

Chronique

de la recherche sur le web

par Marc Duval

2 avril 2001, vol.1, n° 6

AltaVista est-il un automate de recherche ?

A chaque étude du web, la capacité des automates de recherche généraux à couvrir le web est mise en doute. Elle est mise en doute parce que ces études étendent leur champ au-delà du web. Elle est mise en doute parce que l'on oublie que les automates de recherche ont pour fonction principale d'indexer les pages web. Elle est mise en doute parce que l'on pense non en terme de « site de recherche », mais en terme d'automate de recherche.

Mais AltaVista n'est plus un automate de recherche, mais un site de recherche composé d'une banque de données de documents web - l'automate -, d'une banque de données d'images, d'une autre de documents audio et d'un dernier de documents vidéos. Celles de pages web comprend un langage d'interrogation puissant grâce à sa recherche par champs et par l'exploitation de la recherche par les domaines éducatifs et gouvernementaux. A cela s'ajoutent les banques de données partenaires d'AltaVista pour l'actualité avec Moreover, les entreprises avec Hoover's, et son association avec le répertoire Looksmart. Aussi, la banque de données web est assistée par d'autres banques de données satellites comme RealNames, Foxsports.com, Cars.com. Elle complète sa couverture du web avec ses sites nationaux

qui ne contiennent pas nécessairement les mêmes pages web.

De la même façon, le site de recherche iWon est alimentée par la banque de données web d'Inktomi, la banque de nouvelles Moreover, les banques de Fact City et de RealNames, les répertoires Looksmart et Direct Hit.

Lycos, Fast Search, Excite, Northern Light, Voila... suivent le même modèle selon la clientèle visée.

Dire qu'AltaVista, Google ou HotBot sont des « automates de recherche » est inadéquat. Leur développement fait en sorte qu'ils sont devenus des sites de recherche composés de plusieurs banques de données dont la principale est leur « automate ». Mais pour plus de commodité, conservons « automate de recherche » pour les distinguer des répertoires.

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