

Google : ou l'on reparle de Dmoz? et on l'oublie

Dmoz (dmoz.org), c'est l'

Open Directory Project, l'annuaire collaboratif édité par des êtres humains, développé et maintenu mondialement par une vaste communauté d'éditeurs bénévoles. Les moteurs de recherche ayant balayé les annuaires, hormis dans une approche de verticalisation de la recherche par métier, on avait quelque peu oublié cet acteur historique de la recherche, et en particulier ses liens avec Google. Certes ces annuaires généralistes semblent tombés en désuétude, quand ils n'ont pas disparus ! Qui se souvient des annuaires de MSN, Lycos, ou encore de Nomade ? Pourtant, certains d'entre eux résistent – Dmoz, le Guide Yahoo ou le Guide de Volia ? et conservent des adeptes pour qui rien ne remplace l'expertise de l'humain pour qualifier un site Web. Sauf qu'aujourd'hui un site ne se qualifie plus sur un titre et une description, et qu'en dehors d'une recherche métier, donc concentrée, les moteurs de recherche sont devenus autrement plus performants, et la fréquentation des annuaires? anecdotique ! Pourtant, ces annuaires continuent parfois de traîner leurs guêtres dans les résultats des moteurs, à l'image de Dmoz qui renseigne certains résultats de Google. Tout d'abord, si un site est indexé par un éditeur de Dmoz dans l'annuaire, son indexation dans Google sera facilitée et accélérée. Mais surtout Google affichera en priorité la description de Dmoz sur le site proposé en résultat. Et c'est là que rien ne va plus ! En effet, la description dans un annuaire est un instantané de la généralité du site. Or, cette dernière n'est pas représentative des activités proposées en ligne par le site, et celles-ci peuvent évoluer dans le temps. Résultat, même si Google rafraîchit régulièrement sa base, le moteur peut fournir en réponse à une requête une description devenue rapidement obsolète. Afin de corriger ce défaut, Google a fait appel à une bonne vieille méthode, une balise meta, à placer sur les pages d'un site, indiquant à Googlebot, le '*crawler*' (robot d'indexation) du moteur, de ne pas prendre en compte les informations fournies par Dmoz : *meta name= »googlebot » content= »noodp »* Une métadonnée pour renseigner le robot, un retour aux sources bienvenu?