses travaux d'élaboration de politiques sur les nouveaux gTLD et a approuvé un ensemble de recommandations. L'aboutissement de ce processus d'élaboration des politiques a été la décision du conseil d'administration de l'ICANN d'adopter la politique proposée par la communauté en juin 2008, lors de la conférence de l'ICANN à Paris. Un résumé détaillé du processus et de ses conclusions est disponible à l'adresse <http://gns0.icann.org/issues/new-gtlds/>.

Ce document fait partie d'une série de notes explicatives publiées par l'ICANN pour aider la communauté Internet à mieux comprendre l'appel d'offres, également appelé *Guide de candidature*. Une période de commentaires publics sur ce Guide de candidature permettra à la communauté Internet d'effectuer une révision détaillée et de faire part de ses remarques. Ces commentaires seront alors utilisés pour réviser les documents visant à préparer un Guide de candidature final. L'ICANN publiera ce guide final et ouvrira les candidatures au premier semestre 2010. Pour connaître les dernières informations, les activités et les délais concernant le programme des nouveaux gTLD, consultez la page <http://www.icann.org/en/topics/new-gtld-program.htm>.

Notez qu'il s'agit uniquement d'une discussion préliminaire. Les candidats potentiels ne doivent pas s'appuyer sur les détails présentés dans le programme relatif aux nouveaux gTLD, ce programme restant soumis à modification suite aux différents commentaires qui seront reçus.

Résumé des points clés exposés dans le présent document

- Historiquement, les gTLD sont composés d'au moins trois caractères. Les ccTLD utilisent des combinaisons de 2 caractères basées sur la liste ISO3166-1.
- Dans le cadre du processus d'introduction de nouveaux gTLD, le Guide de candidature (version 2.0) stipule que « *les chaînes proposées doivent comprendre au moins trois lettres ou caractères visuellement distincts dans le script* ».
- Plusieurs commentaires reçus soulignent la nécessité d'autoriser des gTLD composés de moins de 3 caractères.
- Ce document décrit une proposition de solution émanant de la communauté, qui autoriserait des chaînes de gTLD de moins de 3 caractères dans certains cas, et en présente les avantages et les inconvénients.

Le Guide de candidature (version 2.0) stipule que « *les chaînes proposées doivent comprendre au moins trois lettres ou caractères visuellement distincts dans le script* ». Il précise également que l'ICANN a reçu de nombreux commentaires sur cette question, certains demandant que les gTLD de moins de trois caractères soient autorisés dans des cas particuliers. Ces cas concerneraient les scripts contenant des idéogrammes dans lesquels des combinaisons d'un ou deux caractères expriment un seul et même mot (terme générique ou géographique).

Ce commentaire avance que la présence obligatoire d'au moins trois caractères dans un nom de domaine gTLD ne fonctionne pas pour de nombreux scripts car une ou deux « syllabes » peuvent représenter un mot. Par exemple, ces questions se posent pour le chinois, le japonais et le coréen. Sans URL composées d'un ou deux caractères au premier niveau, l'introduction de gTLD IDN n'aura aucun intérêt pour les utilisateurs d'Internet dans ces scripts. L'une des solutions possibles serait de créer une exception cohérente pour certains scripts comme le chinois, le japonais et le coréen.

Les problèmes rencontrés lors de l'élaboration de la version 2.0 du Guide, concernant ces exigences et le choix d'une méthode de mise en œuvre appropriée, sont examinés dans le document suivant <http://www.icann.org/en/topics/newgtlds/agv1-analysis-public-comments-18feb09-en.pdf>.

Comme l'explique cette analyse, le principal problème est que l'autorisation des chaînes de moins de 3 caractères pourrait créer de la confusion avec les ccTLD existants ou futurs dans l'éventualité d'une extension de la liste ISO-3166-1. Toutefois, il est clair que certains caractères (dans différents scripts) ne pourraient générer aucune confusion.

La recommandation du GNSO à l'ICANN était d'attribuer des chaînes à un ou deux caractères dans certains scripts au cas par cas. Pendant l'élaboration des

modalités de mise en œuvre, déterminer un jeu de règles bien définies permettant de décider quels scripts seraient autorisés en fonction de leur nature idéographique s'est révélé compliqué. Les règles seraient appliquées individuellement aux chaînes proposées comptant moins de 3 caractères. C'est pourquoi l'ICANN a reconsulté la communauté afin d'élaborer un jeu de règles permettant d'attribuer des chaînes à un ou deux caractères, sans entraîner de problèmes de stabilité ou de sécurité pour le DNS.

Ce document décrit les arguments pour et contre l'assouplissement de la règle des 3 caractères dans les chaînes de gTLD. Le commentaire en faveur de l'assouplissement de cette règle dans les scripts chinois, japonais ou coréens, où un ou deux caractères expriment souvent un seul et même mot, proposait également un test. L'ICANN publie ces informations pour dynamiser les discussions au sein de la communauté en préparation de la réunion de l'ICANN prévue à Sydney (Australie) du 21 au 26 juin 2009. Pour être acceptable, la solution devra définir des principes clairs expliquant dans quelles conditions des chaînes de moins de 3 caractères seront autorisées comme TLD. La solution doit être techniquement réaliste et répondre aux questions légitimes concernant le risque de confusion chez les utilisateurs.

Solution proposée pour autoriser des chaînes de gTLD de moins de 3 caractères

Les propositions et commentaires suivants ont été recueillis dans des consultations et commentaires publics lors de la réunion de l'ICANN à Mexico City en mars 2009 et suite à celle-ci.

« L'ICANN devrait respecter les principes recommandés par le GNSO et lever la restriction sur la longueur d'un TLD IDN (c'est-à-dire plus de deux caractères) ou modifier cette clause afin de la rendre moins contraignante pour certains scripts. Le maintien de cette restriction pénaliserait considérablement les TLD chinois car les mots chinois les plus porteurs de sens sont composés de 2 caractères. »

Des commentaires similaires ont été collectés concernant les caractères japonais et coréens.

Le collège d'experts sur les registres de gTLD a avancé que « les scripts chinois, japonais et coréens devraient être autorisés à déroger à la règle des 3 caractères minimum. »

Plusieurs approches ont été suggérées afin de voir si les enregistrements à un ou deux caractères doivent être autorisés pour certains scripts.

Une recommandation précisait :

Si les caractères Unicode du système d'écriture d'une langue donnée possèdent un sens propre, la restriction sur le nombre minimum de 3 caractères devrait être levée. Le demandeur doit spécifier à quelle catégorie de systèmes d'écriture appartient la chaîne qu'il souhaite utiliser, c'est-à-dire logographique, syllabique, alphabétique, abugida, abjad ou caractéristique. Cette règle doit s'appliquer au cas par cas et par langue (par ex., les chaînes hiragana en japonais peuvent rester soumises à la règle de 3 caractères minimum, mais pas les chaînes kanji). J. Seng (13 avril 2009).

Autre commentaire :

La règle des 3 caractères requis doit être levée pour les chaînes dont le système d'écriture emploie des éléments de base ayant recours à des associations sémantiques généralement acceptées, lorsque des séquences d'un ou deux caractères représentent des concepts à part entière qu'il n'est pas possible d'abrèger. Ces systèmes étant très éloignés du latin, la confusion visuelle n'est pas un problème (les équipes chargées de la vérification des chaînes y veillent). Le répertoire de caractères de ces scripts est bien plus étendu que celui des scripts alphabétiques ou syllabiques (par ex., 71 442 caractères han en Unicode version 3.2 contre 26 caractères anglais). W. Tan (13 avril 2009).

Et voici le commentaire le plus substantiel que nous avons reçu :

Pour les chaînes de gTLD de moins de 3 caractères, la règle doit être la suivante :

- 1. Les caractères alphabétiques majuscules ne sont pas autorisés. Toute chaîne contenant un ou plusieurs caractères associés à la propriété Unicode « lu » ne sera pas autorisée comme TLD.*
- 2. Les caractères alphabétiques minuscules ne sont pas autorisés. Toute chaîne contenant un ou plusieurs caractères associés à la propriété Unicode « ll » ne sera pas autorisée comme TLD.*

Les règles 1 et 2 éliminent pratiquement tous les caractères Unicode qui peuvent être confondus visuellement avec les caractères ASCII utilisés dans les chaînes de ccTLD. Toutefois, certains cas plus délicats méritent un examen plus approfondi.

- 3. Les chaînes qui ne sont pas éliminées par la règle 1 ou 2, mais qui comptent moins de 3 caractères, doivent passer un test de confusion avec les caractères latins de base (a,b,...,z), afin d'éviter tout conflit avec les ccTLD existants ou les futurs ajouts à la liste ISO3166-1.*

La même règle peut être exprimée de différentes manières, comme suit :

- 1. Les chaînes contenant des caractères appartenant à l'une des catégories ci-dessous des propriétés Unicode sont considérées comme relevant d'un « système d'écriture alphabétique » et restent soumises à la règle des 3 caractères minimum. Dans le cas contraire, le système d'écriture est qualifié de non alphabétique et peut déroger à cette règle.*
 - Valeurs de catégorie générales = L& (c'est-à-dire Lu, Ll et Lt) [ceci englobe les scripts latins] ; ou*
 - Liste de propriétés [<http://unicode.org/Public/UNIDATA/PropList.txt>] = Other_Alphabetic ou Other_Lowercase; [ceci englobe les scripts alphabétiques non latins, c'est-à-dire arabiques, hébreux, sanscrits, etc.].*
- 2. Si une chaîne contient des caractères ayant l'attribut « PropertyList [<http://unicode.org/Public/UNIDATA/PropList.txt>] = Ideographic », elle n'est pas soumise à la limite des 3 caractères.*

Ce commentaire public suggère que limiter les enregistrements à 3 caractères minimum par chaîne pénaliserait considérablement les utilisateurs dans ces régions, et présente l'avantage de proposer une solution concrète.

Arguments en faveur de gTLD d'au moins 3 caractères

Ces arguments peuvent être regroupés dans les catégories suivantes :

Équité de traitement

Outre les commentaires émanant des communautés chinoise, japonaise et coréenne, l'ICANN a reçu des arguments de l'Europe expliquant que certaines associations d'un ou deux caractères dans les langues européennes représentent un mot, un concept ou, dans certains cas, des identificateurs géographiques. Ces arguments ont été avancés pour contrer ceux en faveur des chaînes de moins de 3 caractères, dans la proposition ci-dessus. Si les chaînes de moins de 3 caractères sont autorisées pour les langues chinoise, japonaise et coréenne, elles doivent également l'être pour les autres langues, en vertu du principe d'équité de traitement.

Déclarations concernant les mots chinois et le nombre de caractères

D'autres arguments observent que peu de caractères chinois sont des mots. En majorité, les mots chinois associent plusieurs caractères. Si l'on excepte les implications phonétiques spécifiques (1 caractère = 1 phonème), la caractérisation des caractères chinois en tant que syllabes est plus pertinente. De plus, d'autres avis font remarquer que la plupart des mots chinois comptent plus d'un caractère.

Test de quelques gTLD avec moins de 3 caractères

D'aucuns ont suggéré que l'ICANN effectue un test sur un nombre réduit de gTLD de moins de 3 caractères. Les données collectées permettraient de mieux cerner le processus d'élargissement de l'attribution de ces chaînes.

Conversion des TLD

Des commentaires notent également qu'il n'existe aucun modèle de conversion des TLD. Par conséquent, les codes ISO à 2 lettres et les autres TLD qui sont des « abréviations » ne peuvent pas être convertis en chaînes IDN de moins de 3 caractères, conformément à cette définition. Non seulement il ne s'agit pas d'abréviations standardisées, mais les abréviations ne sont pas un concept commun à toutes les langues et à toutes les cultures. Les ccTLD notamment sont un système de codage standardisé, mis en place pour plusieurs raisons, dont la facilité d'identification et le caractère distinctif des caractères latins sans empatement.

Fonction de délégation de ccTLD de l'ICANN

Actuellement, la fonction de délégation de l'IANA gère le nombre réduit de combinaisons ASCII de 2 caractères disponibles. Dès lors qu'elles figurent dans la liste ISO3166-1, celles-ci sont traitées comme des ccTLD. Les discussions menées au sein de la communauté de l'ICANN, ainsi que lors des réunions ISO MA ces dernières années, ont porté sur la faisabilité de l'extension de la liste ISO3166-1 aux combinaisons à 2 caractères pour les autres scripts, représentant des noms de pays et de territoire. La tâche équivaudrait à plusieurs années de travail pour l'ISO (ce qui pourrait intervenir après le PDP d'IDN ccNSO). Le résultat du « Fast Track Process » permettra d'orienter la discussion sur la faisabilité et les modalités d'extension de la liste ISO3166-1, la fonction de délégation de ccTLD de l'ICANN, ainsi que le PDP à long terme du ccNSO pour les ccTLD d'IDN.

La délégation des chaînes à un ou deux caractères risquerait de mettre en péril la forme future du mécanisme de délégation des ccTLD.

Conclusion

La communauté de l'ICANN continue de travailler sur les solutions proposées, en conciliant les besoins de la communauté en faveur d'un élargissement de l'utilisation avec ceux des tenants d'une utilisation conforme aux normes de l'ISO. Cette question pourrait faire l'objet d'une discussion élargie et approfondie.

Ce débat devrait se poursuivre pour préserver l'obligation incontournable de maintenir l'ouverture d'Internet à l'innovation.

L'ICANN invite la communauté à prolonger le dialogue pour dégager des solutions potentielles à ce problème.