

Comment les moteurs de recherche établissent-ils l'ordre de leurs propositions de pages en réponse à une équation de recherche ?

Une fois la requête formulée comment les moteurs de recherche rangent-ils les pages trouvées ?

Mieux utiliser les moteurs de recherche, c'est également essayer de mieux comprendre comment ils fonctionnent et notamment comment ils hiérarchisent la liste des liens vers les pages trouvées en réponse à une équation de recherche. Cet ordre est important pour les sites, surtout les sites commerciaux qui comptent apparaître dans les premiers : en effet l'utilisateur ne consulte bien souvent que les premiers liens de la première page de résultats...

L'ordre d'affichage n'est pas le fruit du hasard ! En ce qui concerne les moteurs de recherche, qui effectuent leur investigation dans une base de données de plusieurs milliards de pages Web, la méthode de rangement des pages trouvées est complexe. Pourquoi telle page est-elle placée avant telle autre dans la liste affichée ? Les critères retenus ne sont jamais indiqués clairement à l'utilisateur (il y a donc dans ce qui suit des suppositions...).

Critères sur lesquels peut être calculé l'ordre d'affichage des liens proposés par un moteur de recherche en réponse à une équation de recherche comportant un ou plusieurs mots-clés :

- * un **accord commercial** entre le moteur et le site : certains liens figurent en haut des pages de résultats non pour des raisons de pertinence mais simplement parce qu'ils renvoient à des pages de sites qui ont "acheté" auprès de prestataires (annuaires, moteurs) tout ou partie des mots clefs que vous avez entrés dans votre équation de recherche : généralement ce type de positionnement « commercial » est indiqué, soyez attentifs !

- * la **présence et la place du mot-clé demandé dans le titre** de la page
(attention, « titre » est à prendre au sens informatique, balise html < title >, c'est-à-dire ce qui figure dans le titre de la fenêtre de la page dans le navigateur et donc dans le bouton associé dans la barre des tâches) ;

- * la **présence du mot-clé dans l'adresse (URL)** de la page
(exemple : fr.wikipedia.org/wiki/Machine_à_laver) ;

- * l'**indice de popularité de la page qui contient le mot-clé**, cet indice est lié au nombre de liens "pointant" - depuis d'autres sites web - vers cette page (pour être comptées ces pages tierces doivent elles aussi être référencées et donc « connues » par le moteur) ;

- * l'**indice de popularité des pages « tierces »** qui pointent vers la page contenant le mot-clé ;

- * la **présence du mot-clé dans le texte de la page et le nombre d'occurrences de ce mot-clé**

- * si la requête est composée de plusieurs mots-clefs : l'**ordre d'apparition des mots-clés dans la page** (les pages comportant les mots clefs dans l'ordre indiqué dans l'équation seront privilégiées) ;

- * si la requête est composée de plusieurs mots-clefs : la **proximité des mots-clefs dans la page** (les pages comportant les mots clefs rapprochés voire juxtaposés seront privilégiées) ;

- * la **place du mot-clé dans la page** ; (une page contenant le mot demandé en haut du document sera mieux "notée" qu'une autre le présentant en bas de page) ;

- * la **façon dont le mot-clé est mis en exergue** : style html, style de police : gras, taille des caractères, etc. ;

- * la **présence du mot-clé dans des champs qui n'apparaissent pas sous forme de texte sur la page** mais qui sont présents dans le code « source » HTML et lus par certains outils de recherche. C'est le cas des « balises Meta » : paragraphe de description de page, mots-clefs attribués par le rédacteur à sa page, texte alternatif destiné aux navigateurs ne prenant pas en charge l'affichage des cadres ;

- * la **présence du mot-clé dans les infobulles d'une image présente sur la page** ;

- * la **présence du mot-clé dans le texte support d'un lien pointant vers la page** : le mot-clef ne figure donc pas dans la page indiquée par le moteur (!) mais dans le texte d'un lien situé sur une autre page tierce, pointant vers cette page (et donc faisant potentiellement référence à son contenu dans le texte du lien) ;