

## L'astuce du mois : Comment éviter le "duplicate content" sur Google ?

[Retour au sommaire de la lettre](#)

*Nous vous proposons chaque mois une série de "trucs et astuces" qui devraient vous simplifier la vie dans le cadre de vos recherches sur le Web. Ce mois-ci : comment éviter que Google ne prenne en compte vos pages - notamment dans le cas d'un site dynamique - comme du "duplicate content", considérant leur contenu comme trop proche...*

Certains sites se voient affligés, sur Google, d'une "pénalisation" due à un phénomène de "duplicate content" (ou "contenu similaire"). En fait, lorsque certaines pages sont trop proches les unes des autres au niveau de leur contenu, Google n'en affiche que quelques-uns avec un lien de ce type pour visualiser les autres :

*"Pour limiter les résultats aux pages les plus pertinentes (total : XX), Google a ignoré certaines pages à contenu similaire.*

*Si vous le souhaitez, vous pouvez relancer la recherche en incluant les pages ignorées."*

C'était, par exemple, le cas de notre site "Boutique Abondance" (<http://www.boutique-abondance.com/>) lors de la mise en place de sa dernière version fin 2004 :

The screenshot shows a Google search interface. The search bar contains 'site:boutique-abondance.com'. Below the search bar, it says 'Rechercher dans : Web Pages francophones Pages : France'. The search results show 'Web Résultats 1 - 2 sur un total d'environ 19 provenant de boutique-abondance.com'. The first result is 'La Boutique Abondance / article : Lettres d'information, livres ...' with a description: 'Lettres d'information, livres, logiciels et études sur les moteurs de recherche et le référencement. Tous nos produits...'. Below the result, there is a link to 'boutique-abondance.com/' and a warning: 'Pour limiter les résultats aux pages les plus pertinentes (total : 2), Google a ignoré certaines pages à contenu similaire. Si vous le souhaitez, vous pouvez relancer la recherche en incluant les pages ignorées.'

Sur la requête site:boutique-abondance.com, seule une page (sur 19) était affichée sous une forme complète par Google, les autres étant considérées comme similaires ("duplicate content")... (Rappelons que le deuxième résultat affiché sur cette copie d'écran - ayant pour titre "boutique-abondance.com" - correspond à une page non encore indexée par Google mais "connue" du moteur, un lien dans son index pointant vers cette adresse). La syntaxe site:, sur Google, est très utile pour détecter les pages considérées comme similaires par le moteur de recherche.

Cette situation est dommageable pour plusieurs raisons :

- Les différentes pages, présentant pourtant parfois un contenu différent, sont considérées comme similaires. Nous verrons plus loin pourquoi.

- L'expérience et les tests que nous avons effectués pour plusieurs sites nous poussent à estimer que les pages considérées comme "duplicate content" sont le plus souvent moins bien classées que les autres par Google. Le positionnement de vos pages est donc d'autant plus difficile que ce phénomène est important pour votre site.

Pourquoi, alors, toutes les pages du site ont-elles été considérées comme "duplicate content" par Google ? En cliquant sur le lien "relancer la recherche en incluant les pages ignorées", on obtient plus d'informations à ce sujet :

Web Images Groupes Annuaire Actualités

Google

site:boutique-abondance.com

Rechercher Recherche avancée Préférences

Rechercher dans : Web Pages francophones Pages : France

Web Résultats 1 - 10 sur un total d'environ 19 provenant de boutique-abc

[La Boutique Abondance / article : Lettres d'information, livres ...](#)  
Lettres d'information, livres, logiciels et études sur les moteurs de recherche et le référencement. Tous nos produits...  
[www.boutique-abondance.com/cgi-bin/pg-shoppro.cgi?ORD=viewproduct&id\\_product=9&id\\_category=9](#) - 12k - [En cache](#) - [Pages similaires](#)

[La Boutique Abondance / article : Lettres d'information, livres ...](#)  
Lettres d'information, livres, logiciels et études sur les moteurs de recherche et le référencement. Tous nos produits...  
[www.boutique-abondance.com/cgi-bin/pg-shoppro.cgi?ORD=viewproduct&id\\_product=1&id\\_category=1](#) - 12k - [En cache](#) - [Pages similaires](#)

[La Boutique Abondance : Lettres d'information, livres, logiciels ...](#)  
Lettres d'information, livres, logiciels et études sur les moteurs de recherche et le référencement. Tous nos produits...  
[www.boutique-abondance.com/](#) - 12k - 4 jan 2005 - [En cache](#) - [Pages similaires](#)

[La Boutique Abondance / catégorie : Lettres d'information, livres ...](#)  
Lettres d'information, livres, logiciels et études sur les moteurs de recherche et le référencement. Tous nos produits...  
[www.boutique-abondance.com/cgi-bin/pg-shoppro.cgi?ORD=viewcat&id\\_category=2](#) - 11k - [En cache](#) - [Pages similaires](#)

[La Boutique Abondance / catégorie : Lettres d'information, livres ...](#)  
Lettres d'information, livres, logiciels et études sur les moteurs de recherche et le référencement. Tous nos produits...  
[www.boutique-abondance.com/cgi-bin/pg-shoppro.cgi?ORD=viewcat&id\\_category=9](#) - 14k - [En cache](#) - [Pages similaires](#)

[La Boutique Abondance - livres et logiciels sur les moteurs de ...](#)  
Lettres d'information, livres, logiciels et études sur les moteurs de recherche et le référencement. Tous nos produits...  
[www.boutique-abondance.com/on-a-aime.html](#) - 9k - 5 jan 2005 - [En cache](#) - [Pages similaires](#)

[La Boutique Abondance / recherche : Lettres d'information, livres ...](#)  
Lettres d'information, livres, logiciels et études sur les moteurs de recherche et le référencement. Tous nos produits...  
[www.boutique-abondance.com/cgi-bin/pg-shoppro.cgi?ORD=search](#) - 10k - [En cache](#) - [Pages similaires](#)

[La Boutique Abondance / coordonnées : Lettres d'information ...](#)  
Lettres d'information, livres, logiciels et études sur les moteurs de recherche et le référencement. Tous nos produits...  
[www.boutique-abondance.com/cgi-bin/pg-shoppro.cgi?ORD=checkout](#) - 10k - [En cache](#) - [Pages similaires](#)

Chaque page, si elle avait un contenu différent, proposait quasiment le même titre (site dynamique en cours de création oblige) et le même descriptif (absence de balise "meta description"). Or il semblerait que Google prenne en compte, dans son algorithme de détection du "duplicate content", le début des pages (mais pas seulement comme nous le verrons ci-après) et plus précisément de leur code source. Si ces débuts sont trop proches les uns des autres, les pages sont considérées comme similaires et donc classées en "duplicate content".

Comment, alors, faire en sorte que la situation évolue ? Voici trois conseils essentiels si vous êtes "victime", comme de nombreux autres sites, de ce phénomène...

### 1. Affichez un titre réellement différent par page

Chaque page de votre site doit afficher un titre très différent afin d'être différenciée par les algorithmes de "duplicate content" du moteur. Notez bien que les titres doivent être assez différents pour être facilement différenciés. En effet, si vous regardez attentivement la copie d'écran ci-dessus, vous vous apercevrez que la plupart des titres sont proches, mais pas exactement équivalents. Pourtant, les pages ont été considérées comme étant du "duplicate content"... De plus, contrairement à ce que certains pourraient croire, proposer des titres différents et différenciés ne suffit pas pour éliminer tout risque de "duplicate content". Les autres conseils ci-dessous sont également essentiels. Tenez-en compte.

Pour notre site, nous avons inclus, dans chaque titre, le nom de la catégorie ou du produit présenté sur chaque page.

Exemple :

*Page d'accueil (titre générique) :*

"La Boutique Abondance : Lettres d'information, livres, logiciels et études sur les moteurs de recherche et le référencement"

*Exemple de titre pour une catégorie :*

"La Boutique Abondance - Lettre 'Recherche et Référencement' - Premier Abonnement"

*Exemple de titre pour un produit :*

"La Boutique Abondance - Abonnement 'Fidélité' (24 mois) à la lettre professionnelle 'R&R' du site Abondance"

Chaque page a ainsi un titre très différent et différenciant. Cette première phase était achevée. Le site étant dynamique, cette étape a été particulièrement simple à mettre en œuvre (simple insertion d'un paramètre correspondant au nom de la catégorie ou du produit dans le titre)... Comme quoi, les sites dynamiques ont également de bons côtés pour le référencement ;-)

## **2. Affichez une balise "meta description" différente par page**

Même chose que pour le titre des pages (voir ci-dessus) : insérez dans vos pages une balise "meta description" différente pour chaque page. Pour notre part, nous avons décliné dans chaque balise "description" le titre de la page en insérant, de la même façon, le paramètre correspondant au nom de la catégorie ou du produit au niveau du site dynamique et le tour était joué en quelques minutes...

Notez également que la balise "meta keywords" ne semble pas avoir d'incidence dans ce cas sur Google (on s'en doutait un peu...)

## **3. Réduisez au maximum le code similaire en haut de page dans votre source**

Plusieurs rumeurs, de source sérieuse, indiquent que Google prend en compte, pour détecter le "duplicate content", non seulement le nombre de caractères du début des pages, mais également un certain nombre de lignes à partir du début du code source (ce qui peut sembler - à raison - assez étonnant...)...

Bref, il vous faut, autant que possible :

- Proposer un début de code HTML différent pour chacune des pages de votre site. Plus ce "début de code" sera différent, moins vos pages auront de chances (ou de malchance) d'être considérées comme du "duplicate content". Difficile de dire, en revanche, combien de caractères sont pris en compte par Google... Petite astuce : vous pouvez également ajouter la date, notamment pour les sites dynamiques, en haut de page :-).
- Eliminer les "retours chariot" au maximum afin de "contracter" le plus possible le début du code pour que la zone de contenu réellement différenciante de votre page soit accessible à la ligne la plus haute possible du fichier-source.
- Prendre en compte les deux conseils ci-dessus également pour le bas de page, car certaines sources font également état du fait que Google ne prendrait pas uniquement en compte le début des pages mais également d'autres zones de la page, et notamment le bas du document HTML pour distinguer les pages dupliquées.

## **Résultat sur le site "Boutique Abondance"**

Nous avons appliqué les quelques conseils ci-dessus à notre site "Boutique Abondance". Résultats, quelques jours plus tard :

- Google affichait, pour la même requête, beaucoup plus de pages maintenant considérées comme différentes. La contrainte du "duplicate content" avait en grande partie "sauté".
- Bonne nouvelle également : le moteur de recherche avait indexé un nombre plus important de pages.
- Deuxième bonne nouvelle : le site est maintenant beaucoup mieux positionné sur la plupart des requêtes en rapport avec les produits et services vendus...

The screenshot shows a Google search interface. At the top, there are navigation links: Web, Images, Groupes, Annuaire, and Actualités. The search bar contains the text 'site:boutique-abondance.com'. Below the search bar, it says 'Rechercher dans : Web Pages francophones Pages : Fr'. The search results are displayed under the heading 'Web' and show 'Résultats 1 - 10 sur un total d''.

The first result is titled 'La Boutique Abondance - Réabonnement 'Fidélité' (24 mois) à la ...'. The description reads: 'La Boutique Abondance - Réabonnement 'Fidélité' (24 mois) à la lettre professionnelle 'R&R' du site Abondance - moteurs de recherche et référencement.' The URL is 'www.boutique-abondance.com/cgi-bin/pg-shoppro.cgi?ORD=viewproduct&id\_product=9&id\_category...'. The second result is titled 'La Boutique Abondance - Abonnement 'Découverte' (6 mois) à la ...'. The description reads: 'La Boutique Abondance - Abonnement 'Découverte' (6 mois) à la lettre professionnelle 'R&R' du site Abondance - moteurs de recherche et référencement.' The URL is 'www.boutique-abondance.com/cgi-bin/pg-shoppro.cgi?ORD=viewproduct&id\_product=1&id\_category...'. The third result is titled 'La Boutique Abondance : Lettres d'information, livres, logiciels ...'. The description reads: 'La Boutique Abondance : Lettres d'information, livres, logiciels et études sur les moteurs de recherche et le référencement.' The URL is 'www.boutique-abondance.com/' - 12k - 8 jan 2005 - En cache - Pages similaires. The fourth result is titled 'La Boutique Abondance - Etudes'. The description reads: 'La Boutique Abondance - Etudes - moteurs de recherche et référencement.' The URL is 'www.boutique-abondance.com/cgi-bin/pg-shoppro.cgi?ORD=viewcat&id\_category=2 - 11k - En cache...'. The fifth result is titled 'La Boutique Abondance - Nous avons aimé...'. The description reads: 'La Boutique Abondance - Nous avons aimé... - moteurs de recherche et référencement.' The URL is 'www.boutique-abondance.com/cgi-bin/pg-shoppro.cgi?ORD=viewcat&id\_category=4 - 10k - En cache...'. The sixth result is titled 'La Boutique Abondance - Lettre 'Recherche et Référencement' ...'. The description reads: 'La Boutique Abondance - Lettre 'Recherche et Référencement' - Réabonnement - moteurs de recherche et référencement.' The URL is 'www.boutique-abondance.com/cgi-bin/pg-shoppro.cgi?ORD=viewcat&id\_category=9 - 14k - En cache...'. The seventh result is titled 'La Boutique Abondance - livres et logiciels sur les moteurs de ...'. The description reads: 'La Boutique Abondance - tous les livres et tous les logiciels sur les moteurs de recherche, le référencement et la recherche d'information.' The URL is 'www.boutique-abondance.com/on-a-aime.html - 9k - 8 jan 2005 - En cache - Pages similaires'. The eighth result is titled 'La Boutique Abondance - livres sur les moteurs de recherche, le ...'. The description reads: 'Lettres d'information, livres, logiciels et études sur les moteurs de recherche et le référencement. Tous nos produits... ...'. The URL is 'www.abondance.com/livre/ - 35k - En cache - Pages similaires'.

Conclusion : si votre site est "victime" de "duplicate content" par Google, n'hésitez à retravailler le début - notamment - du code source de vos pages pour les différencier le plus possible, les quelques minutes à effectuer ce travail en valent réellement la chandelle...

Cette optimisation est souvent intéressante pour des sites comme des forums Web qui fonctionnent sur la base d'une page "template" dont le contenu est modifié dynamiquement... Nous avons appliqué la même "recette" sur le site "Forums Abondance" (<http://www.forums-abondance.com/>) en indiquant dans le titre et la balise meta description de chaque page l'intitulé de la discussion en cours, avec un succès intéressant (la plupart des pages de ce site étaient, auparavant, considérées comme du "duplicate content" par Google). Autant de visibilité et de trafic gagnés en quelques minutes pour le forum avec plus de 1600 pages indexées :

Web Images Groupes Annuaire Actualités

Google site:forums-abondance.com Re

Rechercher dans : Web Pages francophones Pages : France

---

**Web** Résultats 1 - 10 sur un total d'environ

[Mon index n apparait pas dans la recherche. \(Référencement de ...](#)  
Mon index n apparait pas dans la recherche. -  
Référencement de site web - Forum Abondance.  
[www.forums-abondance.com/referencement/open.php?mess=3113](http://www.forums-abondance.com/referencement/open.php?mess=3113) - 10k - Résultat complémentaire - [En ca](#)

[VOILA A CHANGE \(Référencement de site web - Forum Abondance\)](#)  
VOILA A CHANGE - Référencement de site web - Forum Abondance.  
[www.forums-abondance.com/referencement/open.php?mess=3119](http://www.forums-abondance.com/referencement/open.php?mess=3119) - 10k - Résultat complémentaire - [En ca](#)

[recherche de logiciels \(Recherche d'information et veille - Forum ...](#)  
Abondance : recherche d'information, référencement et promotion  
de sites Web : recherche de logiciels.  
[www.forums-abondance.com/recherche/open.php?mess=389](http://www.forums-abondance.com/recherche/open.php?mess=389) - 12k - Résultat complémentaire - [En cache](#) -

[Anciennes versions explorer \(Référencement de site web - Forum ...](#)  
Anciennes versions explorer - Référencement de site web - Forum Abondance.  
[www.forums-abondance.com/referencement/open.php?mess=3027](http://www.forums-abondance.com/referencement/open.php?mess=3027) - 13k - Résultat complémentaire - [En ca](#)

[referencement bilingue spécifique \(Référencement de site web ...](#)  
Abondance, recherche d'information, référencement et promotion  
de sites Web : referencement bilingue spécifique.  
[www.forums-abondance.com/referencement/open.php?mess=2338](http://www.forums-abondance.com/referencement/open.php?mess=2338) - 11k - Résultat complémentaire - [En ca](#)

[Nom des url des pages \(Référencement de site web - Forum ...](#)  
Nom des url des pages - Référencement de site web - Forum Abondance.  
[www.forums-abondance.com/referencement/open.php?mess=3059](http://www.forums-abondance.com/referencement/open.php?mess=3059) - 12k - Résultat complémentaire - [En ca](#)

[Tableau dans des tableaux \(Référencement de site web - Forum ...](#)  
Tableau dans des tableaux - Référencement de site web - Forum Abondance.  
[www.forums-abondance.com/referencement/open.php?mess=2287](http://www.forums-abondance.com/referencement/open.php?mess=2287) - 13k - Résultat complémentaire - [En ca](#)