

IDN

Stéphane Bortzmeyer
AFNIC
bortzmeyer@nic.fr

10 février 2011

Ça veut dire quoi ?

IDN = *Internationalized Domain Names*

Ce sont simplement des **noms de domaines** utilisant toutes les écritures possibles, pas seulement l'alphabet latin réduit.

Exemple de noms IDN :

- ▶ café.fr
- ▶ maçonnerie-générale.fr
- ▶ académie-française.fr
- ▶ Mais cela marche aussi pour des écritures non latines comme l'écriture arabe.

Pourquoi un nom spécial ?

Pourquoi tous les noms de domaine ne sont-ils pas des IDN depuis le début ?

Raisons historiques

Si le DNS lui-même accepte tous les caractères, une partie de son environnement avait des problèmes. D'où la restriction initiale aux caractères latins non-composés.

Donc, maintenant, on peut ?

La norme technique qui permet les IDN (« RFC 3490 ») date de 2003.

Mais attention, norme publiée ne signifie pas déploiement dans **tous** les logiciels. Exemple : Firefox qui reconnaît ces noms mais les affiche sous une forme bizarre (`www.véliplanchiste.com` affiché `www.xn-vliplanchiste-bhb.com`).

En pratique, en 2011, le titulaire d'un IDN ne pourra pas l'exploiter exactement comme un nom traditionnel.

Au fait, pourquoi a t-on besoin des IDN ?

Pour écrire le français correctement, tout simplement.

`congres.fr` concerne les congrès ou bien le poisson ?

Et pourquoi refuser les IDN ?

- ▶ Par refus de l'internationalisation (« De toute façon, tout le monde parle anglais. »)
- ▶ Par crainte des confusions possibles (« Il y a davantage de caractères dans le monde, Horatio, que n'en rêve ta philosophie. »)

Hameçonnage : ces faux sites Web qui se font passer pour votre banque et essaient de vous faire taper votre numéro de carte de crédit.

Ce problème a parfois été cité par les anti-IDN. Mais il existe depuis longtemps (paypal.com et paypa1.com, regardez bien).

En pratique, largement un faux problème.

M. Tout-le-monde ne lit - hélas - pas les URL. Il ne sait pas analyser un nom de domaine. Les hameçonnages dans le monde réel ne pratiquent donc pas la confusion (ou alors très simpliste, genre secure-societe-generale.com au lieu de secure.societe-generale.com).

De la norme technique à la politique

Chaque registre de noms de domaines est libre de sa politique.

Certains acceptent des IDN depuis longtemps (com, org), d'autres plus récemment (la racine seulement en 2010).

Le premier domaine de tête francophone à accepter les IDN a été ch (Suisse) en mars 2004.

Aujourd'hui, la majorité des registres européens (13 sur 22 répondants, janvier 2009) accepte les IDN. Le nombre d'enregistrements est de l'ordre de quelques %, cohérent avec le pourcentage de mots ayant au moins un caractère composé.

Quels caractères ?

La plupart des registres n'acceptent qu'une partie des caractères possibles :

- ▶ Raisons pratiques : le personnel du registre, par exemple le support, doit pouvoir les comprendre et les taper (« Vous avez dit que votre domaine est beethoven-straße.de ? Avec un eszett ? »)
- ▶ Raisons juridiques : certains pays ont des langues officielles, des listes de caractères officielles.

Pour la France : les caractères du français (liste déjà établie et validée), des langues régionales (le breton utilise ñ), des langues de France (arabe, chinois, portugais...), des langues d'Europe ?

Quelles règles d'enregistrement ?

palais-congres.fr et palaiscongres.fr ne sont pas liés. Ils peuvent avoir des titulaires différents (exception : le privilège des noms de commune).

Faut-il que palais-congres.fr et palais-congrès.fr soient liés ? Doit-on gérer un **lot** qui fasse que le titulaire de palais-congres.fr soit le seul à pouvoir déposer palais-congrès.fr ?

Mesures temporaires ?

On peut aussi envisager des mesures temporaires (« lever de soleil », disent les registres de noms) comme de ne permettre, **au début** que le titulaire du nom réduit.

Ensuite, on passerait en mode « pas de lot ».

Dans ce cas, qui bénéficierait du « lever de soleil » ? Les marques ?
Les titulaires actuels ?

Que font les collègues ?

Aucun de nos collègues du CENTR (association européenne des registres) n'a signalé de cas de hameçonnage lié aux IDN.

La totalité des registres européens a décidé de ne **pas** utiliser de lots.

Les registres d'autres pays, utilisant des écritures non-latines, ont parfois choisi de gérer des lots (exemple en cn : le nom en sinogrammes traditionnels est dans le même lot que le nom en sinogrammes simplifiés).

Elles ne sont pas à l'ordre du jour.

Mais peuvent être posées près d'un café.