

Cours de référencement : optimisation

Mise à jour du 20 octobre 2017

Rémi Bachelet

Dernière version de ce cours est ici [optimisation du référencement](#)

Cette formation est également enregistrée [en vidéo](#)

Cours distribué sous licence **Creative Commons**, selon les conditions suivantes :



Optimisation “Onpage”

- 1. Texte : visible sur la page**
- 2. Code : non visible sur la page**

Cours de gestion de projet
[gestiondeprojet.pm/](#) ▼
Cours de gestion de projet : 8 h de formations en vidéo, 18 cours téléchargeables en ppt, pdf. 9 modèles de documents (compte-rendu, Cahier des charges...)
[Le MOOC Gestion de projet - Management de projet - Organisation de projet](#)

- Le titre est :
 - L'« image de marque » qui apparaît sur le snippet de la page de résultats de recherche.
 - C'est le texte du lien,
 - Le reste de la description étant le meta tag meta name="description" .. ou sinon les premiers mots de la page, sa date..
 - Le mot-clé le plus fort pour le robot indexeur
- Entre 7 et 10 mots
 - Éviter les termes « tarte à la crème » ("page", "bienvenue"...)
 - Un titre propre à chaque page
- Le titre doit reprendre les mots clés importants
 - Il donne un résumé de la page (éviter les pages multi thèmes)
 - De plus en plus détaillé quand on descend dans l'arborescence du site
 - Pas de titres/pages multilingues

Rédaction du texte

(voir aussi le chapitre sur les mots-clés)

Le cloaking (texte caché visible pour le robot indexeur, mais invisible par le lecteur) a suscité beaucoup de tentatives pour leurrer les moteurs de recherche

- Techniques : Meta tags, identification des images, texte blanc sur fond blanc, caché derrière une image, balise <div>, noframes, noscript, display:none ..

Le code : *Meta Tags*

Meta = des “données sur les données”

```
<HEAD>
  <TITLE>titre de la page</TITLE>
  <META name="description" content="description du contenu, snippet">
  <META name="keywords" content="mots clés séparés par des espaces">
</HEAD>
```

Balise des images

- `<IMG SRC="url_de_l_image"`
- `ALT="Texte remplaçant l'image si elle ne s'affiche pas"` =mot-clé pour google Images"
- `TITLE="Texte à afficher"`, =mot-clé pour google Image">

- `<meta name="robots" content="valeurs">`
 - all : Le robot indexe tout (= défaut)
 - none : Le robot n'indexe rien du tout

 - noindex : Votre page n'est pas indexée mais les liens sont suivis
 - nofollow *pour toute une page ou comme attribut de lien* : voir diapo suivante

rel="nofollow"

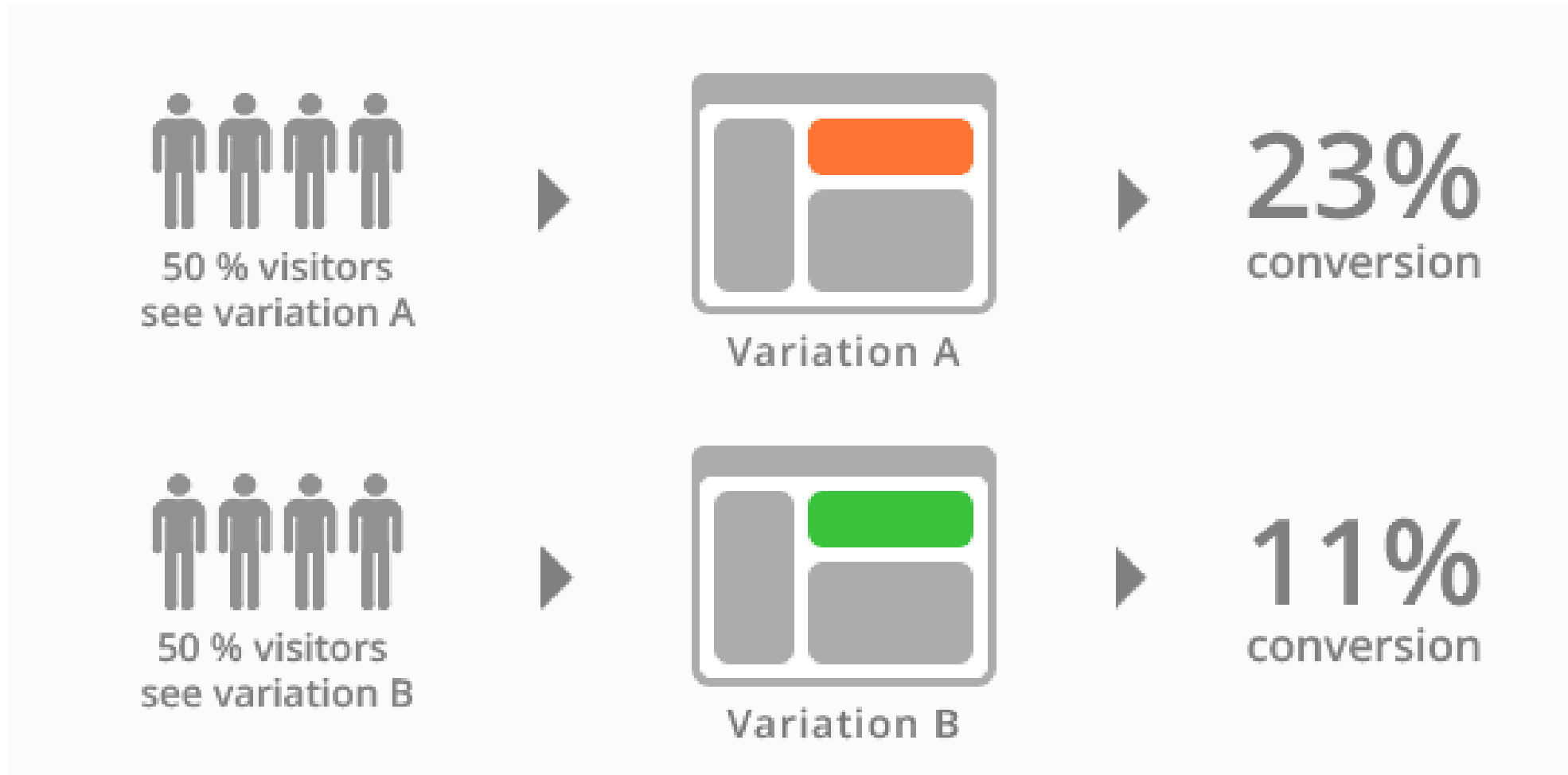
- Permet au webmestre de spécifier que ce lien ne doit pas compter dans le PageRank
 - Balise inventée en 2005, notamment par un "responsable antispam" de Google
 - Les autres moteurs de recherche se sont ralliés au système
 - Devrait plutôt s'appeler rel="nopagerank" , car le lien est suivi par les indexeurs
 - L'utilisation de cette balise a donné lieu à des tentatives de manipulation
- Prévient le *spamdexing*, notamment pour les sites :
 - Ouverts aux commentaires
 - Forums de discussion, commentaires de blogs
 - De crowdsourcing
 - Wikis, Twitter, réseaux sociaux...



This page links to another page using rel=nofollow

This page receives no credit for the link

A/B testing



A/B testing => optimisation de page multivariables

Ou Multivariate landing page optimization

- On teste divers éléments visibles d'une page
 - x Images différentes, y dispositions du texte, z formulations
 - On obtient $x*y*z$ combinaisons possibles
 - On évalue leur taux de conversion
 - Google Website Optimizer + Plan d'expérience

Optimisation “Onsite”

1. Structure du site
2. Fichiers spéciaux
 - *Robots.txt*,
 - *Sitemap.xml*
 - *.htaccess*
3. Choix techniques

Eventail d'opérations d'optimisation d'un site

- Règle des "3C" :
 - Contenu
 - Code
 - Conception,

Structure du site

- Ergonomie = Facilité de lecture du site pour un internaute
 - Plan du site
 - Navigation
 - Breadcrumbs / fil d'Ariane
France > Normandie > Calvados > Deauville
- Facilité de lecture du site pour un indexeur
 - Fonction "Analyser comme Googlebot«
- Intérêt des CMS pour mettre à jour la structure
 - Wordpress, Joomla, Magento...

Fichiers spéciaux

- Robots.txt
 - Spécifie aux robots indexeurs le comportement qu'il doivent adopter
 - Indexer ou pas, suivre les liens etc..
- Sitemap.xml
 - Cartes sémantisée des pages du site à visiter

```
<loc>http://www.example.com/?id=what</loc>
<lastmod>2009-09-22</lastmod>
<changefreq>monthly</changefreq>
<priority>0.5</priority>
```
- Fichier d'identification
 - Pour valider par Google qu'on est bien le responsable d'un site.

Que faire en cas de changement d'adresse d'une page ?

- Installer un fichier .htaccess à la racine du site
 - [How to setup a 301 Redirect - Htaccess](#)
 - Redirect
 - 301 ("*permanently moved* »)
 - 302 ("moved temporarily)
 - 303 ("see other »)

Choix techniques

- Réécriture d'URL
 - <http://www.discounteo.com/tv-video/televiseur/tv-lcd>
 - <http://www.cdiscalcount.com/electromenager/four-cuisson/samsung-bq1d4t007-four/f-1102304-samsusbq1d4t007.html>
- Cas du Flash
 - Mauvaise réputation, en voie de disparition, mais indexé par Google dans certains cas.
- Cas des sites dynamiques
 - Facteurs pénalisants et de leurs solutions (frames, formulaires, AJAX, JavaScript, sessions, contenus dupliqués, etc.)

Extensions utiles pour auditer un site

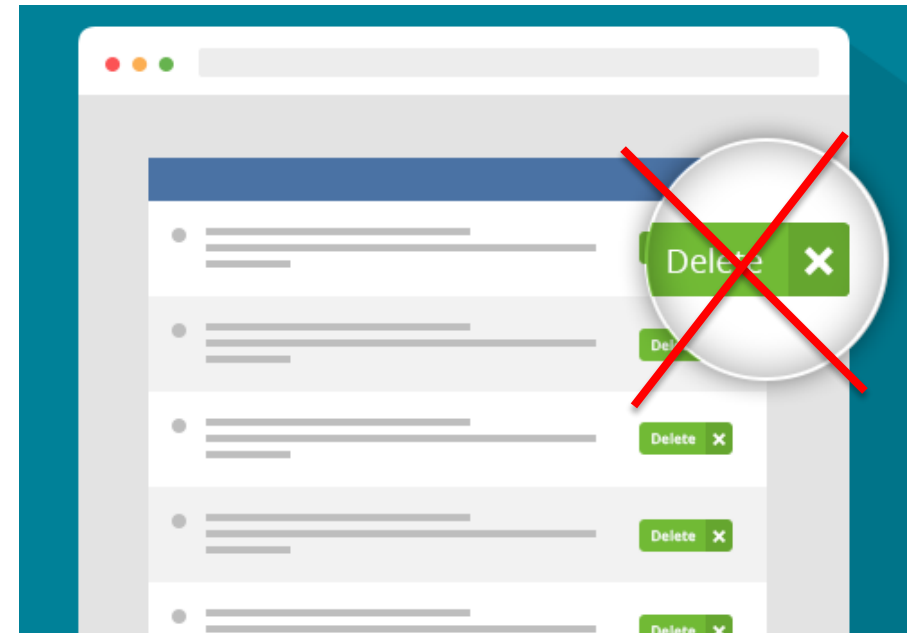
- Chrome
 - SEOquake
- Firefox
- Firebug
 - Vérifie l'exécution du code, notamment pour les outils de web analytics en java
- IE tab
 - Aspect de la page sous IE
- .. et beaucoup d'autres...

Mais le plus important c'est ...

CONTENT IS KING

« *le contenu c'est ce qui compte avant tout* »

Contenu = texte, mais aussi les médias (images, vidéo), la facilité de navigation, l'ergonomie/affordance, l'interactivité, *au global l'expérience du site..*



Symptômes d'une faiblesse de contenu :

- Taux de rebond = les internautes ne restent pas sur le site
- % élevé de nouveaux visiteurs = pas de retour
- Peu d'entrées directes (pas de mémorisation du site en marque-pages)
- Taux de conversion inférieur aux attentes
- ...

Optimisation “Offsite” / Netlinking

- **Le site dans l'écosystème du web**
 - **Backlinks**
 - **Social bookmarking**



CONTENT IS KING



LINKS ARE QUEEN

Les backlinks

- Les liens vers votre site sont le facteur le plus important du référencement naturel
 - Sauf via des sites de spam, de TrustRank faible...
- Ces liens sont plus efficaces s'ils viennent de sites
 - Couvrant le même domaine
 - À fort PageRank
 - Échange de liens
 - Plus efficace s'ils ne sont pas directement réciproques



Efficacité des liens

Les mots-clés dans les liens sont plus ou moins pondérés

– Moins efficace :

- Cours de référencement `ici`.

– Efficace :

- on y propose un ` cours de référencement `.

– Encore plus efficace :

- on y parle de la ` cours de référencement `.

Le Netlinking

- Rechercher des partenaires pour augmenter le ranking du site.
 - Cet aspect prend de plus en plus d'importance, la gestion des partenariats est un élément clé pour ce poste.
 - Les fonctions du référenceur s'approchent ici de celles de *community manager*
- Types de liens :
 1. Liens de ressources : pointent vers des ressources utiles
 2. Liens réciproques : p.e. le site A pointe vers B, qui pointe vers A en échange. Des annuaires d'échanges de liens peuvent aider à la mise en contact.
 3. Liens de spam : p.e. p.e. Un robot insère des commentaires dans les site web 2.0 pour augmenter le PageRank d'un site (blogs, wikis..). *Stratégie neutralisée par les balises nofollow.*
 4. Liens internes automatiques : p.e. toutes les pages du site affichent un lien vers la page principale.

La déclaration aux moteurs de recherche

- **Ajouter l'URL de votre site à Google**
 - Seule la page de premier niveau d'un site est à soumettre (« *to submit* »)
 - Il n'est pas nécessaire d'indiquer les liens mis à jour ou obsolètes. Les liens désactivés disparaissent au cours de l'exploration suivante
 - Voir aussi dans le chapitre "outils"
 - ... Google Search Console.

White Hat ou Black Hat ?

White Hat : Contenu de qualité, site agréable et pertinent pour l'utilisateur

- Optimisation des mots-clés raisonnable
- Stratégie de netlinking honnête (=non rémunérée)
- Ne pas exiger de résultats immédiats (effet sandbox)
- Rien à cacher

Rien qui cherche à tromper l'internaute ou les moteurs de recherche



Black Hat : Manipuler les moteurs de recherche... au risque du bannissement des résultats de recherche

- **Occultage** (transmettre une page différente au moteur de recherche et au visiteur)
- **Texte caché** (Cloaking)
- **Répéter abusivement un mot-clé, mots-clés sans rapport avec le contenu réel** (Keyword stuffing)
- **Pages en double / contenu copié-collé d'autres sites** (Scraper site)
- **Achat abusifs de noms de domaine** (Cybersquatting, Domain spam)
- **Utilisation ciblée des balises nofollow** (PageRank sculpting).
- **Globalement, le spamdexing**

Voir également le lexique en fin de cours...

Le Web 2.0 : nouvelle frontière du SEO ?



Web 2.0 websites and tools such as social networking sites, blogs, wikis, video sharing sites, hosted services, web applications, [mashups](#) and [folksonomies](#) may allow users to interact and collaborate with each other in a [social media](#) dialogue as creators of [user-generated content](#) in a [virtual community](#), in contrast to websites where people are limited to the passive viewing of [content](#).

Questions ?

- Rôle des URL raccourcies de Google

Les thèmes et chapitres du cours

1. Origine du SEO, Google ... et ses concurrents
2. La fréquentation d'un site : les fondamentaux
3. Dans la peau d'un moteur de recherche : le PageRank
4. Web Analytics et liens commerciaux
5. Optimisation du référencement
 - “Onpage”
 - “Onsite”
 - “Offsite”
6. Trouver et optimiser les mots-clés
7. Méthodologie de référencement et avenir du SEO