

Recherche d'information : quelques trucs et astuces

[Retour au sommaire de la lettre](#)

Nous sommes tous des utilisateurs assidus des moteurs de recherche. Maëis les utilisons-nous tous à bon escient ? De plus, il existe de nombreux sites et outils qui peuvent vous faciliter la vie au quotidien. Marianne Dabbadie, experte du domaine, nous livre ici quelques-un de ses secrets...

Nous vous présentons dans cet article quelques savoir-faire – ou savoir surfer - ainsi que quelques "trucs" qui font appel à votre créativité ou bien à votre capacité de contournement pour trouver l'information utile sur le web. A découvrir et à suivre...

Les archives du Web

Parmi les originalités fort utiles du web on trouve **Web Archive** (<http://web.archive.org/>) C'est une « way back machine » ou machine à remonter le temps, qui peut vous permettre d'accéder aux archives d'un site jusqu'à dix ans plus tôt s'il y a lieu. A noter que Web Archive est également accessible au travers du moteur de recherche Mozbot (<http://www.mozbot.fr/>) en cliquant sur le lien "historique" qui figure sous chaque résultat. Une telle fonctionnalité peut s'avérer fort utile lorsqu'on cherche à retrouver une information qui n'est plus disponible sur la Toile, mais dont on a gardé le souvenir.

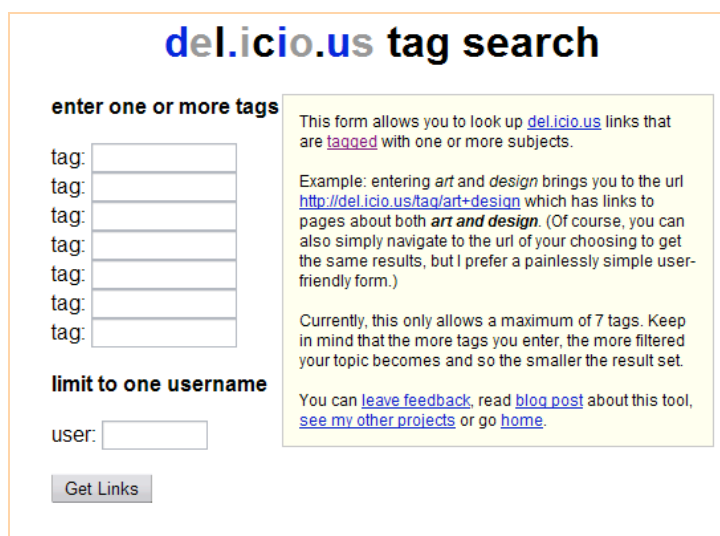
Ainsi, l'exemple ci-dessous, retrouvé sur webarchive.org est une copie d'écran d'une page du site du CNRS en 1997. A cette époque, le CNRS proposait une classification – indexation des principaux sites web dans le monde, à une époque où Google n'existait pas encore.



La recherche par tags

La classification de l'information de façon structurée sous forme de **tags** est également une façon plus souple de trouver l'information. Le *social bookmarking* a fait naître des moteurs de folksonomie comme Yoono (<http://www.yoono.com/>), Wink (<http://www.wink.com/>) ou Wikio (<http://www.wikio.com/>) pour n'en citer que quelques-uns. La classification proposée par ces moteurs a pour avantage d'être issue de chercheurs spécialistes et/ou passionnés puisque ce sont les internautes eux-mêmes qui partagent leurs sources. Nous renvoyons pour plus de détails à notre article paru dans cette lettre au mois de février 2006 ainsi qu'aux nombreuses informations disponibles sur le site Abondance.

Pour exploiter au mieux la puissance des tags, le site de Jarrod Trainque propose une recherche en syntaxe avancée à partir d'une combinaison de tags sur le moteur de folksonomie Del.icio.us (réservée aux anglophones). On peut tester ce type de recherche à l'adresse suivante : http://trainque.com/code/delicious_lookup.php



Ainsi la requête "art" combinée à "design" renvoie à l'url <http://del.icio.us/tag/art+design> qui dirige l'internaute vers les pages parlant à la fois d'un sujet et de l'autre. Pour bien faire la différence avec une recherche avancée sur un moteur classique comme Google, il faut bien avoir à l'esprit que les tags renvoient vers le thème d'une page et non simplement à la présence d'un terme ou chaîne de caractères dans le texte d'une page. Ainsi la recherche effectuée à partir de tags est-elle beaucoup plus pertinente qu'une recherche effectuée sur les mots ou chaînes de caractères. C'est pourquoi cette combinaison de tags permet d'obtenir des résultats d'autant plus intéressants qu'ils font réellement sens.

Le clustering et la cartographie

Nous ne présenterons plus les moteurs qui clusterisent l'information et en proposent une catégorisation sous forme de menu thématique sur la page de résultats. Une telle présentation de l'information déclinée sous forme de thèmes et sous-thèmes, permet d'affiner une recherche de façon non sollicitée et en fonction des contenus. Parmi ces moteurs on trouve notamment Vivisimo (<http://www.vivisimo.com/>), Clusty (<http://www.clusty.com/>), le Hongrois Polymeta (<http://www.polymeta.com/>) et bien sûr le Français Exalead (<http://www.exalead.com/>) qui allie clusterisation et analyse de l'information en langue naturelle. Nous renvoyons également le lecteur à notre article paru sur ce thème dans la lettre R&R de mars 2006.

Parmi les moteurs qui clusterisent, on trouve aussi le français Kartoo (<http://www.kartoo.com/>), métamoteur qui a pour originalité tout à la fois de cartographier l'information et de la clusteriser. La cartographie d'information propose une vision non linéaire des pages de résultats, rendant ainsi plus évidents les liens entre les différents thèmes de la page de résultats. D'une certaine façon cependant, on peut dire que l'accès au sens, par des représentations linéaires ou graphiques n'est pas la problématique la plus importante en soi, au regard de la recherche d'information. L'important, ce sont les processus mentaux que sous-tendent les différents modes d'accès à l'information et la façon dont les modes de représentations véhiculent le message auquel on veut avoir accès. Et de ce point de vue, la représentation graphique est intéressante, car elle fait appel à des facultés cognitives que les représentations purement linéaires ont laissé de côté.

La recherche géographique

Parmi les fonctionnalités originales qui permettent de trouver l'information sur le net on trouve aussi la recherche par entité géographique. Ainsi le moteur de recherche Mirago

(<http://www.mirago.fr/>) vous permet-il de rechercher l'information par région, en cliquant sur une carte de l'hexagone.

La recherche avancée et plus

Pour en savoir encore plus sur les moteurs et leurs particularités nous renvoyons le lecteur vers le site (en anglais) de Tara Calishain : <http://www.researchbuzz.org/>

Pour les technos fans qui s'intéressent à Google nous renvoyons le lecteur vers la deuxième édition du livre de Tara Calishain « Google à 200% » (édition 2005 – O Reilly France Editeur) ; disponible à l'achat en ligne sur le site Amazon.fr. Tara Calishain est l'auteur d'une douzaine d'ouvrages sur Internet.

Ce livre, destiné initialement à un public de documentalistes, étudiants et webmestres, parle des nouveaux services de Google mais également de toutes les fonctionnalités offertes par son interface de programmation Google API, qui utilise des langages tels que XML, PHP, Python, Perl, C# ou .NET.

Les recherches programmables sont basées sur des *hacks*, clés permettant de gagner du temps lors d'une recherche. Des scripts sont proposés dans l'ouvrage. Les trucs et astuces présentés dans ce livre permettent également d'utiliser les possibilités des services proposés par Google pour la recherche d'images, les *newsgroups* ou les blogs. Les programmeurs les plus avertis pourront également adapter ces scripts aux nouvelles fonctionnalités proposées par ce moteur, en constante évolution.

Un outil très riche de recherche avancée conseillé par MSN Search permet aux technos mordus d'affiner la recherche en manipulant de façon très fine les options de recherche avancée de MSN. Un descriptif complet est disponible à l'adresse suivante : http://search.msn.fr/docs/help.aspx?t=SEARCH_REF_AdvSrchOperators.htm

Cet outil permet notamment d'exclure des termes d'une recherche, de la limiter à un domaine précis (ou plusieurs) ainsi que de déterminer une langue ou une zone géographique précise sur laquelle est effectuée la recherche. Il permet également, par exemple, de trouver tous les sites hébergés sous une même adresse IP et bien plus de recherches encore.

La recherche orientée résultats

Antonio Fontes sur son blog supercoach évoque la recherche orientée résultats. <http://supercoach.over-blog.com/article-964137.html>

L'idée est simple : c'est une question de posture mentale vis-à-vis du sens. Posez-vous une question au moteur de recherche ou cherchez-vous une réponse ?

Plaçons-nous dans le second cas de figure. Imaginez alors quels mots pourraient contenir les réponses que vous cherchez. : ainsi, si vous voulez chercher une information sur la date de naissance de Zidane ne tapez pas : "Zidane date naissance" (sans guillemets) mais plutôt avec les guillemets "zidane est né le" et vous trouverez tous les sites qui proposent cette portion de phrase en *exact match*.

Si vous cherchez une location de vacances dans les Alpes-Maritimes pour 5 personnes, ne formulez pas votre requête ainsi, mais imaginez plutôt le résultat : "maison 5 pièces Cannes vacances prix semaine"...

Nous avons essayé ces deux requêtes dans Google : dans le premier cas le moteur a certes renvoyé des possibilités de location mais allant du studio au 5 pièces et il ne s'agissait pas de maisons.

Dans le second cas les réponses étaient toutes pertinentes comme en témoigne la copie d'écran ci-dessous et indiquaient même un ordre de prix :

The screenshot shows a Google search interface with the following elements:

- Search bar: "maison 5 pièces Cannes vacances prix sema" with a "Rechercher" button and links for "Recherche avancée" and "Préférences".
- Location filter: "Rechercher dans : Web Pages francophones Pages : France".
- Section header: "Web" with "Résultats 1 - 1" on the right.
- Search results:
 - Vacances Discount** (www.promovacances.com) + de 3000 offres en promotion ! Séjours, Week-Ends, Circuits, Vols.
 - Vacances Cannes** (www.maeva.com/cannes) Offres spéciales web : Locations d'été à prix mini avec Maeva !
 - LOCATION CANNES Maison. Location Vacances Cannes. 5 personnes ...**
Maison 5 personnes Cannes. Vacances d'été au calme sur la Côte d'Azur. ... Prix semaine hors saison : (Pas de location hors saison) ...
www.vacances.com/annonce.php?id=46576 - 37k - En cache - Pages similaires
 - LOCATION GRASSE Maison. Location Vacances Grasse. 6 personnes ...**
Location d'une maison individuelle 5 pièces avec piscine pour 6 ... Prix semaine hors saison : 1655.00 € (28.04.2007 - 25.05.2007 22.09.2007 - 26.10.2007) ...
www.vacances.com/annonce.php?id=53373 - 37k - En cache - Pages similaires
 - Cannes > Chalet / Maison > 1 - 20**
Prix par logement et par semaine (Frais de dossier inclus) ... Cannes Résidence Pierre et Vacances Cannes Villa Francia 2 Pièces 4/5 Pers. ...
www.location-vacances-express.com/comparer-cannes-chalet-maison,0,31,9,0.html - 176k - En cache - Pages similaires
 - Cannes > Chalet / Maison > 21 - 40**
Cannes Résidence Pierre et Vacances Cannes Beach 2/3 pièces 5/6 pers. ... Logements, Prix/Semaine. 2 pièces cab 4/5 pers. 745 € ...
www.location-vacances-express.com/comparer-cannes-chalet-maison,0,31,9,20.html - 167k - En cache - Pages similaires
 - Revue de presse**
4 semaines disponibles à partir du 09/04/2007 - Prix semaine : 420 Euro [... Maison tout

Pour trouver l'information utile, la liste des trucs et astuces pourrait être illimitée. Nous vous avons livré ici pêle-mêle quelques technologies, façons subtiles d'exploiter ces technos, postures mentales, modes de représentation qui sont à notre avis autant de points de départ utiles pour aboutir à des résultats toujours plus précis. Nous vous en livrerons d'autres dans ces colonnes, au fil du temps, car avec les moteurs de recherche, la créativité conjuguée des éditeurs comme des utilisateurs va bon train et l'œuvre n'est jamais achevée.

Marianne Dabbadie
Directrice Innovation i-KM
Laboratoire GERIICO – EA 1060