

Comment éviter l'enfer de l'index secondaire de Google...

[Retour au sommaire de la lettre](#)

Savez-vous que Google utilise, encore aujourd'hui, deux index – principal et secondaire – pour rechercher ses résultats et répondre aux requêtes de ses utilisateurs ? On s'aperçoit rapidement, à l'analyse, qu'être dans l'index secondaire revient à une quasi invisibilité de la page en termes de positionnement. Combien de vos pages sont dans l'index principal et combien sont situées dans l'"enfer de l'index secondaire" ? Et, dans ce cas, comment les en faire sortir ? Voici comment faire...

Bien sûr, c'est une évidence, comme tous les moteurs de recherche, Google utilise un index qui contient les pages web dans lesquelles il va effectuer ses investigations. Selon les sources, plus ou moins officielles, la taille de cet index varierait actuellement entre 40 et 100 milliards de pages. Peu importe finalement ce chiffre, à partir du moment où l'index contient les pages qui répondent à nos requêtes...

Mais savez-vous que Google utilise en fait deux index depuis 2003 ? Le premier, l'**index principal**, contient les pages que Google considère comme "essentielles", les plus importantes donc. L'**index secondaire**, pour sa part, contient ce qu'on pourrait appeler "un deuxième choix", contenant notamment de nombreuses pages considérées comme étant du "*duplicate content*" (voir nos articles à ce sujet dans les trois lettres R&R précédentes). De plus, les pages présentes dans cet index secondaire sont "*crawlées*" (visitées par les robots du moteur) bien moins souvent que celles de l'index principal.

Longtemps, Google a indiqué le fait qu'une page était issue de cet index secondaire au travers de la mention "**Supplemental Result**" ou "**Résultat complémentaire**" dans ses résultats :

Fatal error: Call to undefined function: add_action() in /home ...
... function: add_action() in /home/ljccinf/public_html/wp-content/plugins/
post_schedule_ping_optimizer/post-schedule-ping-optimizer.php on line 186.
[www.ljcc.info/wp-content/plugins/post_schedule_ping_optimizer/post-schedule-ping-optimizer.php](#) - 1k - **Supplemental Result** - [Cached](#) - [Similar pages](#)

[AMEN.FR : votre fournisseur de présence sur Internet : noms de ...](#)
... Re: Référencement. Auteur: Vincent GERMAIN (82.216.175.---) Date: 29-04-2004 21:29 Moi
mon référencement ca m'a pris du temps. ...
[forum.amen.fr/read.php?f=4&i=38742&t=38734](#) - 34k - **Résultat complémentaire** -
[En cache](#) - [Pages similaires](#)

Et puis, fin 2007, Google a communiqué sur le fait que cet index secondaire n'existait plus (<http://blog.abondance.com/2007/12/rsultats-complimentaires-sur-google-cest.html>) et que le moteur de recherche n'utilisait plus qu'un seul index.

Les deux index cohabitent pourtant encore...

Pourtant, il semble bien qu'aujourd'hui encore, cette différence existe bien encore, entre deux sous-ensembles de l'index, quelle que soit la forme que prenne cette dichotomie... Peut-être ne s'agit-il pas de deux "index", au sens technique du terme, mais toujours est-il qu'il est clair que toutes les pages ne sont pas placées au même niveau par Google dans son index de départ. Dans cet article, nous resterons donc sur

notre vision initiale d'index "principal" et "secondaire", peu importe finalement ce qui se passe sous le capot du moteur...

Le fait est simple à vérifier : faites une recherche sur un site web donné, par exemple actu.abondance.com sur Google.fr grâce à la requête [site:actu.abondance.com](http://www.google.fr/search?hl=fr&q=site%3Aactu.abondance.com) (<http://www.google.fr/search?hl=fr&q=site%3Aactu.abondance.com>). Google renvoie ici 4 090 résultats :



The screenshot shows the Google search interface. The search bar contains the text "site:actu.abondance.com". Below the search bar, there are radio buttons for "Web" (selected), "Pages francophones", and "Pages : France". The search results are displayed under the heading "Web". The first result is "Actu Abondance - Actualité des moteurs de recherche et du ...". The second result is "Actu Abondance: mars 2008 - Actualité des moteurs de recherche et ...".

Google

site:actu.abondance.com

Rechercher


Rechercher dans : ☒ Web ☐ Pages francophones ☐ Pages : France

Web Résultats 1 - 10 sur un total d'environ 4 090 sur actu.abondance.com (0,30 secondes)

[Actu Abondance - Actualité des moteurs de recherche et du ...](#)
Actu Abondance - Retrouvez ici toute l'actualité des moteurs de recherche majeurs et du référencement en France et dans le Monde, jour après jour ...
actu.abondance.com/ - 72k - Il y a 51 minutes - [En cache](#) - [Pages similaires](#) - [À noter](#)

[Actu Abondance: mars 2008 - Actualité des moteurs de recherche et ...](#)
Actu Abondance: mars 2008 - Les archives de l'actualité des moteurs de recherche majeurs et du référencement en France et dans le Monde, jour après jour ...
actu.abondance.com/2008_03_01_archive_actu.html - 292k - [En cache](#) - [Pages similaires](#) - [À noter](#)

Effectuons la même requête maintenant sur un autre site ayant passé un accord de partenariat avec Google, comme AOL (<http://recherche.aol.fr/>) ou Mozbot (<http://www.mozbot.fr/>) :



The screenshot shows the AOL search interface. The search bar contains the text "site:actu.abondance.com". Below the search bar, there are radio buttons for "Web" (selected), "Pages francophones", and "Pages : France". The search results are displayed under the heading "Web". The first result is "Actu Abondance - Actualité des moteurs de recherche et du". The second result is "Actu Abondance: mars 2008 - Actualité des moteurs de recherc...".

AOL search

Web Images Vidéo Shopping

site:actu.abondance.com

Rechercher

Rechercher dans : ☒ Web ☐ Pages francophones ☐ Pages : France

Web Résultats 1 - 10 sur un total de 1,540 de actu.abondance.com

[Actu Abondance - Actualité des moteurs de recherche et du](#)
Actu Abondance - Retrouvez ici toute l'actualité des moteurs de recherche majeurs et du référencement en France et dans le Monde, jour après jour ...
actu.abondance.com/ - 75k - Il y a 18 heures - [Pages similaires](#)

[Actu Abondance: mars 2008 - Actualité des moteurs de recherc...](#)
Actu Abondance: mars 2008 - Les archives de l'actualité des moteurs de recherche majeurs et du référencement en France et dans le Monde, jour après jour ...
actu.abondance.com/2008_03_01_archive_actu.html - 292k - [Pages similaires](#)



Ces deux moteurs trouvent aux environs de 1 400 à 1 500 pages. En fait, ces deux outils travaillent uniquement sur l'index "principal" de Google (ou tout du moins, Google ne leur fournit que des résultats issus de cet index)...

Il semblerait en fait que :

- L'index principal contient les pages considérées comme les plus pertinentes par Google.
- L'index secondaire contient des pages considérées comme moins importantes ou comme dupliquées par Google, et il ne les affichera par exemple que lorsque vous le demanderez, en cliquant sur ce message :

*Pour limiter les résultats aux pages les plus pertinentes (total : 7), Google a ignoré certaines pages à contenu similaire.
Si vous le souhaitez, vous pouvez [relancer la recherche en incluant les pages ignorées](#).*

En clair, dans l'exemple ci-dessus, avant que ce message n'apparaisse, vous visualisez les pages issues de l'index principal. Une fois le lien "[relancer la recherche en incluant les pages ignorées](#)" cliqué, c'est tout l'index général (principal + secondaire) qui est pris en compte...

La problématique est donc très importante puisqu'une page qui se trouve dans l'index principal conservera ses chances d'être bien positionnée dans les résultats du moteur de recherche, alors qu'une page qui se trouve dans l'"enfer de l'index secondaire" est quasiment perdue pour un bon positionnement... Etre "connu de Google" ne suffit pas à être bien positionné si la page se trouve dans l'index secondaire... Il vous faut donc bien vérifier quelles sont vos pages qui se trouvent dans l'une ou l'autre zone d'investigation du moteur...

Comment vérifier dans quel index sont vos pages ?

Pour vérifier combien de pages de votre site sont dans chaque index, il existe plusieurs façons, plus ou moins officielles :

- La première, on l'a vu, est d'utiliser la commande "site:" sur Google puis sur un site "affilié" comme AOL.fr ou Mozbot.fr. Google indiquera alors le nombre total de pages indexées (index principal ET secondaire), le moteur affilié ne renverra que les pages

de l'index principal. La différence entre les deux chiffres renvoyés donnera le nombre de pages dans l'index secondaire...

- Autre solution : utiliser la requête "[site:www.votresite.com/](http://www.votresite.com/)*". La mention "/"* après la requête semblerait indiquer à Google qu'il ne doit utiliser que son index principal pour effectuer la recherche :

The screenshot shows a Google search interface. The search bar contains the text "site:actu.abondance.com/*". To the right of the search bar is a "Rechercher" button and links for "Recherche avancée" and "Préférences". Below the search bar, there are radio buttons for "Web" (selected), "Pages francophones", and "Pages : France". The search results section shows "Résultats 1 - 10 sur un total d'environ 1 700 sur actu.abondance.com/* (0,22 secondes)". The first result is titled "Actu Abondance - Actualité des moteurs de recherche et du ..." and includes a description: "Actu Abondance - Retrouvez ici toute l'actualité des moteurs de recherche majeurs et du référencement en France et dans le Monde, jour après jour ...". It also shows the URL "actu.abondance.com/" with a file size of 72k, a date of "Publié depuis 1 heure", and links for "En cache", "Pages similaires", and "À noter". The second result is titled "Actu Abondance: mars 2008 - Actualité des moteurs de recherche et ..." and includes a description: "Actu Abondance: mars 2008 - Les archives de l'actualité des moteurs de recherche majeurs et du référencement en France et dans le Monde, jour après jour ...". It also shows the URL "actu.abondance.com/2008_03_01_archive_actu.html" with a file size of 292k, and links for "En cache", "Pages similaires", and "À noter".

Le nombre de résultats renvoyés par cette syntaxe (notez que la requête "[site:www.votresite.com/](http://www.votresite.com/)&" fonctionne également) s'approche effectivement de celui fourni par AOL et Mozbot. Cependant, Google n'a jamais communiqué sur cette syntaxe spécifique et nous ne vous la fournissons qu'à titre indicatif...

- Une solution plus fiable peut être d'utiliser les Google Webmaster Tools (<http://www.google.com/webmasters/sitemaps/?hl=fr>), espace dédié aux webmasters sur lequel vous devez d'abord avoir ouvert un compte si vous vous intéressez au référencement...

Dans cet espace, choisissez l'option "Liens" qui vous proposera deux choix importants : "**Pages avec des liens externes**", et "**Pages avec des liens internes**" :

The screenshot shows the Google Webmaster Tools interface. At the top, it says "Google Outils pour les webmasters". Below this, there is a navigation bar with "Tableau de bord" and "Liens". The "Liens" section is active. On the left, there is a sidebar with a list of options: "Vue d'ensemble", "Diagnostic", "Statistiques", "Liens", and "Vue d'ensemble" (repeated). The "Liens" option is selected. The main content area is titled "Liens" and shows the domain "actu.abondance.com". Below this, there are two columns of information. The left column is titled "Pages avec des liens externes" and includes the text: "Découvrez les pages de votre site auxquelles il est possible d'accéder par des liens figurant sur d'autres sites." The right column is titled "Pages avec des liens internes" and includes the text: "Découvrez les pages de votre site auxquelles il est possible d'accéder par des liens figurant sur d'autres pages de votre site."

Explorez ces deux zones : il y a de fortes chances pour que les pages qui y sont listées soient celles qui se trouvent dans l'index principal :

Google Outils pour les webmasters

Tableau de bord > Liens > Pages avec des liens internes

Vue d'ensemble

Diagnostic

Statistiques

Liens

Vue d'ensemble

Pages avec des liens externes

Pages avec des liens internes

Liens de site

Sitemaps

Outils

Pages avec des liens internes

actu.abondance.com

Ce tableau fournit une liste de pages de http://actu.abondance.com/ vers lesquelles pointent des liens d'autres pages internes. Cliquez sur un nombre dans la colonne Liens internes pour voir un exemple de liste de liens vers la page correspondante.

[Comment puis-je utiliser ces données ?](#)

1 au 30 sur 1 607 [Suivante](#)

Trouver des informations sur les liens de : Éléments par page : 30

[Afficher les détails](#)

Afficher des pages avec des liens internes provenant de :

Page	Liens internes
Toutes les pages (nombre total de liens)	39 897
http://actu.abondance.com/	1 842
http://actu.abondance.com/2007_08_01_archive_actu.html	530
http://actu.abondance.com/2007_09_01_archive_actu.html	453
http://actu.abondance.com/2007_10_01_archive_actu.html	474
http://actu.abondance.com/2007_11_01_archive_actu.html	200

En effet, Pour répartir les pages d'un site entre ses index principal et secondaire, il semblerait que Google tienne principalement compte des *baklinks* (liens entrants) qui pointent vers elles. Si une page est considérée comme assez "populaire", assez "liée", elle sera plutôt stockée dans l'index principal. Si elle n'a pas assez reçu de liens (internes ou externes), elle sera directement dirigée vers le "purgatoire" de l'index secondaire...

L'avantage des Webmaster Tools de Google est que vous pourrez télécharger la liste de ces pages proposées par l'interface pour webmasters de Google et regarder son contenu de façon exhaustive (alors qu'au travers d'une syntaxe spécifique sur le moteur, comme vu précédemment, Google ne vous renverra toujours que ses 1 000 premiers résultats)... Vous pouvez également utiliser des outils comme GlynC (<http://www.thegooglecache.com/glynC/>) qui rajoute des graphiques temporels aux données fournies par l'interface Google afin de suivre votre stratégie d'obtention de liens dans le temps.

Conclusion

La notion d'index principal et secondaire est importante, voire capitale, en termes de référencement. Le but sera pour votre site web d'avoir le plus possible de pages dans l'index principal... On voit donc ici l'importance du "*deeplinking*" (stratégie d'obtention de liens externes - ou netlinking - vers les pages internes du site) en passant, bien sûr, par des pratiques loyales et honnêtes pour obtenir ces liens... Il faudra bien sûr aussi que tous vos problèmes de "duplicate content" aient également été résolus...

En clair, pour éviter l'index secondaire, vos pages devront :

- Proposer assez de contenu textuel pour être "analysables" par les moteurs (la limite classique des 100 à 200 mots descriptifs comme contenu éditorial au minimum).
- Ne pas connaître de problématique de "duplicate content".
- Être "assez" liées par des liens internes et externes, si possible depuis des pages populaires. Par exemple, le fait de mettre un lien vers de nouvelles pages pendant quelques jours depuis votre page d'accueil peut être une bonne chose, même si cela peut ne pas suffire...

Mettre en place une stratégie de liens vers la **page d'accueil** de son site est une bonne chose pour accroître la popularité de cette dernière. Cela reste un incontournable du référencement. Mais cette stratégie doit s'accompagner d'un travail important pour gagner des liens vers les **pages internes** du site également (*deeplinking*), afin de faire en sorte que le plus possible de pages web se trouvent le plus rapidement possible dans l'index principal, voyant leurs chances de positionnement optimisées... C'est particulièrement crucial pour les sites de contenu (presse, média...) qui ont pour obligation de voir leur pages internes bien référencées, parfois beaucoup plus que pour leur page d'accueil...

Pour en savoir plus :

- *The Ultimate Fate of Supplemental Results*

<http://googlewebmastercentral.blogspot.com/2007/12/ultimate-fate-of-supplemental-results.html>

- *Supplemental goes mainstream*

<http://googlewebmastercentral.blogspot.com/2007/07/supplemental-goes-mainstream.html>

- *Infrastructure status, January 2007*

<http://www.mattcutts.com/blog/infrastructure-status-january-2007/>

- *Supplemental results (vidéo de Matt Cutts)*

<http://video.google.com/videoplay?docid=-3494613828170903728>

- *Supplemental results (aide en ligne de Google)*

<http://www.google.com/support/webmasters/bin/answer.py?answer=73028&query=supplemental+results&topic=&type=>

- *Google - Please Bring Back the Supplemental Results Query; It's Incredibly Valuable*

<http://www.seomoz.org/blog/google-please-bring-back-the-supplemental-results-query-its-incredibly-valuable>

- *What are Google Supplemental Results?*

<http://www.seobook.com/archives/001545.shtml>

- *View All Your Google Supplemental Index Results*

<http://www.seobook.com/archives/002047.shtml>

- *Damned to Google Hell - Supplemental Results*

<http://www.jimboykin.com/damned-to-google-hell-supplemental-results/>

- *Search Engine Size Wars & Google's Supplemental Results*

<http://searchenginewatch.com/3071371>

- *Supplemental Results and Google's Extended Databases*

<http://searchengineland.com/supplemental-results-and-googles-extended-databases-11897.php>

- *Google Dumps The Supplemental Results Label*

<http://searchengineland.com/google-dumps-the-supplemental-results-label-11830.php>

- *Résultats complémentaires Google : nouvelle syntaxe*

<http://s.billard.free.fr/referencement/?2007/07/30/409-resultats-complementaires-google-nouvelle-syntaxe>

Olivier Andrieu

Editeur du site Abondance.com

Réagissez à cet article sur le blog des abonnés d'Abondance :
<http://abonnes.abondance.com/blogpro/2008/11/comment-viter-lenfer-de-lindex.html>