

LA RECHERCHE NUMERIQUE

I- Généralité :

Les moteurs de recherche permettent de trouver des informations dans des pages dont on ne connaît pas l'adresse, voire dont on ignore l'existence. La méthode de recherche appelée référencement naturel se décompose en trois grandes activités, réalisées par les moteurs de recherche : (1) le parcours automatique du Web pour collecter les pages visitées (aspiration des pages Web effectuée par des robots) ; (2) l'analyse du contenu des pages et leur indexation sur les mots qu'elles contiennent (constitution d'un annuaire inversé qui associe à chaque terme les URL des pages où il apparaît) ; (3) la troisième activité, réalisée à chaque fois qu'un internaute fait une requête, construit une liste ordonnée des pages (classement) comportant les mots clés de la requête. Leur ordre dépend notamment de leur popularité (principe des liens), de leur pertinence (aux mots de la requête), et de l'ordre des termes de la requête.

Les concepteurs de site Web peuvent améliorer le référencement de leurs pages en choisissant bien les mots et en les plaçant à des endroits stratégiques dans les pages.

A- Interaction client/serveur :

Le Web s'appuie sur le dialogue entre **clients** et **serveurs**. L'interaction est à l'initiative des clients (les applications qui se connectent au Web, dont les navigateurs), qui envoient des **requêtes HTTP** aux serveurs. Ces derniers renvoient leur résultat : des pages qu'ils ont stockées ou qu'ils créent dynamiquement en fonction de la requête formulée. Les pages reçues par les clients peuvent contenir des codes exécutables (souvent en JavaScript) qui permettent aux clients d'effectuer des traitements en accédant aux ressources de son ordinateur et en interagissant avec les serveurs.

Les applications peuvent être paramétrées pour autoriser ou interdire l'accès à des ressources locales aux programmes téléchargés par les pages.

B- Sécurité et confidentialité :

En formulant des requêtes sur des sites Web dynamiques et en laissant des programmes s'exécuter sur sa machine, l'utilisateur prend des risques : il peut communiquer des informations personnelles à son insu à des serveurs qui en gardent une trace, à distance ou localement par des cookies, ou encore charger des pages contenant des programmes malveillants, par exemple permettant d'espionner en continu les actions de l'utilisateur. Par ailleurs, un navigateur peut garder un

historique de toutes les interactions, et le laisser accessible aux sites connectés. L'utilisateur peut utiliser des services qui s'engagent à ne pas garder de traces de ses interactions, par exemple certains moteurs de recherche. Il peut aussi **paramétrer son navigateur** de façon à ce que celui-ci n'enregistre pas d'historique des interactions. De fausses pages peuvent encore être utilisées pour l'hameçonnage des utilisateurs. Un nom de lien pouvant cacher une adresse Web malveillante, il faut examiner cette adresse avant de l'activer par un clic.