



Les résultats de recherche de Google Première partie : publicité et notices courantes

La page du résultat de recherche de **Google** n'affiche pas que des notices, la notice étant le format d'affichage identifiant un document. On y trouve de la publicité, différentes formes de présentation de notices, des raccourcis et des services comme les actualités nommés **OneBox**. Chacune de ces présentations fera l'objet d'une chronique.

Cette première partie traitera de la publicité, du format classique des notices, des pages en cache, des pages similaires, des pages des domaines, des tables des matières, de la date des notices, des formats de fichiers et de la traduction.

La publicité

La société **Google Inc.** n'est pas la seule entreprise de recherche à tirer profit de la publicité. Cependant, elle est la seule à l'avoir exploitée de brillante façon. En fait, ses revenus publicitaires sont la pierre angulaire de son petit empire. Elles sont aussi son talon d'Achille.

Avant **Google**, la publicité se présentait sous la forme d'encadrés, de bannières et de liens le plus souvent intégrés aux notices. **Google** a

modifié la dynamique en privilégiant les liens publicitaires.

La page du résultat d'une recherche de **Google** contient deux emplacements de liens publicitaires : l'un est entre la barre d'état de la recherche et le résultat de recherche proprement dit, l'autre est dans une colonne à droite de la page.

The screenshot shows a Google search interface with the search term 'Quebec'. Below the search bar, there are navigation tabs for 'Web', 'Images', 'Groupes', 'Annuaire', 'Actualités', and 'plus'. The search results are categorized under 'Web' and show 'Résultats 1 - 10 sur un total d'environ 270 000 000 pour Quebec. (0,25 secondes)'. The results are split into two columns. The left column contains organic search results for 'Quebec', including 'www.canadavision.fr' (Agence de voyage spécialiste), 'Visitez la Belle Province' (www.BonjourQuebec.com), and 'Bonjour Québec' (Le site touristique officiel du gouvernement du Québec). The right column contains 'Liens commerciaux' (Commercial Links) for 'Hôtels Québec' and 'Quebec' (Tous les Vols, Hôtels et Séjours).

Plusieurs auteurs qualifient le résultat de recherche d'organique ou d'algorithmique, c'est-à-dire provenant de l'index du service de recherche. Cela pour le distinguer de la publicité tant la différenciation est parfois ténue.

Les liens publicitaires de **Google** sont nommés *Liens commerciaux*

(*Sponsored links*, de là l'expression «liens sponsorisés »).

Les liens commerciaux au-dessus du résultat de recherche sont dans un encadré bleu pâle, la même couleur que la barre d'état de la recherche. Le rectangle est si pâle que le lien commercial peut être assimilé à une notice lorsque l'on y prend garde.

La publicité latérale (voir la figure plus haut), séparée du résultat de recherche par une ligne verticale, est plus identifiable.

Les notices de **Google**, aussi nommées résultats de recherche, n'ont aucun signe pour les distinguer clairement de la publicité. En comparaison, le lien publicitaire de Yahoo ! Recherche est précédé d'un point et placé dans un rectangle bleu pâle. Ses notices sont numérotées.

Selon l'étude [Internet searchers are confident, satisfied and trusting – but they are also unaware and naïve](#)¹ de **Pew Internet & American Life Project** parue en janvier 2005, 62% des chercheurs ne distinguent pas la publicité du résultat de recherche, tout service de recherche web confondu. Or, le format de page de **Google** a été

conçu pour confondre la publicité supérieure et les notices afin de tromper l'internaute inexpérimenté, distrait ou naïf.

Les liens publicitaires n'offrent pas la sécurité attendue. L'étude [The Safety of Internet Search Engines](#)² de **McAfee SiteAdvisor** fait ressortir le danger des espionciels et des malicieux en enregistrant un fichier provenant d'une page utilisant un lien publicitaire. La barre d'outil des sites de recherche, dont la [barre d'outils de Google](#), est munie d'une fonction d'alerte ou de détecteur de programmes malveillants.

Le format classique de la notice

Le format classique de la notice est composé du titre et d'un extrait du document comprenant généralement un mot de la requête dans l'une ou dans l'autre partie. L'adresse du document et sa taille complètent la notice. La taille permet d'évaluer l'ampleur du document. Elle est indiquée lorsque le document est indexé. Le document n'ayant pas été indexé complètement n'a pas cette information.

Certaines notices (voir l'exemple ci-dessous) sont placées en retrait. Ces notices proviennent du même site

¹ Deborah Fallows. (Page consultée le 30 juin 2006). [Internet searchers are confident, satisfied and trusting – but they are also unaware and naïve](#). (En ligne). Washington, D.C., Pew Internet & American Life Project, 2005. 36 pages. Adresse URL : http://www.pewinternet.org/pdfs/PIP_Searcheng ine_users.pdf

² Ben Edelman et Hannah Rosenbaum. (Page consultée le 30 juin 2006). [The Safety of Internet Search Engines](#). (En ligne). McAfee SiteAdvisor, 2006. Adresse URL : http://www.siteadvisor.com/studies/search_safety_may2006.html

que la notice précédente. Elles sont évaluées moins pertinentes... ce qui n'est pas toujours le cas.

La page en cache

La plupart des notices auront la mise *en cache* qui est une copie du document. C'est dans cette copie que se fait la recherche. C'est en consultant cette copie que vous pouvez trouver les mots de la requête, sauf si le document a été mis à jour entre-temps ou si un hyperlien textuel cible le document. Ainsi, un document peut être pertinent pour la requête, mais inadéquat lors de sa consultation.

Google.fr

Version française, possibilité de limiter la recherche aux sites hébergés sur des serveurs situés...

[www.google.fr/](#) - 4k - [En cache](#) - [Pages similaires](#)

Outils linguistiques

Pour afficher **Google** dans la langue de votre choix (page d'accueil, messages, boutons, etc. ... Au stade actuel, **Google** supporte les langues suivantes : ...

[www.google.fr/language_tools?hl=fr](#) - 46k -

[En cache](#) - [Pages similaires](#)

[[Autres résultats, domaine www.google.fr](#)]

Les pages similaires

Certaines notices auront un lien vers une page de notices similaires. La consultation de cette page est utile lorsque le document primaire est pertinent. Toutefois, la similarité des pages n'est pas toujours au rendez-vous.

Les pages du domaine

Google peut afficher le lien [Autres résultats, domaine ...](#) Ce lien est utile pour explorer un grand site comme les sites gouvernementaux. Ainsi, **Google** affichera tous les sous-

domaines indexés du site. Vous pouvez l'utiliser pour trouver un sous-domaine et y cibler votre recherche.

La table des matières

Parfois, **Google** capture des hyperliens provenant du document formant un extrait de la table des matières du site. L'exemple ci-dessous montre quatre liens (*t'Chat*, *Mail*, *Traducteur* et *Météo*) issus de la page du portail **Voila**.

Bienvenue sur Voila

Service de recherche et de navigation à travers les Newsgroups francophones.
[www.voila.fr/](#) - 39k - 26 juin 2006 - [En cache](#) - [Pages similaires](#)

[t'Chat](#) - [Mail](#) - [Traducteur](#) - [Météo](#)

[Autres résultats, domaine www.voila.fr](#) »

La date d'indexation

Remarquez aussi que certaines notices ont une date à l'exemple de celle de **Voila**. Il s'agit de la date d'indexation et non de la date de création. Cette date apparaît lorsque la page est fréquemment mise à jour.

Le format de fichier

La plupart des notices réfèrent à des pages web généralement identifiées par l'extension de fichier HTML. Or, Internet contient une multitude de formats de fichiers. **Google**, et ses concurrents, ne font pas de distinction. Leurs robots d'indexation aspirent tous les documents disponibles sur Internet et les indexent. **Google** identifie les fichiers en formats les plus communs : Adobe Acrobat PDF, Adobe Proscript, Microsoft Word, Microsoft Excel, Microsoft PowerPoint et RTF. **Google** ne se limite pas à ces formats. A

l'occasion, vous pourrez voir des fichiers WordPerfect, Lotus, etc. C'est pourquoi vous verrez des notices accompagnées de son format de fichier.

[PDF] [AllTheWeb](#)

Format de fichier: PDF/Adobe Acrobat - [Version HTML](#)

AllTheWeb. Fiche technique. Par Marc Duval. Bibliothécaire ... **AllTheWeb**. Recherche avancée. (case de recherche d'accueil). Préfixe de recherche ...
www.dsi-info.ca/moteurs-de-recherche/AllTheWeb/AlltheWeb-Fiche-technique.pdf - [Pages similaires](#)

Le format de fichier [PDF] est à gauche du titre et sous le titre. La *Version HTML* accompagne la mention du format. Identifiez bien le format de fichier du document et assurez-vous d'avoir le lecteur approprié pour ne pas avoir de surprise lors de l'ouverture du fichier.

Traduction, S.V.P.

Les documents du web couvrent toutes les langues modernes. **Google** les détecte puis affiche le lien *Traduire cette page* à droite du titre. Le service de traduction limite le nombre de mots à 1 500. La traduction est souvent imparfaite. Une connaissance rudimentaire de la langue d'origine

permet de pallier cette lacune.

[Yahoo! Search - Web Search](#) - [[Traduire cette page](#)]

Search Your Home Page - Search Services. Copyright © 2006 **Yahoo!** Inc. All rights reserved. [Privacy Policy](#) - [Terms of Service](#) - [Submit Your Site](#) - [Help](#).

search.yahoo.com/ - 11k - 28 juin 2006 - [En cache](#) - [Pages similaires](#)

Conclusion

Google a su marier les liens commerciaux et le résultat de recherche. Pour plusieurs internautes, le mariage est si intime qu'il est parfois difficile de les distinguer. En connaissant l'un et l'autre, l'internaute averti est à même de les reconnaître et d'apprécier son expérience de recherche.

Marc Duval

Service de recherche documentaire DSI
1280-4, boul. de Montarville.
Boucherville. Québec. Canada J4B 8B4
marcduval@dsi-info.ca