

# Evolution des moteurs de recherche sur le Web

URFIST Bordeaux - 5 avril 2011

Olivier Andrieu (Abondance)

[olivier@abondance.com](mailto:olivier@abondance.com)

<http://www.abondance.com/>



## Evolution des moteurs de recherche

### I. 18 ans d'outils de recherche sur le web



## Evolution des moteurs de recherche

### I. 18 ans d'outils de recherche sur le Web

Historique des outils de recherche :

Le crétacé :

1993-1995 : Arrivée des premiers annuaires (Yahoo!) et des premiers moteurs (Webcrawler, Infoseek, Lycos, Excite...)

1995 : AltaVista débarque et change une première fois la donne (prise en compte des balises meta "keywords").



Le début de l'ère moderne :

2000 : Google révolutionne le monde de la recherche d'information avec un nouvel algorithme innovant, en perpétuelle évolution.

2011 : Fin des annuaires et main-mise de Google sur le marché de la recherche (et bien d'autres domaines...)

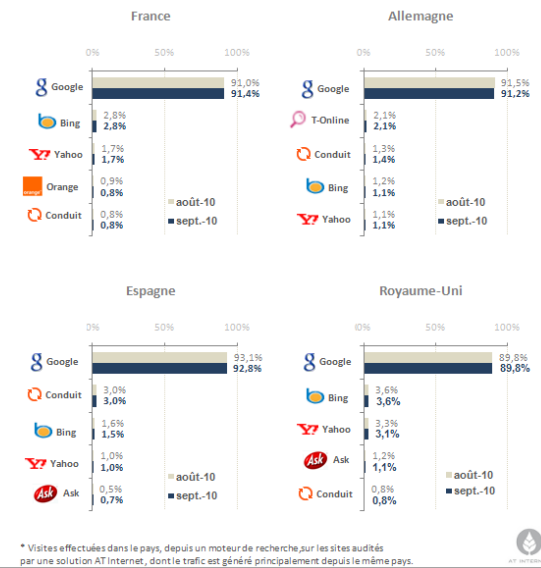
L'avenir ?

## Evolution des moteurs de recherche

### I. 18 ans d'outils de recherche sur le Web

Un marché quasi-monopolistique en Europe

TOP 5 des Moteurs de recherche en part de visites Web\*  
Indicateur moyen par site - ( Septembre 2010 vs août 2010 )



## Evolution des moteurs de recherche

### I. 18 ans d'outils de recherche sur le Web

Un marché plus morcelé ailleurs...

comScore Core Search Report* August 2009 vs. July 2009 Total U.S. – Home/Work/University Locations Source: comScore qSearch			
Core Search Entity	Share of Searches (%)		
	Jul-09	Aug-09	Point Change Aug-09 vs. Jul-09
Total Core Search	100.0%	100.0%	N/A
Google Sites	64.7%	64.6%	-0.1
Yahoo! Sites	19.3%	19.3%	0.0
Microsoft Sites	8.9%	9.3%	0.4
Ask Network	3.9%	3.9%	0.0
AOL LLC Network	3.1%	3.0%	-0.1

## Evolution des moteurs de recherche

### I. 18 ans d'outils de recherche sur le Web

#### Pourquoi le succès de Google ?

- Arrivé au bon moment
- L'anti-portalisation
- Des investisseurs qui ont eu "du nez"
- Des concepteurs visionnaires
- La pertinence
- La rapidité
- La simplicité
- La syntaxe ET par défaut
- La "Google attitude" en plein "funky business"
- L'embauche des meilleurs programmeurs
- L'innovation
- Une parfaite compréhension du Web et une adaptation permanente (au départ : aucune pub sur Google !)



## Evolution des moteurs de recherche

### I. 18 ans d'outils de recherche sur le Web

#### Pourquoi Google est-il un géant aux pieds d'argile ?

- Son chiffre d'affaires repose à 98% sur un seul modèle économique (le lien sponsorisé) qui reste fragile (fraude aux clics)
- Fuite des cerveaux toujours possible (Facebook, Bing, Apple, etc.)
- Pas à l'abri d'un ratage
- Procès multiples
- Risque de dispersion
- Lutte contre le spam (Mayday, Panda, "content spinning"...)
- Usure du pouvoir
- Départ de Larry Page et Sergey Brin
- Nécessité de "splitter" la société en plusieurs entités ?
- Arrivée d'un nouveau "Google Killer" ?



## Evolution des moteurs de recherche

### I. 18 ans d'outils de recherche sur le Web



#### Statistiques Backrub (ancêtre de Google) en 1997 :

Number of Web Pages Fetched: 24 millions  
Number of Email Addresses: 1.7 million  
Number of 404's: 1.6 million  
Source : <http://backrub.c63.be/1997/index.htm>

A la mi-2001, l'index de Google était estimé à 620 millions de pages...

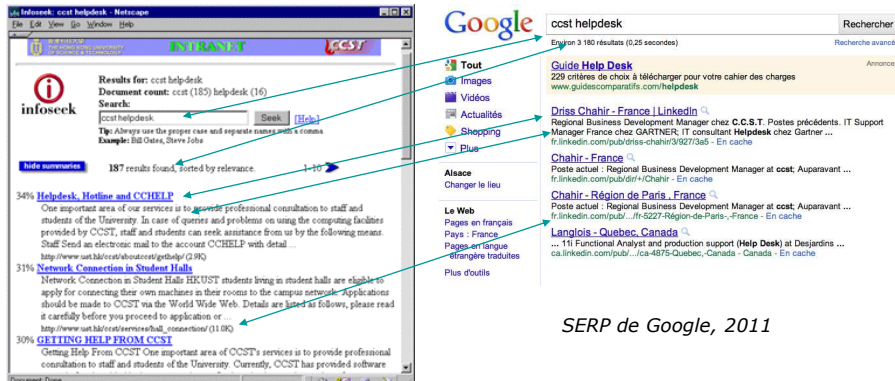
L'index "moyen" des moteurs contient actuellement 20 à 25, voire 40 milliards de pages (et 1 000 milliards de pages "connues")

La taille du web est aujourd'hui un enjeu majeur pour les moteurs de recherche.  
Tout indexer n'est plus possible...

## Evolution des moteurs de recherche

### I. 18 ans d'outils de recherche sur le Web

Les moteurs ont changé (en tout cas pour ce qui est de la mécanique sous le capot).  
La carrosserie est restée très proche de ce qu'on connaissait en 1993...  
Presque 15 ans séparent ces deux copies d'écran :



SERP d'Infoseek, 1997

SERP de Google, 2011

## Evolution des moteurs de recherche

### I. 18 ans d'outils de recherche sur le Web

Les internautes ont changé...

Taille des requêtes saisies par les internautes  
sur les moteurs de recherche :

- 1 seul mot : 13,48 %
- 2 mots : 28,38 %
- 3 mots : 27,15 %
- de 3 à 5 mots : 51,60 %
- de 4 à 10 mots : 30,98 %
- requête contenant plus d'un mot : 86,52 %

(source : rankstat.com, 2007)

Les moteurs se doivent d'être performants sur les requêtes de "Longue Traîne"  
(requêtes longues, mots clés rares) qui deviennent de plus en plus nombreuses...

Percentage of U.S. clicks by number of keywords				
Subject	Jan-08	Dec-08	Jan-09	Year-over-year percent change
1 word	20.96%	20.70%	20.29%	-3%
2 words	24.91%	24.13%	23.65%	-5%
3 words	22.03%	21.94%	21.92%	0%
4 words	14.54%	14.67%	14.89%	2%
5 words	8.20%	8.37%	8.68%	6%
6 words	4.32%	4.47%	4.65%	8%
7 words	2.23%	2.40%	2.49%	12%
8+ words	2.81%	3.31%	3.43%	22%

Note: Data is based on four-week rolling periods (ending Jan. 31, 2009; Dec. 27, 2008; and Jan. 26, 2008) from the Hitwise sample of 10 million U.S. Internet users.  
Source: Hitwise, an Experian company

(source : Hitwise, 2009)

Evolution des moteurs de recherche

## II. Les grandes tendances d'évolution



Evolution des moteurs de recherche

## II. Grandes tendances d'évolution

Essayons de parler d'aujourd'hui et de de demain matin (sans entrer dans la science-fiction)...

3 phases dans une recherche web :

- Phase 1 : saisie de la requête, formulation de la question
- Phase 2 : recherche des réponses pertinentes
- Phase 3 : restitution des résultats

L'avènement d'un futur concurrent de Google pourrait se jouer sur les phases 1 et 3 !!

## Evolution des moteurs de recherche

### II. Grandes tendances d'évolution

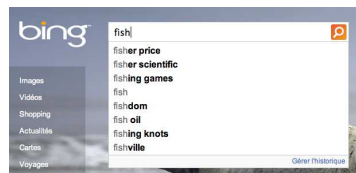
#### Phase 1 : aide à la saisie de la requête, à la formulation de la question

Le règne actuellement de l'auto-complétion



Google Suggest, 2004

Google Instant, 2010



Bing, 2009, 2011



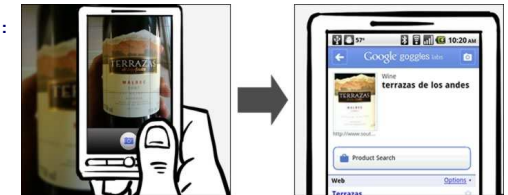
## Evolution des moteurs de recherche

### II. Grandes tendances d'évolution

Phase 1 : aide à la saisie de la requête, à la formulation de la question

D'autres voies de réflexion ?

- Proposition de contenus en "sous-marin"...
- Analyse de blocs de texte "brut" (sélectionné avec la souris ou copié/collé) : Yahoo! Q
- Langage naturel ?
- Questions/réponses ?
- La commande vocale (mobile, Shazam)
- Google Goggles et la reconnaissance d'images\* :
- Assistant automatique de recherche (<http://guru.googlelabs.com/>)
- Assistant de traduction automatique
- Branchement direct sur notre cerveau ?
- D'autres voies de réflexion ?
- Peut-on aujourd'hui sortir de la dictature du mot clé ?



\* Et la résolution des Sudoku :-)

## Evolution des moteurs de recherche

### II. Grandes tendances d'évolution

#### Phase 2 : recherche des réponses pertinentes

Google est aujourd'hui leader.

Un algorithme qui prend en compte 200 critères de pertinence.

Entièrement automatisé (enfin presque...).

Des critères « in page » (recherche des mots clés dans le contenu des pages) : texte, balise TITLE, attributs ALT des images.

Des critères « off page » amenés par Google : PageRank (popularité), réputation, URL, etc.

Dernières nouveautés :

- Le TrustRank ;
- Le temps d'affichage des pages, le temps de réaction du serveur ;
- La personnalisation ;
- Les réseaux sociaux ;
- Transcription automatique des contenus audio ;
- Etc.

Souvent la « loi du plus fort »...

## Evolution des moteurs de recherche

### II. Grandes tendances d'évolution

#### Phase 2 : recherche des réponses pertinentes

Bing est challenger avec son "moteur de décision" qui cherche parfois à anticiper la recherche de l'internaute.



## Evolution des moteurs de recherche

### II. Grandes tendances d'évolution

#### Phase 2 : recherche des réponses pertinentes

Une autre solution pour faciliter le travail des moteurs est peut-être la sémantisation du web.

Exemple : les "rich snippets"

The screenshot shows a Google search for "recette cannellé". The search bar contains the text "recette cannellé" and a "Rechercher" button. Below the search bar, it says "Recherche instantanée" and "Environ 37 300 résultats (0,23 secondes)". The search results are displayed with rich snippets, including star ratings and brief descriptions. The first result is "Cannelés bordelais - Recette de cuisine Marmiton : une recette" with a 3.79 star rating. The second result is "Cannelés (7ème rencontre Marmiton) - Recette de cuisine Marmiton..." with a 4.5 star rating. The third result is "Recette - Cannelés bordelais - Notée 4.6/5 par les internautes" with a 4.6/5 star rating. The fourth result is "Recette - Cannelé au chèvre - Proposée par 750 grammes" with a 4.5 star rating. On the left side, there are navigation options: "Tout", "Images", "Vidéos", "Actualités", "Shopping", "Plus", "Alsace", "Le Web", "Pages en français", "Pays : France", "Pages en langue étrangère traduites", and "Plus d'outils".

## Evolution des moteurs de recherche

### II. Grandes tendances d'évolution

#### Phase 2 : recherche des réponses pertinentes

Les "rich snippets" permettent de mettre en place des filtres de recherche pertinents :

L'HTML 5 peut-il apporter un "plus" dans ce domaine ?

Un web de plus en plus sémantique...

Les avis des autres internautes sont une voie d'exploration intéressante mais très spammée...

The screenshot shows a Google search for "cinco de mayo". The search bar contains the text "cinco de mayo" and a "Search" button. Below the search bar, it says "About 24,700 results (0.15 seconds)". The search results are displayed with rich snippets, including star ratings and brief descriptions. The first result is "Frozen Cinco De Mayo Dessert" with a 4.5 star rating. The second result is "Real Homemade Tamales" with a 4.5 star rating. The third result is "Cinco de Mayo Sombrero Cake" with a 4.5 star rating. The fourth result is "Mexican Whole Wheat Flour Tortillas" with a 4.5 star rating. The fifth result is "Cinco De Mayo Guacamole" with a 4.5 star rating. On the left side, there are navigation options: "Everything", "Images", "Vidéos", "News", "Shopping", "Recipes", "More", "Ingredients", "Yes No", "avocado", "cranberry juice", "pineapple juice", "jalapeño", "lime", "scallions", "peas", "coriander", "Any cook time", "Less than 15 min", "Less than 30 min", "Less than 60 min", "Any calories", "Less than 100 cal", "Less than 300 cal", "Less than 500 cal".

## Evolution des moteurs de recherche

### II. Grandes tendances d'évolution

#### Phase 2 : recherche des réponses pertinentes

La personnalisation des recherches en fonction de :

- votre localisation géographique ;
- votre historique de recherche ;
- la langue de votre navigateur ;
- l'historique de recherche de votre cercle social ;
- votre profil ;
- etc.

Actuellement, plus de 20% des résultats renvoyés par Google sont personnalisés pour l'internaute qui utilise le moteur.

La recherche sociale ou la "socialisation de la recherche" est sans doute l'un des enjeux majeurs pour l'avenir (Facebook "J'aime" repris par Bing, Google "+1", etc.).

## Evolution des moteurs de recherche

### II. Grandes tendances d'évolution

#### Phase 2 : recherche des réponses pertinentes

Temporalité et "chaleur" de la requête.

Détection des recherches "chaudes" (QDF) : affichage des résultats Twitter et algorithme spécifiques pour l'actualité.

Analyse des liens dans les tweets pour détecter les sujets et sites "chauds"



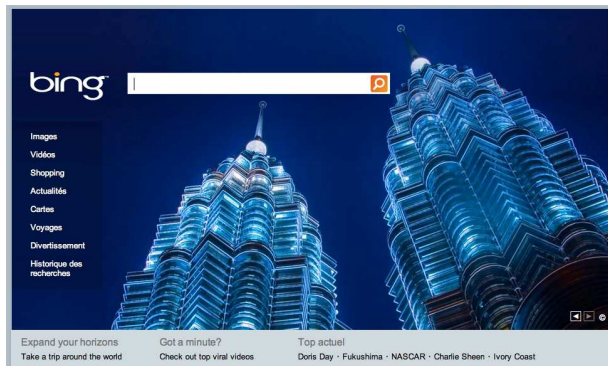
## Evolution des moteurs de recherche

### II. Grandes tendances d'évolution

#### Phase 3 : restitution des résultats, expérience utilisateur

C'est peut-être l'étape qui propose le plus de possibilités, mais aussi la plus complexe à changer (la force des habitudes...)

Des choses simples (certes un peu « gadget ») : image en fond d'écran d'accueil (avec vidéo pour bientôt)



## Evolution des moteurs de recherche

### II. Grandes tendances d'évolution

#### Phase 3 : restitution des résultats, expérience utilisateur

Moteurs graphiques : KartOO

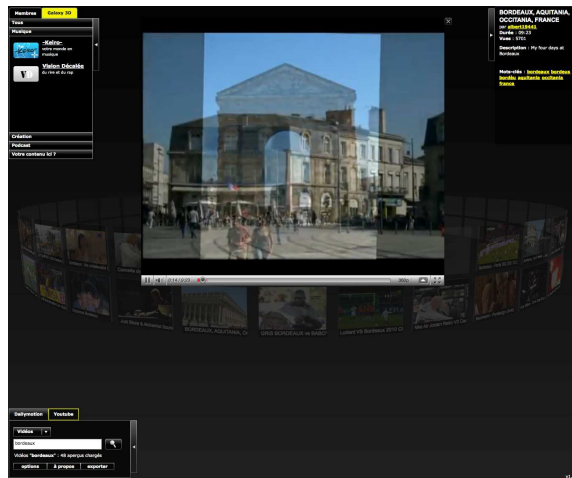


## Evolution des moteurs de recherche

### II. Grandes tendances d'évolution

Phase 3 : restitution des résultats, expérience utilisateur

Moteurs graphiques : Galaxy (<http://www.tkaap.com/galaxy/>)



## Evolution des moteurs de recherche

### II. Grandes tendances d'évolution

Phase 3 : restitution des résultats, expérience utilisateur

Moteurs graphiques : Search Cube (<http://www.search-cube.com/>)



## Evolution des moteurs de recherche

### II. Grandes tendances d'évolution

Phase 3 : restitution des résultats, expérience utilisateur

Moteurs graphiques : SearchMe, CreativeSpace

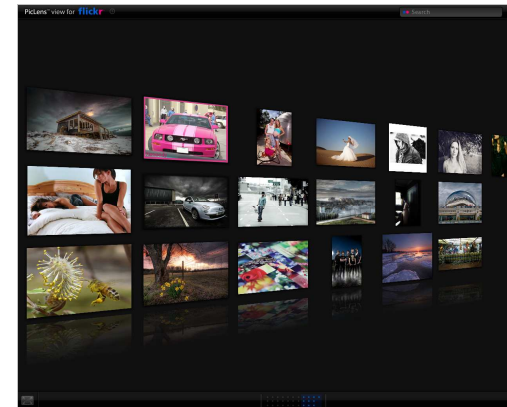


## Evolution des moteurs de recherche

### II. Grandes tendances d'évolution

Phase 3 : restitution des résultats, expérience utilisateur

Moteurs graphiques : PicLens : <http://www.cooliris.com/>

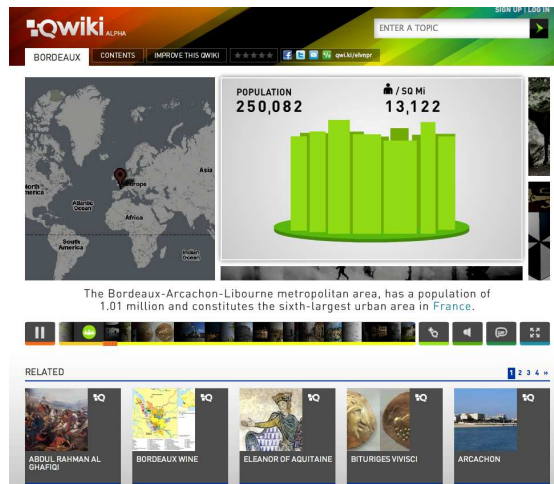


## Evolution des moteurs de recherche

### II. Grandes tendances d'évolution

Phase 3 : restitution des résultats, expérience utilisateur

Moteurs graphiques : Qwiki (<http://www.qwiki.com/>)



## Evolution des moteurs de recherche

### II. Grandes tendances d'évolution

Phase 3 : restitution des résultats, expérience utilisateur

Moteurs graphiques : Exalead Labs (<http://labs.exalead.com/>)

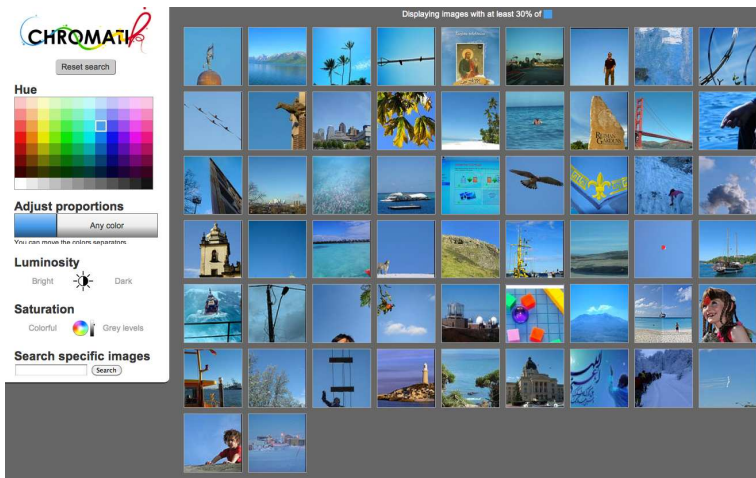


## Evolution des moteurs de recherche

### II. Grandes tendances d'évolution

Phase 3 : restitution des résultats, expérience utilisateur

Moteurs graphiques : Exalead Labs (<http://labs.exalead.com/>)

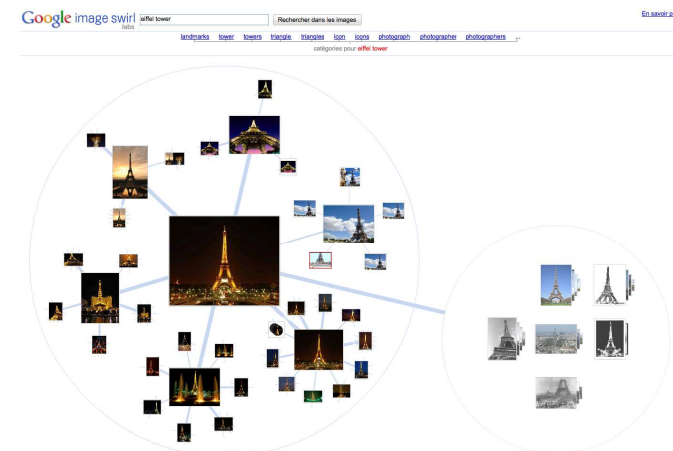


## Evolution des moteurs de recherche

### II. Grandes tendances d'évolution

Phase 3 : restitution des résultats, expérience utilisateur

Moteurs graphiques : Google Image Swirl (<http://image-swirl.googlelabs.com/>)



## Evolution des moteurs de recherche

### II. Grandes tendances d'évolution

Phase 3 : restitution des résultats, expérience utilisateur

Moteurs graphiques : Google Fast Flip (<http://fastflip.googlelabs.com/>)



## Evolution des moteurs de recherche

### II. Grandes tendances d'évolution

Phase 3 : restitution des résultats, expérience utilisateur

Moteurs graphiques : Body Browser (<http://bodybrowser.googlelabs.com/>)  
 Un coup d'œil aux expérimentations Chrome et WebGL (<http://www.chromeexperiments.com/webgl/>) peut être intéressant...



# Evolution des moteurs de recherche

## II. Grandes tendances d'évolution

Phase 3 : restitution des résultats, expérience utilisateur

Regroupement par date : Google News Timeline (<http://newstimeline.googlelabs.com/>)

The screenshot shows the Google News Timeline interface. At the top, there's a search bar with 'apache' entered. Below it, a grid of news articles is displayed, organized by date from March 28, 2011, to April 3, 2011. Each article includes a small thumbnail image, a headline, and a brief snippet of the text. The articles cover various topics, including political events, local news, and international reports.

# Evolution des moteurs de recherche

## II. Grandes tendances d'évolution

Phase 3 : restitution des résultats, expérience utilisateur

Clustering (<http://search.yippy.com/>)

The screenshot shows the Yippy search engine interface. The search bar contains the word 'apache'. Below the search bar, there are several tabs for different search categories: 'clouds', 'sources', and 'sites'. The search results are displayed as a list of categorized items, including 'Foundation, Software (15)', 'Open source, Apache (20)', 'PHP (13)', 'Download (10)', 'Wiki, Wikipedia.Org (3)', 'Indians (5)', 'Modules, Supporting (6)', and 'Mod (4)'. A detailed view of the 'Dell™ Web Server' result is shown on the right, including a description and a link to the Dell website.

## Evolution des moteurs de recherche

### II. Grandes tendances d'évolution

Phase 3 : restitution des résultats, expérience utilisateur

Interface Jazz de Google

- Tout
- Images
- Vidéos
- Actualités
- Shopping
- Adresses
- Plus

Alsace

Changer le lieu

Le Web

Pages en français  
Pays : France  
Pages en langue étrangère traduites

Date indifférente

Les plus récentes  
Moins de 24 heures  
Moins de 2 jours  
Moins d'une semaine  
Moins d'un mois  
Moins d'un an  
Période personnalisée

Tous les résultats

Sites avec des images  
Roue magique  
Recherches associées  
Chronologie  
Déjà consultés  
Pas encore consultés  
Moins d'outils

Abondance - référencement et moteurs de recherche - toute info... 48 visites - 31 mars 15 mars 2011... Abondance d'infos sur le référencement et les moteurs de recherche: description des moteurs, schéma, logs, outils d'audit, méthodologies ...  
Outils - Bibliographie - Ebooks - Le blog d'Abondance  
www.abondance.com - En cache - Pages similaires  
[Abondance]

Abondance - outils - sommaire 0 0  
Abondance.com - Recherche d'information, référencement et promotion de sites ...  
outils.abondance.com - En cache - Pages similaires

Abondance - Actualité des moteurs de recherche et du référencement 0 0  
Abondance - Retrouvez ici toute l'actualité des moteurs de recherche majeurs ...  
actu.abondance.com - En cache - Pages similaires

Le blog d'Abondance - Référencement et moteurs de recherche 10 0 - 3 visites - 3 mars 1 av. 2011... Le blog d'Abondance - Quelques réflexions sur les moteurs de ...  
blog.abondance.com - ... - Accueil - Accueil - En cache - Pages similaires

Plus de résultats de abondance.com

Adresses pour abondance près de Alsace

Abondance 0 1 avis - Page Google Adresses  
www.abondance.com - 3 Rue des Châteaux, Helgenstein - 03 88 08 63 26

Abondance Immobilier 0 - Page Google Adresses  
www.abondance-immobilier.com - 23 Rue Gutenberg, Witzelshausen (LA) - 03 88 96 67 67

Au Cœur d'Abondance 0 - Page Google Adresses  
maps.google.fr - 36 Rue Mar Foch, Blasheim - 03 88 68 86 81

Donnez votre avis sur des adresses pour améliorer nos recommandations »

Office de Tourisme d'Abondance - village abondance - haute savoie... 0  
Office de Tourisme d'Abondance - 16 - 13 094 50 73 02 00 - Mentions légales | Contact | E-mail | Documentation | Newsletter | Livre d'or | Cartes Postales ...  
Site / Activités / Musée - Le Village - Gastronomie - Tarifs des forfaits  
www.abondance.org - En cache - Pages similaires



## Evolution des moteurs de recherche

### III. Conclusion



## Evolution des moteurs de recherche

### III. Conclusion

#### Un attrait élevé pour les fonctionnalités relatives à la pertinence, l'assistance et l'interactivité



Les 3 types de fonctionnalités jugées les plus attractives sont celles qui :  
- améliorent la présentation des résultats pour aider à trouver les renseignements les plus pertinents  
- aident les utilisateurs d'Internet à affiner leurs requêtes pour obtenir la meilleure réponse  
- permettent une interaction entre les utilisateurs d'Internet et le moteur de recherche

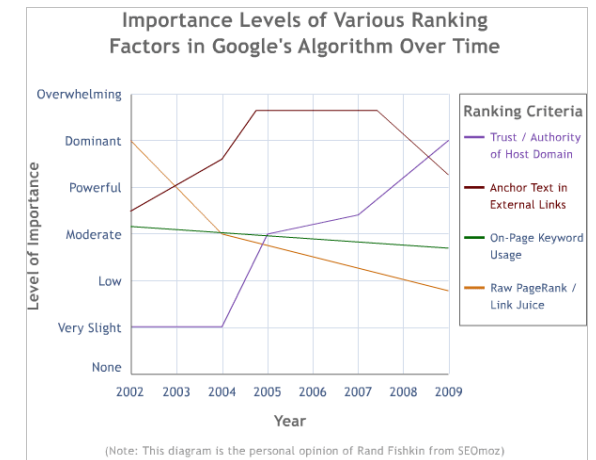


Etude Google "Search of the future"  
[http://blog.abondance.com/2010/01/google-et-le-moteur-de-recherche-ideal\\_21.html](http://blog.abondance.com/2010/01/google-et-le-moteur-de-recherche-ideal_21.html)

## Evolution des moteurs de recherche

### III. Conclusion

Google est un organisme vivant...



## Evolution des moteurs de recherche

### III. Conclusion

Le marché des moteurs généralistes risque de rester aux mains de Google pour quelques temps encore...

Les annuaires ont du mal...

Mais de très nombreux marchés de niches sont disponibles. Exemples :

- Moteurs thématiques (pros ou grand public) basés sur des annuaires / moteurs.
- Moteurs régionaux (pays - région - ville).
- Moteur européen ?
- Moteur spécifiques (PDF, image, vidéo, audio, etc.).

Les cartes ne sont pas toutes distribuées encore !

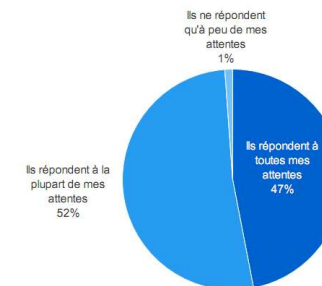
## Evolution des moteurs de recherche

### III. Conclusion

Mais les internautes sont globalement assez satisfaits des moteurs actuels, ce qui limite l'innovation...

#### Satisfaction vis-à-vis des moteurs de recherches actuels

Dans quelle mesure les moteurs de recherche web actuels (google, bing, Yahoo...) répondent-ils aujourd'hui à vos attentes ?



Etude Google "Search of the future"

[http://blog.abondance.com/2010/01/google-et-le-moteur-de-recherche-ideal\\_21.html](http://blog.abondance.com/2010/01/google-et-le-moteur-de-recherche-ideal_21.html)

## Evolution des moteurs de recherche

### III. Conclusion

De qui peut venir l'innovation et la concurrence pour Google ?

- De start-ups ? Non, ce sont des B2BBGOBB !!
- De Bing ? Peut-être... Mais les stats sont contre lui...
- D'Apple ? :-)
- De Facebook ? Possible, mais le veulent-ils ?
- Autres idées ? Quelqu'un veut-il relever le défi ?

## Evolution des moteurs de recherche

### III. Conclusion

La recherche d'information est un champ d'investigation immense !

Nous n'avons envisagé aujourd'hui que quelques hypothèses...

L'utilisation de nouveaux terminaux (téléphones mobiles, iPad et tablettes) peut induire de nouveaux modes de consultation et de recherche de l'information...

Le cimetière des moteurs de recherche se remplit (Cuil, Wikia Search, AllTheWeb, AltaVista, etc.) mais il faut persévérer (Blekkio, Qwiki, etc.) !

Il reste tant de choses passionnantes à faire !!

## Rendez-vous dans 10 ans !!

**Evolution des moteurs de recherche**

**CONTENT IS KING, LINK IS HIS QUEEN !**

**Merci :-)**



*Support proposé par Olivier Andrieu - [olivier@abondance.com](mailto:olivier@abondance.com)*

