

Où vont les métamoteurs ?

[Retour au sommaire de la lettre](#)

L'offre de métamoteurs est pléthorique sur le Web. Pourtant, à l'heure où le marché de la recherche est de plus en plus dominé par trois ou quatre grands moteurs, de nombreuses questions se posent quant à l'avenir de ces outils. Tandis que les internautes tendent de plus en plus à utiliser un nombre restreint de moteurs, que vont faire les métamoteurs pour préserver leurs utilisateurs et en gagner de nouveaux ? Doit-on s'attendre à une évolution de leurs modèles économiques ? L'avenir de la métarecherche passera-t-il par une plus forte verticalisation des offres ? Ou bien, ces outils vont-ils mettre l'accent sur la qualité de restitution des résultats ? Nous vous proposons ici un tour d'horizon des solutions existantes et de leurs perspectives.

Les métamoteurs ou métachercheurs sont des applications en ligne ou sur PC qui permettent d'interroger simultanément différents moteurs et annuaires. Ils n'ont pas d'index en propre mais interrogent les bases de données des autres outils. Les résultats obtenus sont ensuite traités et restitués suivant des critères qui diffèrent considérablement d'un métamoteur à l'autre.

Le nombre élevé d'outils disponibles s'explique en partie par la relative simplicité d'implémentation et d'administration des métamoteurs par rapport aux moteurs de recherche "classiques". En effet, contrairement aux moteurs, ils n'utilisent pas de *crawlers* pour indexer les pages Web. Ils n'ont pas non plus besoin des mêmes bases de données ou de s'occuper de la qualité des contenus indexés (qui sont là aussi gérés par les moteurs).

Un peu d'histoire...

Les premiers métamoteurs sont apparus sur le Web en même temps que les moteurs, au milieu des années 90. Beaucoup d'entre eux sont issus de projets universitaires ou ont été développés par des chercheurs "passionnés".

Il en existe aujourd'hui des centaines ! Apparu début 2004, WebMetaSearch (<http://webmetasearch.com/>) – également appelé Ungoogole – fera, par exemple, plaisir aux "Google addicts" qui souhaitent diversifier leurs sources d'information. Ce métamoteur permet d'effectuer une recherche sur la plupart des outils de recherche majeurs "à l'exception de Google". Lancé dernièrement par l'Université américaine John Hopkins, un autre métamoteur, Dumbfind (<http://www.dumbfind.com/>), permet lui de transmettre une requête sur plusieurs moteurs et de la restreindre à une thématique en saisissant un deuxième mot-clé dans le champ "Topics". Le tout nouveau Polymeta (<http://polymeta.com/>), de la société hongroise K-PROG Bt, catégorise en outre automatiquement les résultats de cinq grands moteurs.

Chronologie des principaux lancements

Année	Nom - URL	Auteur(s) - Pays
Métamoteurs généralistes		
1994	Webcrawler - http://www.webcrawler.com/ : Moteur devenu un métamoteur suite à son rachat par InfoSpace en 2002	University of Washington - US
1995	ISleuth (Internet Sleuth) - http://www.isleuth.com/	Internet Business Connection - US
	SavvySearch : Racheté par CNET en 1999. Devient : Search.com - http://www.search.com/	Colorado State University - US
	ProFusion - http://www.profusion.com/ : Racheté par IntelliSeek en 2000	University of Kansas - US
1996	Mamma - http://www.mamma.com/	University of Ottawa - CA
1997	Ari@ne - http://ariane6.com/	Guy Himmelsbach -FR
	CurryGuide - http://web.curryguide.com/	CurryGuide - UK
	Dogpile - http://www.dogpile.com/ : Racheté par Infospace	US
	metaEUREKA - http://www.metaeureka.com/	4C Software - NL

Année	Nom - URL	Auteur(s) - Pays
1998	IxQuick - http://www.ixquick.com/	David Bodnick - US
	SurfWax - http://www.surfwax.com/	Tom Holt - Larry Burke - US
2000	Fazzle - http://www.fazzle.com/	Anvil Development US / Roumanie
	iBoogie - http://www.iboogie.com/	CyberTavern - US
	SmartEngine - http://www.smartengine.com/	Groupe de développeurs - US
	Vivisimo - http://vivisimo.com/	Carnegie Mellon University - US
2001	KartOO - http://www.kartoo.com/	Laurent et Nicolas Baleyrier -FR
	MapStan Search - http://search.mapstan.net/	MapStan - FR
	Turbo10 - http://turbo10.com/	Turbo10 - UK
	2loq - http://2loq.com/	Holomedia
2002	Killou - http://www.killou.com/	Brioude Internet -FR
2003	ZapMeta - http://www.zapmeta.com/	Groupe de développeurs - US
2004	Clusty - http://www.clusty.com/	Vivisimo - US
	DonBusca - http://www.donbusca.com/	4C Software - NL
	IceRocket - http://www.icerocket.com/	IceRocket - US
	Jux2 - http://www.jux2.com/ (Revenu en 2005)	US
	Magellan Metasearch - http://sourceforge.net/projects/magellan2/	Pierre Caron -EGE - FR
	Meceoo - http://www.meceoo.com/	Abondance - FR
2005	Ujiko - http://www.ujiko.com/	KartOO
	Gnosh - http://webtools.allegheeny.edu/gnosh/	Middlebury College - US
	TurboScout - http://www.turboscout.com/	William Chee - Singapour
	Metadone - http://www.metadone.net/	Robert Viseur - BE
	Polymeta - http://polymeta.com/	K-PROG Bt - Hongrie
Métamoteurs spécialisés		
2004	GoFish (multimedia) - http://www.qofish.com/	GoFish Technologies - US
2005	Kebberfegg (50 moteurs RSS) - http://www.researchbuzz.org/tools/kebbberfegg.pl	Tara Calishain - Research Buzz - US
2005	MammaHealth (santé) - http://www.mammahealth.com/	Mamma - CA
Métamoteurs en mode client		
1997	Copernic - http://www.copernic.com/	Agents Technologies (Copernic) - CA
1998	WebFerret - http://www.ferretsoft.com/	FerretSoft - US
1999	Strategic Finder - http://www.strategicfinder.com/	Digimind - FR
2003	Grokker - http://www.grokker.com/	Groxis - US

Avantages et inconvénients des métamoteurs

Si ces outils peuvent parfois sembler superflus pour des requêtes qui donnent déjà des résultats trop abondants avec les moteurs "traditionnels", ils s'avèrent en revanche très utiles pour rechercher "une aiguille dans une botte de foin". Une seule requête peut en effet être transmise en quelques secondes à plusieurs moteurs (de 3 à 1 000) ou à des sources spécialisées (presse, petites annonces, Web invisible...). Les métamoteurs sont également très intéressants pour le défrichage d'un sujet. Certains offrent la possibilité d'identifier les principales ressources et les principales thématiques des pages comprenant un mot-clé...

Pour les professionnels du Net, ils offrent également souvent d'excellents outils de comparaison du positionnement d'un site ou d'une page sur différents moteurs. Lancé dernièrement, Myriad Search (<http://www.myriadsearch.com/>) permet ainsi d'interroger simultanément Google, Yahoo!, MSN et AskJeeves et de comparer les résultats de chaque moteur (qui figurent dans des "onglets" distincts).

Selon Robert Beens, PDG d'Ixquick, ces outils seraient aussi moins sensibles au spam de certains référenceurs que les moteurs. D'après lui, *"De nos jours, beaucoup de sites Web sont optimisés par des référenceurs (SEO). Ce phénomène entraîne une 'pollution' toujours plus importante des résultats des moteurs avec des pages Web commerciales. Les moteurs comme Google sont pénalisés. Cela diminue considérablement leur pertinence. Ixquick a lui l'avantage de 'l'index pesé'. Pour obtenir un ranking élevé avec Ixquick, un site Web doit avoir un ranking élevé sur la moyenne de tous autres les moteurs de recherche. Ainsi, un site optimisé sur Google uniquement ne dupera pas Ixquick !"*.

Principaux inconvénients des métamoteurs : les fonctionnalités de recherche ne sont pas toujours aussi "avancées" que celles des moteurs de recherche "classiques". La syntaxe d'interrogation utilisée par les métamoteurs se réduit en effet le plus souvent au plus petit dénominateur commun des syntaxes respectives de chaque moteur.

Les métamoteurs ont également parfois souffert d'une distinction floue et quelque peu aléatoire entre liens payants et liens naturels dans leurs listes de résultats. Pour contrer ce problème, la plupart des acteurs majeurs ont aujourd'hui adopté des règles plus transparentes de présentation des résultats payants.

Mode de fonctionnement technique

On distingue trois grands types de métamoteurs : les métamoteurs online généralistes, les métamoteurs online spécialisés, et les métamoteurs offline ou hors ligne qui nécessitent l'installation d'un logiciel sur le poste client. Leur fonctionnement se décompose en quatre grandes étapes :

Interrogation, de façon simultanée, de plusieurs annuaires et moteurs

Les requêtes des utilisateurs sont traduites dans la syntaxe propre à chaque moteur. Les moteurs généralistes les plus connus sont en général interrogés, ainsi que des outils moins connus ou le Web invisible (dépêches d'agences...). Pour accroître le nombre de sources interrogées, certains métachercheurs invitent les moteurs et les sites spécialisés à se faire connaître. C'est notamment le cas de Turbo10 (<http://turbo10.com/>) qui propose depuis 2001 un formulaire de soumission sur son site pour un référencement payant.

Récupération des résultats

Pour collecter les résultats des moteurs, les métamoteurs ont trois possibilités :

- **Utiliser les API** proposées par les grands moteurs de recherche.

Google et Yahoo! proposent, par exemple, des API (interface de programmation ou "Application Programming Interface") qui permettent de développer des solutions d'interrogation directe de leurs services Web pour recevoir des résultats "condensés". Inconvénient : l'utilisation de ces API est limitée en termes de nombre de requêtes quotidiennes. De plus, les index interrogés par les APIs ne sont pas toujours les mêmes que ceux "attaqués" en réel par le moteur.

- **Signer d'accords** avec ces mêmes moteurs pour distribuer les résultats de recherche.

Le groupe Infospace a, par exemple, conclu des partenariats avec Google, Ask Jeeves et Yahoo! (en 2002) puis avec MSN (en 2005) pour proposer leurs résultats de recherche sur ses trois métamoteurs.

- **"Parser" les listes de résultats** des moteurs.

Utilisé entre autres par KartOO, le "parsing" (ou "parsing") consiste à utiliser des masques en XML pour "aspérer" les pages de résultats des moteurs et les "découper" automatiquement en unités (document, titre, lien...) afin de les restituer sur une même page. En clair, il s'agit d'une "capture" des informations présentes sur la page de résultats du moteur de recherche.

Un certain flou juridique entoure le droit de collecte des résultats des moteurs

Les métamoteurs ont-ils le droit de "parser" les résultats des moteurs, et par là-même un droit d'usage de ces données fournies par les outils de recherche ? Il est très difficile d'apporter une réponse "tranchée" à cette question.

Un droit de requête des métamoteurs en ligne semble le plus souvent accordé par les moteurs. Ce droit semble être accordé d'autant plus facilement que les métamoteurs sont souvent partenaires des moteurs sur les liens sponsorisés et qu'ils les aident à augmenter leur trafic. L'accès aux données ne pose apparemment aucun problème lorsque le nombre de requêtes est limité. Cependant, d'après nos informations, pour les métamoteurs qui génèrent un nombre élevé de requêtes (et présentent donc l'inconvénient d'utiliser une quantité importante de ressources), un blacklisting des adresses IP concernées est susceptible d'intervenir. Les administrateurs peuvent aussi se voir signifier dans ce cas qu'ils ne sont plus autorisés à effectuer des requêtes sur un moteur via leurs systèmes. Yahoo!, de son côté, coupe - momentanément - l'accès à une adresse IP qui effectuerait trop de requêtes dans un délai donné.

Pour les métamoteurs sur PC, la question du droit d'usage des données des moteurs ne se pose pas. En effet, ces outils ne sont pas identifiables par les moteurs car l'adresse IP qui effectue la requête est celle de l'utilisateur et pas celle du système.

Analyse et traitement des réponses

Le métamoteur récupère les résultats fournis par chacun de ces outils, supprime les doublons et les classe suivant des critères qu'il a pré-définis.

Classement et présentation des résultats

La restitution des résultats est l'un des principaux critères sur lesquels les métamoteurs cherchent à se différencier. Les résultats peuvent ainsi être restitués de manière aléatoire, par pertinence ou encore être "clusterisés", c'est-à-dire regroupés par thématiques. C'est le cas avec les outils Vivisimo ou iBoogie, par exemple. D'autres métamoteurs comme KartOO offrent aussi une présentation cartographique de ces résultats.

A noter que si nombre de métamoteurs de la première génération se contentaient de lister les résultats offerts par différents moteurs pour une requête, la plupart des métamoteurs proposés aujourd'hui disposent de modules permettant d'analyser les résultats pour les restituer dans une liste unique, synthétisée et organisée.

Modèles économiques

Les métamoteurs exploitent plusieurs sources de revenus :

- **l'affichage de liens sponsorisés d'autres moteurs dans leur liste de résultats.**
- la **commercialisation de liens sponsorisés via leur propre régie.** Le métamoteur canadien Mamma distribue notamment sa propre offre de liens sponsorisés "Mamma Classifieds".
- la **vente de licences aux entreprises.** Webfetch (du groupe Infospace) fournit, par exemple, sa technologie "private-label search" au quotidien britannique The Daily Mirror, qui l'utilise sur son site. Kartoo distribue également aux entreprises sa technologie de cartographie et différents moteurs de recherche "clés en main", notamment sur des intranets.
- les prestations d'**intégration de solutions de recherche dans les entreprises.** Vivisimo réalise 90% de son chiffre d'affaires en déployant des solutions chez ses clients (voir Lettre R&R #61). Son métamoteur en ligne lui procure surtout une visibilité sur le marché.

- les **ventes d'abonnements** pour les métamoteurs spécialisés. SurfWax propose un métamoteur payant de recherche sur les documents juridiques (LawKT - <http://www.lawkt.com/>).

Pour Pierre Caron (à l'origine de l'outil Open Source Magellan Metasearch), l'ajout de capacités d'analyse à ces outils pourrait aussi leur ouvrir des portes sur le marché des PME. "La démocratisation de ces outils, couplée aux avancées technologiques du domaine de l'analyse de l'information, permettra aux métamoteurs offline de toucher un plus large public (notamment les PME qui ont des budgets moins importants). La demande pour ces outils existe bien dans les PME, mais il existe encore peu d'offres dimensionnées pour une telle clientèle".

Les acteurs majeurs et leurs offres

La galaxie Infospace

Au fil des ans, le groupe Infospace (<http://www.infospaceinc.com/>) – très actif dans les solutions mobiles - s'est taillé la part du lion du marché de la méta-recherche. Via différentes acquisitions, il est aujourd'hui à la tête de quatre des outils les plus populaires (Dogpile, Webfetch, Webcrawler et Metacrawler). Le groupe a fusionné en 1999 avec Go2Net, mettant ainsi la main sur Dogpile et Metacrawler (qui appartenaient à Go2Net depuis 1998 et 1999 respectivement). En 2001, il a repris Excite et Webcrawler pour 10 M \$ avec pour objectif de transformer ces moteurs en métamoteurs.



Infospace a autrefois été vivement critiqué par les spécialistes qui lui reprochaient de fournir essentiellement des liens sponsorisés sur ses outils. Son DG, Richard Pelley, avouait d'ailleurs en 2002 que les outils du groupe "en étaient arrivés à un stade où ils étaient devenus des métamoteurs de liens sponsorisés" (voir <http://searchenginewatch.com/searchday/article.php/2160781>). Depuis plusieurs années, InfoSpace semble avoir rectifié le tir et s'est concentré sur l'amélioration de la qualité de ses résultats et la performance de ses outils en terme d'analyse des requêtes.

Dogpile (<http://www.dogpile.com/>)

Dogpile permet d'effectuer une recherche simultanée sur les quatre principaux moteurs : Google, Yahoo!, AskJeeves et MSN (depuis l'été 2005). La recherche peut porter sur les contenus Web, les images, l'audio, la vidéo, les news et les Pages Jaunes et Blanches.



Avantage de cet outil : il est possible d'afficher à côté de la liste de résultats agrégée les résultats trouvés par chaque moteur dans un encart à droite de l'écran.

Il est toutefois regrettable que les liens sponsorisés, qui sont mélangés à la liste de liens organiques, soient relativement difficiles à identifier. Seuls les mots "Sponsored by" et "Ads", ajoutés au petit texte indiquant la provenance du résultat, signalent un lien sponsorisé.

En Europe, le métamoteur Dogpile a adopté en 2004 un nouveau nom : Webfetch (<http://www.webfetch.com/>), le terme "dogpile" étant utilisé pour évoquer les excréments canins en Grande Bretagne !

Metacrawler (<http://www.metacrawler.com/>)

La première version de Metacrawler a été développée par deux universitaires de Washington en 1994 (Erik Selberg et Oren Etzioni). Metacrawler fait donc figure de pionnier sur le marché.



Tombé dans l'escarcelle d'Infospace, il permet lui aussi d'effectuer une recherche par défaut sur Google, Yahoo!, AskJeeves et MSN. L'utilisateur peut également, s'il le souhaite, ajouter les outils About et Looksmart à la liste des moteurs interrogés.

Les fonctionnalités de recherche sont assez similaires à celles de Dogpile et la liste de résultats présente elle aussi l'inconvénient de n'apporter qu'une distinction assez "limitée" entre liens sponsorisés et résultats naturels. Il est possible de personnaliser l'interface et les modalités d'affichage de résultats.

Webcrawler (<http://www.webcrawler.com/>)

Né, lui aussi, en 1994, WebCrawler était à l'origine un logiciel pour PC. Il est successivement passé entre les mains d'AOL (de 1995 à 1997), d'Excite (de 1997 à 2001) avant d'être racheté par Infospace en 2001. Les fonctionnalités de recherche de cet outil sont aujourd'hui quasiment les mêmes que celles de Metacrawler.



KartOO

"KartOO a été lancé en Avril 2001 par Nicolas Baleyrier mon cousin, et moi-même. Nous travaillions à l'époque dans le domaine de la formation et du site Web pour payer un développeur et l'hébergement de nos machines. Avec un peu de chance, nous avons réussi le lancement de l'outil et dès le premier jour, le trafic a explosé, avec nos serveurs d'ailleurs !", explique Laurent Baleyrier, co-fondateur du métamoteur. KartOO (<http://www.kartoo.com/>) est depuis devenu le métamoteur français le plus connu. Il interroge les principaux moteurs (14 au total). Son interface client utilise Flash 6 et son serveur est programmé en Java.



Particularité : il s'agit d'un métamoteur de recherche à interface graphique, cartographiant les résultats d'une requête en reliant entre eux les concepts proches (par utilisation des liens sémantiques). La carte ne propose que des URL. Les titres et la description s'affichent en passant la souris sur chaque URL.



Les résultats considérés comme les plus pertinents sont mis en évidence par le biais d'une icône rouge de taille supérieure. Pour chaque requête, KartOO propose également des expressions connexes au terme demandé (dans la rubrique "préciser la recherche" à gauche de l'écran).

"La prochaine version prévue pour le début de l'année permettra de chercher dans l'actualité, les images, les blogs, les forums et de regrouper les résultats sur une même carte", nous a confié Laurent Baleydié. Par ailleurs, selon lui, "Nous sommes capables aujourd'hui de présenter des cartographies d'une nouvelle génération : des cartes surfaciées pour avoir une vue d'ensemble des thématiques d'une recherche, mesurer facilement leur importance et zoomer sur celle qui vous intéresse ; des cartes centrées sur un résultat, avec autour des thématiques et d'autres liens qui vous permettent de passer d'une donnée à l'autre, de naviguer de proche en proche dans l'information. Ces technologies sont pour l'instant proposées en intranet, par dessus des solutions existantes. Mais j'espère que nous pourrons bientôt prouver leur efficacité sur KartOO.com et de nouveaux moteurs".

Mamma / Copernic

Après l'échec d'une première tentative fin 2004, le groupe canadien Mamma Media Solutions (www.mammamediasolutions.com) devrait finalement finaliser l'acquisition du québécois Copernic Technologies d'ici le 31 décembre 2005.

Mamma (<http://www.mamma.com/>)

Herman Tumurcuoglu (un doctorant de l'université d'Ottawa) a créé Mamma (The Mother of All Search Engines) en 1996. Ce métamoteur - qui permet d'effectuer une recherche sur le Web, les news, les images, les Pages Blanches et les Pages Jaunes - revendique plus de 350 000 requêtes par jour et près de deux millions d'utilisateurs par mois aux Etats-Unis.



Les outils Teoma, Findwhat, About, WiseNut, Open Directory, Looksmart, Gigablast et MSN sont interrogés, les résultats étant organisés par thématiques à côté de la liste de résultats.

The screenshot shows the Mamma search engine interface. At the top, there are navigation tabs for 'Web', 'News', 'Images', 'Yellow Pages', and 'White Pages'. Below these is a search bar containing the word 'Mortgage' and a 'Search' button. To the left of the search bar is the Mamma logo. Below the search bar, there are links for 'Help' and 'Power Search'. The main content area displays search results for 'Mortgage', including a 'Sponsored Links' section and a 'Search Results' section. A red box highlights a 'refine your search' sidebar on the right, which contains a list of links: 'Mortgage Lenders', 'Mortgage Calculator', 'Foreclosure Homes', 'Interest Rates', 'Mortgage Loans', 'Mortgage Rates', 'Abn Amro Mortgage', 'Current Mortgage Rates', 'Mortgage 101', and 'Mortgage Search'.

Depuis juin 2005, Mamma propose un outil sectoriel spécialisé dans le domaine de la santé : MammaHealth (<http://www.mammahealth.com/>). Les requêtes sont transmises à la fois sur des moteurs et des bases de données du Web invisible. D'autres métamoteurs verticaux devraient être lancés prochainement...

Copernic (<http://www.copernic.com/>)

Depuis sa création en 1997, Copernic développe des métamoteurs de recherche d'information en mode client populaires et performants. La version gratuite permet uniquement d'accéder aux fonctionnalités de méta-recherche sur 90 moteurs (contre plus de 1 000 moteurs pour les versions payantes).



Ixquick

Créé aux Etats-Unis en 1998 par David Bodnick et racheté en 2000 par la société néerlandaise Surfboard Holding, Ixquick (<http://ixquick.com/>) effectue des recherches dans 17 langues et sur 10 ou 16 moteurs suivant



les pays, parmi lesquels Yahoo, Gigablast, Ask Jeeves et l'Open Directory. Pour l'année 2004, ce métamoteur revendique 120 millions de requêtes.

Ixquick a fait le choix de limiter au maximum la présence des liens sponsorisés. Un maximum de trois liens sponsorisés est affiché en tête de la liste de résultats.

Un système de cotation par étoiles (Star System), particulièrement intéressant, décerne une étoile à chaque fois qu'un résultat apparaît parmi les dix premiers résultats d'un moteur. Ce système aide les utilisateurs à distinguer les résultats les plus populaires.

The screenshot shows the Ixquick search engine interface. At the top, there is a search bar with the text 'crédit' and a 'RECHERCHER' button. To the right of the search bar, there are several search engines listed: All the Web, Altavista, Ask Jeeves/Teoma, France, Go, Google, MSN, Mirago, Open Directory, and Overture. Below the search bar, there is a section titled 'Conseil : Ixquick possède un système d'étoiles unique >> Pour en savoir plus'. This section contains text about the search results and a list of sponsored results. The first result is 'crédit : COFIDIS, le crÃ©dit en ligne' with a star rating of 5 stars. The second result is '2,7 % TEG fixe jusqu'au 31 décembre 2005 sur Cetelem.fr' with a star rating of 5 stars. The third result is 'Groupe CrÃ©dit Agricole, accueil de la banque en ligne' with a star rating of 5 stars. The fourth result is '123crÃ©dit.com : crÃ©dit' with a star rating of 5 stars and a checkmark. The star ratings are highlighted with red boxes in the original image.

Enfin, cet outil offre un accès à un annuaire téléphonique international gratuit, et à un moteur de comparaison de prix (Ixquick Lowest Price).

SurfWax

SurfWax (<http://www.surfwax.com/>) a été créé par Tom Holt en 1998.

"Il s'agissait à l'origine d'un service 'FocusWord' dont le but était d'aider les utilisateurs à mieux 'articuler' leurs interrogations de recherche avant de lancer la recherche. En 1999 Tom s'est associé à Larry Burke dans le but d'élargir SurfWax et d'inclure la méta-recherche. La société proprement dite date de juin 2000", explique Tom Holt, co-fondateur de SurfWax.



Parmi les avantages de SurfWax : il offre de nombreuses possibilités de paramétrage aux utilisateurs. Il est, par exemple, possible de spécifier dès la première page le nombre de résultats à afficher (25, 50, 100 ou le maximum...) et les critères de tri à appliquer pour la restitution des résultats (par pertinence, ou par source...). Une inscription gratuite au service permet également de se configurer un ou plusieurs panel(s) de sources ("SearchSets") sur mesure, parmi environ 1 850 moteurs.

The screenshot shows the SurfWax search engine interface. At the top, there is a search bar with the text 'abondance'. Below the search bar, there are navigation links for 'About', 'Help', and 'Login'. The main content area is divided into two columns. The left column displays a list of search results, including titles like 'Gourmet Cheeses, Books, Kids, Cheese Parties and Chees' and 'R??7?7?encement et moteur de recherche avec Abondance'. The right column shows a detailed view of a selected result, titled 'SurfWax SiteSnaps™'. This view includes an 'Author Summary' section with text about 'Seme Chevalier in The French Alps' and a 'Key Points' section with a bullet point: 'these robust and generous cows give milk that is particularly rich in...'. A red circle highlights the SiteSnaps section.

La page de résultats est divisée en deux, avec possibilité d'affichage d'indications "SiteSnaps" sur chaque résultat sélectionné, comme les principaux thèmes qui y sont abordés ("Key Points")... Un encart en bas de la page de résultats, à gauche, donne également des statistiques sur la provenance des résultats.

Bref, voici un outil qui rivalise avec Vivisimo et est probablement l'un des meilleurs métamoteurs disponibles actuellement.

Vivisimo

Lancé en 2000 aux Etats-Unis, Vivisimo (<http://www.vivisimo.com>) est surtout connu pour ses fonctionnalités de catégorisation ou "clustering" qui regroupent les résultats dans des dossiers thématiques. Il est possible d'effectuer une recherche dans les clusters.



The screenshot shows the Vivisimo search engine interface. At the top, there is a search bar with the text 'abondance'. Below the search bar, there are navigation links for 'company', 'products', 'solutions', 'customers', 'demos', and 'press'. The main content area is divided into two columns. The left column displays a list of search results, including titles like 'abondance', 'Chapelle d'Abondance', 'Val d'Abondance', 'Outils, Moteurs', 'Holidays, Ski', 'Référencement', 'Cheese', 'Google', 'Search, Engine', 'Abonnes, Recherche avancée', and 'Chat, Transcript'. The right column shows a detailed view of a selected result, titled '10 of the top results retrieved for the query: abondance match the sub-query (Details)'. This view includes a list of search results, including titles like 'Abondance exclusivity: Google about to buy the GoGooroa Island', 'Traduction Val d'Abondance', 'Mozbot - Search Engine', 'Google indexe les plans de site (Abondance actu - lundi 6 juin 2005)', and 'Google KAI As Baby Name - Everyone Call him GOOGLE'. A red box highlights the 'Find in clusters' section on the left.

En Septembre 2004, la société a lancé un nouvel outil. Baptisé Clusty (<http://www.clusty.com/>), un métachercheur interrogeant plusieurs moteurs (Gigablast, Wisenut, Open Directory, Lycos, MSN) et permettant, comme son aîné, de catégoriser les résultats par thématique. Il offre, en plus, la possibilité de visualiser une partie de la page dans une fenêtre intégrée à la liste de résultats.

The screenshot shows the Clusty search engine interface. At the top, there are navigation tabs for Web, News, Images, Shopping, Wikipedia, Blogs, Jobs, and Customize. A search bar contains the word 'abondance' and a 'Cluster' button. Below the search bar, it displays 'Top 199 results of at least 443,041 retrieved for the query abondance (Details)'. A sidebar on the left lists various topics such as 'Chapelle (20)', 'Moteurs, Outils (14)', 'Val d'Abondance (16)', 'Search Engine (14)', 'Cheese (18)', 'Référencement (13)', 'Holidays (15)', 'Google (11)', 'L'hôtel-restaurant (10)', and 'Abonnés, Boutique en ligne Zone (8)'. The main content area shows a shopping result for 'Abondance Fermier AOC - 17.6 lbs. whole wheel, 17.60lb' priced at '\$254.94 from Gourmet Food Store'. Below this, there are advertisements for 'Be a Money Magnet now' and 'Abundance Course'. At the bottom, there is a banner for 'ABONDANCE' with a contest announcement: 'Gagnez 1 000 € et de nombreux lots au concours Mozbot ! Avec : atylla éditions du chêne'. The footer includes a search bar and various links like 'Rechercher sur Abondance', 'L'actu Quotidienne', and 'Sur le blog d'Abondance'.

Depuis fin août 2005, une nouvelle fonctionnalité de méta-recherche sur les moteurs de weblogs est disponible sur Clusty. Elle inclut une recherche sur Blogdigger, Daypop, Feedster, Technorati, Blogpulse et IceRocket.

Magellan MetaSearch

Pierre Caron (voir entretien dans cette lettre) a lancé en août 2004 un métamoteur en Open Source ayant vocation à être déployé sur un serveur Web ou sur un intranet d'entreprise. Baptisé Magellan Metasearch (<http://sourceforge.net/projects/magellan2>), cet outil permet d'interroger dix moteurs, parmi lesquels Google, AllTheWeb, Feedster et Gigablast. Il est possible de planifier les requêtes dans le temps afin de surveiller l'apparition de nouveaux résultats.

Jux2

Elu "meilleur métamoteur" à l'occasion des "Search Engine Awards 2004", Jux2 (<http://www.jux2.com/>) a beaucoup fait parler de lui dernièrement ! Courant octobre 2005, ce métamoteur a en effet été mis en vente sur eBay. Il a été racheté en quelques semaines pour 101 100 \$.



Disposant d'une fonctionnalité de recherche avancée, Jux2 restitue de façon graphique, et sous forme d'onglets, les résultats de Google, Yahoo et Ask Jeeves. Ses nouveaux propriétaires annoncent sur le blog de l'outil (<http://www.jux2.com/blog/>) qu'ils veulent l'améliorer et y ajouter un autre moteur. Affaire à suivre...

Enjeux et perspectives

Grandes tendances d'innovation

Compte tenu de la situation du marché, il faut s'attendre à un regain d'innovation de la part des métamoteurs dans les mois à venir. Selon Robert Viseur (administrateur du nouveau métamoteur Metadone) : "Pour un métamoteur Web classique comme Metadone, il faut aller plus loin en terme d'innovation pour attirer l'utilisateur en masse. D'abord, parce que la concurrence est puissante et bénéficie d'une très forte image. D'autre part, parce que avec la généralisation des APIs de moteurs et la démocratisation des solutions d'hébergement, la réalisation d'un métamoteur de base est à la portée de n'importe quel technicien Web un minimum qualifié. Aussi, les métamoteurs prolifèrent. Metadone ajoute au méta classique la suggestion de mots-clés et de sources. C'est assez utile au quotidien mais ce n'est pas assez pour attirer en nombre les utilisateurs ou même justifier une communication importante autour du projet. Cependant, les pistes d'innovation ne manquent pas (méta-recherche spécialisée, Ajax, analyse des sources, outils complémentaires de coopération communautaire, etc)".

Plusieurs tendances d'innovation émergent actuellement :

- Fonctionnalités de clusterisation.

La catégorisation automatique des résultats, développée également sur le moteur de recherche Exalead, est de plus en plus à la mode chez les métamoteurs. Vivisimo, 2loq.com ou KillerInfo se sont spécialisés dans cette technologie.

- Fonctionnalité d'analyse des résultats.

Qu'il s'agisse de cartographie (pour Kartoo) ou d'analyse statistique des résultats (pour SurfWax), l'analyse a posteriori des résultats de recherche est également en vogue.

- Fonctionnalités de personnalisation et de réseaux sociaux.

Dernièrement, des universitaires américains ont mis au point un nouveau métamoteur, Gnosh, qui permet de rechercher parmi les favoris des internautes inscrits à la communauté Gnosh et de se créer un espace personnalisé (avec favoris, historique des recherches et carnet d'adresse).

- Offre d'outils spécialisés.

On voit actuellement apparaître deux types d'offres spécialisées sur des formats ou des secteurs d'activité. Des offres pour des secteurs d'activité spécifiques, tout d'abord, avec par exemple le lancement par Mamma d'un outil Mamma Health destiné aux professionnels de la santé.

On assiste également au lancement de métamoteurs spécialisés sur des formats de fichiers. Les amateurs de DIVX ou autres fichiers vidéos découvriront d'ailleurs avec plaisir le métamoteur Torrent Typhoon (<http://www.torrenttyphoon.com/>), qui recherche des vidéos sur plusieurs sites spécialisés (TorrentBox, TorrentSpy, MiniNova...). Pour les fanatiques des fils RSS, Tara Calishain (responsable du site ResearchBuzz) propose quant à elle Kebberfegg (pour "Keyword Based Rss Feed Generator" - <http://www.researchbuzz.org/tools/kebbberfegg.pl>), un métamoteur innovant qui effectue des recherches dans environ 50 moteurs et restitue aux utilisateurs les résultats sous la forme d'un fil RSS.

Points de vue de plusieurs acteurs

Pensez vous que la consolidation du marché des moteurs de recherche aura un impact sur les métamoteurs ?

Laurent Baleyrier (KartOO)

S'il s'agit uniquement de rassembler et retrier les résultats, ça n'a plus aucune valeur aujourd'hui. Par contre, si on présente l'information de façon originale, si on aide l'internaute à préciser ses recherches, si on lui permet de personnaliser son outil, ces mêmes technologies sont directement applicables aux moteurs, donc ont de la valeur.

Deborah Kilpatrick (Mamma)

La force des méta-moteurs vient de leur capacité à interroger un grand nombre de moteurs de recherche différents et d'annuaires. Plus les sources sont variées, meilleur est le contenu à disposition du méta-moteur en charge de le hiérarchiser. La consolidation s'est stabilisée depuis l'année dernière et les 3 plus grands moteurs de recherche sur le marché actuellement utilisent des technologies de crawling différentes pour trouver et hiérarchiser les sites. On peut également noter une augmentation des moteurs locaux ainsi que des sites verticaux (Wikipedia, imdb.com etc.), ce qui fait de la place pour un nouveau type de moteur de recherche capable de fournir un type différent de résultats. Le type de contenu sur le Web évolue aussi, ce qui crée de nouvelles opportunités pour faire évoluer les moyens de recherche (ex : rss, blogs etc.).

Tom Holt (Surfwax)

Bien sûr. La domination des moteurs de recherche les plus importants a déjà eu un impact sur les outils de méta-recherche. La recherche à l'aide d'un outil unique proposée par les grands moteurs est de qualité mais ce n'est pas nécessairement la meilleure pour l'utilisateur. La méta-recherche va s'imposer à nouveau lorsque des domaines verticaux et spécifiques tels que la loi, le commerce ou la finance seront ciblés.

Comment envisagez vous l'avenir des outils de méta-recherche ?

Laurent Baleyrier (KartOO)

Il y a un créneau pour les entreprises qui se situent en aval de l'indexation : tout reste à faire en particulier au niveau des interfaces de recherche. Qui fait la différence aujourd'hui entre une liste de Google, MSN et Yahoo ? Ce sont les mêmes couleurs, les mêmes typo, la même présentation textuelle.

Pierre Caron (Magellan)

Je pense que vont apparaître des outils de milieu de gamme, entre les grandes plateformes "tout intégré" et les métamoteurs personnels type Copernic Agent. Leur coût sera plus adapté aux PME, même si leurs capacités d'analyse sémantique ne pourront rivaliser avec des produits beaucoup plus chers. Par ailleurs, les métamoteurs devront s'insérer dans les processus de knowledge management de l'entreprise, de manière à alimenter en informations pertinentes et sélectionnées les collaborateurs qui en ont besoin, et capitaliser les documents identifiés pour recherche ultérieure.

Jerome Pesenti (Vivisimo)

Il est improbable qu'une seule société réussisse jamais à dominer toute l'information. C'est pourquoi la méta-recherche jouera toujours un rôle en optimisant ce que fournissent différents éditeurs en effectuant des méta-recherches dans leurs contenus.

Conclusion : les métamoteurs ont encore de beaux jours devant eux !