

# Google complète ses recherches avec l'audio et la vidéo

L'indexation de fichiers audio et vidéo représente une difficulté majeure pour les outils de recherche à laquelle tous se sont heurtés. Les expériences menées jusqu'à présent par des moteurs comme AltaVista ou AOL ont toutes échoué sur la question des droits des ?uvres?

En effet, une indexation 'en dur' des fichiers multimédia par les '*crawlers*' (robots d'indexation des moteurs de recherche) ne permet pas de différencier les fichiers légaux des fichiers illégaux indûment stockés sur les sites Web. On pense bien évidemment à la multiplication des fichiers partagés en '*peer to peer*'. La solution que Google semble vouloir adopter est directement dérivée de ses dernières démarches commerciales : d'après ses fondateurs Sergey Brin et Larry Page, le moteur privilégierait les accords commerciaux avec des sites de téléchargement, une manière comme une autre, aussi, d'adapter le modèle économique des AdWords (les liens sponsorisés vendus aux enchères au mot clé), voire des liens '*shopping*' sur lesquels les portails sont intéressés. On notera cependant que la démarche de Google n'est pas exclusive, puisque AOL, filiale de Time Warner et client de Google sur la recherche en ligne aux Etats-Unis, a racheté voici quelques mois Singinfish, un outil de recherche multimédia en ligne développé par Thomson. **La recherche multimédia, une histoire de marquage**

La difficulté de rechercher des fichiers multimédia en ligne provient bien évidemment du respect des droits intellectuels sur les ?uvres, mais aussi de la qualification des contenus.

Les moteurs de recherche sont actuellement basés sur des technologies textuelles. Leurs *crawlers* indexent le contenu textuel, mais se révèlent incapables de lire un contenu numérique qui ne répond pas aux règles de la sémantique. L'avenir du multimédia en ligne passe donc par la qualification du contenu, et la capacité à associer à un fichier multimédia des informations sous un format 'universel'. Cet avenir passe sans nul doute par XML, le langage de marquage de l'information et ses différentes déclinaisons, mais aussi par les formats de compressions comme les futurs MPEG, qui devraient prendre en compte ces problématiques. Le chantier reste ouvert?