

Google inverse les noms de domaines

Stéphane Bortzmeyer

<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 16 janvier 2020. Dernière mise à jour le 20 janvier 2020

<https://www.bortzmeyer.org/google-inverse-noms.html>

Le 16 janvier au matin, il y avait une bogue amusante (qui a duré quelques jours) dans le moteur de recherche Google. Les noms de domaine étaient dans certains cas inversés, plus exactement mélangés dans l’affichage.

Quand on cherche quelque chose dans ce moteur de recherche (sur une machine de bureau, car c’est différent sur les mobiles), Google affiche pour chaque résultat :

- L’endroit où a été trouvé le résultat, sous une forme vaguement inspirée des “*breadcrumbs*” (Google affichait autrefois un URL, ce qui était plus intelligent mais a sans doute été changé sur demande du marketing car « M. Michu ne comprend rien aux URL <<https://wersm.com/google-rolls-out-a-new-look-for-search-on-desktop/>> » et Google n’aime pas éduquer les utilisateurs),
- Le titre de la page, qui est un lien hypertexte (en promenant son pointeur dessus, on voit l’URL s’afficher en bas à gauche, ce qui est utile pour savoir où on va avant d’y foncer),
- La date et un résumé.

On voit ici un exemple :

L’endroit de la page est fabriqué à partir de l’URL mais n’est pas un URL. Il commence par un nom de domaine et se continue par un chemin dans le site. Mais ce qui est amusant est que cet affichage est bogué. Dans certains cas, le nom de domaine est mélangé. L’exemple précédent est correct mais, si je cherche le département de physique de l’Université de Cambridge :

Le vrai nom de domaine est `www.phy.cam.ac.uk` mais Google affiche `phy.www.cam.ac.uk`. Autre exemple, avec le gouverneur de Californie :

Ici, le vrai nom était `www.gov.ca.gov` mais Google affiche `gov.www.ca.gov`. L’URL vers lequel pointe le résultat est néanmoins correct, comme on le voit avec cet autre exemple (Firefox affiche l’URL de destination en bas à gauche) :

Pourquoi cette bogue? La raison fondamentale est sans doute une incompréhension des noms de domaine, notamment du fait que le nombre de composants d'un nom soit quelconque (contrairement à une légende répandue <<https://www.bortzmeyer.org/parties-nom-domaine.html>>.) Vous noterez que les seuls noms ainsi massacrés sont des noms ayant plus de trois composants. Mais c'est curieux de la part de Google, dont les services ont d'habitude très peu de bogues.

Mais on peut aussi voir dans cette bogue le résultat d'une mentalité très fréquente dans le Web, et particulièrement chez Google, comme quoi les noms de domaine ne sont pas une information pertinente et peuvent être cachés, ou massacrés <<https://www.wired.com/story/google-wants-to-kill-the-url/>>, sans problème.

Merci à Jean-Philippe Pick pour avoir noté le phénomène. Le problème a été réparé le 19 janvier, sans nouvelles de Google ou sans information publiée.