

Tutorial Référencement naturel sur moteurs (Google...)

par Gilles GRÉGOIRE ([Referencer et positionner en TOP10 à 30 sur Google](#)) (Blog)

Date de publication : juillet 2007

Dernière mise à jour :

Initiation, pour dirigeants et décideurs - Les bases.

Un zeste plus de technique : le référencement passif - On ne modifie que le site Internet.

Le référencement actif - On ne touche pas au site Internet lors de cette phase.

I - Initiation et compréhension.....	5
I-A - Un peu d'histoire.....	5
I-A-1 - La paranoïa en système de conduite d'entreprise.....	5
I-A-2 - Bien comprendre l'objectif central de Google.....	6
I-B - Pourquoi un référencement ? Le changement de comportement des clients et des prospects.....	6
I-C - Pourquoi référencer son site en mode naturel ?.....	7
I-D - Initiation au référencement en termes simples.....	8
I-D-1 - Étape #1 - Mots et expressions clefs - La mission du marketeur.....	8
I-D-2 - Étape #2 - Accessibilité technique & sémantique du site - Les contenus voulus - Le rôle de la technique et du marketeur.....	8
I-D-3 - Étape #3 - Être bien positionné.....	9
I-D-4 - Étape #4 - Maintenir son référencement & quelques trucs.....	9
I-E - Adwords et autres liens commerciaux versus référencement naturel.....	9
I-F - Rappel d'évidences.....	10
I-F-1 - Inutile de référencer un site Internet touffu.....	10
I-F-2 - Inutile de référencer des mots clefs inutilisés.....	10
I-G - Définitions.....	10
I-G-1 - Les termes simples.....	10
I-G-2 - Page Rank.....	11
I-G-2-a - Objectif recherché.....	11
I-G-2-b - Une définition plus complète.....	12
I-G-2-c - La formule.....	12
I-G-3 - Important, ne pas confondre PageRank et positionnement !.....	14
II - Le référencement passif.....	14
II-A - Architecture et accessibilité technique.....	14
II-A-1 - Votre site, s'il existe, est-il visible par Google ?.....	14
II-A-2 - Un peu de technique HTML, XHTML et CSS2 - Le minimum nécessaire.....	14
II-A-3 - La liste des interdits et les problèmes.....	15
II-A-3-a - Le robot Google doit pouvoir suivre les liens cliqués.....	15
II-A-3-b - Pas de site entièrement en Flash.....	15
II-A-3-c - Pas de frames si possible.....	15
II-A-3-d - Pas de pages dynamiques ou alors attention à ce que l'on fait.....	15
II-A-4 - Contourner les interdits techniques.....	17
II-A-4-a - Il faut contourner les interdictions ou les contraintes.....	17
II-A-4-b - Le Flash.....	17
II-A-4-b-i - Objet entièrement en Flash.....	17
II-A-4-b-ii - Menu principal en Flash.....	18
II-A-4-b-iii - Pourquoi tant de problèmes avec Flash alors que PDF est indexé par Google.....	18
II-A-4-b-iv - Insérer un objet Flash dans une page HTML.....	18
II-A-4-c - Les pages dynamiques : Ré écriture d'URL.....	18
II-A-4-d - La gestion des sessions.....	20
II-A-4-e - La gestion de session par cookie.....	20
II-A-5 - Une source, multi sources.....	21
II-A-5-a - Les sources.....	21
II-A-5-b - Avantages & inconvénients.....	22
II-B - Architecture et accessibilité sémantique.....	22
II-B-1 - Préambule.....	22
II-B-2 - Identifier les bons mots clefs (français).....	23
II-B-2-a - Un outil intéressant : les mots clefs retenus par Google sur une page donnée.....	24
II-B-3 - Identifier les bons mots clefs (autres langues).....	25
II-B-4 - Combinatoire et choix des mots clefs.....	26
II-B-5 - "Forcer" Google sur vos mots clefs et non l'inverse.....	26
II-B-6 - Augmenter le trafic grâce à la combinatoire.....	27
II-B-7 - Densité, occurrence, racine, thème... Une approche par l'exemple.....	28
II-B-7-a - Les textes sont analysés par Google sous différents angles.....	28
II-B-7-b - Écrire un texte, 100 mots environ, prenant en compte des mots clefs stratégiques.....	28
II-B-8 - Le poids des noms de sources, de fichiers HTML.....	29
II-B-9 - Les autres critères.....	30

II-B-9-a - Les critères importants.....	30
II-B-9-b - Les critères moins importants.....	31
II-B-9-c - Les critères surprenants.....	31
II-B-9-d - Plan du site.....	31
II-B-9-e - Localisation géographique du serveur.....	31
II-B-9-f - Google writing.....	31
II-B-9-g - Les claques DHTML.....	33
II-B-9-h - Augmenter le socle du référencement => Création de trafic qualifié et ciblé.....	33
II-B-9-i - Mettez votre site à jour.....	33
II-B-9-j - Un JPEG à mettre en zoom.....	34
II-B-9-k - Gérer les pages trop longues en texte.....	34
II-B-10 - Les mots clefs utilisés par vos concurrents... Ou par vous-même !.....	34
II-C - htaccess et canoniser un point d'entrée.....	34
II-D - Héritage, refonte de site, migration.....	35
II-D-1 - Protéger ses investissements.....	35
II-D-2 - .htaccess.....	35
II-D-3 - Liens en dur.....	35
II-D-4 - Redirection en HTML.....	36
II-E - Illustrations des calculs de PR (Page Ranking).....	36
II-E-1 - Quel intérêt à comprendre puis à mettre en œuvre les calculs de PR ?.....	36
II-E-2 - Prenons le Web dans un cas théorique : deux pages existent - A pointe vers B et B pointe vers A.....	36
II-E-3 - L'efficacité de la formule pour accélérer les calculs de PR d'une page.....	38
II-E-4 - Un exemple avec plus de pages.....	38
II-E-5 - Confirmation.....	39
II-E-6 - Ce que nous venons de voir est partiellement faux - Le test.....	39
III - Le référencement actif.....	40
III-A - Le principe.....	40
III-A-1 - Un bombing.....	40
III-A-2 - Utiliser les BL.....	41
III-B - Bien construire un backlink.....	41
III-C - Cohérence contextuelle.....	41
III-C-1 - Explication globale simplifiée.....	41
III-C-2 - Explications plus complètes sur les BL.....	42
III-C-2-a - Dans l'espace.....	42
III-C-2-b - Dans le temps.....	43
III-C-3 - Les backlinks internes.....	43
III-D - Les limites du système des BL.....	43
III-D-1 - Pièges dans l'espace.....	43
III-D-2 - Pièges dans le temps.....	44
III-D-3 - Conclusion.....	45
III-E - Échanges de liens pour pousser des BL vers vous et les autres.....	45
III-E-1 - Que du PR.....	45
III-E-2 - Mais aussi des mots clefs.....	45
III-E-3 - Les différents types de BL.....	45
III-E-4 - Combien cela coûte-t-il ?.....	46
III-E-4-a - En temps homme.....	46
III-E-4-b - Le site loueur de backlinks.....	46
III-E-4-c - Le locataire de liens.....	47
III-F - Améliorer le positionnement - Les BL internes.....	47
III-G - Améliorer son Page Rank : le Web marketing relationnel.....	47
III-G-1 - Troc de liens entre sites ou entre sources.....	47
III-G-1-a - Lien échangés.....	47
III-G-1-b - Quelquefois, les liens disposés dans des messages de forum.....	47
III-G-1-c - Disposer de pub (payante).....	48
III-G-2 - Conséquences.....	48
III-G-3 - Mesurer le retour sur investissement.....	48
III-G-3-a - Rythme de passage de Google bot sur votre site.....	48

III-G-3-b - Gérer les pages trop longues en texte.....	49
III-H - Les outils.....	49
III-H-1 - Google Analytics - Survol.....	49
III-H-1-a - Mise en œuvre.....	49
III-H-1-b - Informations clefs liées au référencement uniquement.....	49
III-H-1-c - Interprétation.....	51
III-H-2 - Google Webmestre.....	51
III-H-2-a - Mise en œuvre.....	51
III-H-2-b - Utilisation.....	51
III-H-2-c - Interprétation.....	52
III-H-3 - Google AdWords.....	52
III-H-3-a - Introduction.....	52
III-H-3-b - Mise en œuvre.....	52
III-H-3-c - Utilisation.....	52
III-H-3-d - Interprétation.....	54
III-H-4 - Google Sitemap.....	55
III-I - Informatiser et automatiser le référencement : niveau 1, le plus simple.....	55
III-I-1 - Chaque page.....	55
III-I-1-a - Un passage d'argument - Cela donne quelque chose du genre.....	56
III-I-1-b - Des descriptions title etc. en table MySQL.....	56
III-I-2 - Liste des éléments à automatiser.....	56
III-J - Informatiser et automatiser le référencement : niveau 2, expert.....	56
IV - Approche résumée et simplifiée.....	57

I - Initiation et compréhension

I-A - Un peu d'histoire

I-A-1 - La paranoïa en système de conduite d'entreprise

Google est une entreprise US paranoïaque ne donnant aucune ou presque indication sur son fonctionnement technique.

Pourquoi ? Outre la théorie de Andy Grove (le fondateur de Intel) il faut se pencher sur la jeune histoire de la recherche d'informations sur le Net pour comprendre.

Les moteurs (Altavista essentiellement) dominent la recherche d'informations sur la première génération du WEB commercial (jusqu'en 1998)

Ces moteurs donnent des résultats de plus en plus approximatifs avec l'arrivée massive et exponentielle de sites web au fil du temps.

Les annuaires prennent le relais et balayent les moteurs en moins de un an (Yahoo !) en 1998/1999

Les résultats de recherche via annuaires sont plus pertinents, les recherches plus faciles.

Quelques inconvénients, mineurs au début, existent dès le début de la domination de Yahoo !

- Il faut payer ou être patient pour avoir son site indexé par l'annuaire Yahoo.
- Un petit site sans intérêt bien *référéncé* peut aisément sortir devant un site réellement pertinent lors de recherches
- Vous avez de l'argent ? vous serez bien *référéncé* au final. Yahoo empoche et les internautes usagers sont floués.

Bref, les annuaires représentent un progrès mais les quantités impressionnantes de nouveaux sites arrivant chaque jour mettent en exergue les défauts de qualité des recherches via annuaires. 2001/2002 : le moteur Google balaie tout le monde - part de marché > 98% en 2003 !!! Google offre de sérieux avantages sur les annuaires tels que Yahoo !

- C'est gratuit pour le référencement. Les mots clef sont payants, comme chez tout le monde (les autres moteurs et annuaires). Mais le référencement de qualité et gratuit est accessible à tous. Que le meilleur gagne, car même lors de la présence de loueurs de mots clefs, il reste une zone où afficher les sites pertinents n'ayant rien payé !
- Il est facile d'avoir son site référéncé par Google. Il n'y a pas de politique de filtrage manuel comme chez Yahoo en 2000.
- Les algorithmes originaux et fort peu documentés de Google assurent de meilleurs résultats (plus pertinents) aux internautes que les autres technologies (à ce jour)

Les internautes aiment la qualité des résultats de recherche de Google : les résultats sont plus pertinents que le résultat de ces même recherches via annuaires. Le basculement se fait très vite. Google règne désormais en maître.

Conséquence : pour ne pas se faire doubler, pour protéger son avance, il n'y a pas de réelles informations publiées sur le fonctionnement des algorithmes par Google. Les référenceurs sont donc obligés de chercher par eux-mêmes, ce qui coûte cher. Donc cela incite chacun d'entre eux à protéger son savoir faire. D'où une rareté importante sur de l'information de qualité pour référencer et surtout positionner efficacement son site.

I-A-2 - Bien comprendre l'objectif central de Google

Même si les 2 fondateurs sont "originaux" et priment quelquefois l'idéologie sur le business, globalement Google est une société qui doit générer beaucoup de profit. Elle sera donc impitoyable avec tout obstacle sur ce chemin.

Comment un site WEB (le vôtre) devient-il un obstacle sur le chemin de Google ???

Tout ce qui peut polluer l'index de Google est un ennemi de Google :

- spam indexing : faire référencer des pages bidons qui renvoient ensuite vers le site principal
- cloaking : Google lit des pages inaccessibles aux internautes (et vice versa)
- duplicate content : un même contenu existant dans 2 pages différentes.

Si votre site fait cela volontairement ou non, il sera plus ou moins sanctionné par Google.

Pourquoi autant d'attention sur les contenus et la pollution de l'index ?

Regardez l'histoire de Google et des moteurs sur le net. Les internautes font le moteur prépondérant. Qu'ils soient déçus, par AltaVista puis par Yahoo!, alors massivement les internautes quittent leurs habitudes et investissent un autre outil.

L'objectif premier de Google, objectif qui passe même devant les résultats financiers, est de satisfaire les internautes en leur donnant des résultats pertinents.

Polluer volontairement ou non l'index de Google plombe la crédibilité des résultats et peut potentiellement dégoûter des internautes de Google. Donc pas de quartier pour les pollueurs d'index.

I-B - Pourquoi un référencement ? Le changement de comportement des clients et des prospects

Tout type d'acheteur, entreprise ou particulier, passe désormais par un moteur sur Internet (Google...) pour :

- Rechercher des offres alternatives à un prestataire en place "des fois qu'il y ait mieux ailleurs"
- Comparer les prix, les prestations, les produits avant tout achat
- Consulter le BUZZ sur une entreprise, un produit... avant tout achat

Être invisible sur Google ou autres moteurs c'est perdre des prospects qualifiés et ciblés. Donc du business. Demain, la valorisation d'une entreprise inclura la qualité de son référencement

I-C - Pourquoi référencer son site en mode naturel ?

[Connexion](#)

Web Images Groupes Annuaire Actualités Desktop plus »

développez vos ventes formation | Rechercher [Recherche avancée](#) [Préférences](#)

Rechercher dans : Web Pages francophones Pages : France

Web Résultats 1 - 100 sur un total d'environ 179 000 pages en français pour **développez vos ventes formation.** (0,61 secondes)

Formation vente Lien commercial
www.cegos-formation.com 100 formations dédiées à tous les métiers commerciaux

508 résultats stockés sur votre ordinateur - [Masquer](#) - [À propos](#)

FaxMiling Orkis ppt - d'idées déco /marketing pour **développez vos ventes** - 12:36
 FaxMiling Orkis V2-10 ppt - pour **développez vos ventes** Nous matérialisor - 12:36

Développez Vos Ventes - Formation Conseil - référencement sur ...
 Formation, assistance, conseil pour le développement des **ventes** - référencement sur Google, Commercial, marketing pour développer les **ventes**, la prospection ...
www.developpezvosventes.com/ - 19k - [En cache](#) - [Pages similaires](#)

[référencer - référencement sur Google - Développez - Développer ...](#)
 Développement des **ventes** - formation & assistance marketing commerciale. Formations pour mieux prospecter, vendre ,manager - Développement de **vos ventes** ...
www.developpezvosventes.com/referencer_-_referencement_sur_google.htm - 11k - [En cache](#) - [Pages similaires](#)
 [[Autres résultats](#), domaine www.developpezvosventes.com]

Formation : Développez vos ventes dans l'Union Européenne
 Formation : **Développez vos ventes** dans l'Union Européenne. Pour connaître les dates et le lieu, cliquez sur le lien ci-dessous.
formation.journaldunet.com/formation/
 5422/développez_vos_ventes_dans_l_union_europeenne/ - 77k - [En cache](#) - [Pages similaires](#)

Formation : Développez vos ventes dans l'Union Européenne
 Organisme de **formation**. Consultez notre catalogue et inscrivez-vous en ligne.
formation.journaldunet.com/formation_etablissement/5422/développez_vos_ventes_dans_l_union_europeenne/ - 78k - [En cache](#) - [Pages similaires](#)
 [[Autres résultats](#), domaine formation.journaldunet.com]

Développer vos ventes
 >**Développez vos ventes.** A la conquête de nouveaux marchés ! ... >Documentation et **formation**. Boutique en ligne: Toute la boutique; **Formation** continue: Toute ...
www.nantes.cci.fr/CCI44/Page/Page.aspx?PUB_ID=185&PUB_VERSION=0&ARBO_ID=40&ARBO_VERSION=0 - 98k - [En cache](#) - [Pages similaires](#)

[Gérer un point de vente](#)
 Agir pour l'environnement; Vous **développez** à l'international ... >Gérer un point de vente

Liens commerciaux

Formation à distance
 Guide des formations : une sélection de 7 écoles spécialistes.
www.formations-distance.com

Master Logistique Achats
 Acquérir une Expertise - 3e cycle en alternance après Bac+3 ou Bac+4
masters.insee-france.com

Formation en Vente
 Management Organisation Négociation **Formation** sur mesure - Efficacité
www.m-formation.fr
 Île-de-France

Formation pro rémunérée
 Filière distribution de boissons, bac+2 des débouchés ! inscrivez-vs
www.distech.fr

Formation professionnelle
 Conseil / Expert en relation client -20% jusqu'au 30/06 pour sept 06
www.yoxia.fr

Formation vente
 Réservé aux entreprises. Devis gratuits en 48h00 sur Companeo.com
www.companeo.com

Formation achat
 Mettez le cap sur l'expertise Retrouvez toute l'offre Demos
www.demos.fr

Formation Professionnelle
 Boostez votre potentiel commercial avec une **formation** en vente !

Objectif : être dans les 30 premiers

Référencement **naturel** – le site est interprété par Google (jusqu'à 1000 critères pris en compte) qui le positionne dans son index

Les internautes qui cliquent sont des prospects ciblés et qualifiés

Vendre en ligne c'est aussi inciter à la prise de contact, pas simplement vendre et livrer à distance.

- Créer du trafic dans ses points de ventes
- Inciter au contact par téléphone, fax ou autre ...

Référencement **payant** chez Google – **Adword** – peu importe votre site, si vous payez vous serez là

I-D - Initiation au référencement en termes simples


I-D-1 - Étape #1 - Mots et expressions clefs - La mission du marketeur

- Identifier les mots clefs du "métier" de l'entreprise (et donc de son site WEB)
- Les tester, en rechercher d'autres, combiner, analyser
 - Donc aucun trafic qualifié
- **Profil de cette mission clef : WEB marketeur**
 - Des capacités à analyser un marché, les comportements, les attentes priment sur la technicité (HTML, PHP etc.) du site.
 - Donc perte de prospects

I-D-2 - Étape #2 - Accessibilité technique & sémantique du site - Les contenus voulus - Le rôle de la technique et du marketeur

L'accessibilité se divise en deux parties :

- **L'accessibilité technique** : le robot doit pouvoir accéder à la page et à son contenu.
- **L'accessibilité sémantique** : les bons termes destinés à devenir des expressions clefs à positionner doivent être efficacement mis en valeur.

 *Répéter, "spammer", un mot dans une page n'est certainement pas efficace. Au contraire. Disperser cette expression clef partout sur le site n'est pas efficace non plus. Une accessibilité sémantique est bien plus subtile et elle sera abordée en détail plus avant dans ce document.*

Le site WEB qui accueille les prospects et clients doit :

- Être accessible aux moteurs (un % important de sites bloquent tout ou partie des robots de Google pour cause de maladresses techniques)
- Pousser les bons mots et expressions clef au bon endroit
 - Dernier exemple en date : "espace crédit" (l'expression clef référencée en réalité par le site aux yeux de Google) n'a aucun rapport avec "Assurance crédit" (l'expression clef recherchée) sur le site d'un courtier en assurances.
 - Éparpiller un mot clef dans un site peut pénaliser le référencement du site :
 - Il est nécessaire de le faire
 - Mais fait sans comprendre inversera la performance. Google aura une vision brouillée de la maîtrise de ce mot clef par le site et le positionnement en sera pénalisé.
 - Conclusion : il faut savoir équilibrer les contenus et donc avoir une bonne compréhension de Google.
- Ne pas tricher et faire attention à ce que l'on fait
 - Sanctions Google : de black liste à Sand Box => site WEB invisible ou presque sur Google

Profil type pour cette étape : Webmestre ou développeur comprenant Google

- Maîtrise, de correcte à excellente, des techniques du WEB (HTML, PHP, Java script, Photoshop...)
- Comprendre Google en accessibilité demande du temps, du travail, des essais, du temps, de la patience, du temps, de la persévérance, du temps, de l'opiniâtreté, du temps, de la réflexion et puis encore et toujours du temps, du temps, beaucoup de temps...
- À 90%, comprendre les mécanismes de référencement de Google n'est pas compliqué. Juste long. Comprendre comment bien optimiser voire jusqu'à tricher sans se faire prendre exigent beaucoup de temps et de tests.

I-D-3 - Étape #3 - Être bien positionné

Référencer son site en le modifiant techniquement ne suffit pas à avoir un bon positionnement

- Définition : Être dans les 30 premières réponses à une requête correspond à l'expression "être bien positionné"
- Référencer son site en le modifiant pour qu'il soit accessible avec les bons mots clefs rend possible d'être bien positionné "un jour" dans le futur. Pas plus ! Il n'y a aucune garantie.

Il faut avoir des 'liens nommés entrants' (backlinks dans le jargon) sur son site

- Stratégique et indispensable, long, difficile...
- Techniquement simple quand on comprend Google sinon il est facile d'avoir un travail massacré et même contre productif

Profils types :

- "Petite main" : pour démarcher les échanges de liens
- "Traffic manager" nouveau métier en cours d'apparition ; il gère les opérations payantes ou non pour booster le positionnement du site dont la gestion des liens.

Piège(s) : les liens dits "annuaires" - Ex : "350 € HT l'inscription dans 5000 annuaires du net" - Une alternative un zeste "provoc" : prenez 7 billets de 50 € et jetez-les par la fenêtre de votre bureau. Ce type de liens ne participe que fort modestement au référencement efficace d'un site.

I-D-4 - Étape #4 - Maintenir son référencement & quelques trucs

- Sans suivi du référencement, ce dernier se dégrade vite car la pression est grande sur le Net
- Un nouveau métier va apparaître : positionneur de site
- Ne pas se décourager - nombre d'expressions clefs ne sont pas bien gérées, mal défendues par les "bien positionnés" actuels - il est tout à fait faisable de bien se placer.
 - Marketing & Réflexions bien menées peuvent faire très fort rapidement en positionnement

I-E - Adwords et autres liens commerciaux versus référencement naturel

	Adword - AdSense	Référencement naturel
Prestataire à payer	0 € ou petit budget pour marketing sur mots et expressions clefs.	Budget moyen à élevé en € : analyse mots clefs, accessibilité, positionnement de base
Coût de l'exploitation	Payer Google ou Yahoo au fil des clics sur les liens commerciaux - de 20 à 1000 € mensuel voire plus	Éventuel positionnement payant (louer de vrais et bons backlinks), suivi du référencement - budget largement inférieur à Adword ou AdSense
Escroquerie(s) & fraude(s)	Malgré les efforts de Google et Yahoo, sabotage et faux clics entre concurrents ou par	Si prestataires incompétent ou arnaquant le client sur la qualité.

	escrocs en ligne faussent ces outils	
Délai	Immédiat	De quelques jours à quelques semaines ou mois
Conclusion	Rapide mais cher si il n'y a que cela pour créer du trafic sur le site. Les 2 approches se complètent Gare aux arnaques dans les 2 approches	Investissement de départ élevé mais plus d'efficacité dans le temps et moins cher sur la durée

I-F - Rappel d'évidences

I-F-1 - Inutile de référencer un site Internet touffu

Comment évaluer si votre site est acceptable ou si il doit être révisé, retouché voire même refondu complètement ? Comparez le à ses concurrents, demandez à des "prospects" réels ou tests leur avis et leurs raisons. Questionnez aussi un spécialiste.

Réfléchir avant d'agir, ne pas hésiter devant un budget "réflexion" et analyse pourra vous faire économiser du temps, de l'argent et gagner en qualité.

I-F-2 - Inutile de référencer des mots clefs inutilisés

Personne ne les trouvera ! Donc pas de visiteurs.

Nous verrons comment utiliser les outils gratuits de Google et Yahoo ! pour évaluer approximativement un potentiel de fréquentation.

Un outil de Google payant (une poignée d'euros) permet d'affiner.

I-G - Définitions

I-G-1 - Les termes simples

Google référence des **pages HTML** et non des sites !

Le référencement passif : c'est le travail à réaliser sur le site tant en mot clef qu'en modifications techniques pour référencer le site

Le référencement actif : ce sont les actions menées hors du site pour référencer ledit site.

le Page Rank (PR) : la définition fait l'objet d'un paragraphe dédié plus loin dans le présent document.

Un **annuaire** Internet fonctionne comme les "Pages Jaunes" de Wanadoo : les sites y sont classés par thèmes ou par ordre alphabétique. Une intervention humaine côté annuaire est généralement nécessaire pour contrôler la qualité de l'inscription. Cela en explique le coût facturé par l'annuaire à l'entreprise qui s'y inscrit.

Un **moteur** est automatisé. Une flotte de robots logiciels (spider, Google Bot...) parcourent Internet, suivent des liens de sites en sites et ils capturent des pages, qu'ils stockent dans les espaces disques du moteur (15 000 serveurs

en 2004 chez Google pour vous donner une idée) et dont ils indexent tous les mots significatifs dans une immense base de données centralisée.

Le **positionnement** traduit le fait d'être présent ou non dans un résultat de recherche sur mots clefs existants dans une des pages de votre site. Il traduit plus finement votre position dans la réponse par rapport à la 1ère réponse (la meilleure place bien sur). Les places qui comptent sont celles des 3 premières pages de 10 réponses unitaires (10 réponses correspond au format standard de réponse de Google. Ce format est bien sur modifiable par l'internaute mais c'est un autre aspect de Google)

Un **cookie** est un code de marquage envoyé par un site Internet sur votre disque dur. Ce code est unique. Si vous revenez sur le site Internet auteur du cookie, seul lui peut le lire... et vous identifier. De nombreux sites commerciaux utilisent cette technique pour donner du confort à leurs usagers, contrôler l'accès et... comprendre vos habitudes.

Un **jeton de session** est un genre de cookie (un code de marquage) utilisé comme code obligatoire d'accès à des pages du site. Cette technique est utilisée pour contrôler l'accès à des contenus sensibles, pour lutter contre les intrusions, pour permettre le cheminement dans des transactions financières en ligne etc.

Critères page ON : ce sont les critères de chaque page qui ont un impact sur l'aspect visuel, au sens large. L'internaute perçoit une modification d'un tel critère. Exemple : gérer les balises `<hx></hx>` (h1, h2...) pour la mise en page plutôt que du ``.

Critères page OFF : ce sont les critères de chaque page qui n'ont AUCUN impact sur l'aspect visuel. La balise meta DESCRIPTION par exemple.

Critères ON et OFF PAGES : il y a environ 1000 critères pris en compte par Google... Tous n'ont pas le même poids bien sur. Les plus importants :

- En local : Title, Description, URL de la page, nom de fichier de la page HTML, le nombre de mots significatifs
- En Backlink : le contexte et la thématique, la sémantique du paragraphe, de la page, du site. Le positionnement du lien BL au sein de la page et du texte. La qualité du texte en rédactionnel.

I-G-2 - Page Rank

I-G-2-a - Objectif recherché

Google veut satisfaire ses usagers (et les fidéliser par leur contentement)

Comment détecter une page ayant un contenu et une pertinence de bonne qualité par rapport à une autre moins pertinente ?

Une des (multiples) réponses consiste pour Google à se baser sur l'avis des internautes.

Le principe est le suivant : plus les internautes vont sur un site, plus des sites WEB mentionnent un site, alors plus ce site aura de visiteurs et donc il dispose potentiellement de contenus de qualité susceptibles de répondre efficacement à la requête d'un internaute.

Comment faire ? En créant le page rank, unité de mesure basée sur :

- Le nombre de liens pointant vers une page à indexer
- La qualité en page rank de chaque page contenant un lien pointant vers la page à indexer

Plus un site populaire a un lien pointant vers la page à indexer, plus elle communiquera un page rank élevé à cette page (principe des vases communicants)

Plus il y a de sites qui ont des pages pointant vers cette page, plus le page rank sera élevé.

Voilà pour la définition simple et l'objectif recherché par Google.

Afin de limiter les échanges de liens non pertinents, Google "facture" les pages qui possèdent des liens sortants vers d'autres pages externes au site. Il retire un % de PR de la page appelante pour le transférer vers la page pointée. (principe des vases communicants)

I-G-2-b - Une définition plus complète


Concrètement : à quoi sert le PR, pourquoi en avoir c'est mieux ?

Un PR élevé est garant des points suivants :

- La régularité du passage de Google sur le site - vous mettez votre site à jour ? avec un PR élevé les nouveautés seront prises en compte plus vite
- Le volume de pages prises en compte. Vous avez plusieurs niveaux de profondeurs dans votre arborescence ? plus le PR est élevé, plus Google ira profond.
- Les pages dynamiques : plus vous avez de PR, plus Google accepte des pages à URL dynamique non "rewrité"
- Bref, avoir du contenu, de l'accessibilité efficace et du PR conduit à avoir un site **plus facile à positionner**.

La base du Page Rank, noté PR dans la suite des pages de ce document, est une formule mathématique, d'aspect complexe mais finalement simple à appréhender quand on en comprend les principes fondamentaux.

Comprendre le PR est important pour pouvoir améliorer son positionnement et sa visibilité sur le net via Google et pour limiter au maximum ses propres pertes de PR pour les liens sortants de ses propres pages.

 **ATTENTION** : l'application de cette formule comporte des exceptions non documentées par Google.

Abordons ces exceptions avant même d'étudier la formule, c'est important :

- Un site n'ayant pas de liens sortants et donc aucune fuite de PR se verra sanctionner par un fort retrait de PR. Vous ne jouez pas le jeu du WEB ni surtout le jeu de Google (si tout le monde fait pareil alors comment ferait Google pour fonctionner correctement ? Crime de lèse majesté !) alors forte sanction.
- Un site ayant beaucoup de liens sortants devrait voir son PR s'effondrer d'après la formule du PR. Pas du tout, au contraire ! Google ajoute un fort bonus de PR à tout site ayant de la fréquentation et disposant de nombreux liens commentés et donc potentiellement informatifs de qualité pour les usagers.

I-G-2-c - La formule

Nous assumons qu'une page A reçoit des liens (ou "votes") émis par les pages T1. à Tn.

Le paramètre d est un facteur d'amortissement pouvant être ajusté entre 0 et 1. Nous donnons généralement à d la valeur 0.85.

De même, C(A) est défini comme le nombre de liens émis par la page A, (liens sortants). Le PageRank de la page A est défini comme suit :

$$PR(A) = (1-d) + d (PR(T1)/C(T1) + \dots + PR(Tn)/C(Tn))$$

Le PageRank peut être calculé en utilisant un simple algorithme itératif et il correspond au vecteur propre principal de la matrice normalisée des liens du Web.

PR(A)	Le Page Rank de la page A
PR(Tn)	Le Page Rank de la page Tn
C(Tn)	Le nombre de liens émis depuis la page Tn
d	Tous les "liens entrants" sur la page A sont additionnés, mais pour en limiter l'importance, le total est multiplié par ce coefficient d'amortissement (0.85)
1 - d	Un petit peu de "magie mathématique" qui permet de garantir que la moyenne des Page Rank de l'ensemble des pages du Web sera de 1.

L'examen de cette formule permet de voir que le PageRank d'une page n'ayant aucun lien entrant sera de 0.15.

Soit :

$$(1 - 0.85) + 0.85 * (0) = 0.15$$

Plusieurs sites, forums pensent que l'échelle du Page Rank est logarithmique, sans que ceci ne soit officiellement admis par Google. Personnellement j'ai mesuré que oui, l'échelle du PR est logarithmique. La base utilisée est estimée. Il est probable que cette base évolue dans le temps. Plus complexe encore, je soupçonne que la base de log ne soit pas la même tout le long de la courbe. Mais bon, restons sur l'hypothèse que la base de log est stable pour simplifier la compréhension de ce mécanisme.

Prenons une échelle logarithmique de base 10 pour simplifier nos calculs et pour bien illustrer notre propos.

La communauté des référenceurs étudiant Google estime qu'actuellement, la base de log utilisée oscille dans la plage 5 à 8. Notre exemple de 10, rappelons le, permet de simplifier le cours. Il n'est d'ailleurs pas exclus que Google ne passe un jour à la valeur 10 devant l'augmentation continue du nombre de sites et de pages à indexer.

PageRank Affiché (log base 10)	PageRank réel (calculé)
PR0	$0 \leq PR < 1$
PR1	$1 \leq PR < 10$
PR2	$10 \leq PR < 100$
PR3	$100 \leq PR < 1000$
PR4	$1000 \leq PR < 10000$

Chaque niveau de Page Rank est 10 fois plus élevé que le niveau précédent. En clair il est 10 fois plus ardu de passer de PR4 à PR5 que de passer de PR3 à PR4.

Une première conséquence de cette formule : si un site n'évolue pas, son PR se dégrade au fil du temps. En effet, le nombre de pages, de sites augmentant continuellement, le nombre de liens C(Tn) augmentent et le PR transféré diminue.

Chaque fois que la base augmente pour faire face à la quantité de pages à indexer, la plage de chaque PR augmente et tend à faire diminuer le PR des pages d'un site non maintenu à jour.

Une autre conséquence de cette formule : elle tend à faire converger le PR moyen d'une page vers UN.

La manière dont Google a présenté cette formule indique qu'il n'est pas nécessaire de disposer de tous les PR des pages. En itérant la formule, le calcul converge vers la valeur finale. Bref, cela ne consomme pas de ressources si importantes.

I-G-3 - Important, ne pas confondre PageRank et positionnement !

Un bon page Rank, c'est-à-dire une bonne popularité d'après Google de votre site n'est pas l'assurance d'être bien positionné en résultats de recherche.

Certes un bon page Rank aide à avoir un bon positionnement mais il est très facile d'avoir un site de page Rank 1 ou 2 et d'être mieux positionné sur des mots clefs qu'un site de page Rank 4 ou 5. mal géré sur ces même mots clefs.

Pourquoi ? Parce que Google veut publier des résultats pertinents à ses internautes et les sites à fort pouvoirs financiers sont ainsi limités en efficacité de l'argent face à la qualité. des contenus.

Comment ? en travaillant bien la sémantique puis quelques backlinks, un site à petits moyens financiers pourra être mieux positionné qu'un site disposant d'un budget référencement (et donc des moyens d'obtention d'un bon page rank) mais relativement pauvre en contenu ou mal géré en accessibilité (technique ou sémantique)

Conclusion : le rédactionnel prime sur les astuces techniques. Mais sans une bonne accessibilité, le rédactionnel, le contenu aura beaucoup de mal à émerger.

II - Le référencement passif

II-A - Architecture et accessibilité technique

II-A-1 - Votre site, s'il existe, est-il visible par Google ?

3 outils simples vous permettent de faire une vérification :

- Dans la barre d'outil Google : "site:www.lesiteAtester.com" => Google affiche toutes les pages indexées.
 - Si vous n'avez rien : problème !
 - Si vous n'avez QUE la page d'accueil : des choses sont à revoir dans la gestion de vos liens.
 - Le Page Rank (via la barre Google) de vos pages est à 0 : problème manifeste de référencement (ou le site est jeune)
- L'autre outil vous notifie les défauts techniques dont certains déplaisent à Google :
 - W3C : <http://validator.w3.org/>
 - Outre la disponibilité d'un site lisible par tous, Google lira sans blocage technique basique vos pages. Ce qui ne signifie pas qu'il les indexera ou arrivera à suivre vos liens. Simplement, il n'y aura pas de blocage technique.
- Lynx : ce navigateur pour mal voyant et aveugle vous montre le site tel qu'il est pour un bot Google. Si il manque des liens, vous le verrez ! Rappel : plus le texte et les liens sont hauts dans une page, plus ils sont pris en compte par Google. Il ya une légère priorité de haut en bas.

II-A-2 - Un peu de technique HTML, XHTML et CSS2 - Le minimum nécessaire

Il existe de nombreux cours gratuits sur HTML, CSS etc. ainsi que des ouvrages payants.

CSS version 2 représente une grande amélioration sans aucun inconvénient. Faisons un rapide tour de cet aspect technique qui facilitera la tâche clef de cette étape : disposer d'un site lisible par Google.

Voici quelques raisons d'utiliser les feuilles de style CSS pour gérer la mise en forme de vos pages HTML

- Compatibilité avec les versions antérieures des principaux navigateurs.
- Compatibilité assurée avec les navigateurs de nouvelles générations et leurs futures versions.
- Compatibilité avec les normes et recommandations du W3C.
- facilité des mises à jour du site.
- Permet de séparer le fond et la forme du site ce qui facilite la maintenance (plus rapide, plus facile, moins de bogue)
 - La gestion de l'affichage est confiée à la (aux) feuille(s) de style
 - Spécialiser les feuilles de style permet d'assouplir et de simplifier la maintenance du site, cela facilite la ré utilisation de feuilles de styles pour d'autres sites WEB également.
 - Les données sont dans les pages HTML gérées par la (les) feuille(s) de style
- Facilite la relecture du code.
- Diminution des coûts de développement et de maintenance du site au final après l'investissement initial.
- Limiter voire supprimer les erreurs et maladroresses techniques sur le référencement.

II-A-3 - La liste des interdits et les problèmes

II-A-3-a - Le robot Google doit pouvoir suivre les liens cliqués

Sinon, aucun travail de référencement ne pourra donner de résultat.

Pour permettre un accès technique aisé aux robots logiciels de Google (et des autres moteurs concurrents) voici quelques conseils clefs :

II-A-3-b - Pas de site entièrement en Flash

Vous pouvez bien sûr agrémenter votre site d'exposés et autres outils d'aide à la vente développés en flash, mais il est très déconseillé de réaliser toute le site en flash.

Google, pour différentes raisons, n'indexe pas correctement le flash et pour certaines raisons non techniques mais conceptuelle, cette situation n'est pas près de changer.


II-A-3-c - Pas de frames si possible

Une frame, c'est-à-dire une partie de la page HTML, est vue comme une page indépendante. Cela perturbe (négativement) le référencement. Les performances actuelles des réseaux publics, dopés par ADSL, permettent de se passer de cette optimisation (l'usage des frames).

II-A-3-d - Pas de pages dynamiques ou alors attention à ce que l'on fait

Une page dynamique est une page HTML dont les contenus textes et visuels sont issus en temps réel d'une base de données

- Exemple : Combien y a-t-il de places pour la séance de tel film tel jour, telle heure, tel ciné ?
- Le logiciel de réservation de place de cinéma en ligne affiche la page HTML de réponse vide (le logo, le cadre, le nom du ciné, heure de la séance, adresse etc.) puis va chercher l'information en temps réel et l'affiche au bon endroit dans la page qui sera lue par l'internaute cinéphile.

 Une page contenant des animations, certes dynamiques, n'est pas concernée.

Google sait identifier les liens qui renvoient des pages dynamiques

- Toute référence de page (URL) ayant plus de 2 caractères variables liés à la gestion d'une page dynamique est plutôt ignorée.
 - Le robot de Google évalue ainsi la quantité. Si elle est faible (2 digit maximum pour coder les différentes pages de résultats, alors il accepte de référencer plus ou moins bien ces pages, une par une, en fonction du PR de la page "mère"
 - La profondeur de son accès dans ces pages dynamiques est fonction du PR de la page mère. Si le PR est élevé il ira assez profondément. Sinon, il évitera plus ou moins ces pages dynamiques.
- Si le nombre de pages dynamiques est trop élevé selon ses critères (à partir de 3 digit variables par URL de page et selon le PR de la page "mère"), alors il ignore le lien et n'indexe rien.
 - C'est volontaire ! Sinon toute société ayant 50.000 articles en base de données avec un site à faible PR verrait 50000 pages indexées encombrant inutilement (au sens Google du terme) l'index du moteur.
- Si le lien cliqué par le robot Google aboutit à un formulaire à remplir pour disposer de l'information, alors le robot Google est bloqué.

Google est perturbé ou bloqué par la gestion des cookies, des session ID

- Par définition, le robot ne saisit aucune donnée, ne possède pas d'espace où stocker un cookie ou un jeton de session.
- Même si Google désormais passe ce type de blocage plus ou moins bien, se faire référencer avec des QSA (ex : "[?article=12&sid=12eadb1258](#)") est très négatif pour le site car un click depuis l'index Google entraîne aussitôt un 404 et les accumulations de 404 ne sont pas appréciées par Google pour différentes bonnes raisons
- Recommandation : bloquer à Google les zones où un SID est nécessaire et si vous désirez faire indexer ces pages, alors laisser un "passé" invité où aucun SID ou autre n'est géré. L'invité n'aura bien sûr le droit à aucun formulaire.

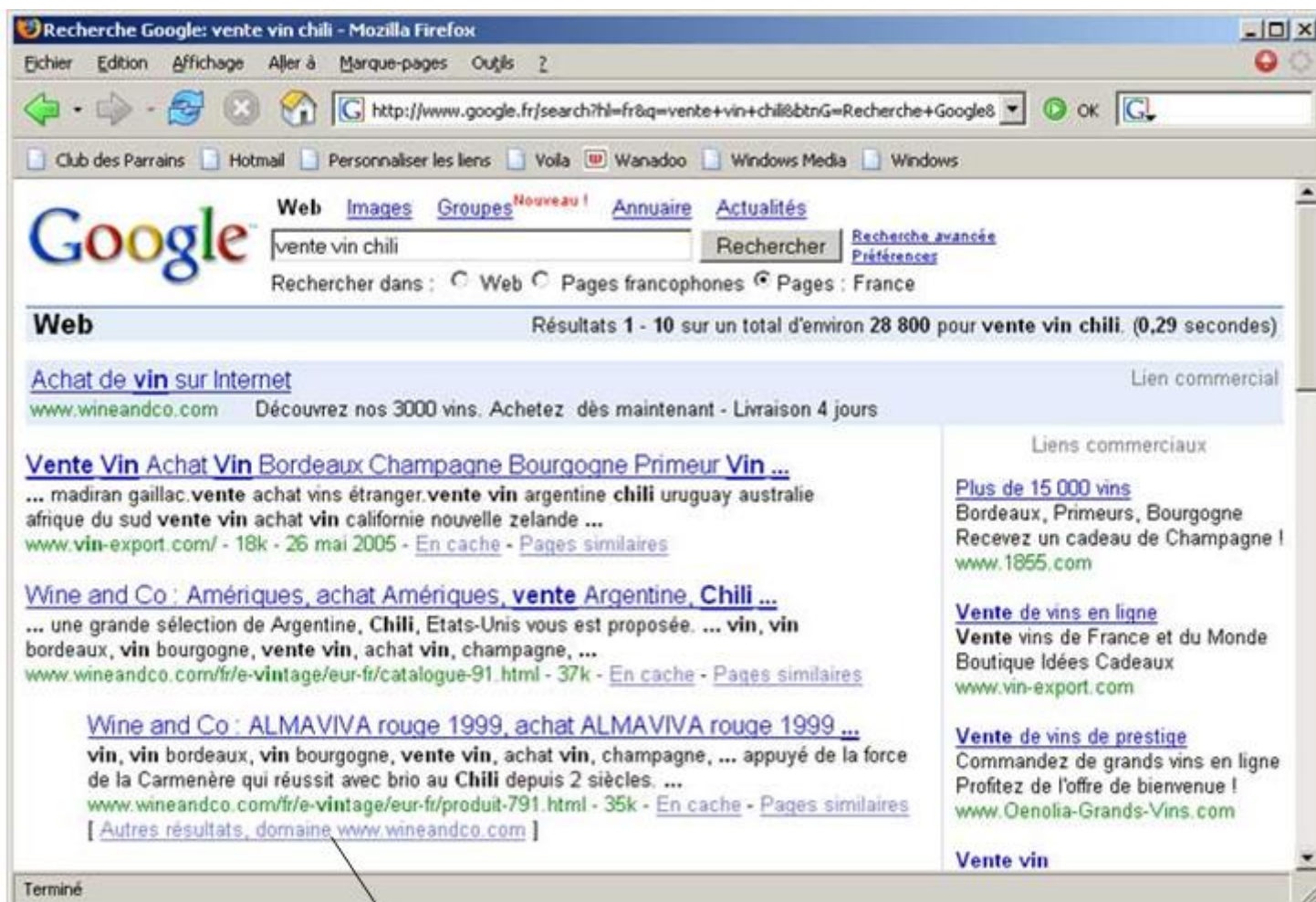


Google ne sait pas simuler une souris sur des liens : les liens cachés en Javascript sont illisibles pour lui. Ainsi que les listes déroulantes, les cases à cocher et autres techniques d'interfaces interactives en Javascript.



Sauf sur les pages de 10 résultats maximum, Google limite à 2 pages le résultat d'une recherche d'un site. Si une requête d'internaute correspond à plus de 2 pages de votre site, les 2 plus adéquates selon Google seront présentées et les autres seront regroupées sous une appellation du style :

[Autres résultats, domaine <http://www.wineandco.com/>]



Par source, 2 citations maximum, jamais plus.
Les autres pages sont proposées en option
sous la forme de [Autres résultats ...]

Si la page de résultats est à 10 maximum, alors une seule citation du site est faite.

II-A-4 - Contourner les interdits techniques

II-A-4-a - Il faut contourner les interdictions ou les contraintes

- Le marketing de votre site peut exiger de disposer de Flash, de boîtes à cocher, de gestion de cookie, d'un catalogue en ligne *etc.* : toutes technologies déconseillées pour un bon référencement.
- Il faut alors contourner ces interdits ou contraintes techniques.

II-A-4-b - Le Flash

II-A-4-b-i - Objet entièrement en Flash

Faire un site entièrement en Flash exige un budget important. Une telle dépense doit être justifiée par un avantage compétitif important.

Les deux principales justifications d'un tel usage de flash sont :

- Image de marque luxe, haut de gamme à présenter absolument
- Besoin d'une ergonomie exceptionnelle

Si vous disposez d'un tel budget, alors pour un budget modique en comparaison, vous pouvez faire développer un site WEB HTML répondant à vos besoins marketing et disposant de multiples liens vers le site Flash non référencé.

Autre cas : votre site, HTML, doit présenter des animations, des jeux, des schémas animés *etc.* pour illustrer vos propos et explications.

Doublez-les d'une page HTML dotée de liens pointant vers des pages contenant uniquement ces animations.

II-A-4-b-ii - Menu principal en Flash

Les menus Flash sont plus agréables et plus ergonomiques.

Doublez le menu Flash par une barre de menu classique en bas de page

Si toute la page contient un objet Flash, doubler le menu en bas et y ajouter un texte référençable.

II-A-4-b-iii - Pourquoi tant de problèmes avec Flash alors que PDF est indexé par Google

Probablement à cause des algorithmes de Google qui ne sont pas adaptés à la complexité de Flash.

En Flash, des textes peuvent être vectorisés, bitmap, avec des politiques de mise en page (titre1, 2 *etc.*) tellement différentes qu'il est difficile d'identifier pour Google les bons mots clefs. Cela rompt avec sa stratégie très évoluée d'indexation et d'évaluation des textes, donc il ne fait pas.

II-A-4-b-iv - Insérer un objet Flash dans une page HTML

Voici un exemple de code à mettre dans une page HTML, page qui sera référençable mais pas le Flash bien sûr.

```
<object id="numanuma" width="320" height="240">
  <param name="movie" value="fichier.swf"/>
  <param name="quality" value="high"/>
  <param name="bgcolor" value="#006633"/>
  <embed src="numanuma.swf"
    quality="high" bgcolor="#000000" width="320" height="240" type="application/x-shockwave-flash"
    pluginspage="http://www.macromedia.com/shockwave/download/index.cgi?
    P1_Prod_Version=ShockwaveFlash">
  </embed>
</object>
```

II-A-4-c - Les pages dynamiques : Ré écriture d'URL

La méthode consiste à ré écrire l'URL des pages dynamiques. En effet, les pages écrites avec des ?, des &, des .php *etc.* ne sont pas toujours bien indexées par Google pour différentes raisons (déjà abordées précédemment)

Plusieurs possibilités existent pour que le bot Google indexe ces pages. Le principe est le suivant :

Pour être accessibles aux Google bot votre catalogue doit disposer, dynamiquement ou autre, d'un accès en "râteau" : une arborescence permet d'accéder à chaque article de votre catalogue, directement ou via des pages intermédiaires :

Tête de catalogue listant en liens toutes les rubriques :

- Vins de France
 - vins de Champagne
 - Vins de Bourgogne
- Vins du Chili
- Vins de Californie etc.

Cela donnerait quelque chose du genre : `75cl.com/page_vin_3_7_11.html` où **3** représenterait le code pays, 7 pour la région, 11 pour le château.

D'un côté on a donc une arborescence de liens que Google peut suivre, d'un autre côté ces liens pointent vers des pages statiques fabriquées depuis le contenu de la base de données.

Cette approche a une limite théoriquement peu gênante : la mise à jour des pages statiques conçues depuis un contenu dynamique en base de données. Si l'information est modifiée en base de données, tant que la page statique publiant cette donnée dynamique n'a pas été reconstruite, l'information publiée sera erronée.

Pour un catalogue de produits ayant une mise à jour chaque nuit, ce n'est pas gênant. On met à jour les pages statiques chaque nuit.

Mais si cette page indique un délai de livraison fonction de stock, alors il peut y avoir problème. Le site afficherait par moment des délais de livraison faux.

Pour pallier cet inconvénient, deux pistes complémentaires :

- Si un délai est modifié, alors toutes les pages statiques concernées sont régénérées avec la nouvelle information et elles remplacent les anciennes pages sur le site.
- Aucune information sur le délai n'est affichée dans une page statique construite de puis des informations dynamique. "Délai" est désigné via un lien qui pointe vers une vraie page dynamique calculée en temps réel par exemple. Peu importe que cette page affichant le délai ne soit pas indexable par Google..

La ré écriture d'URL offre au passage quelques avantages non négligeables :


- Accéder à une page statique pré calculée est beaucoup plus rapide que de faire mouliner un moteur de base de données. Les temps de réponse sont meilleurs
- la ré écriture peut cacher la technologie utilisée (.asp, .cfm, .php3) qui sont autant d'informations clefs pour un pirate en ligne.
- Les variables en ligne sont résumées à des nombres séparés par des "_". Il est là aussi plus difficile à un pirate de comprendre l'architecture du programme et donc de tenter de le hacker.

Comment ?

Plusieurs approches existent :

- Serveurs http Apache et langage PHP : paramétrage de Apache et il y a du travail d'administration à faire. Consulter la documentation Apache sur le module "mod_rewrite" et aussi "RewriteRule". (dans le .htaccess)
- Serveurs IIS / langage ASP de Microsoft : il existe des filtres payants via ISAPI.
- Serveurs IIS / technologie .NET : utiliser la fonction RewriteURL() et autres qui gèrent la ré écriture d'URL.
- Approche "maison" : écrire vous-même les scripts PHP, Java ou autres qui transforment vos pages dynamiques en pages statiques. Les avantages et inconvénients sont les suivants :
 - Vous pouvez enrichir la ré écriture d'URL avec d'autres fonctionnalités. Ex : génération de critère off pages adéquats pour Google totalement automatisée.
 - Écriture automatisée et normée du nom de fichier HTML, du Title de la page, des meta Keyword, description, de certains titres en balises H, de certains ALT de visuels de la charte graphique, de title de liens etc.
 - Vous pouvez utiliser les technologies de votre choix : Java J2E, CFM etc.

En fonction de vos budgets, de vos contraintes informatiques et de vos besoins, vous avez plusieurs pistes.

 *Si vous êtes en hébergement mutualisé, il y a risque de conflit avec votre hébergeur qui peut refuser les réglages Apache dont vous avez besoin. Il vous faudra changer d'hébergement ou passer à un hébergement dédié.*

Le choix du séparateur : Google considère que le tiret (signe moins) "-" est un séparateur de mots.

Je recommanderai donc quelque chose du style :

http://www.75cl.com/ventes-vins-champagne-dom-perignon-12_7_17.htm : des tirets pour séparer les mots, des soulignés "_" pour séparer les nombres.

La ré écriture d'URL vous permet de gérer des noms au lieu de nombres (facilite la pertinence) ou de mixer les deux. Google retiendra surtout les noms séparés correctement.

II-A-4-d - La gestion des sessions

Cette technique est utilisée pour contrôler finement l'accès à des informations.

Par définition, si vos données sont confidentielles, elles n'ont pas à être accessibles à Google.

Exemple : tant qu'un internaute surfe sur les produits de votre catalogue en ligne sans acheter, nul besoin de gérer sa session. A partir du moment où il rentre dans une phase d'achat, via le dépôt d'articles en panier, alors seulement à ce moment le site déclenche la gestion de session.

Cela permet à Google d'indexer les pages du catalogue ou en laissant la possibilité au site de gérer en mode session les zones critiques de la phase d'achat et de paiement


Ma recommandation :

- Si possible essayer de séparer complètement la partie publique de la partie sensible.
- Ne pas changer le contenu de fond selon la session. Google semble disposer ou disposera d'une politique de vérification des contenus et il sanctionnera le cloaking réel (cf. Trusted Rank qui remplacera ou a peut être déjà remplacé Page Rank).

II-A-4-e - La gestion de session par cookie

Identifier l'internaute revenant sur votre site est un atout intéressant. Cela vous procure des informations marketing intéressantes, cela vous permet de donner un confort d'utilisation à vos internautes. Mais Google ne se prête pas à cela.

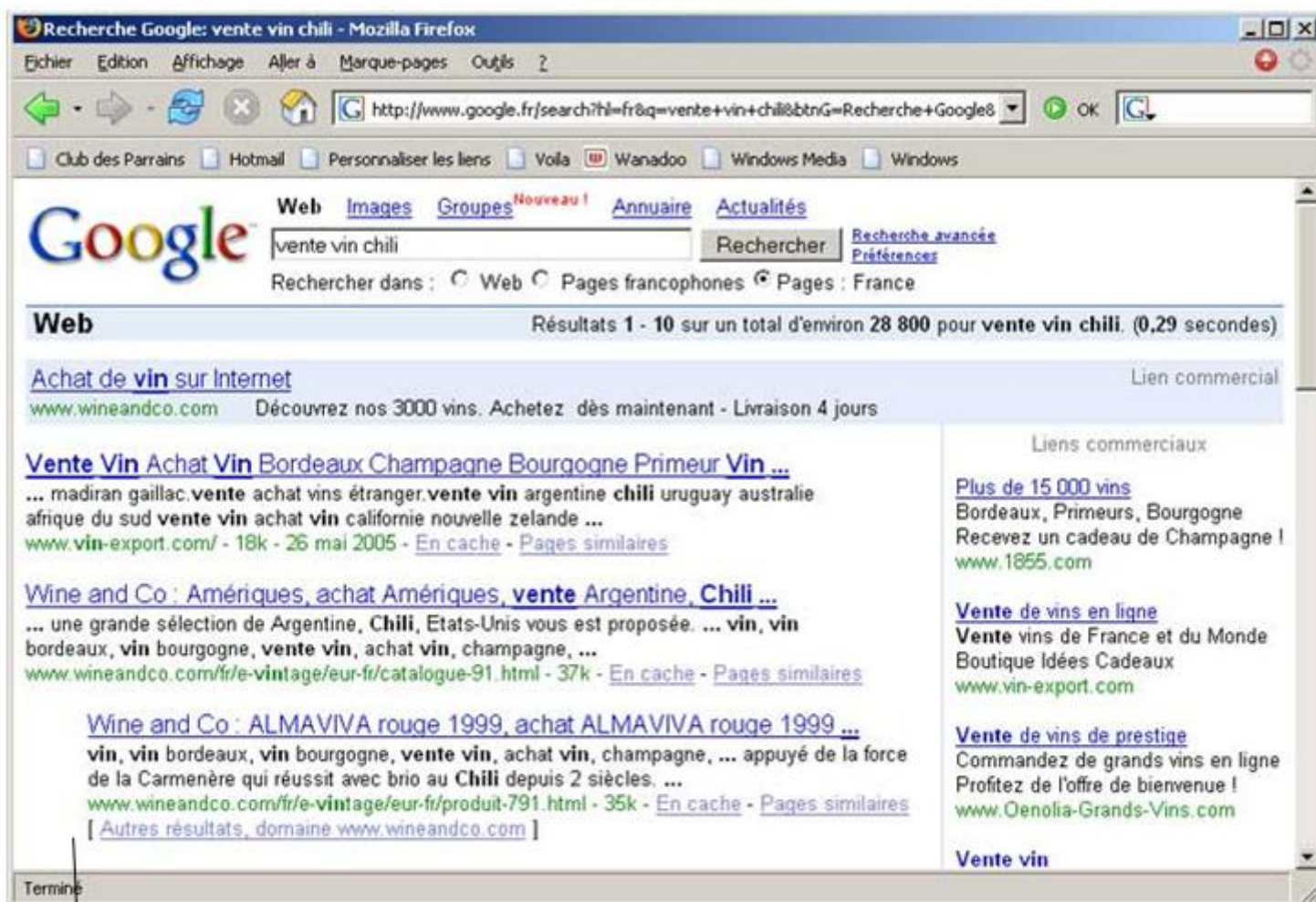
La piste la plus intéressante pour contourner ce problème est de gérer les cookies à la "demande". Si votre internaute est un humain, vous le suivez via un cookie. Si vous l'identifiez comme un robot, ou un humain refusant les cookies pas de gestion des cookies. Attention, ne changez pas le contenu de la page en fonction de l'un ou de l'autre : ce serait du cloaking sanctionné par Google.

 *Cette approche doit être réalisée uniquement sur des pages sans importance financière. En effet, un pirate peut se faire passer pour un robot, ne pas être soumis à la gestion de session du type cookie et donc accéder à ce qu'il veut sans contrôle.*

Les phases d'achat, de paiement et autres procédures critiques doivent être protégées systématiquement. De toute façon, ces pages ne sont pas intéressantes à référencer.

II-A-5 - Une source, multi sources

II-A-5-a - Les sources



Rappel : Google limite le résultat d'une recherche à 2 pages maximum par **SOURCE**. Les autres pages pertinentes sont rangées sous un lien du style : [[Autres résultats, domaine www.wineandco.com](#)]

Une SOURCE selon Google correspond à deux grands domaines :

- Des sites : <http://www.wineandco.com/>, <http://www.wineandco.org/> seraient 2 sources différentes
- Des sous domaines : <http://chili.75cl.com/> / http://champagne.75cl.com sont 2 sources différentes

Mieux, le mot clef Chili étant inclus dans le nom du sous domaine, Google prend un peu mieux en compte une telle source.

ATTENTION : il ne faut pas abuser du système et éviter absolument les erreurs suivantes :

- 2 noms de domaines différents pointent vers le même site. Google le détectera inévitablement et un de ces 2 noms de domaines sera black listé (donc perte d'investissement pour votre site)

- Trop de sous domaines avec très peu d'informations par sous domaine : idem, sanction inévitable de Google. Il black liste.
- Des sous domaines pointent vers les mêmes informations : même sanction.
- Pas assez de contenu dans un sous domaine. Il aura du mal à peser en PR et cela nuira donc au positionnement et à l'efficacité de cette source "trop légère"

La bonne méthode : gérer honnêtement les sous domaines en leur affectant des contenus réels et correspondant à leurs noms. Ex : tout sur les vins du Chili pour le sous domaine <http://chili.75cl.com>

II-A-5-b - Avantages & inconvénients

Un site peut être divisé en plusieurs sites, par ex une URL par pays, par langue :

- <http://www.captor.fr/>
- <http://www.captor.be/>
- <http://www.captor.de/>
- *etc.*

Un site peut être divisé en sous domaines. Ex : <http://www.75cl.com>, <http://chili.75cl.com>, *etc.*

Cela revient au même.

Les avantages du multi sources :

- Chaque source a des liens vers les autres : cela produit du backlink ! **Très précieux** comme nous l'abordons plus avant dans le présent document. Néanmoins, ce type d'échange entre sources est moins bien poussé qu'un réel backlink d'un site étranger vers une des pages du site.

Les inconvénients du multi sources :

- Chaque source doit être référencée : cela coûte du temps et du budget. Les efforts sont donc dispersés. Cela peut déclencher un mauvais positionnement tant que les backlinks ne montent pas en puissance.
- Chaque source doit réellement être autonome : index.html, .htaccess, canoniser la page d'accueil, un réel contenu en qualité et en quantité *etc.* : encore du temps à y consacrer.

Un site unique, permet de concentrer tous les efforts sur un seul nom de domaine. Cela se traduit par une montée en positionnement plus rapide au début avec un budget temps et argent moindre. Durant le passage du temps, la gestion des backlinks va primer dans le positionnement à accessibilité identique. Le multi domaines alors l'emportera en positionnement.

II-B - Architecture et accessibilité sémantique

II-B-1 - Préambule

L'architecture sémantique est un des points les moins abordés dans les publications spécialisées.

C'est une des clefs du référencement passif et en plus il a un impact sur la qualité de la gestion des backlinks.

Dans ce chapitre nous allons voir :

- Comment identifier et vérifier les bons mots clefs
- Comment traquer les oublis possibles en expressions clefs
- Comment combiner les mots clefs en mettant en place une architecture sémantique
- Constituer un ensemble de pages socles destinées à être retenues par Google, pages appuyées par un ensemble de pages "terreau" destinées à pousser ces pages socles en avant vers Google.

II-B-2 - Identifier les bons mots clefs (français)

Il est nécessaire de sélectionner les mots clefs liés au métier du site, du produit *etc.* C'est le point de départ.

Les mots clefs peuvent être composés. Exemple : "Airbus A340" en les mettant entre guillemets

L'étape n°2 consiste à vérifier la pertinence de ces mots clefs en étudiant les enchères sur les moteurs.

Yahoo ! et Google mettent 2 interfaces intéressantes pour réaliser cette étude :

- <http://inventory.overture.com/d/searchinventory/suggestion/?mkt=fr>
- <https://adwords.google.fr/select/Login> puis cliquer sur *Trouver des idées de mots clés*

En utilisant ces deux outils (il y en a d'autres mais ceux-ci sont les deux plus intéressants et ils sont complémentaires) on peut se faire une idée du trafic sur des mots ou expressions clefs.

Google offre en plus une recherche par synonymes et vous permet de rechercher d'autres mots clefs saisis dans des situations proches. La plupart des réponses de Google sont hors sujet mais très régulièrement, une expression clef "oubliée" sort de cette liste !

Exemple : lors d'une recherche de mots clefs pour un fabricant de système robotisé, le webmestre avait carrément oublié les mots clefs de son métier : électronique robotique et informatique robotique !


Bref, cette liste est à étudier de près systématiquement.

Les expressions clefs ainsi identifiées pourront à leur tour être auditées *via* ces outils.

On a alors une liste d'expressions clefs.

Il peut être nécessaire de les vérifier ? Comment ? Deux pistes :

- Prendre un compte Google Adword avec peu d'argent dessus et tester. On a alors accès à des outils Google beaucoup plus fins et précis que ceux de l'aide générique. On peut avoir une mesure au mot clef près saisi au lieu de vagues barres verticales et colorées en estimation de trafic !
- Analyser les logs du site en place. Il contient une rubrique : les mots clefs saisis sur tel moteur de recherche a orienté un internaute sur votre site. À étudier de près, on y fait des trouvailles intéressantes et... oubliées du propriétaire du site !

 **Google fait la différence entre singulier et pluriel, féminin et masculin : Vin est différent de Vins mais Vin = vin = VIN. La casse n'est pas interprétée. Fenêtre = FENETRE : les accents ne sont pas interprétés.**

Fenêtre est équivalent plus ou moins à FENETRE : selon les outils Google, les accents ne sont pas interprétés (ou le sont faiblement). Un site sera en rang 4 sur une requête avec les bons accents correctement positionnés et rang 2 sans accent par exemple. Les écarts de rang sont faibles. Google, comme bien d'autres outils, semble éliminer la problématique de traitements des accents en n'en tenant que faiblement compte. Nous avons constaté le même comportement dans Adwords. Il est logique que Google harmonise le traitement des caractères accentués car ils ne sont pas disponibles aisément sur un clavier portugais, anglais *etc.*

Cette approche des moteurs et de Google renforce le besoin de créer un solide socle afin :

- De pouvoir disposer de multiples variantes autour de peu de mots orientés métier : Vin, vins, vente, ventes, achat, achats, vignoble, vignobles, chardonnay, pinot noir, château, châteaux *etc.*
- De pouvoir tester différentes expressions clefs : "vente de vins" ou "achat de vins" - quelle est la combinaison la plus utilisée pour chercher un site de vente de vins en ligne par des internautes amateurs ?

- D'exploiter au mieux la combinatoire de mots clefs. Si les deux expressions ramènent chacune un trafic intéressant et différents, il faut donc architecturer son site de manière à disposer de ces 2 expressions clefs.

Exercice : les bons mots clefs pour un site de services dans les travaux du bâtiment, des maisons etc.

II-B-2-a - Un outil intéressant : les mots clefs retenus par Google sur une page donnée

Rappel : sur Google, la commande "site:http://www.votresite.com" permet de vérifier l'accessibilité technique des pages à étudier. Inutile de travailler votre page si Google ne peut la lire !

Saisir les expressions clefs d'une page visée jusqu'à l'obtenir en résultat chez Google :

- Truc & astuce : au besoin saisir dans Google 3, 5, 10 mots de votre page, mots que vous estimez clef.
- Exemple : "Développez Vos Ventés" donne en sept. 2006 : <http://www.developpezvosventes.com/> en rang 1



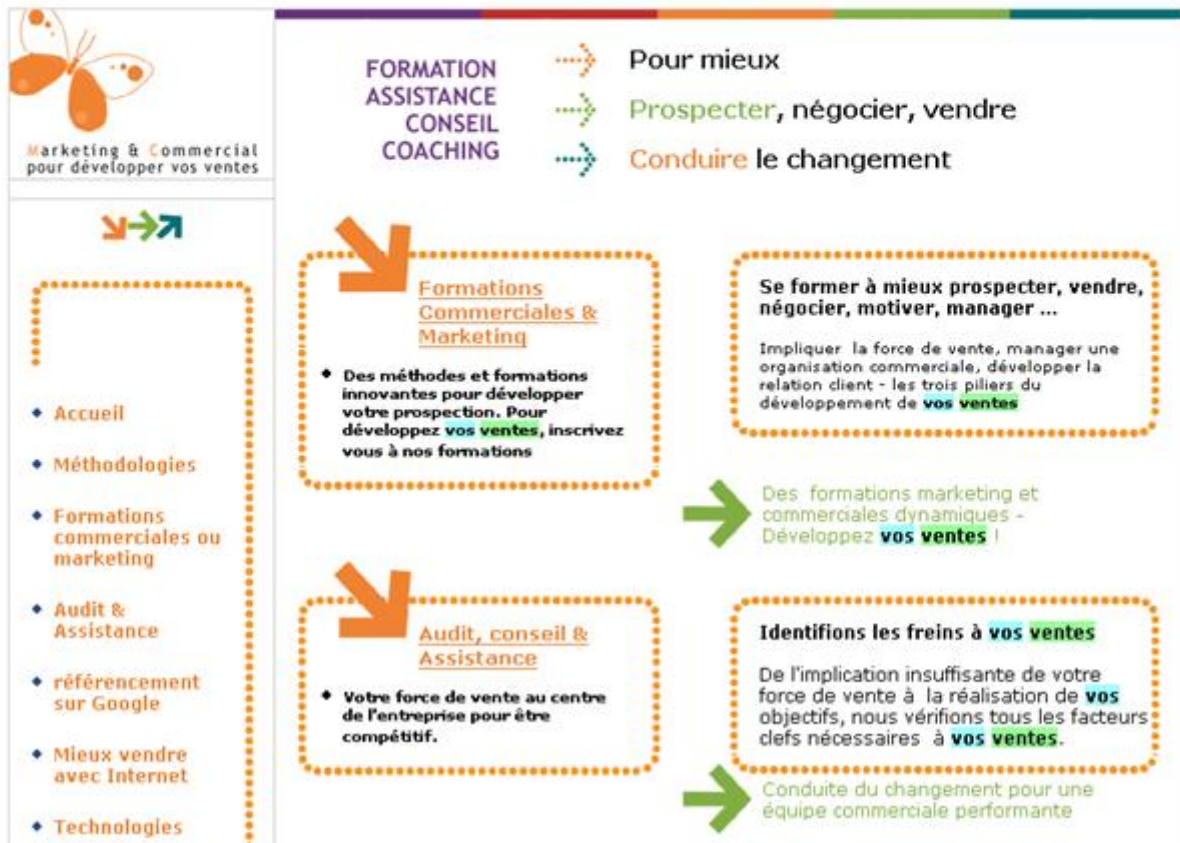
Vous obtenez alors quelque chose comme :

Voici la version **G o o g l e** de la page mise en cache de <http://www.developpezvosventes.com/> extraite le 22 sep 2006 09:11:25 GMT. La version « En cache » proposée par **G o o g l e** correspond à la page telle qu'elle se présentait lors de la dernière consultation effectuée par Google. Il se peut que la page ait été modifiée depuis cette date. Cliquez ici pour consulter [la page actuelle](#) (sans mises en valeur). Cette page mise en cache peut renvoyer à des images qui ne sont plus disponibles. Cliquez ici pour obtenir uniquement le [texte mis en cache](#). Pour créer un lien avec cette page ou l'inclure dans vos favoris/signets, utilisez l'adresse suivante :

http://www.google.com/search?q=cache:_eIq3Fe2B0kJ:www.developpezvosventes.com/+developpez+vos+ventes&hl=fr&gl=fr&ect=clnk&cd=1

Google n'est ni affilié aux auteurs de cette page ni responsable de son contenu.

Les termes de recherche suivants ont été mis en valeur : **vos ventes**
Ces termes apparaissent uniquement dans les liens pointant sur cette page : **developpez**



Pour savoir ce que Google a retenu dans le corps de page, cliquez sur en cache et après sur texte mis en cache. Tous les mots clefs seront surlignés en jaune ou en vert etc.

Surprise : j'ai saisi "developpez" et Développez n'est pas mis en sur lignage. Conclusion, Google a évolué et désormais il prend en compte les accents !!! il ne m'a trouvé que sur vos ventes ?? non. Les backlinks (nous verrons plus loin ce que c'est) poussent nombre de mots clefs "developpez" (sans accent), Google a identifié de nombreuses racines proches : développement développez (avec accent) et il en tient compte.

Ce type d'outil permet de vérifier la qualité de ses mots dans la page (fréquence, densité), leurs réelles présences (chasse aux oublis)

II-B-3 - Identifier les bons mots clefs (autres langues)

C'est le même principe. Mais les adresses des outils Google, Yahoo et autres sont différentes.

Voici une liste - attention, les sociétés changent l'URL régulièrement pour différentes raisons.

http://searchmarketing.yahoo.fr/fr_FR/rc/srch/oyr.php# nouvelle page d'accueil ouverture depuis mars 2006.

Overture

- Générateur International : <http://inventory.overture.com>
- Générateur Allemagne : <http://inventory.de.overture.de>
- Générateur Autriche : <http://inventory.overture.at>
- Générateur Corée : <http://inventory.overture.kr>
- Générateur Danemark : <http://inventory.overture.dk>
- Générateur Espagne : <http://inventory.overture.es>
- Générateur Finlande : <http://inventory.overture.fi>
- Générateur France : <http://inventory.overture.fr>
- Générateur Grande Bretagne : <http://inventory.overture.co.uk>
- Générateur Italie : <http://inventory.overture.it>
- Générateur Japon : <http://inventory.overture.jp>
- Générateur Pays-Bas : <http://inventory.overture.nl>
- Générateur Suède : <http://inventory.overture.se>
- Générateur Suisse : <http://inventory.overture.ch>

Espotting

- Générateur Allemagne : <http://www.espotting.de>
- Générateur Danemark : <http://dk.espotting.com>
- Générateur Espagne : <http://es.espotting.com>
- Générateur France : <http://www.espotting.fr>
- Générateur Grande Bretagne : <http://www.espotting.com>
- Générateur Irlande : <https://ie.espotting.com>
- Générateur Italie : <http://www.espotting.it>
- Générateur Norvège : <http://no.espotting.com>
- Générateur Suisse : <https://ch.espotting.com>
- Générateur Suède : <http://se.espotting.com>

II-B-4 - Combinatoire et choix des mots clefs

Une analyse permet donc de disposer des différentes listes de mots clefs nécessaires.

Exemple sur <http://www.75cl.com/> : des mots clefs sont plus intéressants pour la home page que pour une page sur un champagne. Vente ou achat est intéressant pour les pages en haut d'arborescence. Le mot Champagne et les mots Epernay, France, bulle / bulles sont plus intéressants comme mots clef sur une page d'un champagne.

Plusieurs dizaines de mots clefs sont potentiellement intéressants pour le groupe de pages consacré au Champagne. Il faut les lister et réaliser les rédactionnels des différentes pages pour les grouper, les factoriser, tous si possible 3 à 3, à raison de 3 mots clefs poussés en exergue par page HTML. Remarques :

- Plus il y a de pages, plus c'est facile. Une photo d'une bouteille devra disposer d'un zoom en page HTML avec un peu de texte et bien sur les mots clefs adéquats : balises alt et title, textes adéquats, liens nommés vers la page zoomant la bouteille, nom du fichier photo (évités : dv000235.jpg et préférez plutôt : bouteille-champagne-dom-perignon-epernay-France.jpg)
- **Factoriser** : Exemple d'URL mappant un répertoire : "vin-france/champagne/nom domaine - eprenay.html"

II-B-5 - "Forcer" Google sur vos mots clefs et non l'inverse

En fonction du contenu, Google fixe les mots clef de la page.

Comment influencer Google pour que ce soit VOS mots clefs décidés par vous qui soient retenus :

- Afin d'avoir la maîtrise de votre combinatoire de mots clefs
- Afin de choisir VOS mots clefs et être meilleur en positionnement que vos concurrents par exemple.

En les gérant au mieux grâce à la répartition sémantique : Un mot clef stratégique devrait être présent dans une page :

- Title de la page
- Description
- Balise H
- ALT de visuel
- P : paragraphe
- Title de lien
- Dans les textes, des backlinks pointant vers la page

II-B-6 - Augmenter le trafic grâce à la combinatoire

La combinatoire de ces mots vous permettra de "pêcher" plus d'internautes saisissant des combinaisons non prévues de mots clefs.

Exemple :

La page tête de rubrique 75cl.com/vin-France/Champagnes.html donne accès à toutes les pages liées au champagne. Chacune a un lien partant de cette page. Nulle part ailleurs on ne parle de Champagne SAUF :

- Pour indiquer qu'il y a du champagne sur le site et ce texte est associé à un backlink vers la page champagnes.html ou une des pages de cette rubrique.
- En saupoudrant un mot par ci par là pour bien conforter Google que la thématique des pages de ce site traite bien du champagne.

Voici un exemple complet :

Thème : les vins du Chili, vente de vins. L'objectif est de couvrir tous les mots clefs centraux, Vin, vins, achat, achats, ventes, vente etc. avec les mots clef propre aux vins du Chili : Chili, chilien, chiliens, Amérique du sud,

===== sur <http://chili.75cl.com/> Un extrait de ce site daté de avril 2005 =====

Edito

Nous sommes aujourd'hui fier de partager avec vous notre passion pour les vins du Chili.

Un internaute témoigne sur Tarapaca

Il a dégusté le Viña Tarapaca 1995, Reservado, et nous fait part de son enthousiasme.

Le climat et le vignoble chilien

Des conditions bénies... Le Chili est situé sur la côte ouest de l'Amérique du sud, s'étirant sur plus de 5000 kilomètres du nord au sud, mais seulement entre 90 et 400 kms de d'est en ouest.

Les cultures de la vigne sont comprises entre les latitudes 27°5 et 39°5. La situation géographique procure des conditions favorables au développement de ces cultures.

Dans le menu de gauche nous trouvons :

=====

Les grandes maisons du Chili

- Cousino Macul
- Viña Tarapaca
- Viña Undurraga
- Viña Francisco de Aguirre
- Viña San Pedro
- Viña Concha y Toro

Connaître le vignoble chilien

- Présentation générale du vignoble
- Tableau récapitulatif des vallées
- Descriptif des grandes vallées
- Comparaison des vignobles chiliens et argentins
- Petrus a fait des emules !
- In vino fatalitas
- Connaître le Pisco !

Les cépages :

- Les cépages rouges
 - Cabernet-Sauvignon
 - Merlot
 - Cabernet-Franc
 - Carménère
- Les cépages blancs
 - Sauvignon
 - Chardonnay

[Conditions de vente](#) | [Articles de presse](#) | [Vente de vins - Vin](#) - [Recettes de cuisine](#) | [Qui sommes-nous ?](#) | [Galaxie de sites](#)

====

Pour résumer : cette page est sémantiquement formidable : tous les mots clefs ont été abordés en texte, en liens. Des liens vers des pages spécialisés vont permettre de répéter tout ou partie des mots clefs de cette page avec d'autres mots clefs.

II-B-7 - Densité, occurrence, racine, thème... Une approche par l'exemple

II-B-7-a - Les textes sont analysés par Google sous différents angles

La densité des mots, des mots clefs. Combien de fois sur un nombre de mots donnés, tel mot clef revient.

Le nombre d'occurrences brut d'un mot donné dans la page.

II-B-7-b - Écrire un texte, 100 mots environ, prenant en compte des mots clefs stratégiques

Proximité

Répéter plusieurs fois les 2 ou 3 mots clefs stratégiques que VOUS voulez :

- Ex 1 : Mots clefs : casque et combat : Notre gamme de casques protège les sportifs pour tous leurs entraînements en arts martiaux et en sport de combat. Le casque ADIDAS est bien positionné pour les sports de combat type Boxe Thaï, combat où les coups portés entre combattants sont plus particulièrement violents. Le casque NIKE est adapté aux combattants en arts martiaux où les coups portés entre combattants sont moins portés La tête, crâne et visage, sont bien protégés.. Casque ADIDAS à 57 € / Casque NIKE à 47 €, tours de tête enfants à adulte.
- Ex 2 : pour une agence de styliste concevant ou prévoyant les évolutions de la mode : <http://www.space-maker.com/>
 - **Cahiers des tendances** : nos cahiers des tendances permettent à nos clients de consulter les tendances de la mode homme - femme pour les prochaines années. Ces cahiers sont publiés par Space Maker et son équipe de stylistes.
 - **C'est un peu lourd comme rédactionnel mais** : sur ce paragraphe seul, on a 3 fois "tendances" et "cahiers" ; plus que tout autre mot - Google est clairement guidé vers ces 2 mots clefs.

Différentes déclinaisons de chacun de ces 2 mots clefs et des mots connexes :

- Casque, casques
 - *Tête* (un casque va sur une tête)
 - Un *visage* et un *crâne* sont protégés par un casque
- Combat, Combats
 - *Combattants*
 - *Casque et combattant*,
 - *Coups violents* (donc *protections, protège*)

II-B-8 - Le poids des noms de sources, de fichiers HTML

Prenons l'exemple de <http://www.DeveloppezVosVentes.com/>

Une recherche sur Développez Ventes donne ce site en rang 5 de la page 1 (12/2005)

Google isole dans les noms de domaines les mots clefs accolés les uns aux autres.

La page d'accueil contient un texte utilisant plusieurs fois Développez et ventes

De même les noms de fichiers html appuient le référencement. Faites des tests sur les mots clef "Antilles" et "croisières" et vous verrez que nombre de pages utilisant ces mots clefs comme noms de fichier et leurs rangs dans la page de résultats.

Extrait ci après (12/2005): on remarquera que la plupart des sites cités ont croisières et Antilles dans le chemin d'accès donnant l'URL de la page HTML

Sail Pilot - Antilles - CROISIERES

Sail Pilot, guide pratique de croisière et de navigation sur les **antilles**.
www.sailpilot.com/antilles/croisieres/int_crois.htm - 18k - [En cache](#) - [Pages similaires](#)

Sail Pilot - Antilles - CROISIERES

ANTILLES. Escales · **Croisières** ... **LES CROISIERES AU DEPART DE GRENADNE NE SONT PAS ENCORE PRETE, REVENEZ PLUS TARD S'IL VOUS PLAIT** ...
www.sailpilot.com/antilles/croisieres/grenade/cr_gen.htm - 7k - [En cache](#) - [Pages similaires](#)
[[Autres résultats, domaine www.sailpilot.com](#)]

Croisiere Maritime avec MSC croisiereset Costa croisière

Agence de voyages spécialiste de la **croisière** maritime et de la **croisière** fluviale. Toutes destinations (Méditerranée, Caraïbes, **Antilles**, Mer Rouge, ...
www.aarcroisiere.com/ - 47k - [En cache](#) - [Pages similaires](#)

Antilles, croisières, voile

Antilles et Pacifique, **croisières** sur voiliers propriétaires (catamarans, goélette, sloops..)!
www.b3aocean.com/ - 38k - [En cache](#) - [Pages similaires](#)

Croisière Antilles. Votre croisière aux Antilles

Croisière Antilles. Votre **croisière** aux **Antilles**. Comparez et réservez votre **croisière** aux **Antilles** en 2006 avec Vacances.com.
www.vacances.com/sejour-week-end/croisiere/croisiere-antilles.php - 62k - [En cache](#) - [Pages similaires](#)

voyages Antilles croisières Antilles séjours Dom Tom

Voyage **Antilles**, **croisière Antilles**, séjour **Antilles**, Saint Martin Online , portail de l'île de St Martin vous propose un guide touristique, ...
www.saint-martin-online.com/voyages_croisieres_antilles.htm - 2k -

II-B-9 - Les autres critères

II-B-9-a - Les critères importants

Nous venons de voir que l'URL (les mots la composant) et le nom de fichier sont très importants dans le référencement.

Au sein de la page HTML, les critères suivants sont clefs :

- Title : limiter à moins de 10 mots
- Description : limiter à moins de 20 mots si possible. Mais dépassement possible si mots clefs en début.
- Présence de Hx reprenant **des mots clefs** abordés dans Title et Description

- Présence de fichiers images ayant des noms et balises ALT / TITLE contenant **des mots clefs** abordés dans Title et Description
- Présence dans le texte **des mots clefs** abordés dans Title et Description si possible appuyé par quelques balises .

II-B-9-b - Les critères moins importants

Gestion des balises Hx qui utilisent au mieux les mots clefs

Utilisation des balises STRONG et EM sur des mots clefs dans le texte.

Construire des images : plutôt que de mettre un tableau contenant des références, des dates de disponibilités et des sommes en euros, faites en une image JPEG et insérer la dans la page.

- Du texte inutile et polluant la page et noyant les mots clef est ainsi retiré.
- La densité des mots clefs augmente

Balises ALT et TITLE des visuels et autres objets. Les renseigner avec des mots clefs.

II-B-9-c - Les critères surprenants

La position dans chaque page de chaque mot ou lien ou balise (title et alt d'un visuel) . Plus le mot clef est placé vers le haut, mieux il est pris en compte.

Les noms des fichiers des visuels : table-jardin-blanche.jpg est nettement plus efficace que img00678.jpg directement issu de votre appareil photo numérique.

II-B-9-d - Plan du site

Il est important car il permet au référenceur d'en profiter pour réguler les flux de liens et mieux gérer ainsi les PR des pages du site. Il peut ainsi forcer sur des pages socles en jonglant avec des liens en dur ou en javascript. (cf plus loin les exercices sur les transferts théoriques de PR).

II-B-9-e - Localisation géographique du serveur

- Un site en .fr peut être hébergé n'importe où dans le monde. Il restera considéré comme "pages France" par Google.
- Un site en français et en .com hébergé en France sera considéré comme "pages France"
- Un site en français et en .com hébergé HORS de France sera considéré comme "pages francophones - langue française"

Attention, certains hébergeurs "français" ne sont pas en France - exemple en 2007 : 1 & 1 ce qui perturbe profondément le référencement et donc le trafic du site !

II-B-9-f - Google writing

Le rédactionnel pèse dans le référencement. N'oublions pas que nous avons à faire avec des robots. Avantages pour Google - grande puissance de traitement, faible coût d'exploitation. Inconvénient majeur : on peut tenter de le rouler (un peu) pour mieux arriver à ses fins.

Une page doit traiter d'un seul sujet - il faut acquérir la discipline pour éclater en arborescence une problématique traitée à raison de un sujet par page.

Les phrases doivent être correctes en syntaxe, grammaire et orthographe. Le spam de mot est aisément découvert par Google. Inutile de le tenter. Mais mettre en page en insistant un peu sur les mots clefs aide (la mise en exergue via les balises adéquates). Il suffit de mettre un peu plus de titres dans la page, de reprendre les mots clefs même si le lecteur humain trouve la mise en page un peu surprenante par moment.

Quelques astuces :

LA METHODE CLASSIQUE	NOTRE METHODE
 <p>Vous frappez à toutes les portes sans ménager vos efforts. Quand on est à la recherche d'une location, on fait comme tout le monde: on élargit au maximum son périmètre de recherche et on essaie de déposer des dossiers dans le plus d'agences possible. Les visites sont souvent faites à des heures fixes. Il faut donc pouvoir jongler avec ses disponibilités pour pouvoir visiter les biens qui se révèlent malheureusement trop souvent inintéressants.</p>	 <p>Vous déposez en 5 mn une demande détaillée. Notre méthode réside dans le rapprochement de l'offre et de la demande tout en privilégiant le contact utile et les échanges efficaces.</p> <p>Vous ne visitez que ce qui correspond réellement à vos attentes. Une recherche ciblée du bien et des échanges approfondis avec nos partenaires sont à la base du succès de cette offre. Les loueurs sélectionnent en fonction de critères très</p>

Ce texte mis en exergue, en bleu gras, ne véhicule pas de mot clef pour Google.

En effet, c'est une balise qui le passe sous cette forme.

Il faudrait plutôt l'appel en CSS à une balise du type H (h2 ou h3) définie en bleu, 12px, gras.

SPECIAL ETUDIANTS



Trouvez facilement votre logement.
 Location, co-location, cautions et garanties, les agents partenaires ont mis au point des offres spécial étudiants qui vont vous faciliter les démarches pour trouver rapidement votre futur logement.

L'intérêt du site ici est de s'adresser simultanément à Google et aux humains surfant sur la page.

"SPECIAL ETUDIANT" est une image qui contient un texte séduisant pour les humains mais guère intéressant en référencement sur Google.


Ce visuel peut recevoir un texte en ALT et en TITLE qui sera pas ou peu lu par les internautes mais lu par Google qui ne sait pas lire un texte dans un visuel :

Plein de mots clefs intéressants :

```
alt="logement, appartement, chambre pour étudiants en location, co-location et à louer"
```

Sous ce visuel SPECIAL ETUDIANT, on pourrait mettre en bleu entre balise h2 "étudiant, trouve ton logement à louer". On remarquera que j'ai mis étudiant au singulier volontairement. Et re mots clefs mais sans saouler le lecteur.

Facilement serait à la ligne et dans dans le même style visuel mais avec des balises span + style. En effet, aucun intérêt à pousser en exerque vers Google le mot "facilement" mais intérêt à le maintenir visuellement fort pour le lecteur humain.

 *Le texte "Location, co-location etc." doit être mis entre balise <p>*

II-B-9-g - Les claques DHTML

Cette technique permet de mettre dans une page des textes qui apparaissent uniquement sur événement mais ces textes, quand la technique est en css permet de les avoir lissibles par Google, toujours, rarement par les humains ce qui présente quelques intérêts.

II-B-9-h - Augmenter le socle du référencement => Création de trafic qualifié et ciblé

En étudiant les statistiques de votre site et la gestion des mots clef sur le net, vous déterminerez de nouveaux mots clefs, moins usités mais néanmoins utilisés qui pourraient vous conduire en première place sur certaines requêtes.

Voici un exemple simple :

Vous êtes un jeune cabinet comptable. Le mot comptable est relativement disputé. Le mot Essonne aussi. Mais la combinaison Comptable - Essonne - 91 sur une même page est peu utilisée par les experts comptables de l'Essonne. En accentuant sur différentes pages du site cette combinaison de trois mots, vous arriverez en tête de page de résultats sur un internaute recherchant un expert comptable dans l'Essonne.

Puis vous ajoutez de nouveaux mots clefs sur ces pages : Expert, commissaire aux comptes, normes IFRS etc.

Puis vous ajoutez de nouvelles pages avec les nouveaux mots clefs : commissaire aux comptes, normes IFRS etc.

Des pages Rank se transfèrent très vite.

Une recherche sur "IFRS 91" vous liste en tête de cette nouvelle combinaison peu usitée... pour le moment.

II-B-9-i - Mettez votre site à jour

Nouveaux liens entrants ou sortant, nouveau textes etc. un site maintenu est mieux noté par Google

II-B-9-j - Un JPEG à mettre en zoom

Cela consomme du temps mais si vous en avez, créez une page HTML qui contient le zoom de la photo et profitez-en pour mettre un peu de texte et des Meta et autres critères. Le nombre de pages d'un site influe sur son positionnement indirectement (plus de BL internes, plus de terreau etc.)

II-B-9-k - Gérer les pages trop longues en texte

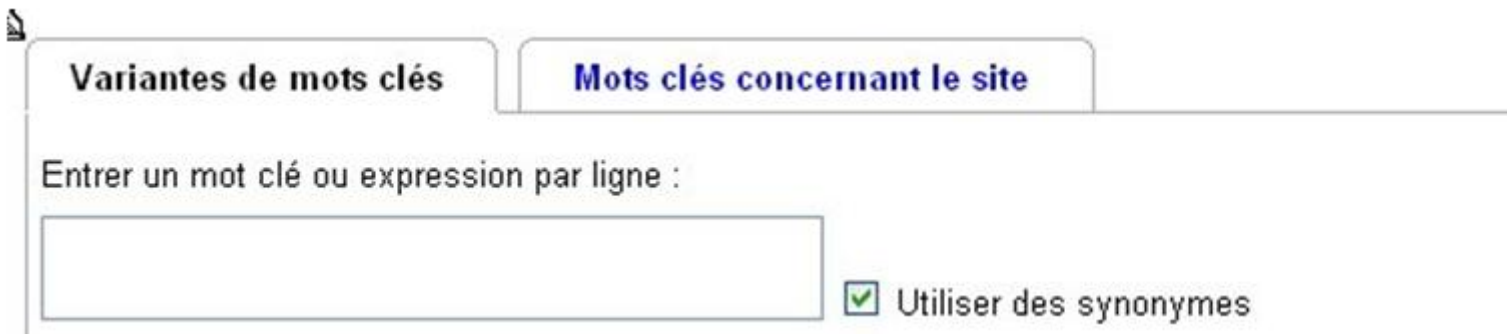
Vous avez une page de plusieurs centaines de mots. Scindez-la en plusieurs pages via une navigation locale. Ce sera l'occasion pour vous de créer des liens nommés internes au site pour expliciter leur navigation... et pousser des mots clefs via des BL internes (cf. chapitre référencement actif).

Une page doit faire un minimum de 300 mots significatifs. A partir de 700 à 1200 mots, il est temps de songer à scinder la page en plusieurs pages reliées par la navigation.

II-B-10 - Les mots clefs utilisés par vos concurrents... Ou par vous-même !

Étudier la concurrence fait partie du travail du référenceur. Selon les moments, il y a plus ou moins de méthodes.

- Utiliser les outils de Google (ou d'autres !) sur les sites de vos concurrents



- L'outil d'aide Google pour identifier des mots clefs (générateur de mots clefs)
- <https://adwords.google.fr/select/KeywordToolExternal?defaultView=2>
- L'onglet "Mots clefs concernant le site" propose de lister les mots clefs en rapport avec une page d'un site WEB
 - Google analyse la page et liste les mots clefs pertinents de cette page selon lui
- Limite du système : attention, nombre de sites sont à côté du sujet de l'entreprise. En clair, le site ne reflète pas l'entreprise. C'est très bien pour vous, un concurrent incompetent sur Internet favorise vos propres affaires. Il va de soi qu'il ne faut donc pas tenir compte des résultats de tels outils pour concevoir votre propre stratégie de référencement et positionnement.
- On peut aussi étudier le site et ses meta, textes, titre pour évaluer les mots clefs réellement poussés et ceux qui seraient... nécessaires à cette entreprise et à son site.

II-C - htaccess et canoniser un point d'entrée

Google doit avoir un seul point d'entrée sur le site :

- <http://www.lesite.com/> : c'est l'idéal. Choisir autre chose, c'est possible, mais cela demande un niveau expert pour bien comprendre ce qu'il se passe et pourquoi et comment faire !!
- Et non pas <http://www.lesite.com/index.html>
- Et non pas <http://www.lesite.com/accueil.html> et autres du même genre.

Cela permet de concentrer sur un seul point d'entrée les propriétés (positives) liées au référencement. Sinon, le PR va être étalé, les définitions vont rentrer en conflit, bref, de la dispersion, ennemie du bon référencement, surtout quand le site est en compétition féroce sur des mots clefs.

La gestion adéquate du fichier .htaccess permet cette optimisation.

Certains hébergeurs ne supportent pas ou mal le canonical : OVH par exemple.

II-D - Héritage, refonte de site, migration

II-D-1 - Protéger ses investissements

Google attache un PR et des propriétés à chaque page qu'il indexe. Une page est identifiée par un code unique sur la planète WEB : son URL. Il ne peut en exister 2 identiques !

Que se passe-t-il si vous changez l'URL de votre page ?

- Car vous changez la page de répertoire pour des raisons d'administration de fichier (passer de la racine vers Fra les pages en français, vers ENG les pages en anglais etc.)
- Car vous changez le nom du fichier : de "page154.html" vous la renommez "table-de-jardin-blanche.html"
- Car vous changez le nom de domaine

=> Vous perdez tout. En clair, vous jetez des mois de travail, d'ancienneté, d'investissement à la poubelle.

Comment limiter la perte ? En gérant un héritage. Il faut faire comprendre à Google que la page a déménagé.

Plusieurs techniques peuvent permettre un tel héritage.

II-D-2 - .htaccess

L'utilisation du fichier .htaccess indique à Google et aux autres les redirections 301 (définitive) ou 302 (provisoire). La page "page154.html" est alors mappé par le serveur Apache vers la page "table-de-jardin-blanche.html"

- Tout internaute appelant cette page sera redirigé en temps réel vers la nouvelle adresse.
- Google "interprète" .htaccess en analysant les headers http et en tiendra compte lors de la prochaine Google Dance ou de son équivalent en fil du temps.

Cette ligne dans un fichier .htaccess redirige l'internaute de index.html vers accueil.php :

```
Redirect 301 /index.html /accueil.php
```

II-D-3 - Liens en dur

Vous laissez la page en ligne et vous mettez un lien en dur vers la nouvelle avec un texte explicatif.

ATTENTION, multiples pièges et astuces :

- Duplicate content n°1 : le texte d'accompagnement est le même pour toutes les pages ainsi re-routées et vous avez vidé le contenu des pages anciennes : les nouvelles pages seront en black liste ! Donc texte très court et très changeant.
 - Faisable que sur quelques pages, après forte consommation de temps homme !
- Duplicate content n°2 : malin comme un singe, vous laissez les textes dans les anciennes pages. Oui mais vous avez copié ces textes dans les nouvelles ! Elles seront mises en black liste !! Perdu !
- Les pages anciennes et nouvelles sont dans le même site. Faites un .htaccess. Plus simple et plus sur.

- Les pages sont dans deux noms de domaines différents : pour avoir du backlinks "externe" cela peut valoir le coup de garder les anciennes pages (ou une partie de celles-ci) et de personnaliser page par page un lien nommé (cf. le chapitre consacré à ce sujet important) vers la page correspondante.

II-D-4 - Redirection en HTML

Cette méthode, plus simple mais bien moins propre et efficace que .htaccess, utilise les documents HTML eux-mêmes, pour indiquer leur déplacement.

Cette méthode ne devrait être utilisée que lorsque les redirections http via .htaccess sont impossibles.

- <http://ancienne.adresse.fr/dossier1/page1.html> vers <http://nouvelle.adresse.fr/dossier2/page2.html>

page1.html

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
<head>
  <meta http-equiv="content-type" content="text/html; charset=ISO-8859-1" />
  <meta http-equiv="refresh" content="0; url=http://nouvelle.adresse.fr/dossier2/page2.html" />
  <title>Redirection</title>
  <meta name="robots" content="noindex, follow" />
</head>
<body>
  <p><a href="http://nouvelle.adresse.fr/dossier2/page2.html">Redirection</a></p>
</body>
</html>
```

II-E - Illustrations des calculs de PR (Page Ranking)

II-E-1 - Quel intérêt à comprendre puis à mettre en œuvre les calculs de PR ?

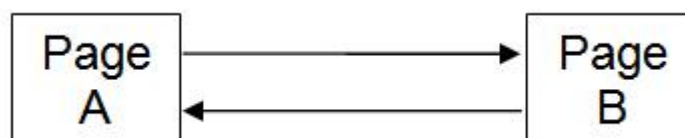
Nous avons vu dans ce document comment le PR est géré (approximativement) par Google.

Nous avons vu qu'il est intéressant d'architecturer le site en sémantique :

- En pages de socle destinées à devenir des pages à sortir sur Google
- En pages "terreau" destinées à pousser les pages socles en évidence

Il est de l'intérêt du site de pousser les pages socles par du transfert de PR depuis les pages "terreau & engrais"

II-E-2 - Prenons le Web dans un cas théorique : deux pages existent - A pointe vers B et B pointe vers A



Donc, $C(A) = 1$ et $C(B) = 1$

Le PR est inconnu. Une 1ère itération avec une valeur de départ de 1 donne :

$$PR(A) = (1 - d) + d(PR(B)/1) \quad \text{ou} \quad PR(A) = 0.15 + 0.85 * 1 = 1$$

$$PR(B) = (1 - d) + d(PR(A)/1) \text{ } \hat{=} PR(B) = 0.15 + 0.85 * 1 = 1$$

Le PR converge vers 1.

Prenons une valeur de départ différente : 0

Première itération

$$PR(A) = 0.15 + 0.85 * 0 = 0.15$$

$$PR(B) = 0.15 + 0.85 * 0.15 = 0.2775$$

Deuxième itération

$$PR(A) = 0.15 + 0.85 * 0.2775 = 0.385875$$

$$PR(B) = 0.15 + 0.85 * 0.385875 = 0.47799375$$

Troisième itération

$$PR(A) = 0.15 + 0.85 * 0.47799375 = 0.5562946875$$

$$PR(B) = 0.15 + 0.85 * 0.5562946875 = 0.622850484375$$

Nous remarquons que les valeurs augmentent à chaque itération. Google a précisé que le PR moyen de toutes les pages du WEB doit être égal à 1. Nous constatons que dans notre exemple, nos deux pages constituant le WEB mondial voient leurs PR moyens converger vers 1.

Prenons une valeur de départ différente : 2 (nous aurons ainsi testé 1, moins que 1, plus que 1)

Première itération

$$PR(A) = 0.15 + 0.85 * 2 = 1.85$$

$$PR(B) = 0.15 + 0.85 * 1.85 = 1.7225$$

Deuxième itération

$$PR(A) = 0.15 + 0.85 * 1.7225 = 1.614125$$

$$PR(B) = 0.15 + 0.85 * 1.614125 = 1.52200625$$

Troisième itération

$$PR(A) = 0.15 + 0.85 * 1.52200625 = 1.4437053125$$

$$PR(B) = 0.15 + 0.85 * 1.4437053125 = 1.377149515625$$

Les PR de nos 2 pages constituant le WEB mondial convergent vers 1. Cela fonctionne. La moyenne normalisée des pages du WEB tend vers un dans notre exemple, comme indiqué par Google.

II-E-3 - L'efficacité de la formule pour accélérer les calculs de PR d'une page

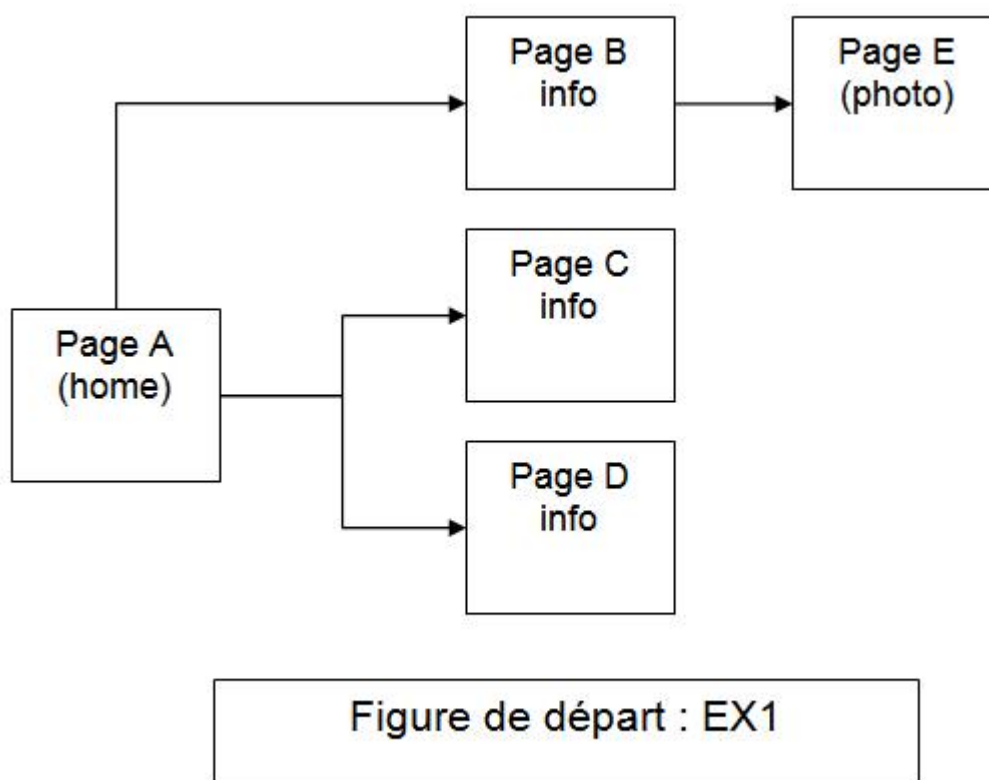
Le facteur d'amortissement permet de diminuer le nombre d'itérations nécessaires. S'il est trop élevé, cela exige beaucoup d'itérations pour permettre la convergence. Cela consomme trop de ressources.

Si le facteur est trop bas, la convergence approche le résultat escompté. C'est insatisfaisant.

Refaites l'exemple ci-dessus pour appréhender cela en remplaçant le coefficient 0,85 par 0,2 puis par 0,95.

La valeur 0,85 est la valeur idéale pour effectuer cette mission.

II-E-4 - Un exemple avec plus de pages



EX1 : Nous avons un site cas d'école : Une home qui pointe vers 3 pages secondaire d'informations, une page d'info B qui pointe vers une photo incluse dans une page HTML plutôt que un fichier JPEG.

La home reste à une valeur mini, peu intéressante pour nous. B, C, D monte un peu. La photo récupère le PR le plus élevé du site, ce qui est peu intéressant pour nous.

EX2 : Les pages B, C, D pointent vers la page A : nous ajoutons une navigation sommaire à EX1.

Les PR augmentent. Bien. Mais la home augmente peu. Car elle favorise les pages de niveau inférieur. La photo possède un PR supérieur à la home. Sans lien de retour vers la home, cela implique qu'un internaute pourrait arriver sur cette page sans possibilité aisée de remonter sur la home du site (ou sur une autre page du site). Pas terrible.

EX3 : Ajoutons un lien de la photo (page E) vers la home pour pallier le défaut constaté en EX2.

Le PR augmente fortement. Bien. Le focus de notre référencement en PR est la home page. Cela peut être intéressant si la stratégie globale du site consiste à disposer des messages marketing clef sur la home page et à vouloir attirer un maximum d'internautes vers la home.

EX4 : Déplaçons le lien de la photo (page E) vers une autre page info du site, la page B, d'où est issu le lien vers la photo.

La home reste la page en PR le plus élevé et une page de rubrique est favorisée aux dépens des autres

Premier commentaire : On peut faire monter le PR d'un site pour obtenir un PR moyen de un. On peut, grâce à la gestion des liens internes, influencer sur la répartition du PR pour favoriser la ou certaines pages clefs du site. A mots clef équivalent, les pages à PR élevées seront favorisées.

EX5 : Mettons en place une vraie navigation. La home pointe vers les têtes de rubriques B, C, D. Chaque tête de rubrique pointe vers la home et vers les têtes de rubrique autres. B pointe vers C et D, C vers B et D, D vers B et C.

La page photo améliore le socle de référencement du site (une page en plus harmonisée au thème de son groupe de page avec textes, mots clefs etc.) mais elle étale le PR sur les pages autres que la home (sauf à lui coller un lien de elle vers home).

Second commentaire : Une gestion fine des pages de photo (ou autre) permet de distribuer le PR à sa convenance tout en augmentant le socle.

Page	EX1	EX2	EX3	EX4	EX5
A	0.15	0.18	2.20	1,79	1,04
B	0.19	0.48	0.77	1,23	1,42
C	0.19	0.48	0.77	0,66	1,04
D	0.19	0.48	0.77	0,66	1,04
E	0.31	0.36	0.45	0,62	0,45

II-E-5 - Confirmation

En juin 2007, Google via son blog officiel (Matt Cutts) confirme que le transfert de PR est toujours présent dans ses algorithmes.

II-E-6 - Ce que nous venons de voir est partiellement faux - Le test

Le Page Rank est beaucoup plus complexe que ce que nous venons de voir (qui est simplement incomplet et non pas faux).

Voici un témoignage issu du web :

Lors de la modification du code source de son forum, Dan, le webmestre de Webmaster Hub expose un contexte unique pour mieux comprendre Google.

Le nouveau programme informatique du forum de Webmaster Hub mis en place le 15 mai 2004 entraîne une génération massive et soudaine de 30000 à 40000 liens sortants depuis Webmaster Hub.

Étudier l'évolution du Page Rank de ce forum et site lors des prochaines Google dance, permettra de mieux comprendre la fuite du Page Rank exposé dans les équations vues ci-dessus.

12.6.1.1 Mise à jour suite à la Google dance du 31 mai 2004

Webmaster Hub est passé de PR6 à... PR7. D'autres pages du site montent à PR6 ou PR5. l'arrivée brutale de backlinks n'a donc pas fait diminuer son PR, au contraire.

12.6.1.2 A suivre...

Les liens sont la base du web. Les premiers résultats de ce test montrent que les liens sortants massifs n'ont pas eu d'effet négatif sur le PR, au contraire.

Des liens sortants faits judicieusement auraient un effet positif sur le positionnement dans Google, élément beaucoup plus important pour un site que le PageRank de ses pages.

De même, des liens sortants maladroits ont un effet négatif sur le site. Les équations vues plus haut ne sont pas fausses. Simplement elles changent selon l'importance et le poids du site.

Je ne recommanderai pas à un jeune site de se doter de milliers de liens sortants !

Autre contre exemple : les annuaires de liens. Ils ont de forts PR et pourtant des milliers de liens en sortent !

III - Le référencement actif

III-A - Le principe

III-A-1 - Un bombing

C'est un jeu (de mauvais goût ?) qui consiste à référencer un site à son insu sur des mots clefs non désirés.

Exemple : dans Google, saisir "miserable failure" (traduction : *échec lamentable*) puis cliquer. Le résultat sera... la biographie de Georges W Bush sur le site fort respectable de la maison blanche.

Mise à jour de juin 2007 : ce bombing a été bloqué, probablement par une mise à jour de Google qui bloque désormais la montée en pertinence s'il n'y a pas de cohérence complète ou partielle entre le lien et les mots clefs de la page visée par le lien.

Ce bombing montre quelque chose de fort intéressant pour créer du trafic sur un site - on peut référencer un site de l'extérieur sans toucher au site lui-même !!! Il est fortement recommandé d'avoir une cohérence entre les pages et les mots clefs du lien. Cela deviendra obligatoire dans le futur avec les évolutions de Google.

Comment ? Grâce aux backlinks (les BL dans le jargon)

C'est un lien qui part d'une page d'un site hébergé sur un serveur dans un pays et pointe vers une autre page.

Tout lien peut être nommé :

```
<a href="daube-provence.html">recette de cuisine : la daube provençale</a>
```

La page 'daube-provence.html' se voit indexée par Google pour les mots clefs :

- recette
- cuisine
- daube
- provençale
- provençale

III-A-2 - Utiliser les BL

Les BL sont très puissants dans les algorithmes de Google. Une fois votre site rendu accessible techniquement et sémantiquement, vous pouvez faire des campagnes de BL ! Et doubler vos concurrents trop lents, trop maladroits sur le positionnement de votre site dans Google.

III-B - Bien construire un backlink

Le lien doit être précédé de texte et suivi de texte, contextuellement correct.

- Si le lien pointe vers une page gastronomie avec mots clefs "recette cuisine daube provençale", c'est mieux si du texte connexe à la gastronomie encadre le lien : vins, repas, restaurant etc.
- Si possible, la page de départ doit avoir du contenu contextuel, pas seulement le paragraphe contenant le lien
- Le lien doit être nommé, contenir des mots clefs : "recette cuisine daube provençale"
- La page visée doit traiter de ces mots **si possible** dans les meta TITLE, DESCRIPTION, balises Hx, texte.

Exemple :

```
<p>
Vous désirez partir en vacances au soleil ? Nous vous recommandons l'Espagne sur la <a href="http://
www.locations-costa-del-sol.fr">Costa Del Sol,
en location près de la mer et des plages</a> pour disposer d'appartements et maisons à louer en
Andalousie, agréables et bénéficiant de 300 jours
d'ensoleillement annuel et de températures clémentes pour le farniente et les jeux sur la plage ou en
mer.
</p>
```

Disposer un tel paragraphe le plus haut possible dans la page.



Si de tels liens vous posent des problèmes marketing : prendre une disposition multi colonnes dans votre page et disposer de tel paragraphes dans la colonne de droite. En effet, grâce à la feuille de style, vous pourrez disposer le conteneur (la boîte) en position haute, définir la boîte très tôt dans la page et donc aux yeux de Google elle est en haut donc lourde en charge (beaucoup plus qu'en bas de page !)

Avec un tel paragraphe, vous pouvez tenter de le mettre sur une page qui traite d'un sujet totalement différent. Le changement de thématique est amorti par le texte environnant. Google est nettement moins sévère avec ce type de présentation. En effet, est ce un piège de référenceur ou un réel point intéressant d'un site ayant un webmestre un peu maladroit ? Il ne peut le savoir et il ne peut pas sanctionner tous les webmestres maladroits de la planète, son index serait vide ! Donc il ne pénalise que peu ou pas un tel procédé ! Néanmoins, il ne faut surtout pas abuser de cette astuce. En effet, trop de liens "bizarres" sortant d'un site pourront alerter les algorithmes de Google. User sans abuser, est une règle d'or.

Moralité : toujours mettre Google dans l'embarras pour tricher.

III-C - Cohérence contextuelle

III-C-1 - Explication globale simplifiée

Google considère plus ou moins mal tout ce qu'il considère comme une tentative de plus ou moins leurrer et donc de berner les internautes utilisant Google.

La notion de thématique est impliquée dans cette chasse par Google

Google classe toute page par rapport à une thématique :

- gastronomie
- emploi
- etc.

Regarder la partie *annuaire* de Google renseigne sur l'organisation thématique qu'il respecte lui-même.

Notions Importantes :

- Thème central
- Thème connexe
- Autre thème

Chaque lien d'une page A (d'un site A) vers une page B (d'un site B) se voit affecté d'une note de pertinence qui prend la forme suivante :

couple (pertinence thématique centrale de A vers B, pertinence thématique connexe de A vers B)

Exemple pour bien faire comprendre cette notion fondamentale :

Un lien du site de vin 75cl.com vers champagne.75cl.com est noté comme totalement thématique centrale : une source de vin est liée à une source de vin spécialisée mais de vin quand même.

75cl a aussi des pub (annonces gooooogle AdSense), gérées dynamiquement par Google donc, pleines d'enseignements. Il y a une annonce pour un site d'électroménager ayant un rayon dédié aux caves à vins !

75cl a aussi des liens vers des sites de gastronomie. Nous remarquerons que ces liens existent via des pages qui parlent vin et nourriture - gastronomie. Ces thèmes sont connexes pour le site (75cl) et pour Google. Google note ces pages avec une forte pertinence Thématique Connexe et une faible pertinence thématique centrale

Si 75cl avait un lien vers, disons un site de pièces détachées automobiles, Google noterait ce lien à (0,0) sur thématique centrale et connexe :

- Dans le meilleur des cas, Google ignore le lien ou le note faiblement en PR
- Dans le pire, la note de PR est négative ! Google peut interpréter ce lien sans signification comme une tentative de le leurrer.

Si vous voulez faire un lien avec une note connexe, alors il faut passer par des pages de transition :

- 75cl, site de vins, pointent vers sécurité routière (automobile, ne pas abuser de l'alcool et du vin) et de cette page sécurité routière pointer vers site de pièces détachées automobile. Là la pertinence est bonne, Google attribuera une pertinence élevée.

III-C-2 - Explications plus complètes sur les BL

La cohérence d'un backlinks s'étudie dans l'espace et dans le temps :

III-C-2-a - Dans l'espace

- Page de départ : sa thématique, ses mots clefs
 - Mots clefs dans TITLE de page et dans DESCRIPTION
 - Mots clefs dans les titres entre balises h

- Mots clefs dans texte et entre balises ALT
- Mots clefs dans nom de fichier, URL de la page.
- Idem pour page d'arrivée
- Mots clefs véhiculés par le lien

III-C-2-b - Dans le temps

- Date de pose estimée du backlink dans la page
- Date de dernière modification de la page de départ
- Date de création puis de dernière modification de la page visée
- Évaluer les modifications thématiques et mots clefs de ces 2 pages au fil du temps
- Date modification des mots clefs du backlink.

III-C-3 - Les backlinks internes

Un site issu d'une ou plusieurs sources.

<http://www.75cl.com/>

<http://chli.75cl.com/>, qui a sa propre page [index.html](#) ou [index.php](#).

Les liens échangés entre sources d'un même site (ici 75cl.com) sont évalués par Google, et pris en compte.

Les liens échangés entre pages d'une même source sont également pris en compte.

D'où l'intérêt de mailler correctement son site et correctement afin de doper le positionnement sans faire appel à des liens externes loués ou échangés.

III-D - Les limites du système des BL

De nombreux pièges se cachent dans le déploiement de backlinks non naturels. En voici une liste non exhaustive.

III-D-1 - Pièges dans l'espace

Google n'étudie pas tous les BL entre 2 sites à chaque fois. Cela lui consommerait trop de ressources. Il préfère les statistiques et investir de la ressource de ses serveurs sur les anomalies.

Voici un exemple :

- Un site est spécialisé dans l'automobile.
- 3000 liens quittent ce site.
- 95 % sont échangés avec des sites plus ou moins liés au monde de l'automobile.
- 5% sont échangés avec des sites hors automobile.
- Cela fait-il des 5% de liens "anormaux" des liens non naturels à éplucher ou à sanctionner ?
- Réponse : non. Si le taux de liens hors contexte passait un seuil que seul Google connaît, disons 10%, alors là il y aurait une reconsidération de la situation par les algorithmes de Google, voire une vérification humaine par quelqu'un de l'équipe Adwords qui est en compétition avec les liens non naturels.

Google évalue de très nombreux critères caractérisant une page et son site :

- Classe C (cela donne l'hébergeur, le pays *etc.*)
- n° IP : identification plus ou moins fiable du serveur
- Le whois : il donne le propriétaire, son email, l'administrateur, le contact technique.

La théorie des graphes est une branche des mathématiques très intéressante. Elle permet de poser un ensemble de points comme des sites WEB par exemple et de les relier par des arcs, des liens en d'autres termes. Sur un tel ensemble, on peut appliquer des algorithmes mathématiques fort intéressants et ainsi cartographier la toile *via* des lois mathématiques.

On peut par exemple identifier les interactions entre sites et y détecter des anomalies comme par exemple : pourquoi ce site A aurait-il 100 backlinks vers ce site B ?

Il y a de multiples raisons possibles :

- Signature d'un designer qui a signé en pied de page et non pas en "page crédit graphique" par exemple
- Ou un petit malin qui a mis des liens non naturels

Autre exemple classique :

- Un lien relie le site A au site B et inversement
- L'échange de liens entre site est non naturel, donc Google commence à annuler ce type de liens.

Les exemples sont nombreux.

III-D-2 - Pièges dans le temps

Une page peut évoluer et le lien qui la pointe n'évolue pas. Peu importe que ce soit la page de départ ou la page d'arrivée du lien qui évolue et pas les autres membres du triplet (page départ, lien, page arrivée). Cela peut rendre le triplet obsolète et inciter Google à annuler le lien.

Si le triplet n'évolue pas, l'ensemble devient obsolète et le lien est donc annulé.

Le triplet devient quartet quand on y ajoute la thématique de la rubrique de la page (rappel : la théorie des graphes permet de détecter le groupe de pages d'une même rubrique rien qu'en parcourant les liens de navigation)

Et puis on peut au final y ajouter la thématique du site, puis des rubriques connexes

Enfin, on peut ajouter la thématique du paragraphe contenant le lien, la thématique du bloc de page contenant le paragraphe etc.

Mes mesures montrent que Google pondère tout ceci dans la note donnée au lien.

Autre dimension du temps, la vitesse d'établissement des liens et de leurs qualités.

Illustration :

- Un site A voit soudainement après des mois d'inactivité et aucune mise à jour sérieuse du site et de ses contenus, ce site A donc voit quelques pages soudain pointées par de multiples pages de bonne qualité.
- La théorie des graphes plus un différentiel temporel mettront ainsi en lumière l'arrivée d'un référenceur et l'irruption de liens non naturels.
- Il faut y aller :
 - Progressivement
 - Et en noyant le poisson.
 - Et en restant le plus possible dans le contexte. Il est normal que quelques liens soient "hors contexte", ce serait une anomalie statistique sinon.

Exemple : pour monter le mot clef d'un site A, ce mot clef étant à l'origine de 90% du new business sur les ventes de petits objets voici les actions que j'ai réalisées successivement sur 6 mois :

- Ré écriture Google writing des pages qui vont être visées par la campagne de backlinks.
- Re référencement des pages en question

- Pose des premiers liens, de la faible qualité et surtout longues à être lues par les bot de Google.
 - Google détecte : mise à jour d'une partie du site
 - Un peu plus tard arrivée de quelques liens
 - Logique quoi !
- Montée en puissance de la qualité des liens étalée sur 5 mois et issus de multiples sites
- A N + 6 mois, le mot clef a le rang 7 face à un million de pages éligibles. Le CA de mon client a sensiblement augmenté.
- Ne pas être pressé, faire attention, être méthodique, connaître et comprendre Google : ce sont les clefs du positionneur efficace.

La cohérence globale est notée : mots clefs de la page de départ, paragraphe, bloc, page (voir l'utilitaire Google de Adword qui nous les indique gratuitement, Google Webmaster tools le propose aussi), mots clefs du lien, mots clefs de la page d'arrivée, mots clefs de la page tête de rubrique, mots clefs de la page d'accueil.

III-D-3 - Conclusion

Pour les 2 parties, loueurs et locataires il faut diversifier. Peu de liens allant d'un même site vers un même site. Contextualiser au maximum. Surveiller les évolutions des pages interconnectées et les mots clefs véhiculés par les liens.

Le loueur de liens aura intérêt à louer peu de liens à chaque locataire et inversement. Et pas trop vite.

III-E - Échanges de liens pour pousser des BL vers vous et les autres

III-E-1 - Que du PR

```
<a href="http://www.developpezvosventes.com"/>
```

Ce BL ne transfère le nom de domaine (qui est unique ou presque !!) comme mots clefs. Seul intérêt, il y a du PR de transféré.

Dans certains cas, certaines inscriptions en annuaire, il est impossible de nommer le BL.

III-E-2 - Mais aussi des mots clefs

```
<a href="http://www.developpezvosventes.com">développement des ventes grâce au référencement sur  
Google</a>
```

Ce BL est nommé et pousse "développement" (même racine que développez) ventes, référencement, Google.

Il y a aussi bien sur du transfert de PR.

III-E-3 - Les différents types de BL

- **Liens internes** à votre site. Exemple : Riesling d'une page de <http://www.75cl.com/> pointe vers la page <http://www.75cl.com/riesling.html> : le mot clef Riesling est poussé en pertinence vers cette page.
 - Le poids d'un tel BL est faible mais réellement existant. À faire donc.
- **Liens externes** à votre site mais le lien pointe vers une page dans un de vos sous domaine (ex : lien Vins du Chili d'une page de <http://www.75cl.com/> pointant vers la page [chili.75cl.com](http://www.chili.75cl.com/)). Les mots clefs vins et Chili sont poussés en pertinence vers la page d'accueil [index.html](http://www.chili.75cl.com/index.html) du sous domaine [chili.75cl.com](http://www.chili.75cl.com/)
- **Liens externes** à votre site : Riesling pointe vers l'office de tourisme adéquat en Alsace par exemple. Et ce site disposera d'un lien "ventes de vins Riesling" vers la page adéquate dans [alsace.75cl.com](http://www.alsace.75cl.com/)

- **Lien dans la même langue** ou dans un langue étrangère : à priori Google interprète les thèmes de chaque page quelque soit la langue.
- **Lien sortant et entrant sur un même serveur** : Google pénalise car il soupçonne une "collaboration hors contexte"
- Lien sortant et entrant **de serveurs différents d'un même hébergeur** : la pénalité est moins forte que ci-dessus
- Lien sortant et entrant de serveurs totalement différents : **note maximum**.
- **La circulation des liens**. $A \Rightarrow B \Rightarrow A$: note faible voire pas de prise en compte. // $a \Rightarrow b \Rightarrow c \Rightarrow a$ met Google dans l'embarras. Note maximum. Vous n'avez qu'un seul site / source. Créez vous un blog ! ou un sous domaine ou un site connexe avec une URL différente chez un autre hébergeur !
- **Les liens pourris** : un site est black listé pour une cause très sérieuse : grosse triche contre Google. Pas une maladresse, un réel cloaking massif doublé de spam sur l'index Google. Des liens qui pointent vers ce site partent de pages et de sites qui pourront avoir de mauvaises surprises. Bien que Google se défend de sanctionner des pages et sites liés à des black listés, différents écho et tests du petit monde du référencement montrent que Google sanctionne les "complices" de ces sites.
- **Les liens en or** : si vous connaissez le webmestre d'un site en .org, en .gov ou gouv et autres du même genre (site de mairie, de collectivité territoriale, ambassades etc.) et que vous pouvez en obtenir un BL, alors vous voilà fort riche en PR et en tranquillité ! Google ne sanctionne jamais ces sites et il considère avec respect ceux qui sont pointés par eux.
- **Les liens échangés 2 à 2** : vous échangez vos liens avec un site copain. Depuis février 2007, ce type d'échange ne serait plus pris en compte par Google. En revanche, les échanges non reliés eux seront pris en compte. De A vers B et de C vers A, et rien entre B et C directement.
 - **Vu la difficulté à gérer ce type de troc** : il est recommandé de passer par un professionnel pour gérer cela et d'avoir plusieurs sites.
 - **Exemple** : Gestion des échanges sans interconnexion : Site K backlink vers site B ; site C backlink vers site K. Aucun lien direct entre B et site C. J'ai indemnisé le site C en louant un lien du site X vers C. Le lien X vers C m'est facturé directement. Un modèle en théorie des graphes ne lève aucun lien. Parfait.

III-E-4 - Combien cela coûte-t-il ?

III-E-4-a - En temps homme

Il faut contacter les sites intéressants, leur proposer un échange de liens nommés, argumenter etc. Cela prend du temps :

- Le temps brûlé à ces opérations n'est pas consacré à d'autres travaux.
- Beaucoup de temps pour obtenir un lien de qualité inconnue
- Un lien qu'il faut rendre (encore du temps) et suivre : réciprocité, pas de 404, suivi de l'évolution des contenus mais ne coûte rien.

Vous pouvez sous-traiter ce travail à une société spécialisée.

Vous pouvez louer à l'année des BL issus de pages en PR 0, 1, 2, 3, 4 ou 5, liens simples pas ou faiblement nommés ou issus d'une page dédiée (le plus efficace !). c'est une variante de la publicité ! Moins cher et si bien géré dans le temps et techniquement, bien plus efficace. !

III-E-4-b - Le site loueur de backlinks

Il tire des revenus du type publicitaire. Le prix de location d'un lien en dur est géré par l'offre et la demande.

Pour les raisons vues ci-dessus, les liens sont caractérisés par :

- La thématique de la page de départ, des paragraphes de cette page, de l'emplacement de la "boîte à liens" au sein de cette page. Plus c'est bas, moins cela vaut.
- Le PR ou page Rank : la petite barre verte dans la Google bar. Plus il est élevé, plus la page a de la valeur.

- La taille de la boîte à liens : 10 mots, 20 mots, 30 mots ; Plus elle est grande, plus elle sera furtive et plus elle sera efficace. Plus elle est grande, moins la page peut en supporter.

III-E-4-c - Le locataire de liens

Son site est boosté dans le top 30 voire plus haut. À lui de transformer un tel avantage en recrutement de clients, en ventes etc.

Il faut donc choisir ses mots clefs stratégiques avec soin.

III-F - Améliorer le positionnement - Les BL internes

Une fois le référencement passif fini, seule une stratégie de BL pourra améliorer votre positionnement.

Maillez vos pages de liens internes.

Sur une page vous discutez des bijoux en général dont ceux en jade. Transformer ce mot "jade" en "bijoux en Jade de Chine" lien pointant vers la page du Jade sur votre site Le maillage interne paie un peu. C'est mieux que rien et surtout très facile à faire.

Un conseil : ne pas en abuser et bien rester dans la cohérence contextuelle ! Google n'est pas idiot !

III-G - Améliorer son Page Rank : le Web marketing relationnel

La méthode clef pour augmenter la composante PR consiste à disposer de liens venant de sites externes. Comment en obtenir ?

III-G-1 - Troc de liens entre sites ou entre sources

Un troc de lien signifie : un lien relie A à B, et un lien relie C à A, le troc porte sur ce que reçoit C grand perdant de l'affaire car il n'y a aucun lien entre B et C.

III-G-1-a - Lien échangés

Situés sur des serveurs différents (physiquement), voire dans des lieux géographiquement distants

- Annuaire acceptant de poser un backlink
- Offre d'emploi sur des sites visibles par tous
- BLOG et plus spécialement ceux de TYPEPAD
 - Cette technologie de BLOG permet d'automatiser des échanges de textes et de liens entre BLOG et donc de multiplier très vite les backlink vers un site.
 - Piste : faites un blog personnel disposant d'un réel contenu et disposer ainsi de liens multiples et indirects vers votre blog et le site.
- Croiser des textes et liens entre BLOG ou avec des blog
- ATTENTION : La thématique de site doit être respectée (centrale ou connexe), sinon c'est inutile.

III-G-1-b - Quelquefois, les liens disposés dans des messages de forum

Google est sensé ignorer les liens sortant de forums mais on en voit régulièrement qui sont pris en compte. Manifestement, Google accepte les liens de certains forums et ignore ceux d'autres forums.

⚠ Piège : Certains forums génèrent automatiquement une balise **rel="nofollow"** dans le lien : ``. Cela indique à Google (et aux autres moteurs) : ignore l'existence de ce lien, indexe juste le texte du lien mais ne suis pas ce lien. De tels liens ne génèrent aucun BL au sens Google par définition. Mais ils peuvent générer du trafic d'internautes humains, bien sûr.

Disposer d'un trafic important. Si vous mettez à disposition du contenu intéressant, le trafic va augmenter, des sites vont parler de votre URL et le PR de votre home page va monter. Vous pouvez signer avec votre URL de home page des documents mis à dispositions pour partage, mise en ligne etc. Vous diffusez ainsi du contenu avec un lien vers votre URL ce qui fait monter votre PR à chaque rencontre de ce type de lien avec un Google bot.

III-G-1-c - Disposer de pub (payante)

À condition de respecter la pertinence de liens (thématique centrale, thématique connexe) - sinon, aucun intérêt, Google ignore ces liens ou pire, les note négativement. Il doit y avoir un BL en dur (pas de javascript - Google l'ignorera)

III-G-2 - Conséquences

L'effet de levier est non négligeable (un lien transporte en moyenne un % de PR d'une page à PR de 1) :

Nombre de liens externes entrants sur une base de log oscillant entre 4 et 8 :	PR Google Bar correspondant
De 1 à quelques uns	PR1
Une quinzaine	PR2
Moins de 70	PR3
Moins de 400	PR4
Moins de 2500	PR5
Moins de 15000	PR6
<i>etc.</i>	

III-G-3 - Mesurer le retour sur investissement

III-G-3-a - Rythme de passage de Google bot sur votre site

Noter le rythme de passage de Google bot sur votre site et mesurer son éventuelle accélération en mettant en face vos travaux de référencement.

- Date 1 - Page Rank, positionnement d'une liste de combinaisons de mots clefs
 - Travaux de référencement effectués depuis date 1, mises à jours incluses
 - Statistiques de fréquentation via mots clefs de moteur de recherche
 - Courbes de votre activité commerciale
- Date 2 - Page Rank, positionnement de la liste de combinaisons de mots clefs
 - Travaux de référencement effectués depuis date 2, mises à jours incluses
 - Statistiques de fréquentation via mots clefs de moteur de recherche
 - Courbes de votre activité commerciale
- *etc.*

Aligner ensuite sur le même chronogramme la fréquentation de votre site (depuis vos log et autres stats)

Aligner vos autres actions de marketing : emailing, envoi postal, bannières pub dans le même chronogramme.

Caler enfin vos CA et autres sources de revenus liées au WEB (fréquentation magasin(s), commande(s) en ligne, date d'envoi des coupons de réduction...)

Vous aurez alors un tableau de bord pointant la rentabilité de chaque action ou chaque groupe d'actions.

III-G-3-b - Gérer les pages trop longues en texte

Rien n'y fait, vous avez une page de plusieurs centaines de mots, il n'est pas possible de la scinder en plusieurs pages via une navigation ergonomique.

Concentrer les mots clefs, un résumé stratégique de 100 mots par exemple, en début de cette page.

Vous scindez cette page en deux - un lien unique relie la partie 2 du texte à la partie 1. Exemple : 1 2 **La suite** informe l'internaute de cliquer sur 2 ou sur la suite pour continuer. La page s'ouvre en Open Window (une nouvelle fenêtre, pop up) Le début de cette page peut commencer avec une balise demandant à Google de passer immédiatement son chemin.

Ainsi, la note de pertinence de la page ne sera pas affectée par la longueur du texte. La plus grande partie de celui-ci est regroupée sur une page dont on sacrifie le page rank via un refus d'indexation ou tout simplement en la laissant s'indexer comme elle peut.

Comment empêcher Google de lire une page ?

- La meta <norobot> informe Google de ne pas indexer cette page (robots.txt) ;
- La gestion de : 1 2 La suite est en JavaScript, cela bloque Google.

Pourquoi demander à Google de ne pas indexer une page ?

- Pour ne pas avoir une page techniquement impropre à un référencement de qualité indexée par Google
- **Pour ne pas avoir des informations confidentielles mises en ligne par Google** qui a réussi à lire une page mal protégée par le firewall et autres systèmes de sécurité (très courant !) Une des principales sources de piratage ! Découvrir des sites Intranet privés reliés au site Internet via des liens cachés mais visibles à Google qui remonte le fil et indexe des sites... confidentiels !

III-H - Les outils

III-H-1 - Google Analytics - Survol

III-H-1-a - Mise en œuvre

S'inscrire, déposer sur toutes les pages à traquer un script Javascript en bas de page.

III-H-1-b - Informations clefs liées au référencement uniquement

Analytics couvre de nombreux points. Je me focaliserai sur ceux ayant un rapport avec le référencement naturel. Je ne m'occuperai pas de la partie Adword par exemple.

Sources de trafic > mots clés > afficher le rapport complet

	Visites	Pages par visite ↓	Temps moyen passé sur le site	Nouvelles visites (en %)
biométrie	2	11,00	00:03:17	50,00 %
	1	9,00	00:00:00	100,00 %
vi des activités	1	9,00	00:07:46	100,00 %
liser	3	9,00	00:04:34	66,67 %
oces autonome	1	9,00	00:01:49	100,00 %
ométriques contrôle d'accès	1	9,00	00:16:42	100,00 %
biométrique	1	8,00	00:01:54	100,00 %
0	1	8,00	00:03:50	100,00 %
mps de travail+logiciel	1	8,00	00:03:00	100,00 %
ation des carrières	1	8,00	00:03:31	100,00 %

Ces mots clefs amènent un faible trafic mais hautement ciblé ou presque.

Captor biométrie ne vaut pas grand chose. La personne cherchait l'entreprise. Donc le référencement, hors un fonctionnement annuaire à la "page jaune" n'a rien fait.

Les autres expressions clefs indiquent peu de visiteurs mais beaucoup de pages lues => le contenu du site suscite un fort intérêt donc prospect potentiel.

	Pages par visite	Temps moyen passé sur le site	Nouvelles visites (en %)	Taux de rebond
en % : 42,37 %	1,84 Moyenne du site : 1,82 (0,72 %)	00:01:14 Moyenne du site : 00:01:20 (-7,50 %)	85,09 % Moyenne du site : 86,13 % (-1,21 %)	67,70 % Moyenne du site : 65,41 % (3,29 %)
	Visites ↓	Pages par visite	Temps moyen passé sur le site	Nouvelles visites (en %)
	268	2,78	00:01:55	63,06 %
des ressources humaines	65	1,35	00:00:57	92,31 %
le badge	33	1,55	00:00:55	96,97 %
at	31	1,71	00:02:19	90,32 %
du temps	24	1,29	00:00:10	91,67 %
us rh	24	1,04	00:00:00	79,17 %
rance	23	3,00	00:02:10	60,87 %
les heures	23	1,35	00:00:38	82,61 %
du temps contrôle	17	1,00	00:00:00	5,88 %
n des ressources humaines	13	1,00	00:00:00	92,31 %
	11	1,18	00:00:05	100,00 %
le temps	11	1,45	00:00:49	63,64 %
badge	10	1,50	00:01:07	100,00 %
des temps	9	2,00	00:05:13	55,56 %
	9	1,44	00:00:15	77,78 %
le gestion du temps	9	1,22	00:00:08	88,89 %
du temps de travail	8	1,12	00:00:24	87,50 %
suivi activité	8	1,75	00:00:14	100,00 %

Ici on voit que le mot clef "gestion temps" sous différentes formes collecte du trafic ciblé : de nombreuses visites mais un ratio assez faible en pages / visite.

Cela signifie que de nombreux internautes ne se sentent pas concernés par la vocation du site :

- Le site s'exprime mal
- L'internaute ne recherche pas ce type de gestion des temps.

Dans les 2 cas, il faut creuser pour voir si on n'a rien oublié.

"Gestion du temps - contrôle" : 17 visites, 1,00 indique un échec total. Là il faut creuser le pourquoi.

"Gestion des temps" avec - 9 visites et 2 pages vues en moyenne + 5 mn de temps de visite moyen - est plus intéressant. Cela indique, courbe de gauss oblige, que certains visiteurs ont creusé le sujet sur le site => prospect potentiel.

III-H-1-c - Interprétation

Ces tableaux contiennent des statistiques clefs pour comprendre ce qui se passe dans le site vis à vis du référencement.

Les mots ayant un fort taux de pages vues / visite sont à étudier de près. Ils génèrent du trafic de qualité.

Les mots ayant un taux de 1 page vue / visite indique une totale inadéquation du contenu face aux mots clefs choisis et utilisés par des internautes. Là aussi, il faut creuser pourquoi.

Le nombre de pages vues par visiteur est élevé mais aucun formulaire de prise de contact n'est saisi.

Il manque quelque chose au site :

- Un formulaire contact pas assez visible ;
- Un contenu qui inciterait le visiteur à finaliser une décision de prise de contact : exemples de devis, exemples de prix, liste de ce qui est disponible le 1er de chaque mois etc. Une information très concrète lui permettant de caler sa demande face à un budget.

III-H-2 - Google Webmestre

III-H-2-a - Mise en œuvre

Elle est simple. Il suffit de s'inscrire chez Google, de suivre les instructions. L'ajout d'une meta dans la page index permet de se faire authentifier comme possesseur du site.

III-H-2-b - Utilisation

Ce panel d'outils est fort simple et pratique :

- Diagnostic d'incidents sur le site (lien cassé, page en 404 etc.)

III-H-2-c - Interprétation

All searches: All locations

<u>Top search queries</u>	<u>Average top position</u>
1. adresse paris	8
2. "www pages" "jaunes fr"	8
3. abts	5
4. domiciliation	18
5. pages "jaunes fr"	19
6. domiciliation fiscale	5
7. domiciliation sociale	2
8. domiciliation paris	12
9. domiciliation commerciale	16
10. secretariat telephonique	7
11. paris adresse	6
12. adresse fiscale	3
13. secretariat a distance	5
14. decret n "85 1280" du "05 12 1985"	10
15. paris adresses	4

<u>Top search query clicks</u>	<u>Average top position</u>
1. abts	1
2. adresse de domiciliation à paris	2
3. centre des impots paris 75008	2
4. impot paris 75008	2
5. reunica adresse	2
6. réunica adresse	2
7. adresse fiscale	3
8. domiciliations commerciales paris	5
9. domiciliation chez le gerant	6
10. centres "d affaires" paris	11
11. domiciliation	18

[Download this table](#)

Colonne de gauche, le "ratio" mots saisis dans Google, mots clefs "calculés" dans vos pages, position moyenne selon Google API (c'est très moyen en effet ! mais bon cela donne une tendance).

Colonne de droite les requêtes ayant entraîné un maximum de clics vers le site.

III-H-3 - Google AdWords

III-H-3-a - Introduction

L'objet n'est pas de donner un cours de Adword, ce serait hors sujet, mais de savoir dépenser 65 € habilement pour en déduire, interpréter de multiples informations clefs pour le référencement naturel

III-H-3-b - Mise en œuvre

Il faut s'inscrire chez Google Adword. Prévoir une CB internationale, le nom de domaine.

Une fois inscrit, on peut lancer une campagne de mots clefs.

III-H-3-c - Utilisation

Le but est de disposer de statistiques plus précises sur l'usage de mots clefs donnés dans une zone géographique précise (France, Europe, français, anglais etc. Tout est combinable)

Pour chaque expression clef, Google donne le nombre de fois où votre annonce a été affichée (impr). Donc un impr = une requête sur ce mot clef si vous êtes affiché en page #1 . Pour cela, où le mot clef est peu disputé et vos petites

enchères vous permettent d'être dans les 11 premiers annonceurs donc en page 1 (il y a maximum 11 annonces adword / page)

III-H-3-d - Interprétation

1 juil. 07 à 10 juil. 07 [Modifier la période](#)

+ Ajouter des mots clés : [Ajout rapide](#) | [Générateur de mots clés](#) | [Modifier des mots clés](#) | [Rechercher dans cette liste](#) | [Personnaliser les colonnes](#)

Suspendre | Annuler la mise en veille | Supprimer | Modifier les paramètres des mots clés

1 - 12 sur 12 mots

Mot clé	État	Enchère actuelle CPC max	Clics	Impr. ▼	CTR	CPC moyen	Coût	Pos.
Total	Activé	Par défaut €0,08 [Modifier]	1	872	0,11%	€0,08	€0,08	
<input type="checkbox"/> "robien recentré"	Actif	€0,08	0	478	0,00%	-	-	
<input type="checkbox"/> défiscalisation	Actif	€0,08	1	303	0,33%	€0,08	€0,08	
<input type="checkbox"/> "borloo populaire"	Actif	€0,08	0	22	0,00%	-	-	
<input type="checkbox"/> exemple robien	Actif	€0,08	0	17	0,00%	-	-	
<input type="checkbox"/> défiscalisation investissement immobilier	Actif	€0,08	0	14	0,00%	-	-	
<input type="checkbox"/> défiscalisation de robien	Actif	€0,08	0	13	0,00%	-	-	
<input type="checkbox"/> robien borloo	Actif	€0,08	0	12	0,00%	-	-	
<input type="checkbox"/> exemple borloo	Actif	€0,08	0	7	0,00%	-	-	
<input type="checkbox"/> rentabilite investissement locatif	Actif	€0,08	0	4	0,00%	-	-	
<input type="checkbox"/> simulation investissement locatif	Actif	€0,08	0	2	0,00%	-	-	
<input type="checkbox"/> de robien immobilier locatif	Actif	€0,08	0	0	-	-	-	
<input type="checkbox"/> -gilles de robien								

478 impressions avec une position moyenne de 11,5 c'est à dire page 1 ou page 2 d'une requête

- ⇒ on est dans les 600 requêtes sur cette période de 10 jours
- ⇒ soit un potentiel de 2000 visiteurs / mois environ sur cette expression clef en référencement naturel.

Défiscalisation : 303 impr avec une position moyenne de 42,3 ! soit page 4 ou 5 !

- ⇒ des milliers (2 à 3000 ?) affichages en 10 jours.
- ⇒ C'est le mot clef à avoir ! sauf Robien recentré, tous les autres pèsent moins

Deux choix marketing possible.

- Proposer en défiscalisation du Robien / immobilier face aux assurances vie en se battant sur défiscalisation.
- Se battre face aux autre Robien etc.

Selon le mot clef retenu, le positionnement n'est pas du tout le même.

Dans un cas on vante Robien contre les produits financiers.

Dans un autre cas on vante « son » Robien face aux autres.

Donc deux campagnes distinctes à mener en Adword et en référencement naturel avec impact possible sur les pages du site.

III-H-4 - Google Sitemap

Auparavant, il fallait avoir un hébergement acceptant Python. Au début de SiteMap, un programme (un script fourni par Google) devait être installé sur le serveur. Désormais, il existe des développements indépendants en PHP ou autre langage générant un SiteMap de site Internet.

Pour exécuter le script en langage Python de Google, il faut avoir des droits par le propriétaire du serveur (cas des serveurs WEB mutualisés).

 **Tutoriel :**  *Création et optimisation d'un google Sitemap, par Brice Franzoia.*

III-I - Informatiser et automatiser le référencement : niveau 1, le plus simple

III-I-1 - Chaque page

Chaque page possède des caractéristiques qui doivent être uniques ou personnalisées :

- TITLE
- Description
- Son contenu texte (pas de duplicate content)
 - Les titres entre balises h
 - Les titres en
 - Ils sont destinés aux humains. En span, Google n'y prête aucune attention mais le span permet de mettre en valeur aux yeux des humains ;
 - On peut ainsi avoir un jeu de mots comme titre *via* une mise en valeur par le tout doublé d'un titre contenant des mots clefs, discret visuellement mais en balise H1.
 - Les mots en
- Les balises ALT et TITLE de chaque visuel y compris celui de la charte graphique.
- La liste des mots clefs stratégiques qui sont affectés à cette page
- La liste des mots clefs basiques à combiner à cette page

 *Chaque page HTML, dynamique ou statique, doit avoir un TITLE unique et un meta DESCRIPTION unique.*

Le non respect de cette clause a plusieurs conséquences négatives sur le référencement dans Google :

- Perte de mots clefs significatif. Google sélectionne quelques pages du site et ignore plus ou moins les autres. Pour lui elles sont répétitives.
- Perte de poids des mots clefs situés dans le corps de page.
- Au niveau de compétition où se situe le site, cela signifie la non présence d'une quelconque page du site, sur quasiment tous les mots retenus dans le TOP 30 voire même le TOP100. En clair, 0 trafic ou presque.

Pour les pages statiques, le webmestre les mettra à jour via un CMS ou Dreamweaver ou autre éditeur HTML du même genre.

Pour les pages dynamiques, plusieurs possibilités :

- Directement en base de données via PHP MyAdmin ;
- *Via* un programme de mise à jour des données ;
- *Via* un import depuis un fichier .csv ;
- *Via* un fichier .ini ;
- *Via* des données gérées en table dans MySQL.

Il reste à diffuser ces données dans la page dynamique à construire à la volée.

III-I-1-a - Un passage d'argument - Cela donne quelque chose du genre

Le passage d'argument permet de lire une table (BDD) puis de lancer la construction de la page depuis l'extérieur de la page.

```
require_once("php_inc/MiseEnPage.inc.php");
$title_page = "Prestalya Immobilier - Achat, Vente, ...";
$meta_description_content = "Immo-devis est le spécialiste de ...";
$map = MiseEnPage::haut($title_page, $meta_description_content, "./");
```

La fonction "haut" construit le haut de toute page dynamique.

Il faut penser à y ajouter d'autres arguments possibles :

- Le contenu des balises ALT TITLE du ou des visuels constituant le haut de la page. On peut ainsi avoir des textes en balises ALT personnalisé à chaque page

III-I-1-b - Des descriptions title etc. en table MySQL

Dans cette approche, la page qui se construit va chercher elle même ses propres données.

Un exemple pour illustrer le propos : un produit, disons une table de jardin blanche, en plastique etc.

En table, ce produit a un champ le décrivant. Ce row devrait avoir d'autres champs : title de la page qui affichera le détail du produit (drilldown donc), description de cette même page, ALT TITLE du ou des visuels composant le haut de page. DESCRIPTION "table jardin d'été Belle Table de Dupont Plastique et fils en promotion - ref: 123658B" etc.

III-I-2 - Liste des éléments à automatiser

L'objectif est de constituer un arbre d'informations. Pour chaque page on sait ce qui lui est affecté comme mise en page dans le référencement, ses contraintes en cohérence. On peut donc en déduire ou automatiser nombre d'éléments à afficher dans la page.

- Titres de rubriques
- Mise en page des titres de rubriques : un visuel doté d'une balise ALT renseignée ou un texte directement
- Choix des URL et des noms de fichiers PHP ou HTML
- Écriture des titres et chapeaux
 - 1 Pour Google
 - 2 Pour les humains surfant sur le site.
- Liste des mots clefs affectés à cette page

III-J - Informatiser et automatiser le référencement : niveau 2, expert

Le positionnement dépend de la qualité du référencement dans l'espace et dans le temps.

Les pages doivent vivre et évoluer, en partie ou certaines d'entre elles.

Les meta, titres H, title etc aussi, de manière cohérente

- Un outil de suivi capable de lister les mots clefs :
 - Page de départ
 - Paragraphe, son titre, la page entière
 - Site de la page de départ.
 - Page arrivée si elle existe toujours

- Site arrivé
- Véhiculés par le lien
- Une factorisation indiquant ce qui sera prioritairement pris en compte
- Il faut un outil affichant les classes C et n° IP des sites de départ et arrivée
- Une recherche WHOIS pour chaque site.
- Les dates de mise à jour des 2 fichier
- Identifier les différences textuelles des versions précédentes
- Historique des dates
- Corrélation avec le positionnement sur liste de mots clef

Les textes doivent respecter certains critères :

- Qualité de l'orthographe, de la grammaire etc. Les robots logiciels produisant du contenu spam ne sont pas doués en orthographe / grammaire et génèrent régulièrement des fautes.
- L'étendue du vocabulaire. Plus il est étendu, plus il y a de chances que ce soit un humain qui l'ai conçu et donc une vraie page digne d'intérêt..
- => Il est clef de mesurer ces points, de les historiser et de les corrélés par rapport au positionnement dans google sur une liste de mot clef (la même que pour les autres points de mesure serait bien)

Les erreurs 404 sont très mal vues par les moteurs et tout rapatrier via un .htaccess vers la page d'accueil est insuffisant.

Il faut déclencher une alarme sur un accès à une page du type 404.html :

- Il faut tracer la disponibilité des pages du site et historiser incident et durée
- Une corrélation du positionnement dans Google sur une liste de mots clef

Un outil permettant de déterminer l'origine et les éventuels mots clefs

Date création des liens et de leurs évolutions : départ, arrivée, lien

Par des algorithmes de théorie des graphes, traquer les boucles à 1, 2, 3 et plus niveaux : on doit obtenir une corde serrée de liens sans quasiment aucun recoupement entre les brins.

En cas de sites issus d'un outil d'édition identique, les noms des paramètres, fichiers etc. doivent différer afin de ne laisser aucun footprint discernable : Un outil traquera les noms des éléments accessibles à un bot afin d'identifier les noms identiques dénotant une conception commune.

Le trafic créé doit être mesuré et corrélé dans le temps face aux actions de [web] marketing et de coding.

On obtient ainsi un tableau de bord affichant les évolutions du site, le suivi du référencement et des actions de positionnement en référence chronologiquement à :

- Des résultats (trafic, qualité du trafic...)
- Un positionnement dans Google

IV - Approche résumée et simplifiée

Référencer un site passe par un bon coding et marketing ASSOCIÉS.

- Le site doit avoir du contenu : voir son créateur et les notes précédentes
- Le site doit être organisé en thèmes (une arborescence compréhensible par le commun des mortels)
- Les mots clefs doivent être sélectionnés en 2 lots : ceux utilisés par les internautes pour vous trouver et ceux que vous voulez absolument, pour vos raisons à vous, employer sur votre site
- Les mots clefs pour Google doivent être entre balises H, en title, en meta, dans le texte. Les visuels doivent avoir une balise ALT ou TITLE cohérente

- C'est un travail mixant technicité sur XHTML, PHP, MySQL, rédactionnel, maîtrise de Google de niveau moyen. Ce n'est pas compliqué à faire jusque là. Cela prend beaucoup de temps quand on n'a pas l'entraînement ou les outils capables de booster la productivité.

Reste à positionner le site. Là cela se complique et il y a besoin d'un vrai professionnel.

Il faut trouver des liens de qualité, sans retour, pointant vers les pages Internet du site Client. Ces liens doivent véhiculer des mots clefs **cohérents** avec chaque page pointée et depuis chaque environnement de départ.

C'est simple à écrire et à lire. C'est incroyablement compliqué à mettre en œuvre ! Vous allez comprendre pourquoi en lisant la suite.

Les liens doivent être techniquement parfaits bien sûr.

Pour plein de raisons liées aux algorithmes de Google, vous ne pouvez pas mettre plus de 1 ou 2 liens en dur quittant un site pour aller vers le vôtre quand on est un néophyte sur Google. Quand on maîtrise, on peut adroitement pousser jusqu'à une petite dizaine dans des conditions extrêmement précises à respecter. Pas plus.

Sinon, on rend les sites éligibles à un potentiel trop élevé en liens et les algorithmes de Google commencent à faire des choses gênantes comme annuler les liens. Et là on perd la maîtrise de la situation.

Un bon référencier peut influencer des algorithmes Google "courants". Mais influencer adroitement des algorithmes intervenant exceptionnellement devient une opération difficile qu'il vaut mieux éviter.

Donc il faut de nombreux sites fournissant des liens en dur, sans retour, des liens de qualité.

Donc il faut plein de négociations avec des webmestres.

Et quasiment tous les webmestres refusent de négocier de telles pratiques pour plein de bonnes raisons.

Vous êtes donc bredouille sauf à être extrêmement débrouillard et à y consacrer un temps quasiment infini. Ce qui n'est pas très compétitif comme offre commerciale !

Sélectionner un prestataire de référencement / positionnement passe par sélectionner quelqu'un ayant :

- Un savoir faire en référencement côté technique (XHTML, CSS, Google writing...)
- Des contrats lui permettant de gérer habilement la pose cadencée de multiples liens (backlinks) vers le site à référencer

Ce métier de positionnement est déjà très pratiqué aux USA où il y a 20 millions de pages minimum pour le moindre mot clef.

En France il y a une moyenne de 1 million de pages pour tout mot clef intéressant, avec des pointes à 4 / 6 millions.

Donc le positionnement en France est moins crucial pour le moment. Pour une PME ayant son site et ayant absolument besoin d'être visible, cela devient un enjeu clef.