

Chronique

de la recherche sur le web

par Marc Duval

2 avril 2001, vol.1, n° 7

Pages web et automates de recherche

La principale fonction d'un automate de recherche est de répertorier les pages web. Pour un automate, une page web est un document en lui-même. En tant que page d'entrée d'un site, elle doit avoir tous les éléments « recherchables » dans l'index. Sa composition doit s'adapter au médium car elle influence directement sa possibilité d'être trouvée. Les concepteurs des automates ont développé des champs de recherche pour nous faciliter la tâche.

Les principaux champs sont le titre, l'URL, le nom de domaine et l'hôte. Certains ont ajoutés la recherche par l'hyperlien, par le nom de l'hyperlien, par le type de fichier (pdf, images, etc.). Selon l'automate, la recherche se fera sur les méta-données, visibles en utilisant la commande « Source de la page » (Netscape) ou « Source » (Internet Explorer), et sur le texte. Étant donné que le document est indexé *in extensio*, chaque page doit comprendre la carte de visite de l'auteur, soit au minimum le nom de l'auteur et son adresse électronique. Pour un entreprise, la carte de visite complète est recommandée. Ces informations permettent au chercheur de trouver plus facilement un contact. Ainsi, un client d'une entreprise peut la retrouver en tapant le numéro de téléphone dans la fenêtre d'adresse d'un fureteur au lieu de l'adresse web.

La page doit aussi comprendre des

liens internes qui permettent une navigation facile dans le site et surtout elle doit netraiter que d'un seul sujet. Ce sujet doit être répété dans le titre visible, dans le titre des méta-données, dans la méta-description et dans l'adresse URL. Ce sont autant d'éléments qui augmente la possibilité d'être trouvé.

Parce que la page est unique, elle doit aussi avoir une composition classique comprenant une introduction, un développement et une conclusion.

Enfin, une partie de la page doit être consacrée aux mots clés utilisés dans des répertoires tels que Yahoo, Looksmart, ODP, Voila, Nomade, la Toile du Québec.

C'est à partir de ces conditions que l'on pourra trouver directement la page ou indirectement par sa référence dans une autre page du site comme lien de navigation.

Marc Duval
Service de recherche documentaire DSI
1280-4, bd de Montarville.
Boucherville. Québec. Canada
J4B 8B4
marcduval@dsi-info.ca