


Collecter de l'information sur Internet

<ul style="list-style-type: none"> > Rechercher des informations > Veiller des thématiques et des sites 	<ul style="list-style-type: none"> > Rechercher des informations <ul style="list-style-type: none"> - Penser sa stratégie de recherche - Les annuaires - Les moteurs de recherche - Les méta-moteurs de recherche > Veiller des thématiques et des sites <ul style="list-style-type: none"> - Penser sa stratégie de veille - Les listes de diffusion - Les outils de surveillance 	
---	---	--

OUTILS & CONSEILS

" Une question mal cernée, des mots-clés imprécis - ou au contraire trop précis - un outil de recherche choisi sans réflexion préalable... et le netsurfer est noyé sous les réponses, ou n'obtient au contraire qu'une maigre liste de résultats. Pour tirer partie de l'Internet, il est indispensable d'avoir au préalable acquis certaines connaissances sur ce réseau (...) ".

Béatrice FOENIX-RIOU

Guide de recherche sur Internet, Outils et méthodes, ed ADBS, 2002.

Dans cette rubrique, nous nous efforcerons de présenter les principaux outils de recherche, leur utilité et les modalités d'utilisation, ainsi que quelques conseils pour élaborer sa stratégie de recherche.

Penser sa stratégie de recherche

CONSEIL



Objectif

Pour naviguer sur Internet, deux préalables sont nécessaires avant même de commencer la recherche :

- > il faut avoir bien compris le mode de fonctionnement des outils de recherche pour pouvoir choisir l'outil le mieux adapté à la question,
- > il faut avoir pensé sa recherche et prendre le temps de définir sa stratégie de recherche pour ne pas fausser dès le départ la pertinence des résultats proposés par l'outil.

Méthodologie - Démarche

1 – Délimiter le champ de la recherche

Se poser la question :

quels sont les types d'informations que je recherche ?

Cette première étape permet d'identifier plusieurs types d'informations et contraint dans tous les cas l'internaute à préciser sa question, ce qui est toujours utile ; chaque piste doit ensuite être soigneusement explorée.

2 – Définir les types de sources susceptibles de répondre

La définition des types d'informations recherchés met en évidence les sources susceptibles de les offrir : site des acteurs sur lesquels l'internaute veut obtenir des informations, sites de presse, sites officiels, forums, listes de diffusion, ...

3 – Choisir les outils de recherche à interroger

- > Lorsqu'on recherche des pages, les moteurs ou méta-moteurs sont mieux adaptés
- > Lorsqu'il s'agit de rechercher un site, l'annuaire peut permettre d'effectuer une première recherche.

4 – Trouver les bons mots-clés

- > Etablir une liste de mots ou expressions descriptives de l'information recherchée
- > Utiliser des mots-clés primaires ou génériques qui définissent le domaine de recherche si on utilise un annuaire. Ces mots clés permettent d'identifier les catégories qui couvrent les différents aspects du problème.
- > Avec un moteur de recherche utiliser des mots clés plus précis (utiliser de préférence plusieurs mots clés qui affinent la recherche)

Notes

Méthodologie - Démarche (suite)

- > Il existe des générateurs de mots clés, comme le site Abondance d'Olivier Andrieu (www.audit.abondance.com rubrique " Outils / Audit de site " Outiref) : l'utilisateur indique un, deux mots clés se rapportant à sa requête et il obtient alors les mots clés les plus souvent présents dans les pages web qui contiennent les mots demandés.*
- > D'autres outils proposent avec leur liste de résultats, d'autres mots ou expressions connexes (le moteur de recherche Exalead (www.exalead.com) ou KillerInfo (www.killerinfo.com))

Zoom sur : les opérateurs de recherche

L'objectif

Les annuaires et les moteurs de recherche utilisent des langages d'interrogation qui peuvent être différents.

La stratégie de recherche doit respecter certaines règles :

> les guillemets

pour rechercher une expression ou un nom propre, voire une citation, encadrer l'expression recherchée de guillemets (sans espace entre le symbole et le mot) pour lancer la recherche sur l'expression entière et non sur la somme des mots

> le +

pour de nombreuses recherches, il est nécessaire d'utiliser plusieurs mots clés pour préciser sa question et réduire le nombre de réponses non pertinentes. Mais une requête lancée avec deux termes n'est pas interprétée de la même façon par tous les outils de recherche.

Certains considèrent que la présence de deux mots clés signifie que les deux doivent être dans la page (a ET b), et d'autres considèrent que c'est l'un OU l'autre qui est recherché (a OU b).

Pour résoudre ce problème, il suffit d'utiliser systématiquement le symbole +. Ce symbole doit être placé immédiatement devant le mot, sans espace entre le signe et le mot, pour demander la présence impérative du terme dans les pages sélectionnées. (mettre un + devant chaque mot recherché)

Les annuaires

OUTIL

matériel nécessaire :
connexion à Internet



Objectif

Un annuaire est un répertoire de sites, organisé selon un classement thématique ou géographique fait par un éditeur. Chaque site est commenté par l'éditeur et classé par catégorie. L'annuaire est surtout utile pour explorer un sujet général, trouver des sites ressources dans un domaine ou encore trouver des sites similaires sur un même thème.

Méthodologie - Démarche

Deux modes de recherche sont possibles avec un annuaire :

- > par mot clé
- > par arborescence

> Par mot clé, il s'agit de taper dans l'annuaire, le mot clé sur lequel l'utilisateur souhaite faire des recherches. L'annuaire soumet alors la liste des sites répertoriant le mot recherché.

> Par arborescence, l'annuaire propose une liste de thématiques et de sous thématiques. La recherche s'effectue alors en avançant de sous-thèmes en sous-thèmes jusqu'à obtenir une liste de sites correspondant au sujet.

Exemples d'annuaires :

> Aol

www.recherche.aol.fr

> Open Directory

www.dmoz.fr

> Yahoo! Guide Web

<http://fr.yahoo.com>

Notes

*** Niveau de difficulté



Quantité de Matériel requis

Exemple

> Pour effectuer une recherche sur l'expression " *développement durable* " et savoir quels sont les sites référencés dans le domaine.

Par mot clé :

1 – Taper " *développement durable* " dans la barre de recherche (entre guillemets pour que l'annuaire fasse la recherche sur l'expression et non sur les deux mots pris indépendamment l'un de l'autre)

2 – "Aol" propose en tête de page, les expressions associées à " *développement durable* " (par exemple : sommet mondial du développement durable, développement territorial...). Pour affiner la recherche cliquer sur l'expression la plus proche de ce qui est recherché et l'annuaire effectuera un nouvel état des lieux des sites qu'il a répertorié.

Le reste de la page présente les sites les plus pertinents en matière de développement durable. Il est alors possible de choisir :

- > un aperçu du site + nom du site + description en quelques mots du site + adresse internet
- > uniquement la partie texte, sans l'aperçu du site
- > uniquement l'aperçu du site et l'adresse internet

3 – Sous chaque réponse est indiqué, à coté d'un pictogramme représentant un dossier, le chemin à parcourir dans l'arborescence pour atteindre le site proposé.

Notes

Par arborescence :

1 – Choisir parmi la liste des thématiques proposée, celle susceptible de contenir l'information recherchée.

2 – De clic en clic, pour " *développement durable* ", le dossier " société " puis " environnement " puis " développement durable " conduit à la liste des sites traitant du développement durable.



www.recherche.aol.fr



www.dmoz.fr



<http://fr.yahoo.com>

Les moteurs de recherche

OUTIL

matériel nécessaire :
connexion à Internet



Objectif

Un moteur de recherche est un outil de recherche de pages web. A la différence de l'annuaire, c'est un répertoire de pages et non de sites. Le moteur de recherche balaye le réseau Internet sur un mot ou une expression et affiche les pages web, les fichiers (Excel, Word, pdf, PowerPoint...) traitant du sujet demandé. Il peut également répertorier des groupes de discussion, des images, des actualités...

Méthodologie - Démarche

Pour faire une recherche :

- > Saisir dans la barre de recherche du moteur, le mot (ou expression) pour lequel (laquelle) on souhaite obtenir des informations.
- > Deux modes de recherche sont possibles : normal ou avancé. Le mode avancé permet de paramétrer plus finement les critères de recherche : langue, format de fichier, date, etc.

A savoir :

- > Tous les moteurs de recherche n'utilisent pas les mêmes critères de référencement des réponses.
- > Certains traitent les réponses en fonction d'un critère de notoriété (calculée en fonction du nombre de liens qui pointent vers la page pour Google par exemple) alors que d'autres privilégient un critère de pertinence (la pertinence est évaluée en fonction du nombre de fois où le mot recherché est cité dans une page. Yahoo! par exemple utilise ce principe).

Exemples de moteurs de recherche :

- > [Google](http://www.google.fr)
www.google.fr
- > [Exalead](http://www.exalead.com)
www.exalead.com
- > [HotBot](http://www.hotbot.com)
www.hotbot.com

Notes

*** Niveau de difficulté



Quantité de matériel requis

Exemple

> Pour effectuer une recherche sur l'expression " *système d'information géographique* " et obtenir des informations de différentes natures sur ce sujet.

- 1 – Prévoir de faire la requête sur plusieurs moteurs de recherche pour comparer les résultats et compléter les informations obtenues.
- 2 – Choisir un premier moteur de recherche, Google par exemple, et saisir dans la barre de recherche centrale, l'expression " *système d'information géographique* ". Veiller à toujours mettre les expressions entre guillemets pour ne pas lancer une recherche sur la somme des mots de l'expression, mais bien sur l'expression elle même.
- 3 – En plus des pages web, Google propose, à travers ses onglets " Images ", " Groupes ", " Annuaire " et " Actualités ", d'autres types de réponses.
- 4 – Pour affiner la requête, il est possible d'ajouter des mots dans la barre de recherche : par exemple, pour connaître les logiciels de système d'information géographique, taper l'expression "*système d'information géographique*" et ajouter à la suite, hors des guillemets, *logiciels*.

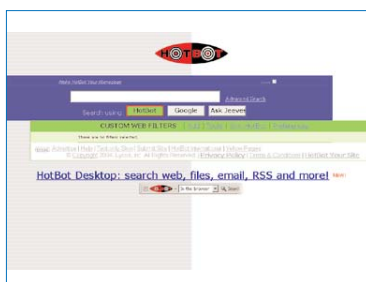
Notes



www.google.fr



www.exalead.com



www.hotbot.com

Les méta-moteurs de recherche

OUTIL

matériel nécessaire :
connexion à Internet



Objectif

Les méta-moteurs fonctionnent en partie comme des moteurs de recherche, mais ils interrogent simultanément plusieurs moteurs. Ils envoient une requête à plusieurs moteurs, sélectionnent les résultats les plus pertinents de chaque outil, les analysent et les classent. Ils vérifient également la validité des liens et dédoublonnent les réponses.

Méthodologie - Démarche

Pourquoi utiliser un méta-moteur ?

- > Les méta-moteurs sont surtout préconisés pour effectuer une recherche sur un thème large. Comme ils interrogent un certain nombre de moteurs de recherche, ils scrutent un plus grand nombre de sites et garantissent alors de trouver soit :
 - des informations sur le sujet à traiter,
 - ou au moins de donner des pistes pour affiner la requête et offrent alors plus de chances d'être pertinent dans la suite des recherches.

Comment fonctionnent-ils ?

- > Le principe de fonctionnement d'un méta-moteur est le même que celui d'un moteur de recherche.
- > Saisir la requête dans la barre de recherche de l'outil et celui-ci affiche les pages qui contiennent des informations sur le sujet.
- > Certains proposent d'effectuer la recherche sur une liste de site prédéfinie (Meceoo par exemple), d'autres représentent les réponses possibles sous forme de carte (Kartoo) ou encore suggèrent des expressions connexes pour poursuivre les recherches.

Exemples de méta-moteurs de recherche

> **KartOO**
www.kartoo.com

> **Meceoo**
www.meceoo.fr

> **Killer Info**
www.killerinfo.com

Notes

*** Niveau de difficulté



Quantité de matériel requis

Exemple

> Pour effectuer une recherche sur l'expression " *aménagement des territoires* " et identifier aussi bien les acteurs que les concepts :

1 – Prévoir de faire la requête sur plusieurs méta-moteurs de recherche pour comparer les résultats et compléter les informations obtenues.

2 – Choisir un premier méta-moteur de recherche, Kartoo par exemple, et saisir dans la barre de recherche l'expression " *aménagement des territoires* ". Veiller à toujours mettre les expressions entre guillemets pour ne pas lancer une recherche sur la somme des mots de l'expression, mais bien sur l'expression en elle même.

3 – Kartoo présente les sites et les pages traitant de cette thématique sous forme de cartes. Chaque réponse est représentée par une page plus ou moins importante en fonction de sa pertinence et des mots clés lui sont associés.

4 – Kartoo propose également des suggestions de recherche sous forme de mots clés à ajouter à la requête.

Notes



www.kartoo.com



www.mecceo.fr



www.killerinfo.com

Penser sa stratégie de veille


CONSEIL

Objectif

Même si la veille représente un atout stratégique pour les entreprises et les territoires, il n'est pas toujours facile de mobiliser les moyens financiers et le temps nécessaire pour mettre en place un système de veille. C'est pourquoi il est préférable de bien réfléchir à ses besoins, à ses attentes, avant de se lancer dans un programme de veille trop coûteux ou peu adapté.

Méthodologie - Démarche

1 – Cibler les besoins

La veille doit correspondre à un besoin qu'il faut définir au préalable, notamment, en identifiant le cœur de métier de l'entreprise et par là même ses besoins en information sur les deux ans à venir, selon des ambitions réalistes. Il faut avoir un but précis à court terme permettant de canaliser les personnes et les efforts sur un but commun.

2 – Identifier les acteurs

- > Lorsque les besoins ont été correctement ciblés, l'entreprise doit identifier les salariés qui sont en contact avec les différentes sources d'information possibles. Ces personnes sont autant de relais d'informations organisés en réseau.
- > Le facteur clé de succès de la veille tient dans la motivation de ces personnes, qu'il faut entretenir de diverses manières.

3 – Récolter l'information

- > L'utilisation d'un outil n'est pas une obligation, surtout si l'effectif est concentré sur un même site.
- " La machine à café reste un lieu d'échange non négligeable pour la veille et l'e-mail un outil quotidien pour faire remonter l'information " Yann, Guilain, directeur consulting chez Digimind.
- > Une surveillance de l'actualité, des publications, des communiqués de presse des " concurrents ", une présence sur les salons, sont des actes de veille.
- > Il faut cibler les thèmes, les domaines ou les types d'information à chercher et définir le rôle de chacun pour éviter de récolter du " bruit ", de l'information inutile ou redondante.

Notes

Zoom sur : Définition de la veille

" La veille est un processus informationnel volontariste à travers lequel l'on recherche des informations à caractères anticipatif concernant l'évolution d'une activité, de connaissances ou d'un environnement particulier."

www.cyber-strategie.com/

Méthodologie - Démarche (suite)

4 – Analyser les données

Les données récoltées sont ensuite analysées par rapport aux besoins de départ. Ce travail est synthétisé dans un document dont le but est de faciliter la prise de décision : tableau de bord, chiffres clés... et préconisations pour le décideur.

5 – Diffuser pour agir

A ce stade, la veille ne sert à rien si le résultat n'est pas diffusé auprès des collaborateurs de l'entreprise qui pourront agir en conséquence. Il faut donc que l'information redescende rapidement vers les personnes intéressées, ce qui demande de faire un travail préalable d'identification des domaines d'intervention de chacun.

Notes

Zoom sur : l'avis de l'ADBS

pour
l'Association des
professionnels
de l'information
et de la
documentation :

“ Dire que l'on " fait de la veille sur Internet " est un abus de langage.

En fait on utilise Internet comme un outil de surveillance des entreprises, des marchés, des technologies, des évolutions de la société...

L'apport d'Internet par rapport à une démarche de veille :

- une information ouverte, disponible à tout moment,
- une information régulièrement actualisée,
- des informations multisources, multidisciplinaires,
- une information numérisée, pouvant être triée et exploitée rapidement.

Mais il ne faut pas oublier les aspects négatifs :

- risque de désinformation : qualité de la source, de l'information..
- risque de se noyer dans l'information
- une information parfois difficilement accessible (barrière de la langue, services payants...)
- une information en perpétuelle évolution et donc instable
- ... "

Source : ADBS, Internet : recherche avancée et outils de veille, supports de cours communs, octobre 2002

Les listes de diffusion

OUTIL

matériel nécessaire :
connexion à Internet



Objectif

Les listes de diffusion permettent à un groupe de personnes de communiquer sur un thème par l'intermédiaire du courrier électronique. Chacun écrit à l'adresse électronique de la liste, un automate se charge de distribuer les messages à l'ensemble des participants.

Certaines listes n'ont pas pour but des discussions, mais juste l'émission d'une information (bulletin, magazine, lettre d'information, revue de presse...). Dans ce cas l'abonné ne peut pas s'exprimer sur la liste.

Méthodologie - Démarche

1 – Tout d'abord, il faut identifier la liste de diffusion la plus pertinente pour l'utilisateur, traitant de la problématique qui l'intéresse.

Pour cela, des annuaires de listes proposent une description de chaque liste qu'ils répertorient ainsi que l'adresse à laquelle s'inscrire.

2 – Ensuite, pour s'abonner à une liste, il faut généralement envoyer un courriel à l'adresse électronique de la liste.

3 – Toutes les informations échangées dans le cadre de cette liste de diffusion sont alors envoyées sur la boîte aux lettres électronique de l'utilisateur.

Exemples d'annuaires de listes :

> [Francopholistes](http://www.francopholistes.com)
www.francopholistes.com

> [Yahoo! Groupes](http://fr.groups.yahoo.com)
http://fr.groups.yahoo.com

> [Google Groupes](http://groups.google.fr)
http://groups.google.fr

Notes

*** Niveau de difficulté



Quantité de matériel requis

Exemple

> Pour s'inscrire à une liste de diffusion sur les problématiques de "veille" :

- 1 – Prévoir de faire une requête sur plusieurs annuaires de liste
- 2 – Aller sur l'annuaire de listes www.francopholistes.com
- 3 – Saisir dans la barre de recherche le sujet recherché, ici, " veille "
- 4 – Francopholistes indique qu'il existe 41 réponses possibles, avec un degré de pertinence qualifié par un système d'étoiles.
Parmi ces réponses, sélectionner la liste qui présente le plus d'intérêt et Francopholistes indique l'adresse mail à laquelle envoyer un courriel pour être inscrit.

Notes

Exemples de listes :

> Liste " Prospective "

prospective-abonnement@yahogroupes.fr
envoyer un courriel vide à cette adresse

> Liste pour les utilisateurs d' " Arcview "
(Logiciel de Système d'Information Géographique) :

SIG-Arcview-Avenue-subscribe@yahogroupes.fr
envoyer un courriel vide à cette adresse

> Liste " Veille "

veille-abonnement@yahogroupes.fr
envoyer un courriel vide à cette adresse

Les outils de surveillance

OUTIL

matériel nécessaire :
connexion à Internet



Objectif

Ces outils permettent de surveiller soit une page précise, soit un site entier. Les résultats sont généralement envoyés par courriel selon une périodicité que l'utilisateur peut déterminer. La majorité des outils de surveillance de sites ou de pages web permettent de définir des critères d'alertes. Il s'agit de définir des termes ou des groupes de termes qui intéressent l'utilisateur : l'alerte ne sera envoyée que si ces termes apparaissent dans la page ou le site surveillé.

Méthodologie - Démarche

Deux méthodes sont possibles, en fonction des besoins :

- > la surveillance de site pour ne pas manquer une nouvelle mise à jour (promotion d'un nouveau produit sur le site d'un concurrent, publication d'une nouvelle étude par un laboratoire de recherche partenaire...)
- > la surveillance des parutions sur un sujet et l'information de toutes les nouveautés dans ce domaine.

Dans chaque cas, le paramétrage des outils est différent :

> Pour effectuer une surveillance de site :

La surveillance de site est préconisée dans le cadre de sites qui bougent peu, ou pas régulièrement. Ces outils simplifient la veille en l'automatisant.

- 1 – Sélectionner les sites à surveiller et leur adresse Internet
- 2 – Créer un compte utilisateur dans l'outil de surveillance avec le courriel sur lequel la page à surveiller sera envoyée
- 3 – Indiquer à l'automate la périodicité de l'envoi des alertes
- 4 – Indiquer à l'outil la forme de la réponse : envoi par courriel de la page qui a changé, envoi par courriel de la liste des changements
- 5 – Saisir la liste des sites à surveiller. Pour chaque site enregistré l'outil offre la possibilité de voir depuis quand ce site est surveillé, de supprimer la surveillance, etc.
- 6 – Une fois toute procédure terminée, certains outils proposent d'ajouter un bouton dans les favoris. Cette fonctionnalité permet, d'ajouter une page que l'utilisateur est en train de consulter à la liste des sites ou pages qu'il souhaite surveiller.

Notes

*** Niveau de difficulté



Quantité de matériel requis

Méthodologie - Démarche (suite)

> Pour effectuer une **surveillance de thème** : il s'agit en fait de surveiller l'évolution d'une requête dans un moteur de recherche.

- 1 – Définir la requête à surveiller
- 2 – Créer un compte utilisateur dans l'outil de surveillance avec le courriel sur lequel les réponses à la requête seront envoyées
- 3 – L'outil propose d'insérer dans les favoris un outil qu'il faudra faire glisser dans la barre d'adresse pour indiquer la requête à surveiller
- 4 – Indiquer à l'outil la requête à surveiller
- 5 – Indiquer à l'automate la périodicité de l'envoi des alertes
- 6 – Indiquer à l'outil la forme de la réponse :
 - envoi par courriel de la requête avec les modifications surlignées,
 - ou envoi par courriel du lien vers la page, sur laquelle l'outil surligne les modifications.

A savoir : la majorité des outils de surveillance qui ne sont pas des logiciels, sont en anglais et proposent une version gratuite avec laquelle il est possible de surveiller plusieurs sites.

Notes

Exemple

faire le test simultanément pour bien identifier chaque étape.

Surveillance de thème :

Pour surveiller les productions (travaux de recherche, articles de presse, interview...) sur un sujet qui peut être évoqué sur de nombreux canaux comme l'*Intelligence Economique* :

- 1 – Aller sur la page d'accueil du site www.trackengine.com
- 2 – Cliquer sur " *New User ? Sign up now !* " pour créer son profil utilisateur : courriel de réception de la requête, nom et prénom
- 3 – Valider en cliquant sur " *Accept terms and register* "
- 4 – Trackengine envoie alors un courriel contenant les coordonnées et le mot de passe
- 5 – Cliquer sur " *Sign In* ", saisir les coordonnées envoyées dans le courriel puis valider de nouveau, toujours en cliquant sur " *Sign In* ".
- 6 – Une nouvelle fenêtre s'ouvre dans laquelle un bouton " *Track Me* " est présenté. Positionner la souris sur ce bouton et la faire glisser jusque dans le dossier " Favoris " dans la barre d'outils.
- 7 – Il suffit ensuite d'ouvrir le moteur de recherche Exalead par exemple, dans lequel l'utilisateur lance la requête " *Intelligence Economique* ". Quand la page de réponses s'affiche, cliquer sur " Favoris ", et faire glisser le bouton " *Track me* " dans la barre d'adresse de la page Exalead.

8 – Automatiquement, Trackengine ouvre une fenêtre dans laquelle il demande de saisir un nom pour la surveillance qui vient d'être sollicitée, ainsi que la fréquence à laquelle il devra envoyer les mises à jour.

9 – L'utilisateur peut consulter l'ensemble des surveillances qu'il a paramétré (maximum 5 dans la version gratuite), en cliquant sur " My Bookmarks ".

Surveillance de site :

Pour surveiller les nouvelles publications du laboratoire LEPONT (*Laboratoire d'Etudes Pour l'Ouverture aux Nouvelles Technologies*), par exemple.

- 1 – Aller sur la page d'accueil du site www.watchthatpage.com
- 2 – Cliquer sur " *Register* " pour créer un profil utilisateur : courriel de réception de la surveillance, mot de passe et fréquence d'envoi
- 3 – Valider et saisir l'adresse Internet du laboratoire LEPONT (<http://lepont.univ-tln.fr/>) puis cliquer sur " *add* "
- 4 – Watchthatpage ajoute automatiquement cette nouvelle adresse dans la liste de sites surveillés et transmettra les mises à jour à la fréquence demandée.